



CARF Luzern 2024

Controlling. Accounting. Risiko. Finanzen.

Stefan Behringer (Hrsg.)

Impressum:

Herausgegeben von
Stefan Behringer

Copyright
©2024 Verlag IFZ – Hochschule Luzern

Verlag:
Verlag IFZ – Hochschule Luzern
Suurstoffi 1
CH-6343 Rotkreuz
www.hslu.ch/ifz

ISBN: 978-3-907379-39-4

CARF Luzern 2024

Controlling. Accounting. Risiko. Finanzen.

Stefan Behringer

Inhaltsverzeichnis

Vorwort und Danksagung	9
Konferenzleitung und Programmkomitee	13
Controlling.....	19
<i>Machine Learning und Big Data im «Management Accounting Ecosystem» - Impulse für die Budgetierungspraxis.....</i>	<i>21</i>
Isabel Lausberg, Arne Eimuth	
<i>An overview of instruments in modern project management and controlling.....</i>	<i>27</i>
Patrick Ulrich, Benjamin Nothacker	
<i>Komplexitätsreduktion im Nachhaltigkeitsreporting durch richtiges Dashboard Design</i>	<i>37</i>
Lisa Perkhofer, Christoph Eisl, Johannes Teufl, Peter Hofer	
<i>Einsatz von Generativer KI in Controlling und Finance: Bedarf nach einer AI Governance?</i>	<i>65</i>
Thomas Rautenstrauch, Anja Lindenmann	
<i>Operative Planung und Budgetierung in Familienunternehmen: Eine empirische Analyse der Professionalisierung und Anwendung</i>	<i>71</i>
Robert Rieg, Patrick Ulrich	
<i>Digitale Technologien in Planung, Budgetierung und Forecasting – es kommt (nicht) auf die Größe an?.....</i>	<i>77</i>
Patrick Ulrich, Robert Rieg, Felix Stockert	
<i>Ergebnisse einer Dokumentenanalyse zur Identifikation möglicher Einflüsse der Corporate Sustainability Reporting Directive auf das Controlling</i>	<i>85</i>
Patrick Ulrich, Jasmina Metzger	
Accounting.....	101
<i>Wertschöpfungsverständnis in der Corporate Governance</i>	<i>103</i>
Björn Baltzer, Patrick Ulrich	
<i>Die Bilanzierung von Crypto Assets nach verschiedenen Rechnungslegungsnormen – Vergleich und mögliche Vorschläge zur Anpassung.....</i>	<i>107</i>
Wladislav Gawenko	
<i>Evaluating the Impact of AI Mention in Key Audit Matters on Stakeholder Perceptions: An Experimental Study on Decision Usefulness and Trust</i>	<i>113</i>
Theresa Herrmann	
<i>Digitalisierung als Antwort auf den Fachkräftemangel? Sich ändernde Prozesse und Anforderungen in Rechnungslegung und Wirtschaftsprüfung und ihre Auswirkungen auf die Hochschulausbildung.....</i>	<i>119</i>
Tim Kampe	
<i>Berichterstattung über immaterielle Ressourcen – Umfang und Grenzen der neuen Berichtspflicht.....</i>	<i>133</i>

Robert Nothhelfer	
<i>The effects of auditor rotation and auditor tenure on audit quality: a meta-analysis</i>	143
Markus Widmann, Nicolas Flohr, Matthias Wolz	
Risk Management	165
<i>Corporate Social Responsibility: Determinanten, Motivation und Implementierung</i>	167
Nicola Alves-Illi	
<i>Ganzheitliches Risikomanagement und Nachhaltigkeit in universitären Forschungsprojekten: Ein integraler Ansatz zur Maximierung des gesellschaftlichen Mehrwerts</i>	175
Christina Angela Groß, Christian Georg Zimmermann, Christian Trapp, Philip Sander	
<i>Integriertes Risikomanagement für Alters- und Pflegeheime</i>	181
Wilfried Lux, Sebastian Müller, Kevin Bieber	
<i>The risk trilemma of how to hedge gas-fired power plants in times of war</i>	199
Cedric Schneider, Michael Römmich, Andreas Heinen	
<i>Einfluss weiblicher Nachfolge auf Risiken in mittelständischen Familienunternehmen</i>	213
Patrick Ulrich, Natalie Sorg	
Finanzen	219
<i>Sustainability Reporting - A Global Analysis of Sustainability Reporting and its Impact on Cost of Capital - A Bibliographic and Content Analysis</i>	221
Guido Sopp, Björn Bunzel	
<i>Optimierungsmöglichkeiten im Finanzbereich: Einsatz von ChatGPT & Co. und Prompt Engineering – Nutzungsverhalten, Anwendungsbereiche & Herausforderungen</i>	245
Sonia Frenkenberger, Susanne Leitner-Hanetseder	
<i>The influence of culture on the capital structure of SME – A pan European quantitative research</i>	261
Markus Wienbreyer, Carola Spiecker-Lampe, Maria Iborra	
<i>Interne Ratings – Einordnung und Ableitung weicher Rating-Faktoren</i>	281
Matthias Kretschmar, Kai Wiltinger	
Lehre	287
<i>Digifit Steuerlehre: Ein Pilotprojekt zur Umsetzung von Blended Learning Konzepten in Modulen mit den Schwerpunkten Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung</i>	289
Katharina Dillkötter, Melanie Frieling	
<i>Transfer anders gedacht – Entrepreneurship Education als Schul-Hochschul-Kooperation in der 3rd Mission</i>	311
Tim Kampe, Ingo Ballschmieter	
<i>Lern-App BWL Champion: Konzeption und Anwendung in der digitalen Lehre/Prüfung</i>	319
Jörn Littkemann, Kristopher Pantani	
<i>Digitalisierung in der Lehre – Projekt zur Neukonzeption von Prüfungsleistungen</i>	329
Janina Matern	

<i>Gamified financial education application for young adults</i>	345
Manuel Szecsenyi, Max Monauni	
<i>Inverted Classroom im Teilmodul Buchhaltung - aus Fehlern lernen -</i>	373
Andreas Stach, Carsten Theile	
<i>Fit für die Nachhaltigkeitstransformation? – Zur Integration von Nachhaltigkeitskompetenzen in der Controllinglehre</i>	395
Ute Vanini, Annika Jacobsen	
Institut für Finanzdienstleistungen Zug IFZ	411

Vorwort und Danksagung



Vorwort und Danksagung

Liebe Teilnehmerinnen, liebe Teilnehmer

Es freut uns ausserordentlich, Sie im Namen der Konferenzleitung bei der CARF Luzern 2024 begrüßen zu dürfen. In diesem Jahr ist es eine besondere CARF: Die Konferenz für anwendungsorientierte Forschung in den Bereichen Controlling, Accounting, Risk & Compliance und Finance feiert ihr 10jähriges Jubiläum. Wir wünschen Ihnen eine anregende, interessante und informative Konferenz.

Im Zentrum der CARF Luzern stehen die Themenfelder Controlling, Accounting, Risiko und Compliance sowie Finanzen. Die Konferenz bietet eine Plattform, um diese Themenfelder aus verschiedenen Perspektiven zu beleuchten und wertvolle Kontakte und Kooperationen zu initiieren oder zu vertiefen. Mit etwas Stolz möchten wir an dieser Stelle darauf hinweisen, dass aus CARF Luzern Konferenzen zahlreiche Forschungsprojekte und Publikationen entstanden sind. Eine Besonderheit unserer Konferenz besteht in der besonderen Würdigung der Hochschullehre in den Konferenzbereichen. Damit haben wir ein einmaliges Format, dass sich ausschliesslich der Lehre in der Finanzierung widmet.

In diesem Jahr dürfen wir Ihnen zahlreiche Beiträge aus den fünf Tracks Controlling, Accounting, Risiko, Finanzen und Lehre vorstellen. Als Keynote-Speaker bringen Dr. Franz Schencking, Geschäftsführer des forstwirtschaftlichen Zusammenschlusses «Die Märkischen», und Dr. Alan Hippe, CFO und CIO von Roche, spannende Einblicke in die theoretische und praktische Betriebswirtschaftslehre.

Die Organisation der Konferenz ist ein Grossprojekt über das ganze Jahr. Unser herzlicher Dank gilt all denen, die ihre Zeit, ihr Wissen und ihre Erfahrung für die CARF Luzern 2024 zur Verfügung gestellt haben. Er gilt allen Track-Verantwortlichen: Prof. Viviane Trachsel, Prof. Dr. Thomas Kurt Birrer, Prof. Dr. Marco Passardi, Prof. Dr. Philipp Lütolf, sowie Prof. Dr. Stefan Hunziker und allen Gutachtern: Prof. Dr. Ute Vanini, Prof. Dr. Gabrielle Wanzenried, Prof. Dr. Max Monauni, Prof. Dr. Sören Dressler, Prof. Dr. Thomas Rachfall und Prof. Dr. Patrick Ulrich.

Ein besonderer Dank gilt natürlich allen Konferenzteilnehmenden. Sie machen durch Ihre Teilnahme, Ihre Beiträge, Ihre Vorträge und Ihre Diskussionsbereitschaft, die CARF Luzern 2024 zu einem ganz besonderen Ereignis.

Wir danken den Konferenzorganisatoren Karolin Babilon, Jonas Miesner und Saskia Wiegandt für ihr Engagement bei der Organisation der CARF Luzern 2024 sowie bei der Erstellung des vorliegenden Konferenzbandes. Ohne ihren Einsatz wäre eine solche Konferenz nicht möglich.

Zum Schluss danken wir unseren Sponsoren, welche mit ihrer grosszügigen Unterstützung einen würdigen Rahmen für die Konferenz ermöglicht haben.

Wir freuen uns auf die nächste Ausgabe der CARF. Im Jahr 2025 (12. und 13.09.2025) findet die 11. CARF statt. Wir freuen uns darauf, Sie 2025 in Luzern begrüßen zu können.

Konferenzleiter und Leiter Programmkomitee



Prof. Dr. Stefan Behringer

September 2024

Konferenzleitung und Programmkomitee



Konferenzleitung und Programmkomitee



Prof. Dr. Stefan Behringer

Konferenzleiter und Leiter Programmkomitee
Verantwortlich für den Track Controlling
Review Controlling, Risk Management & Lehre

Hochschule Luzern – Wirtschaft
Institut für Finanzdienstleistungen Zug IFZ



Saskia Wiegandt

Administrative Konferenzbetreuerin

Hochschule Luzern – Wirtschaft
Institut für Finanzdienstleistungen Zug IFZ



Karolin Babilon

Konferenzorganisatorin

Hochschule Luzern – Wirtschaft
Institut für Finanzdienstleistungen Zug IFZ



Dr. Jonas Miesner

Konferenzorganisator

Hochschule Luzern – Wirtschaft
Institut für Finanzdienstleistungen Zug IFZ



Prof. Dr. Marco Passardi

Verantwortlich für den Track Accounting
Review Accounting

Hochschule Luzern – Wirtschaft
Institut für Finanzdienstleistungen Zug IFZ



Prof. Dr. Stefan Hunziker

Verantwortlich für den Track Risikomanagement
Review Risikomanagement

Hochschule Luzern – Wirtschaft
Institut für Finanzdienstleistungen Zug IFZ



Prof. Dr. Thomas Kurt Birrer

Verantwortlich für den Track Finance
Review Finance

Hochschule Luzern – Wirtschaft
Institut für Finanzdienstleistungen Zug IFZ



Prof. Viviane Trachsel

Verantwortlich für alle Tracks im Bereich Lehre
Review Lehre

Hochschule Luzern – Wirtschaft
Instituts für Finanzdienstleistungen Zug IFZ

Konferenzbeiträge

2024



1

Controlling



Machine Learning und Big Data im «Management Accounting Ecosystem» - Impulse für die Budgetierungspraxis

Extended Abstract

Prof. Dr. Isabel Lausberg

Hochschule Ruhr West, Wirtschaftsinstitut, Mülheim an der Ruhr, E-Mail: isabel.lausberg@hs-ruhrwest.de

Prof. Dr. Arne Eimuth

Hochschule Ruhr West, Wirtschaftsinstitut, Mülheim an der Ruhr, E-Mail: arne.eimuth@hs-ruhrwest.de

Abstract

In der Unternehmenspraxis wird zumeist noch klassisch budgetiert. Mit zunehmender Diffusion von Künstlicher Intelligenz (KI) und Maschinellem Lernen (ML) stellt sich die Frage, inwieweit Unternehmen diese Werkzeuge einsetzen können, um die Budgetierung zu vereinfachen und zu verbessern. In diesem Beitrag wird das Konzept des „Management Accounting Ecosystem“ (Akroyd et al., 2023) vorgestellt, dass ein Framework bietet, um externe Informationen und moderne Technologien in das Controlling zu integrieren. Big Data und Maschinelles Lernen spielen dabei eine Schlüsselrolle. ML bietet die effiziente Verarbeitung großer, komplexer Datensätze und ermöglicht die Erstellung von Prognosen und Handlungsempfehlungen. Der Einsatz von ML in der Budgetierung steht jedoch noch am Anfang und erfordert weitere Forschung, insbesondere hinsichtlich der Auswahl geeigneter Datengrundlagen und Algorithmen sowie der Integration in Unternehmenssysteme. Ein Paradigmenwechsel zu datengetriebenen Planungs- und Entscheidungsprozessen ist mittel- bis langfristig vorgezeichnet.

1 Problemstellung

Die Budgetierung spielt seit langem eine zentrale Rolle in der Unternehmenssteuerung, ist aber auch Gegenstand kritischer Diskussionen. So werden bspw. der hohe Ressourcenaufwand, eine unzureichende Integration der strategischen in die operative Planung, die Finanzorientierung, mangelnde Flexibilität sowie Akzeptanzprobleme und die Förderung unerwünschten Verhaltens als Kritikpunkte angeführt (Möller, 2021, S. 283f.; Gleich & Kopp, 2001, S. 430; Koch, Form & Priebe, 2020, S. 55).

Trotz dieser Kritik ist die Budgetierung in der Praxis nach wie vor in großer Breite vertreten und erfüllt als umfassendes Steuerungsinstrument wichtige Aufgaben in der Unternehmensführung. Vor dem Hintergrund der Digitalisierung sowie der genannten Kritik wurde die Budgetierung mit Konzepten wie Rolling Forecasts, IT-gestützten Simulationen und agilen Ansätzen weiterentwickelt. Es besteht dennoch weiterer Veränderungsdruck, um sich an die Bedingungen einer zunehmend digitalisierten und volatilen Umwelt anzupassen (Möller, 2021, S. 284ff.; Pedell, 2024, S. 387) und mit innovativen Ansätzen eine effizientere, flexiblere und aussagekräftigere Budgetierung zu entwickeln.

2 Das «Management Accounting Ecosystem» als Framework

Das Konzept des «Management Accounting Ecosystem» (Akroyd et al., 2023) bietet einen gedanklichen Rahmen, um die Forschung stärker auf das Unternehmensumfeld und externe Informationsquellen zu lenken, und damit die Performance von Unternehmen zu verbessern. Das Ecosystem “is a complex (and self-organized) network that includes an **information source layer** and an **information technology layer** (including computer science, statistics, math, machine learning, artificial intelligence, etc.) through which data and information move into the management accounting system where it is transformed into knowledge that is useful for decision-making.” Das Ecosystem integriert Daten- und Informationsbeschaffung, -zugriff und -fluss, um durch moderne Technologien eine effiziente Verbindung zwischen (externen) Informationsquellen und dem Management-Accounting-System zu schaffen (Akroyd et al., 2023, S. 1f.). Das Konzept knüpft an den Kompetenzrahmen des IMA® Institute of Management Accountants (2019) an, der die wesentlichen zukunftsgerichteten „knowledge, skills and abilities“ für Finanz- und Rechnungswesenprofessionals definiert.

3 Der Einbezug von Big Data und Maschinellern Lernen in der Budgetierung

3.1 (Interne und) externe Daten als «Information source layer» für die Budgetierung

Daten gelten als „Treibstoff“ der digitalen Wirtschaft (Chernyakov & Chernyakova, 2018). Geschwindigkeit, Vielfalt und Volumen der Daten aus unterschiedlichsten Quellen nehmen kontinuierlich zu. Unternehmen können die wachsende Datenflut (Big Data) u.a. durch die Cloud-Technologie, die weit verbreitete Verfügbarkeit von Open-Source-Analysebibliotheken sowie den sinkenden Hardwarekosten bei steigender Rechenleistung heute effektiv und effizient nutzen. Der Wert von Big Data liegt allerdings nicht in der Quantität, sondern in der Fähigkeit, geschäftsrelevante Informationen zu extrahieren und zu nutzen, wofür die Qualität der Daten entscheidend ist (Faccia et al. sowie die dort zitierten Quellen, 2022, S. 38f.).

In der Budgetierungspraxis liegen im Wesentlichen interne, strukturierte Daten als Informationsbasis zugrunde. Nach der Studie von Bley, Giesel & Ruhwedel (2020, S. 46-51) werden zwar externe Informationen zur Planung genutzt, dies sind aber vor allem Informationen von Branchenverbänden oder Marktforschungsinstituten, die i.d.R. nicht unter den Begriff Big Data fallen. Das Interesse an externen Prognosedaten, wie Wettbewerbsentwicklung, Wettbewerbsumfeld und Nachfragetrends, ist bei Controlling-Verantwortlichen groß. Echtzeitdaten könnten zudem die Prognosegenauigkeit erhöhen und unterjährige Strategieanpassungen erleichtern. Koch, Form & Priebe (2020, S. 54-60) kommen in einer Befragung von Controlling-Verantwortlichen zu dem Ergebnis, dass Big Data Analysen potentiell sehr nützlich für transparente und qualitativ hochwertige Budgets sind und erkennen darin einen Mehrwert

für die Budgetierung. Ferner wird für die Budgetierung „eine verbesserte Vorhersagegenauigkeit und eine nachvollziehbare Zielbildung“ erwartet.

3.2 Das Maschinelle Lernen als «Technology layer» für die Budgetierung

Maschinelles Lernen (ML) als ein Teilbereich der Künstlichen Intelligenz (KI) beschreibt die Fähigkeit von Computersystemen, aus Daten zu lernen und ihre Leistung selbstständig zu optimieren, ohne speziell dafür programmiert zu werden (z.B. Sarker, 2021, S. 1). KI und insbesondere ML sind Schlüsseltechnologien, um die wachsenden Datenmengen intelligent zu analysieren und die passenden smarten, automatisierten Anwendungen zu entwickeln. ML hat sich in den letzten Jahren im Kontext der Datenanalyse und Datenverarbeitung rasant entwickelt. Dabei unterscheidet man i.d.R. drei Arten von Lernen: supervised (überwachtes), unsupervised (unüberwachtes) und reinforcement (verstärkendes) Lernen, zusätzlich gibt es Mischformen und das Deep Learning, das mit künstlichen neuronalen Netzen arbeitet (z.B. Frochte, 2020, S. 20ff., Sarker, 2021, S. 1). Abbildung 1 zeigt die verschiedenen Lernalgorithmen mit den jeweiligen Analyseformen.

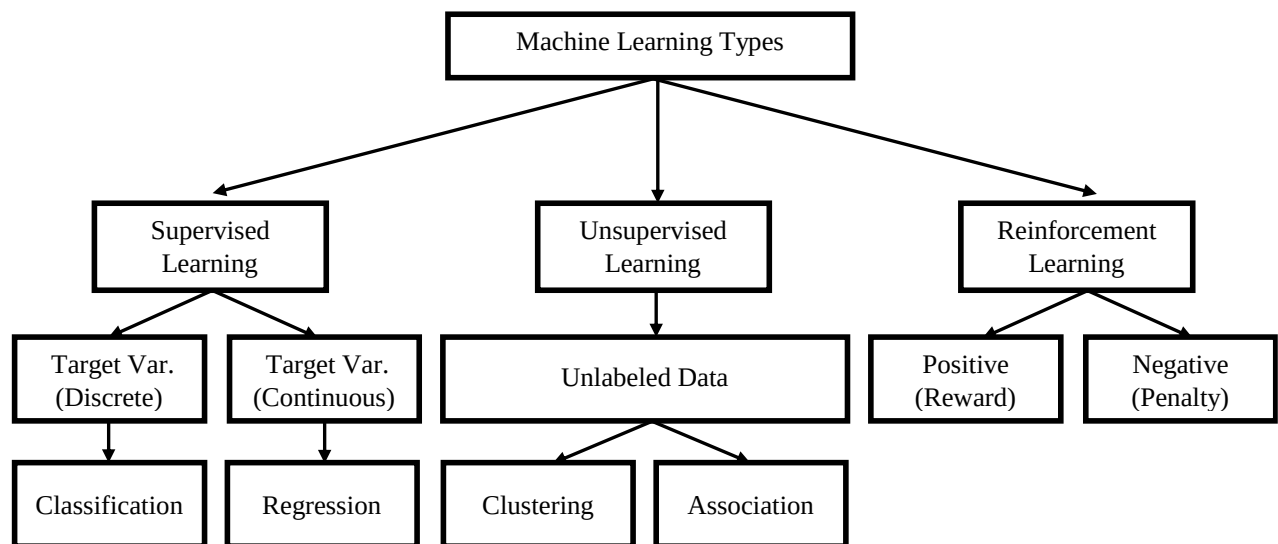


Abb. 1: Typen maschinellen Lernens, in Anlehnung an Sarker, 2021, S. 4

Der Vorteil von ML-Verfahren liegt in ihrer hohen Leistungsfähigkeit und Effizienz bei der Verarbeitung strukturierter und unstrukturierter Daten. Hierdurch können sie präzise Vorhersagen für spezifische Ergebnisvariablen treffen (Ghoddusi, Creamer & Rafizadeh, 2019, S. 2). Verbreitete Anwendungskategorien sind Predictive Analytics, Entscheidungsunterstützung, Bild-, Sprach- und Mustererkennung sowie Verarbeitung natürlicher Sprache und Sentiment Analysen (Sarker, 2021, S. 15ff.).

Die Herausforderungen bei der Nutzung von Machine-Learning-Algorithmen liegen zum einen in der Qualität der Daten und zum anderen in der Auswahl geeigneter Algorithmen für die jeweilige Anwendung. Erfolgreiche Modelle erfordern ein Training der Algorithmen mit realen Daten und anwendungsspezifisches Wissen, um Entscheidungen intelligent unterstützen zu können (Losbichler, 2020, S. 375; Sarker, 2021, S. 17f).

Während sich in der Literatur schon eine größere Anzahl von Anwendungen von ML im Financial Accounting findet (Kureljusic & Karger, 2023, geben einen Überblick), steht die Anwendung in der Budgetierung noch am Anfang und konzentriert sich (noch) auf Kostenprognosen mithilfe von Regressionen und künstlichen neuronalen Netzen (z.B. Tijanić, Car-Pušić. & Šperac, 2019, und Hashala & Andrews, 2023).

4 Ausblick und Forschungsbedarf

Der Einsatz Maschinellen Lernens in der Budgetierung ist eine Frage der Zeit. Dieser Schritt stellt nicht nur eine technologische Weiterentwicklung dar, sondern auch einen Paradigmenwechsel hin zu stärker datengetriebenen Planungs- und Entscheidungsprozessen im Unternehmen. Aus den Ausführungen zum Framework des «Management Accounting Ecosystem» lassen sich dahingehend u.a. folgende Forschungsfragen ableiten:

- Mit welchen externen Daten können Budgetierungen besser fundiert werden?
- Wie gelangen die Daten und Informationen (möglichst in Echtzeit) in die Rechnungswesensysteme der Unternehmen?
- Inwieweit kann ML bei der Analyse der Daten für die Budgetierung von Nutzen sein? Und welche Algorithmen sind dafür am besten geeignet?
- Wie sieht die zukünftige Rollenverteilung zwischen Mensch und Maschine aus?

Zur Beantwortung dieser Forschungsfragen können Use Cases sowie Befragungen von ML- und Controlling-Experten und Expertinnen wichtige Beiträge leisten.

Die technologischen Entwicklungen und potentiellen Anwendungsfälle in der Unternehmenssteuerung geben starke Impulse, um die entsprechenden Instrumente weiterzuentwickeln. So müssen sich Controllerinnen und Controller auch in der Budgetierung noch stärker mit der Unterstützung durch KI auseinandersetzen.

Literaturverzeichnis

- Akroyd, C.; Dow, K.E.; Drake, A.; Wong, J. (2023): "The New Management Accounting Ecosystem: A Retrospective View and Path to the Future", Akroyd, C. (Ed.) *Advances in Management Accounting*, Vol. 35, Emerald Publishing Limited, Leeds, pp. 1-18. <https://doi.org/10.1108/S1474-7871202335>.
- Bley, C.; Giesel, A.; Ruhwedel, F. (2020): Einsatz von Big Data und Predictive Analytics in der Unternehmensplanung – Ergebnisse einer Befragungsstudie. *Controlling*, 32(2), 45–52. <https://doi.org/10.17323/j.jcfr.2073-0438.12.4.2018.99-109>.
- Chernyakov, M.; Chernyakova, M. (2018): Technological risks of the digital economy. *Corporate Finance*, 12(4), pp. 99-109.
- Faccia, A.; Cavaliere, L. P. L.; Petratos, P.; Mosteanu, N. R. (2022): Unstructured over structured, big data analytics and applications in accounting and management. In *Proceedings of the 2022 6th International Conference on Cloud and Big Data Computing* (pp. 37-41). <https://doi.org/10.1145/3555962.3555969>.
- Frochte, J. (2020): *Maschinelles Lernen: Grundlagen und Algorithmen in Python*. Carl Hanser Verlag GmbH Co KG.
- Georgopoulos, A.; Georg, S. (2021): Anforderungen an das Controlling. Auswirkungen von Big Data und Digitalisierung auf das zukünftige Kompetenzprofil des Controllers. Springer. <https://doi.org/10.1007/978-3-658-34938-7>.
- Ghoddusi, H.; Creamer, G. G.; Rafizadeh, N. (2019): Machine learning in energy economics and finance: A review. *Energy Economics*, 81, 709-727. <https://doi.org/10.1016/j.eneco.2019.05.006>.
- Gleich, R.; Kopp, J. (2001): Ansätze zur Neugestaltung der Planung und Budgetierung. *Controlling*, 13(8), 9. <https://doi.org/10.15358/0935-0381-2001-8-9-429>.
- Gleich, R.; Kappes, M.; Leyk, J. (2019): *Planung, Budgetierung und Forecasting*, Haufe, Freiburg.
- Hashala, A.; Andrews, K. (2023): Improving Project Budgeting Systems by Developing Machine Learning Models. Fall Issue, 2023. <https://doi.org/10.47670/wuwijar202371amhka>.
- IMA® Institute of Management Accountants (Hrsg.) (2019): *IMA Management Accounting Competency Framework*, abrufbar unter: <https://eu.imanet.org/research-publications/statements-on-management-accounting/ima-management-accounting-competency-framework> (zuletzt abgerufen am 10.07.2024).
- Koch, S.; Form, S.; Priebe, D. (2020): Die Budgetierung im Kontext der Digitalisierung. Eine wissenschaftliche Untersuchung zu den Potentialen der Digitalisierung für die Budgetierung. *Controlling*, 32(4), 54–60. <https://doi.org/10.15358/0935-0381-2020-4-54>.
- Kureljusic, M.; Karger, E. (2023): Forecasting in financial accounting with artificial intelligence – A systematic literature review and future research agenda. *Journal of Applied Accounting Research*, (ahead-of-print). <https://doi.org/10.1108/JAAR-06-2022-0146>.
- Losbichler, H.; Lehner, O.M. (2021): "Limits of artificial intelligence in controlling and the ways forward: a call for future accounting research." *Journal of Applied Accounting Research* 22.2 (2021): 365-382. <https://doi.org/10.1108/JAAR-10-2020-0207>.
- Möller, K. (2021): Die Budgetierung zwischen Revolution, Degeneration und Evolution. *Die Unternehmung*. Vol. 75, No. 2 (2021), pp. 283-291. <https://doi.org/10.5771/0042-059X-2021-2-281>.
- Pedell, B.; Hofmann, C.; Hofmann, Y.E.; Küpper, H.-U.; Friedl, G.: *Controlling (2024)*, Schäffer-Poeschel, Pöhl, ISBN: 9783791053486.
- Rieg, R. (2015): *Planung und Budgetierung*, Springer: Wiesbaden.
- Sarker, I. H. (2021): Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2(3), 2:160. <https://doi.org/10.1007/s42979-021-00592-x>.
- Tijanić, K.; Car-Pušić, D.; Šperac, M. (2019): Cost estimation in road construction using artificial neural network. *Neural Computing and Applications*, 32, 9343-9355. <https://doi.org/10.1007/s00521-019-04443-y>.

An overview of instruments in modern project management and controlling

Extended Abstract

Prof. Dr. habil. Patrick Ulrich

Aalen University of Applied Sciences, Aalen Management Institute (AAUF), Beethovenstraße 1, 73430 Aalen, Germany, patrick.ulrich@hs-aalen.de

Benjamin Nothacker, M.A.

Aalen University of Applied Sciences, Aalen Management Institute (AAUF), Beethovenstraße 1, 73430 Aalen, Germany, benjamin.nothacker@hs-aalen.de

Abstract

The application of agile methodologies and instruments is becoming increasingly important in a business environment that must adapt to rapidly changing external factors. In practice, traditional and agile approaches to project management and project controlling often exist concurrently. This coexistence can give rise to uncertainty about which approaches to follow. Based on a systematic literature analysis, this article explores which instruments are used in modern project management and controlling.

1 Introduction

The capacity to swiftly and effectively adjust to shifting external circumstances is assuming increasing significance in both academic discourse and practical applications (Magistretti & Trabucchi, 2024). The contemporary business environment is distinguished by significantly higher volatility, uncertainty, and complexity compared to the 20th century (Zhang et al., 2023). While the term VUCA (volatility, uncertainty, complexity, ambiguity) is frequently employed, it is increasingly evident that this concept is inadequate for capturing the complexities of the contemporary reality.

The term BANI (brittle, anxious, non-linear, and incomprehensible) has been introduced to more adequately address these modern challenges (Cascio, 2020). BANI expands the scope of VUCA by incorporating psychological and systemic elements, emphasizing the fragility of seemingly robust systems and the anxiety this can cause. It highlights the difficulty in discerning cause-and-effect relationships in a non-linear world, leading to unpredictable outcomes, and addresses the challenge of comprehending a complexity that exceeds human understanding (Dieffenbacher, 2023; WU Executive Academy, 2022).

The speed of market change is increasing significantly, impeding the predictability of market impacts. These effects are increasingly manifesting in their unpredictability. As evidenced by the work of Bonk & Hedfeld (2019) and Becker et al. (2021), these developments have substantial implications for the design of project management and project controlling.

Similar to the choice of the right organizational form to adapt to changing environmental conditions, the selection of the appropriate project approach and project management methods is a crucial factor. It is imperative to respond flexibly to constantly changing conditions and challenges (Fernandez & Fernandez, 2008), as changing environmental conditions lead to uncertainty and risks for project teams.

The selection of an appropriate project approach and practices, as well as the integration of these with existing controlling instruments, can be challenging. This is particularly the case when these instruments are in conflict with one another. Traditional control instruments such as budgets, reporting, and management by objectives coexist with agile methods like Scrum, check-ins, and objectives and key results (OKRs). This coexistence can give rise to uncertainty among employees and open conflict about which approaches to follow, and raises the question of whether a combination is beneficial at all (Möller & Schmid, 2021).

The existing literature offers no theoretical, conceptual, or empirical evidence on the relationship between agile project management methods and the nature of project control in an agile corporate environment as Ulrich & Rieg, (2020) and Niederman et al. (2018) described.

At this preliminary stage, a precise allocation of the utilization of the employed instruments to a specific approach cannot yet be made. Consequently, this paper serves as an extended abstract. The initial step towards enabling such an allocation is to identify which instruments are being utilized, which forms the core content of this extended abstract. The research question of the paper, which has not yet been examined theoretically or empirically in this form, is as follows:

Which methods and instruments are used in which project approaches?

2 Agile, traditional and hybrid approaches

For decades, project management has been conducted using a collection of practices and methodologies collectively referred to as the traditional approach. The traditional approach is defined by linear and predictable project planning methodologies, with the objective of achieving a predefined set of outcomes (Gemino, 2021). (Špundak, 2014) builds upon prior definitions to suggest that the primary objective of the traditional project management approach

is the achievement of optimization and efficiency by meticulous adherence to the initial comprehensive project plan, with the objective of completing the project within the specified time, cost, and scope parameters.

In traditional project management, a "waterfall" process of planning and execution is typically employed, whereby the client provides a clear indication of the desired outcomes at the commencement of the project (Thesing et al., 2021). The project process is defined by the sequential processing of work packages within individual phases. The initialisation, definition, planning, implementation, and completion of the project represent pivotal phases of project work (Preußig, 2020). In relation to the established target values of the traditional approach, a comprehensive and detailed scope of services for the project is defined at the outset. This scope remains virtually unchanged throughout the project, with adjustments being made to time and resources as necessary. This is provided that the fulfilment of the project objectives requires it (Kuster et al., 2022).

Projects are becoming increasingly complex, with a greater number of tasks and intricate interrelations, in line with the general trend observed in business environments. The traditional project management approach is inadequate for capturing the complexity and dynamics of contemporary projects (Špundak, 2014). It is an inevitable consequence of undertaking any complex project that modifications to the original plan will be required due to the necessity of accommodating unforeseen and evolving circumstances within the project environment or the project itself (Collyer et al., 2010; Olsson, 2006).

The advent of new approaches is closely linked with the field of software engineering and software development. New project management approaches emerge alongside new approaches to software development. These new approaches have emerged with a multiplicity of names, each of which is intended to convey a sense of distinction from more established traditions (Špundak, 2014). The most frequently utilized term is the agile approach (Aguanno, 2005; Highsmith, 2004; Williams, 2005), although a similar concept and methodology can be identified under the labels of lean approach (Williams, 2005), extreme approach (DeCarlo, 2004; Wysocki, 2007) and adaptive approach (Shenhar & Dvir, 2007; Wysocki, 2007).

A key foundation of successful project management is the ability to adjust to change, dividing the work into distinct stages throughout the project (Gemino, 2021). An agile project management approach also depends upon early and continuous customer involvement. Thus, the iterative nature of agile allows for frequent stakeholder interaction, adjustments made on the fly, and re-scoping project requirements in light of new information or customer requests (Serrador & Pinto, 2015). In agile project management, the control parameters of budget and time are fixed. While the content and scope of projects and their ultimate results can be flexibly adapted to the conditions, project management remains a fixed process (Kuster et al., 2022).

A third approach is the hybrid project management. It is a methodology that integrates the fundamental principles of traditional project management with the agile approaches that are characteristic of the Agile Manifesto. The focus is on meeting the customer's requirements and deriving the maximum benefit (Känel, 2020). This approach enhances the plan-driven process model with agile principles by employing a waterfall model to plan the "big picture," while managing suitable sub-projects in an agile manner. Sub-projects employ agile methodologies, such as transparency and adaptability, while adhering to an overarching classical project structure (Thesing et al., 2021).

3 Findings

A systematic review of the literature was conducted for the analysis. These follows Tranfield et al. (2003) systematic review principles that are rigorous and structured approaches to synthesizing and structured approach to synthesizing research evidence. The selection of literature for the Systematic Literature Review (SLR) was based on specific criteria to ensure that the papers analyzed were relevant and up to date. These criteria include:

1. Relevance to the topic: Only books directly related to the topics of project management and project controlling were selected. This was ensured by examining the titles, abstracts, and tables of contents.

2. Reputation of the authors: Only works by authors and editors who are recognized in the professional world and have relevant publications and research experience in the field of project management and controlling were considered.

3. Recency: The books were considered to be relatively new to ensure that the latest developments and trends in the field of project management and project controlling were considered.

4. Citation frequency: The selection was also based on the frequency with which the books were cited in other academic works to confirm their relevance and influence in the research field.

While the primary focus was on books, the analysis also incorporated current scientific articles. To ensure that no relevant and current findings were neglected, a further literature search was conducted in scientific databases, including Google Scholar, IEEE Xplore, and Scopus.

Furthermore, the method of "backward reference searching" was applied. This entailed a systematic search of the reference lists of the selected books and articles to identify additional relevant sources for inclusion in the analysis. This approach facilitated a more comprehensive and nuanced understanding of the research field, ensuring that the most pertinent and widely discussed instruments and methods were addressed.

The following twelve books have been identified through the described systematic process.

Table 1: Identified books

Nr.	Author	Year	Title
1	Drews & Hildebrand	2007	Lexikon der Projektmanagement-Methoden
2	Wieczorrek	2011	Management von IT-Projekten
3	Bergmann & Garrecht	2016	Organisation und Projektmanagement
4	Madauss	2017	Projektmanagement
5	Bohinc	2019	Grundlagen des Projektmanagements : Methoden, Techniken und Tools für Projektleiter
6	Fiedler	2020	Controlling von Projekten
7	von Känel	2020	Projekte und Projektmanagement
8	Meyer & Reher	2020	Projektmanagement : von der Definition über die Projektplanung zum erfolgreichen Abschluss
9	Preußig	2020	Agiles Projektmanagement : Agilität und Scrum im klassischen Projektumfeld
10	Kusay-Merkle	2021	Agiles Projektmanagement im Berufsalltag
11	Kuster et al.	2022	Handbuch Projektmanagement
12	Leyendecker & Pötters	2022	Werkzeuge für das Projekt- und Prozessmanagement

Based on the analysis on reference books and the instruments mentioned in therein, 30 traditional instruments and 23 agile instruments are listed. The analysis shows that traditional instruments are mentioned 90 times and agile instruments 62 times. Since traditional instruments are generally mentioned more often, an average must be calculated to correctly reflect the ratio. This shows that traditional instruments have an average of 3 mentions and agile instruments 2.7 mentions. This shows that there is still a tendency towards traditional instruments.

The three most frequently mentioned traditional instruments are the Work Breakdown Structure (8), Network Planning (8) and Gantt Chart (7). Among the agile instruments, the Product Backlog, User Stories and Kanban were each mentioned five times. The following table presents a comprehensive overview of the detailed results.

Table 2: Excerpt of most frequently named instruments

Instrument \ Author	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Work Breakdown Structure (traditional)	x			x	x	x		x	x		x	x
Network Planning Technique (traditional)	x		x	x	x	x		x	x			x
Gantt Chart (traditional)	x		x		x	x		x	x			x
Product Backlog (agile)		x				x	x		x	x		
User Stories (agile)		x			x	x			x	x		
Kanban (agile)		x				x			x	x	x	

[1] Drews & Hillebrand (2007); [2] Kuster et al. (2022); [3] Wiczorrek (2011); [4] Bergmann & Garrecht (2016); [5] Fiedler (2020); [6] Bohinc (2019); [7] Känel (2020); [8] Meyer & Reher (2020); [9] Preußig (2020); [10] Kusay-Merkle (2021); [11] Leyendecker & Pötters (2022); [12] Madauss (2017)

4 Conclusion

The contingency theory suggests that the implementation of differentiated management systems is advantageous in the context of increasing complexity. It can be reasonably assumed that companies that undertake classic projects also employ classic project control. The same should be true of hybrid and agile forms of project management. The initial step in examining this subject matter in greater depth is to analyze whether certain methods and instruments of project management and project controlling are better suited to the particular type of project.

The objective of this article was to examine the instruments currently in use. It can be observed that the three most frequently referenced traditional instruments are the Work Breakdown Structure, Network Planning, and Gantt Chart. Among the agile instruments, the Product Backlog, User Stories, and Kanban were each mentioned five times. The analysis shows that there is still a tendency towards traditional instruments.

The findings of the systematic literature review were evaluated in comparison with existing sources in order to ascertain their validity and relevance. This revealed that the instruments and methods identified largely correspond to those described in the literature.

1. Traditional Instruments: The most prevalent traditional instruments, including the Work Breakdown Structure, Network Planning and Gantt Charts, are also prominently featured in the extant literature (e.g., Kuster et al., 2022; Patanakul et al., 2010). This serves to confirm that these tools continue to occupy a central role in the practice of traditional project management.

2. Agile Instruments: The most frequently mentioned agile tools, including product backlog, user stories, and Kanban, also correspond to the descriptions provided in the literature (e.g., Gemino et al., 2021; Kuhrmann et al., 2022). This agreement underscores the significance of these instruments in agile project management.

3. Preference for traditional instruments: The analysis revealed a slight inclination to cite traditional instruments with greater frequency, a trend also reflected in the literature. This could signify that, despite the growing prominence of agile approaches, traditional approaches remain prevalent and are frequently employed in conjunction with agile methods.

In accordance with the findings of the analysis and the assessment of the existing literature, the following research hypotheses warrant further investigation:

H1: The type of project (traditional, agile) has a significant impact on project management and project performance.

H2: Projects managed with a project management approach that aligns with respective methods and instruments (traditional, hybrid, or agile) show higher success rates compared to those that do not.

Literature

- Aguanno, K. (2005). *Managing Agile Projects*. Multi-Media Publications Inc.
- Becker, W., Feichtinger, C., Langer, J. & Becker, N. (2021). *Agilität im Supply-Chain-Controlling: Eine empirische Analyse. Europäisches Forschungsfeld angewandte Mittelstandsforschung: Bd. 278*. Otto-Friedrich-Universität Bamberg.
- Bergmann, R. & Garrecht, M. (2016). *Organisation und Projektmanagement*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-32250-1>
- Bohinc, T. (2019). *Grundlagen des Projektmanagements: Methoden, Techniken und Tools für Projektleiter* (1. Aufl.). Whitebooks. Gabal Verlag GmbH. https://www.wiso-net.de/document/GABA,AGAB__9783956238512240
- Bonk, L. & Hedfeld, P. (2019). Agile Transformation und Innovationsfähigkeit. *ERP Management*(1), 41–44. https://www.wiso-net.de/document/BLIS__20190603694
- Cascio, J. (2020). *Facing the Age of Chaos*. <https://medium.com/@cascio/facing-the-age-of-chaos-b00687b1f51d>
- Collyer, S., Warren, C., Hemsley, B. & Stevens, C. (2010). Aim, Fire, Aim—Project Planning Styles in Dynamic Environments. *Project Management Journal*, 41(4), 108–121. <https://doi.org/10.1002/pmj.20199>
- DeCarlo, D. (2004). *EXtreme project management: Using leadership, principles, and tools to deliver value in the face of volatility* (1st ed.). *Jossey-Bass business & management series*. Jossey-Bass. <https://learning.oreilly.com/library/view/-/9780787974091/?ar>
- Dieffenbacher, S. F. (2023). *BANI World: What is it and Why We Need it?* <https://digitalleadership.com/blog/bani-world/>
- Drews, G. & Hillebrand, N. (2007). *Lexikon der Projektmanagement-Methoden*. *Haufe Projektmanagement*. Rudolf Haufe Verlag GmbH & Co. KG. https://www.wiso-net.de/document/HAUF__9783448080520285
- Fernandez, D. J. & Fernandez, J. D. (2008). Agile project management - Agilism versus traditional approaches. *Journal of Computer Information Systems*, 49(2), 10–17. https://www.researchgate.net/publication/265225700_Agile_project_management_-_Agilism_versus_traditional_approaches
- Fiedler, R. (2020). *Controlling von Projekten: Mit konkreten Beispielen aus der Unternehmenspraxis - alle controllingrelevanten Aspekte der Projektplanung, Projektsteuerung und Projektkontrolle* (8., aktualisierte und überarbeitete Auflage). Springer Vieweg. <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6005423>
- Gemino, A. (2021). Agile, traditional, and hybrid approaches to project success. *Project Management Journal*, 52(2), 161–175. https://www.wiso-net.de/document/ECON__1753118719
- Gemino, A., Reich, B. H. & Serrador, P. M. (2021). Agile, Traditional, and Hybrid Approaches to Project Success: Is Hybrid a Poor Second Choice? *Project Management Journal*, 52(2), 161–175. <https://doi.org/10.1177/8756972820973082>
- Highsmith, j. (2004). *Agile Project Management: Creating Innovative Products*. Addison–Wesley. https://www.researchgate.net/publication/234809670_Agile_Project_Management_Creating_Innovative_Products
- Känel, S. von. (2020). *Projekte und Projektmanagement*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-30085-2>
- Kuhrmann, M., Tell, P., Hebig, R., Klunder, J., Munch, J., Linssen, O., Pfahl, D., Felderer, M., Prause, C. R., MacDonell, S. G., Nakatumba-Nabende, J., Raffo, D., Beecham, S., Tuzun, E., Lopez, G., Paez, N., Fontdevila, D., Licorish, S. A., Kupper, S., . . . Richardson, I. (2022). What Makes Agile Software Development Agile? *IEEE Transactions on Software Engineering*, 48(9), 3523–3539. <https://doi.org/10.1109/TSE.2021.3099532>
- Kusay-Merkle, U. (2021). *Agiles Projektmanagement im Berufsalltag. für mittlere und kleine Projekte*. Springer Gabler. https://www.wiso-net.de/document/ECON__174386342X
- Kuster, J., Bachmann, C., Hubmann, M., Lippmann, R. & Schneider, P. (2022). *Handbuch Projektmanagement: Agil – Klassisch – Hybrid* (5. Aufl. 2022). Springer Berlin Heidelberg. <https://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-2068836>

- Leyendecker, B. & Pötters, P. (2022). *Werkzeuge für das Projekt- und Prozessmanagement*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-34724-6>
- Madauss, B.-J. (2017). *Projektmanagement: Theorie und Praxis aus einer Hand* (7. Aufl. 2017). Springer Berlin Heidelberg. <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1596122>
- Magistretti, S. & Trabucchi, D. (2024). Agile-as-a-tool and agile-as-a-culture: a comprehensive review of agile approaches adopting contingency and configuration theories. *Review of Managerial Science*, 1–31. <https://doi.org/10.1007/s11846-024-00745-1>
- Meyer, H. & Reher, H.-J. (2020). *Projektmanagement: Von der Definition über die Projektplanung zum erfolgreichen Abschluss* (2., überarbeitete Auflage). Springer eBook Collection. Springer Gabler. <https://doi.org/10.1007/978-3-658-28763-4>
- Möller, K. & Schmid, J. (2021). Agile Praktiken im Performance Management. *Controlling : Zeitschrift für erfolgsorientierte Unternehmenssteuerung*, 33(4), 48–55.
- Niederman, F., Lechler, T. & Petit, Y. (2018). A Research Agenda for Extending Agile Practices In Software Development and Additional Task Domains. *Project Management Journal*, 49(6), 3–17. <https://doi.org/10.1177/8756972818802713>
- Olsson, N. O.E. (2006). Management of flexibility in projects. *International Journal of Project Management*, 24(1), 66–74. <https://doi.org/10.1016/j.ijproman.2005.06.010>
- Patanakul, P., Lewwongcharoen, B. & Milosevic, D. (2010). An Empirical Study on the use of Project Management Tools and Techniques across Project Life-Cycle and their Impact on Project Success. *Journal of General Management*, 35(3), 41–66. <https://doi.org/10.1177/030630701003500304>
- Preußig, J. (2020). *Agiles Projektmanagement: Agilität und Scrum im klassischen Projektumfeld*. Haufe-Lexware GmbH & Co. KG. https://www.wiso-net.de/document/HAUF_9783648137789270
- Serrador, P. & Pinto, J. K. (2015). Does agile work? *International Journal of Project Management*, 33(5), 1040–1051. https://www.wiso-net.de/document/ECON_82841016X
- Shenhar, A. J. & Dvir, D. (2007). *Reinventing Project Management: The Diamond Approach To Successful Growth And Innovation*. Harvard Business Review Press.
- Špundak, M. (2014). Mixed Agile/Traditional Project Management Methodology – Reality or Illusion? *Procedia - Social and Behavioral Sciences*, 119, 939–948. <https://doi.org/10.1016/j.sbspro.2014.03.105>
- Thesing, T., Feldmann, C. & Burchardt, M. (2021). Agile versus Waterfall Project Management: Decision Model for Selecting the Appropriate Approach to a Project. *Procedia Computer Science*, 181(30), 746–756. <https://doi.org/10.1016/j.procs.2021.01.227>
- Tranfield, D., Denyer, D., Smart & Palminder (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(Nr. 3), 207–222.
- Ulrich, P. & Rieg, R. (2020). Agilität in Projektmanagement und Projektcontrolling - Ergebnisse einer empirischen Studie. *Die Unternehmung*, 74(2), 187–215. https://www.wiso-net.de/document/BLIS_20200701060
- Wieczorrek, H. W. (2011). PROJEKTMANAGEMENT. *Betriebswirtschaftliche Blätter*, 60(10), 592–598. https://www.wiso-net.de/document/BLIS_EA733FD812B03E975901C72DB6077E4B
- Williams, T. (2005). Assessing and Moving on From the Dominant Project Management Discourse in the Light of Project Overruns. *IEEE Transactions on Engineering Management*, 52(4), 497–508. <https://doi.org/10.1109/tem.2005.856572>
- WU Executive Academy. (2022). *BANI vs. VUCA: How Leadership Works in the World of Tomorrow*. <https://executiveacademy.at/de/news/detail/bani-statt-vuca-so-geht-fuehrung-in-der-welt-von-morgen/>
- Wysocki, R. K. (2007). *Effective project management: Traditional, adaptive, extreme* (4th ed.). Wiley Pub; Safari Books Online. <https://learning.oreilly.com/library/view/-/9780470042618/?ar>

Zhang, S., Zhang, F., Xue, B., Wang, D. & Liu, B. (2023). Unpacking resilience of project organizations: A capability-based conceptualization and measurement of project resilience. *International Journal of Project Management*, 41(8), 102541. <https://doi.org/10.1016/j.ijproman.2023.102541>

Komplexitätsreduktion im Nachhaltigkeitsreporting durch richtiges Dashboard Design

Research Paper

Dr. Lisa Perkhofer, MA

MCI Die Unternehmerische Hochschule, Department für Wirtschaft und Management, 6020 Innsbruck,
E-Mail: lisa.perkhofer@mci.edu

FH-Prof. Mag. Dr. Christoph Eisl

FH Oberösterreich, Institut für Controlling Rechnungswesen und Finanzmanagement, 4400 Steyr
E-Mail: christoph.eisl@fh-steyr.at

Johannes Teufl, MA

FH Oberösterreich, Institut für Controlling, Rechnungswesen und Finanzmanagement, 4400 Steyr,
E-Mail: johannes.teufl@fh-steyr.at

FH-Prof. Mag. DI Peter Hofer

FH Oberösterreich, Institut für Controlling Rechnungswesen und Finanzmanagement, 4400 Steyr
E-Mail: peter.hofer@fh-steyr.at

Abstract

Die neuen Nachhaltigkeitsstandards der EU bringen für die Unternehmenspraxis umfassende Berichtspflichten mit vielfältigen und komplexen Inhalten und Kennzahlen. Die Inhalte sind für Berichtsersteller und -empfänger neuartig und müssen unternehmensextern wie intern übersichtlich dargestellt werden. Um diese Anforderungen umzusetzen, bieten sich interaktive Nachhaltigkeitsdashboards an. Dieses Paper untersucht mithilfe eines Quasi-Experiments ausgewählte Gestaltungsaspekte von Dashboards, welche in etablierten Designrichtlinien ungeklärt sind. Darunter fallen (1) Inbar-Diagramme, welche sich zur Darstellung von Ziel- und Schwellenwerten besser eignen als ihre viel verwendete Alternative – das kombinierte Diagramm bestehend aus Säulen und einer Linie, (2) Filterpanels, welche im Vergleich zu statischen Filtern im Berichtskopf zu bevorzugen sind, und (3) die interaktive Visualisierungsart „Analysebaum“ zur Darstellung von Hierarchien, welche sich als effizientere Alternative im Vergleich zur Tabelle herausstellt.

1 Einleitung

In vielen europäischen Ländern werden aktuell schrittweise die *European Sustainability Reporting Standards (ESRS)* aus der bereits beschlossenen *CSRD (Corporate Sustainability Reporting Directive)* abgeleitet. Schon ab 2024 müssen die ersten Unternehmen die ESRS implementieren, um die relevanten Inhalte und Kennzahlen 2025 standardkonform berichten zu können (Vgl. Richtlinie (EU) 2022/2464 (CSRD), 2022, Art. 5 Abs. 2.). Da viele Unternehmen bisher nur freiwillige Standards zur Berichterstattung anwenden (z.B. *GRI – Global Reporting Initiative* oder *SDGs – Sustainable Development Goals*) oder gar keine Nachhaltigkeitsberichte veröffentlicht haben, werden sie durch diese nunmehr verpflichtende, äußerst umfangreiche, Berichterstattung vor große Herausforderungen gestellt. Dies liegt nicht nur darin begründet, dass die zu berichtenden Kennzahlen prüfungspflichtig sind und entscheidend die Informationsversorgung der Stakeholder beeinflussen (Frostenson und Johnstone, 2023, S. 125), sondern auch, weil zahlreiche Kennzahlen bei den Stakeholdern vielfach noch eher unbekannt bzw. ungewohnt sind. So müssen durch die neuen Bestimmungen erstmals Kennzahlen wie der CO₂-Fußabdruck (Emissionen nach Scopes 1-3), verursachter Abfall, Recyclingquoten, verwendete (hoch-) gefährliche Substanzen, Wasserverbrauch, das Gender-Pay-Gap, Zahlungspraktiken, Lobbying Aktivitäten, uvm. berichtet werden.

Neben dem Neuheitsgrad der Kennzahlen an sich, ist die Einführung dieser Kennzahlen zudem verbunden mit einem enormen Aufwand i.Z.m. der Datenbeschaffung und Gewährleistung hoher Datenqualität sowie in der Kommunikation dieser komplexen und teilweise aus vielen Positionen zusammengesetzten Key Performance Indicators (KPIs) (Vgl. Baumüller et al., 2023, S. 7 ff.) Die Masse an Daten, welche beispielsweise ein produzierendes Unternehmen verarbeiten muss, um den CO₂-Fußabdruck von Produkten berichten zu können, ist enorm. Dabei fallen nicht nur die CO₂-Emissionen des Unternehmens an sich ins Gewicht, sondern es muss auch in der Supply Chain jeder Zulieferbetrieb Zahlen zum Endergebnis beisteuern und der Beitrag eines jeden Unternehmens ist offenzulegen (Delegierte Verordnung ESRS Anhang1, S. 13, zukünftig auch *CSDDD – Corporate Sustainability Due Diligence Directive* – also das EU-Lieferkettengesetz).

Während das Nachhaltigkeitsreporting inzwischen vielfach in der Literatur diskutiert wird, gehen nur die wenigsten Beiträge genauer auf das Thema Dashboarding und eine entsprechend der hohen Komplexität gewählte Informationsaufbereitung ein. Autor:innen wie Sailer & Lösken (2023, S. 13 ff.), Baumüller et al. (2023, S. 7 ff.) und Ortner et al. (2022, S. 84) betonen zwar die Notwendigkeit von Dashboards in der Nachhaltigkeitsberichterstattung, behandeln aber nicht, wie solche Dashboards designt und/oder aufgebaut werden sollen. Aus diesem Grund werden in diesem Beitrag, basierend auf einer selbst erstellten Fallstudie, deren Aufbereitung und Interaktion mit Hilfe eines Quasi-Experiments empirisch getestet. Ziel ist es herauszufinden, welche Gliederung, welche Interaktion und welche Informationsaufbereitung dabei hilft, die hohe Komplexität der Nachhaltigkeitsberichterstattung zu reduzieren. Konkret wird folgende Forschungsfrage beantwortet: *„Wie können die ausgewählten Daten und KPIs in einem interaktiven Nachhaltigkeitsdashboard visualisiert werden, um die kognitive Last zu reduzieren und eine hohe Effizienz und Effektivität zu erzielen?“*

Im Folgenden werden zuerst die spezifischen Herausforderungen der Nachhaltigkeitsberichterstattung erläutert sowie erklärt, wie mittels dem Instrument Dashboard diesen Herausforderungen begegnet werden kann. Im dritten Kapitel werden die spezifischen Themen für die empirische Forschung präsentiert, bevor in Kapitel 4 die Datenerhebung und in Kapitel 5 die Analyse vorgestellt werden. Zuletzt werden in Kapitel 6 die Schlüsse gezogen und Limitationen der vorliegenden Untersuchung diskutiert.

2 Eignung von interaktiven Dashboards im Bereich Nachhaltigkeit

Wir argumentieren in diesem Kapitel für die Verwendung von interaktiven Dashboards zum Berichten von Informationen im Zusammenhang mit Nachhaltigkeitskennzahlen. Diese Empfehlung wird durch mehrere Argumente untermauert, bevor im nächsten Kapitel auf relevante Gestaltungsfragen als spezifische Antwort auf die Herausforderungen der ESRS eingegangen wird.

1. Die ESRS sollten nicht rein als Ergänzung des Jahresberichtes gesehen, sondern als steuerungsrelevanter Bestandteil in der Unternehmensstrategie verankert werden

Ein wesentlicher Teil unserer Überlegungen in Richtung Nachhaltigkeitsdashboards ist auf unserem Verständnis der Nachhaltigkeit begründet. Nachhaltigkeit darf nicht nur als Pflicht zur Informationsversorgung der Stakeholder gesehen werden, sondern der Themenkomplex Nachhaltigkeit ist als wesentlicher Teil in der Strategie und den Werten des Unternehmens zu verankern (Frostenson und Johnstone, 2023, S. 125). Die Triple-Bottom-Line – *Planet, People, Profit* (Elkington, 1998, SS. 18-22) – hat sich schon lange als strategischer Ansatz etabliert und ist in der jüngeren Vergangenheit (vor der CSRD und der damit verbundenen Pflicht zur Berichterstattung) vor allem von großen Konzernen verfolgt worden, welche auf ihr Image achten und auf die immer stärkere Nachfrage der Stakeholder nach ökologischer und sozialer Gerechtigkeit reagieren (Dienes et al, 2016, S. 168). Es hat sich gezeigt, dass durch die Nachhaltigkeit eine Win-Win Situation herbeigeführt werden kann, da nachhaltiges Wirtschaften neben positiven Effekten für die Ökologie und das soziale Umfeld des Unternehmens (= Planet & People), auch vielfach langfristig zu einer positiven Ergebniswirkung (=Profit) führt (Isil & Hemke, 2017, S. 1241).

Um strategische Initiativen zu verfolgen und im Unternehmen zu verankern, müssen diese in Ziele übersetzt werden. Es ist dabei wichtig, klare und anspruchsvolle Ziele zu formulieren und diese auch den entsprechenden Anspruchsgruppen zu kommunizieren (vgl. dazu die *Goal Setting Theory*, Locke und Latham, 2006, 265). Damit der Themenkomplex unterjährig die Aufmerksamkeit erfährt, welche es für eine strategische Stoßrichtung und eine erfolgreiche operative Umsetzung braucht und ein aktives Managen der Ziele über Kennzahlen ermöglicht (Kloviene und Speziale, 2014, SS. 50-51), wird die Steuerung an bisherige Berichtspraktiken im Unternehmen angepasst (beispielsweise die Verwendung einer *Balanced Scorecard* – BSC – oder das Verwenden von *Objectives & Key Results* – OKR). Für diesen Zweck und die damit verbundene Nachhaltigkeit der Zielerreichung haben sich interaktive Dashboards bereits etabliert und sind weit verbreitet (Velcu-Laitinen und Yigitbasioglu, 2021, S. 50-52). Der kontinuierliche Vergleich der aktuellen Performance mit den Zielen des Unternehmens erlaubt den PDCA (Plan-Do-Check-Act)-Zyklus fortlaufend zu etablieren und in weite Teile des Unternehmens zu kommunizieren (Eisl et al. 2019, S.4).

Ein Dashboard erlaubt dabei wichtige Kennzahlen zu einem bestimmten Thema auf einer Bildschirmseite zu präsentieren und ist defacto immer als interaktiver Report zu verstehen (Eisl et al. 2018, S. S. 53). Das bedeutet, übergeordnet wird ein Überblick zum Thema (z.B. die ökologische Perspektive) präsentiert und bei Wunsch kann, je nach Informationsbedarf, an der einen oder anderen Stelle ein Vertiefen in Details über Interaktion, wie beispielsweise mittels Tooltip oder Drill-Through Option, ermöglicht werden. Dies entspricht dem *Information Seeking Mantra* von B. Shneidermann (1996) „Overview first, zoom and filter, then details on demand“ (Shneidermann, 1996, S. 337) und erlaubt ein selbständiges und autonomes Auseinandersetzen mit den Inhalten durch den Berichtsempfänger. Dashboards gewähren relevanten Stakeholdern damit einen schnellen Überblick und sind auf die Einhaltung der unternehmensintern gesteckten Ziele ausgerichtet.

Bei den ESRS ist das Setzen und Kommunizieren von Zielen ein integraler Bestandteil. Dabei ist als Meilenstein oft ein zu erreichender Zielwert (Erfolg entsteht durch Erreichen oder Überschreitung des Zieles) bzw. ein zu unterschreitender Schwellenwert (Erfolg entsteht durch Erreichen oder Unterschreitung der Schwelle) für einen langfristigen Zeithorizont, oftmals 2030 bzw. 2050, definiert (wie bspw. in ESRS E1 Klimawandel zum Thema CO₂-Emissionen). Verglichen wird dieses Ziel kontinuierlich mit einem ‘Base Year’ (bspw. das erste Jahr der Berichtspflicht), um die Entwicklung des Unternehmens hin zur Zielerreichung auch als externer Stakeholder mitverfolgen zu können. Es liegt also ein hohes Augenmerk auf Trends, Entwicklungen und Vergleichsinformationen. Dashboards scheinen gerade in diesem Zusammenhang ein gutes Konzept zur Verbreitung und Erreichung der für die ESRS relevanten Ziele und in weiterer Folge hoffentlich zum Gelingen des Green Deals der Europäischen Union beizutragen.

2. Die Nachhaltigkeitskennzahlen bringen eine gewisse Komplexität mit sich, Dashboards reduzieren nachweislich diese Komplexität.

Die übergeordnete Idee der Gesetzgebung hinter den Nachhaltigkeitskennzahlen ist, Unsicherheiten und Risiken im Zusammenhang mit den ESG (*Environment, Social, Governance*)-Kriterien für alle beteiligten Stakeholder (unter anderem Kunden, Lieferanten, Investoren, Eigentümer, Mitarbeiter etc.) zu reduzieren. Dass Investoren schon seit einigen Jahren die Themenbereiche ESG in ihre Investitionsentscheidungen miteinbeziehen, ist bereits durch mehrere Studien belegt (z.B. Hamed et al., 2022, S. 100459-100460). Unternehmen haben damit neben der Pflicht zur Berichterstattung auch ein gewisses intrinsisches Interesse mit diesem Thema transparent umzugehen und soziale und ökonomische Initiativen zu fördern.

Damit die Kennzahlen nachvollziehbar werden und unterschiedliche Unternehmen miteinander verglichen werden können, gibt es detaillierte Vorschriften und Definitionen von Kennzahlen durch die ESRS (Vgl. Sailer & Lösken, 2023, S. 13–14). Wie die normierte Berichterstattung inkl. der detaillierten Definition dieser sehr umfangreichen Kennzahlen auf die Interessensgruppen wirken, bleibt abzuwarten. Zu erwarten ist, dass es einige Zeit dauern wird, bis die neuen und oft durch komplexe Rechenverfahren zu ermittelnden Kennzahlen (z.B. der CO₂-Fußabdruck) von allen Anspruchsgruppen richtig interpretiert werden können. Dies liegt in der Funktion der menschlichen Informationsverarbeitung begründet. Menschen sind nur zu einem gewissen Teil fähig, neue Informationen gut zu verarbeiten, da Grundlagen und Erfahrungen fehlen, die es erst aufzubauen gilt. (Vgl. dazu das *Multi-Store-Model of Memory*, Atkinson und Shiffring, 1968)

Eine Möglichkeit die Interpretationsfähigkeit zu unterstützen bzw. schnelles Lernen zu fördern, ist die Verwendung von Visualisierung und Dashboarding. Komplexe Kennzahlen zu visualisieren ist immer der erste Schritt, um Daten initial verständlich zu machen und um Trends und Zusammenhänge erkennen zu können. (Vgl. dazu die *Cognitive Load Theory*, Sweller 1988; Sweller 2010; Perkhofer, 2019, S. 19; Falschlunger et al., 2016, S. 4) Sind unterschiedliche Fragen durch einen Datensatz zu beantworten, ist es sinnvoll für jede spezifische Frage eine dazu passende visuelle Aufbereitung anzufertigen (Vgl. dazu die *Theory of Cognitive Fit*, Vessey, 1991; Speier, 2006) und alle Visualisierungen auf einer Dashboard-Seite zur Gewinnung eines ersten Überblicks zu präsentieren (Yigitbasiglu and Velcu, 2012, S. 41 ff.). Es sollen also mehrere Ansichten auf ein und denselben Datensatz erlaubt werden, was nachweislich die Komplexität der zu interpretierenden Informationen reduziert. (Falschlunger et al., 2016, S. 4; Baumüller et al., 2023, S. 9)

3. Die Informationen kommen aus unterschiedlichsten Quellen und müssen über Datenmodelle sinnvoll kombiniert werden.

Eine zentrale Herausforderung im Bereich ESG liegt im Thema Datenkomplexität. Dies ist einerseits in der mangelnden Datenqualität aktueller Datensätze begründet und andererseits durch hohe Datenmengen sowie eine hohe Anzahl an unterschiedlichen Quellsystemen bedingt (z.B. Daten von Kunden und Lieferanten zu CO₂-Emissionen, Daten aus Umfragen zur Mitarbeiterzufriedenheit, Daten aus Personalverwaltungssystemen zur Personalstruktur, Daten aus Intranet zum Weiterbildungsangebot und dessen Annahme, ...) (Lösken und Sailer, 2023). Ist man in der glücklichen Lage, all diese Daten zur Verfügung zu haben, braucht man Business-Intelligence (BI)-Systeme, die ein Kombinieren dieser breiten Palette an Daten ermöglichen. In diesem Zusammenhang ist allerdings anzumerken, dass bei vielen Unternehmen die Datenverfügbarkeit noch sehr eingeschränkt ist, also Daten zu den geforderten Inhalten bisher nicht dokumentiert und/oder nachgehalten wurden. Dies gilt es in den kommenden Jahren aufzuholen – einerseits durch Investitionen in die Dateninfrastruktur und andererseits in veränderte Unternehmensprozesse (Vgl. Hofinger & Röhm, 2023, S. 90–91; Waniczek, 2021, S. 52).

Zusätzlich werden die Prozesse komplexer. Allein durch die Anforderungen aus dem ersten Set der ESRS entstehen für das Unternehmen, je nach Wesentlichkeit, potenziell über 1.100 Datenpunkte (ein Datenpunkt ist bspw. die Angabe des gesamten Energieverbrauchs) (Kämmler-Burrak u. a., 2024, S. 13). Die Daten kommen zudem, wie bereits erwähnt, aus unterschiedlichen Quellsystemen und jedes Quellsystem und jede Information bedingt eigene Regeln für die Datenbereinigung und Aufbereitung für den Endbericht (Vgl. Baumüller et al., 2023, S. 7; Sattler &

Schentler, 2022, S. 49). Eine genaue Dokumentation der Prozesse ist unerlässlich, um die Daten nachzuvollziehen und die Qualität der Kennzahl beurteilen zu können. Aus diesem Grund ist auch der Datenerstellungsprozess und die Kennzahlendokumentation, und nicht die Kennzahl selbst, Gegenstand der Prüfung in der Nachhaltigkeitsberichterstattung (Vgl. Sailer & Lösken, 2023, S. 13–14).

Diese große Menge an Datenquellen braucht deshalb eine Regulierung für die Datenablage (oft eigene Data Lakes oder Data Warehouses für Nachhaltigkeit) um einen ‘Single Source of Truth’ für die Berichterstattung herstellen zu können. Zusätzlich braucht es ein geeignetes System, um mit der Masse der Daten arbeiten zu können. Klassische Excel-Kalkulationen sind damit oftmals bereits überfordert (Lösken und Sailer, 2023). Besser eignen sich dafür sogenannte *OLAP (Online Analytical Processing)* -Anwendungen wie beispielsweise Microsoft Power BI, SAP SAC, Tableau, etc. OLAP-Anwendungen erlauben ein Laden relevanter Tabellen, ein Transformieren (z.B. Summieren) falls notwendig, sowie ein systematisches Auswerten der Daten durch die Endverbraucher:in mittels interaktiver Dashboards (Vgl. Patloch, 2019, S. 69).

4. Mit Hierarchien und Interaktionen zu arbeiten erlaubt eine individuelle User-Journey, braucht aber sehr detaillierte und multidimensionale Daten

Werden Datensätze umfangreich, also beinhalten diese viele Fakten/Kennzahlen und viele Dimensionen – wie im Bereich der Nachhaltigkeit üblich – ist nicht nur ein Visualisieren der Daten von hoher Bedeutung, sondern auch ein interaktives Arbeiten mit diesen. Es ist notwendig zu überlegen, welche Informationen sinnvoll gruppiert werden können (Reisinger, 2023, S. 37 ff.), für welche hierarchische Ebenen ein und dieselbe Auswertung verwendet werden kann und wo bei Bedarf Details für ein tieferes Verständnis hinzugefügt werden müssen (Perkhofer et al. 2019). Insbesondere interaktive Dashboards sollen dazu beitragen, eine Informationsüberlastung der Berichtsnutzer:innen zu verhindern. (Vgl. dazu *Information Overload Theory*, Moko et al., 2023, S. 24 ff., Baumüller et al., 2023, S. 7–8)

Ein großer Vorteil interaktiver Dashboards ist, dass weitere Information bei Bedarf durch Klick auf ein Element oder durch Tooltips hinzugefügt werden können. Dies ist insbesondere in der ersten Zeit der Verwendung von neuen Kennzahlen im Bereich Nachhaltigkeit von entscheidender Bedeutung, da sowohl gesetzliche Grundlagen als auch Berechnungslogik und Definition einer Kennzahl mithilfe eines Tooltips in die Berichterstattung integriert werden können. Damit ist eine entsprechende Kontextinformation für die jeweiligen Nutzer:innen permanent greifbar. Zusätzlich erlauben Dashboards bei der Analysetätigkeit vom Überblick ins Detail zu gehen und sich dort zu vertiefen, wo Interesse oder Unklarheiten auftauchen. Das beeinflusst nachweislich positiv die Interpretationsfähigkeit von Menschen (Shneiderman, 1996, S. 336 ff.). Mithilfe von Interaktion wird nämlich ein individuelles Auseinandersetzen der User mit den Daten gefördert und dadurch die Verständlichkeit der Daten erhöht (Brehmer und Munzner, 2013, S. 2376 ff.). Um dies gewährleisten zu können braucht es, wie bereits im vorangegangenen Argument erörtert, eine gute Datenmodellierung und eine solide Datenbasis.

Damit Dashboards von Ihren Anspruchsgruppen aktiv genutzt werden, sind diese zudem auf die Nutzerbedürfnisse anzupassen, um durch Gewährleistung einer hohen Effizienz und Effektivität die Datenexploration zu unterstützen (Eisl et al. 2018; Magdalena et al., 2019, S. 5-6). Während einzelne ESRS mit Vorgaben oder Vorschläge zur Darstellung der geforderten Kennzahlen ausgestattet sind (bspw. tabellarische Darstellungen in ESRS, G1, § AR 8; ESRS, S1, § AR 55), gilt es bei den meisten Kennzahlen selbst eine Form der Darstellung zu wählen. Nachstehendes Kapitel fasst wichtige Erkenntnisse zum Thema Gestaltung zusammen und zeigt, wo noch aktiver Forschungsbedarf besteht.

3 Theoretische Grundlagen für die empirische Forschung

Bisher wurden die Anforderungen bzw. Herausforderungen im Zusammenhang mit der Nachhaltigkeitsberichterstattung vorgestellt, sowie die Idee diese mittels Dashboarding zu mittigeren. Dabei wurde festgestellt, dass Visualisierung und Interaktion wesentlich sind. Was noch fehlt, sind ein konkreter Überblick der unterschiedlichen Gestaltungsmöglichkeiten sowie eine Diskussion welche konkreten Fragen im Zusammenhang mit den spezifischen Herausforderungen von ESG einhergehen.

3.1 Visualisierungsmöglichkeiten und spezifische Anforderungen durch ESG

Für klassische Visualisierungstypen bilden die Gestaltungsempfehlungen des Internationalen Controllervereins (ICV) für eine höhere Benutzerfreundlichkeit bereits eine solide Grundlage (Eisl et al., 2018). Dabei werden Standard-Business Grafiken bzgl. ihrer Hauptaufgabe / ihres Informationsgehaltes kategorisiert sowie Designempfehlungen für ein effizientes und effektives Dekodieren der Information durch die Nutzer:in gewährleistet. Beim Design gilt es, die Information in den Mittelpunkt zu stellen und Dekoration zu vermeiden. Dazu sollte zur Überprüfung der Umsetzung vor allem das Konzept des *Non-Data-Ink* (Tuft, 2001, S. 91 ff.) Berücksichtigung finden. Einen Überblick zu den Empfehlungen zum Einsatz verschiedener Visualisierungsformen ist nachstehender Tabelle 1 zu entnehmen.

Tabelle 1: Überblick Visualisierungsformen und deren Hauptaufgaben

Visualisierung	Zentrale Aufgabe
Säulendiagramm	Darstellung einer zeitlichen Entwicklung, wobei die Höhe der einzelnen Werte von Relevanz ist.
Liniendiagramm	Darstellung einer zeitlichen Entwicklung, wobei der generelle Trend und nicht der einzelne Wert von Relevanz ist
Balkendiagramm	Darstellung von Strukturvergleichen (Aufriss eines Faktes/einer Kennzahl zu einem bestimmten Zeitpunkt auf eine bestimmte Dimension – z.B. Kundengruppen, Produktgruppen etc.)
Inbar-Diagramm	Darstellung eines Vergleichs der aktuellen Werte im Vergleich zum Budget/Ziel. Das Diagramm zeigt zusätzlich zur Entwicklung der Ist-Werte die Veränderung zwischen Ist und Ziel in Form von Abweichungsbalken. <i>Hier gibt es allerdings aus der bisherigen Forschung gemischte Resultate. Inbar-Diagramme bedürfen erfahrene Leser und damit eine Verbreitung der Visualisierungsform unter den Berichtsempfänger:innen.</i>
Wasserfall-diagramm	Darstellung eines Verlaufs. Hier wird ein Ausgangswert durch unterschiedliche Einflüsse verändert und diese Veränderung wird visuell vergleichbar dargestellt.
Small Multiples	Eine Ansammlung gleichartiger Grafiken (z.B. Säulen, Balken, Linien) mit derselben Skalierung zur Gewinnung eines Überblicks der jeweiligen zeitlichen Entwicklungen unterschiedlicher Segmente/Dimensionen des Datensatzes.
Tabelle	Eine tabellarische Darstellung ist immer dann von hoher Relevanz, wenn genaue Werte wesentlich, allerdings Trends und Zusammenhänge in den Hintergrund rücken.
Treemap	Darstellung hierarchischer Zusammenhänge, wobei immer nur maximal zwei hierarchische Ebenen gleichzeitig sichtbar sind. Weitere Ebenen müssen über

Visualisierung	Zentrale Aufgabe
	Interaktion sichtbar gemacht werden. Es soll ein Überblick zu möglichen Gruppierungen in Relation zu einem Fakt/einer Kennzahl hergestellt werden.
Entscheidungsbaum	Darstellung hierarchischer Zusammenhänge, wobei immer alle wesentlichen Ebenen gleichzeitig sichtbar sind. Man kann abhängig vom eigenen Interesse Details in der Baumstruktur hervorheben. <i>Hierzu gibt es im Zusammenhang mit unternehmensinternen Berichten keine relevanten Erkenntnisse. Das Aufzeigen von Hierarchien ist aber in der Nachhaltigkeitsberichterstattung von entscheidender Bedeutung.</i>

Weitere Visualisierungen gibt es noch aus der Forschung im Zusammenhang mit Big Data (unter anderem Sunburst Diagramme, Sankey Diagramme, Parallel Coordinates Plots, Scatterplot Matrizen etc.). (Perkhofer et al., 2020, S. 61 ff.) Diese sind aber nicht geeignet um Einsichten an eine breite Masse an Nutzer:innen zu berichten. Dazu müssen Kenntnisse zum Lesen der Diagramme vorliegen und zudem braucht es Erfahrung, um richtige Schlüsse ziehen zu können.

3.1.1 Vergleichsinformationen (Ziel- und Schwellenwerte) richtig präsentieren

Gemäß den Vorgaben der ESRS sind in vielen Fällen Ziele und Schwellen anzugeben. Um verschiedene Werte miteinander zu vergleichen und darzustellen, finden in der Unternehmenspraxis kombinierte Säulen- und Liniendiagramme Verwendung (Abbildung 1 linke Darstellung). Im Finanzbereich werden zudem Inbar-Diagramm eingesetzt, um IST-Werte im Vergleich zu definierten Planwerten wie dem Budget darzustellen (Abbildung 1 rechte Darstellung). (Vgl. Eisl et al., 2018, S. 80–82). Wie bereits erwähnt gibt es im Zusammenhang mit Inbar-Diagrammen immer wieder widersprüchliche Erkenntnisse. Dieser Diagrammtyp wird zu Beginn von vielen Berichtsleser:innen oftmals als schwierig zu lesen wahrgenommen. Jedoch verbessert sich die Effektivität und Effizienz je nach Erfahrung mit dem Diagrammtyp deutlich. Inbar-Diagramme werden somit für erfahrene Nutzer:innen für den Vergleich von IST- und Plandaten empfohlen. (Vgl. Eisl et al., 2018, S. 80–82). Da der Grad der Erfahrung im Zusammenhang mit Nachhaltigkeitskennzahlen gering ist, gilt es diesen Zusammenhang nochmals zu prüfen.

Neben der Frage nach dem Diagrammtyp gilt es auch die Farbverwendung zu überprüfen. Im deutschsprachigen Raum hat sich rot für negative und grün für positive Abweichungen durchgesetzt. Ob sich diese Logik auf Ziel- und Schwellenwerte in der Nachhaltigkeitsberichterstattung übertragen lässt, ist noch nicht geklärt. Oftmals kann schlechtes Design durch die Fähigkeiten der Nutzer:innen ausgeglichen werden. Dies ist zunehmend schwieriger, je komplexer die zu bewältigenden Aufgaben werden. Bei komplexen Aufgaben sind mehrere aufeinanderfolgende Denkschritte notwendig, um eine richtige Interpretation zu ermöglichen. Bei höherer inhärenter kognitiver Belastung (vgl. dazu Intrinsic Cognitive Load, Perkhofer 2019, S. 17 ff.) ist deshalb das Design und damit die zusätzliche und externe Belastung (vgl. dazu Extraneous Cognitive Load, Perkhofer 2019, S. 17 ff.) von höherer Bedeutung. Kommt auf einer bereits hohen Aufgabenkomplexität eine zusätzliche Belastung durch schlechtes Design, kann dies durch die Berichtsnutzer:in nicht mehr ausgeglichen werden. Die Folge sind höhere Missinterpretationen und ein erhöhter Zeitbedarf. In diesem Zusammenhang soll deshalb auch die Komplexität in der Untersuchung beobachtet werden. Es werden Aufgaben also hinsichtlich ihrer inhärenten Komplexität variiert, um den Effekt des Designs besser isolieren zu können.

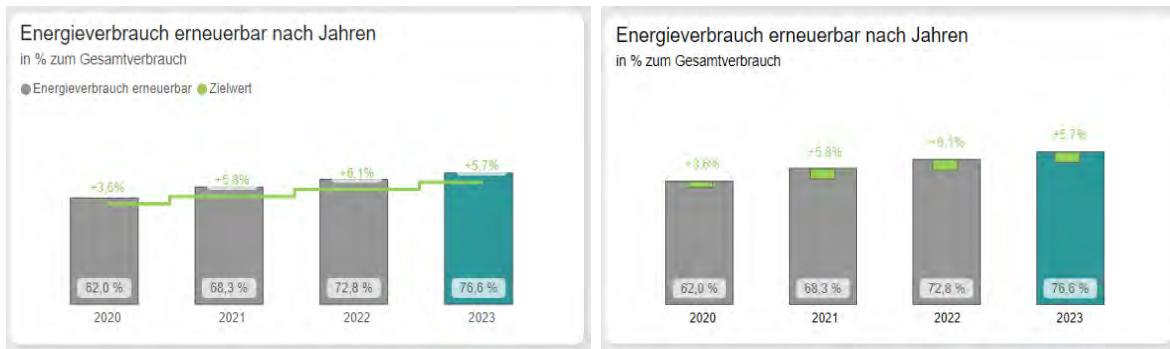


Abbildung 1: Alternative Möglichkeiten zur Visualisierung von Zielwerten (kombiniertes Diagramm vs. Inbar-Diagramm)

Unbeantwortet ist zudem die Frage, ob sich die Linienfarbe von Ziel- und Schwellenwerten durch die Verwendung von Rot und Grün voneinander unterscheiden sollte und ob eine der Varianten zu einer besseren Darstellung im Sinne der Qualitätskriterien *Effizient* (Geschwindigkeit der Informationsverarbeitung) und *Effektivität* (korrekte Interpretation der Informationen) führt. Dazu wird eine neutrale, einheitliche Farbe (Schwarz) mit den Farben Rot (zur Darstellung von Schwellen) und Grün (zur Darstellung von Zielen) miteinander verglichen. Damit weicht das Experiment von der üblichen Bedeutung der Farben Grün und Rot für positive / negative Abweichungen ab und untersucht eine weitere Einsatzmöglichkeit der verkürzten Ampellogik. (Vgl. Eisl et al., 2018, S. 42)

Tabelle 2: Hypothesen zu Visualisierungsformen Ziel- und Schwellenwerte

ID	Thema	Hypothesen
H1	Visualisierungs-typ	Die Visualisierungsform (kombiniertes Diagramm mit schwarzer Linie vs. kombiniertes Diagramm mit farbiger Linie vs. Inbar-Diagramm) hat keinen Einfluss auf a) die <i>Effektivität</i> , b) die <i>Effizienz</i> bei der Informationsverarbeitung.
H2	Komplexität	Die Komplexität (einfach vs. komplex) der Aufgabenstellung hat keinen Einfluss auf die <i>Effektivität</i> bei der Informationsverarbeitung.

3.1.2 Hierarchische Zusammenhänge richtig visualisieren

Aus den Anforderungen der ESRS ergeben sich zudem in manchen Bereichen steile Hierarchien, wie beispielsweise bei der Darstellung der Abfallverursachung, die eingeteilt wird in die Hierarchiestufen gefährliche und ungefährliche Abfälle, Verwertungs- bzw. Entsorgungsart und deren jeweilige Unterkategorien. (Vgl. Delegierte Verordnung EU 2023/2772, Anhang 1, 2023/22.12.2023, S. 157–158) Um solche breite Hierarchien darzustellen, empfiehlt die Literatur die Nutzung von Baumdiagrammen oder auch Sunburst Charts (Vgl. Islam & Jin, 2019, S. 5–6). Die empirische Untersuchung dazu wird in einem Musterdashboard in Microsoft Power BI präsentiert und daher wird auf die dort verfügbaren Diagramme zurückgegriffen. Sunburst Charts können zu diesem Zeitpunkt in Power BI nicht in ihrer vollen Funktionalität verwendet werden. Allerdings bieten sich interaktive Analyseebenen (Decomposition Trees) für große Hierarchien an. Im Rahmen des Quasi-Experiments werden sie für den Einsatz im ökologischen Dashboard getestet und mit ihrer herkömmlichen Alternative, eine Tabelle, verglichen. In Tabellen können ebenfalls Hierarchien dargestellt und mithilfe von +/- Symbolen auf- und zugeklappt werden. (Vgl. Myers, 2024)

Tabelle 3: Hypothesen zu Visualisierungsformen von Hierarchien

ID	Thema	Hypothesen
H3	Visualisierungs-typ	Die Visualisierungsform (Entscheidungsbaum vs. Tabelle) hat keinen Einfluss auf a) die <i>Effektivität</i> , b) die <i>Effizienz</i> bei der Informationsverarbeitung.
H4	Komplexität	Die Komplexität (einfach vs. komplex) der Aufgabenstellung hat keinen Einfluss auf a) die <i>Effektivität</i> b) die <i>Effizienz</i> bei der Informationsverarbeitung.

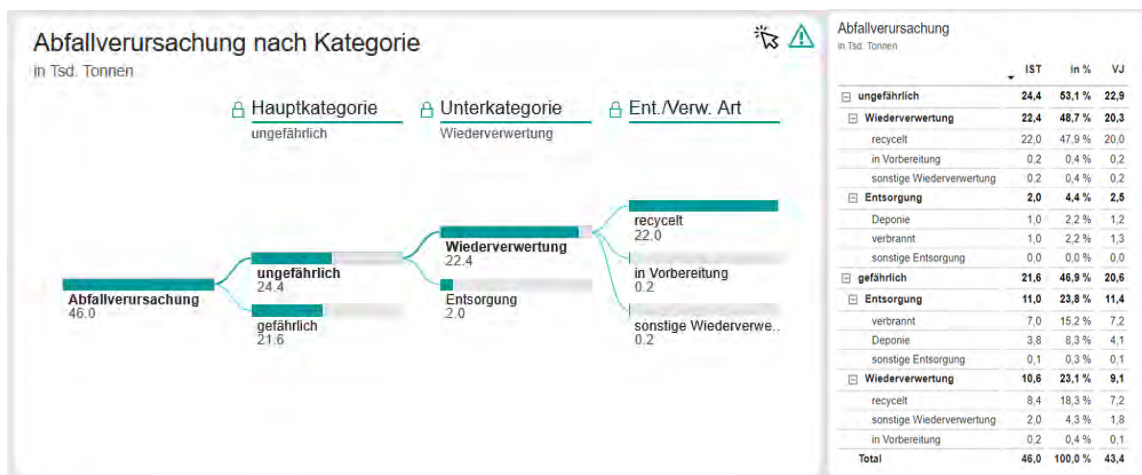


Abbildung 2: Alternative Möglichkeiten zur Visualisierung von Hierarchien (Entscheidungsbaum vs. Tabelle)

3.2 Interaktionsmöglichkeiten und spezifische Anforderungen durch ESG

Die im vorigen Kapitel beschriebenen Visualisierungstypen, welche in Kombination mit großen Datenmengen und Hierarchien eingesetzt werden, entfalten erst durch den Einsatz von passenden Interaktionen ihre volle Wirkung zur Steigerung der Effizienz und Effektivität. Im Analyseprozess können in einem ersten Schritt originäre Such- oder Bewertungsaufgaben des Users mittels einzelner Interaktionen unterstützt werden. Oft kommt man aber erst durch weiteres Explorieren der Daten zu neuen Einsichten. Hierbei können Interaktionstypen wie das Selektieren, Filtern, Navigieren, Anordnen, Ändern oder Aggregieren gezielt im Self-Service-BI eingesetzt und zu Interaktionskonzepten gebündelt werden. (Hofer et al, 2023, S. 165-169)

In Literaturquellen wie Heer & Shneiderman (2012, S. 47–49) und Hofer et al. (2023, S. 151) wird die Interaktionsmöglichkeit Filtern als eine essenzielle Funktion hervorgehoben, um mit großen Datenmengen umzugehen und von einem kompakten Überblick einzelne Details zu separieren. Wie ein Filter wirkt, ist klar. Daten werden nach spezifischen Kriterien hin untersucht und entsprechend der Auswahl ausgewiesen. Allerdings wird nicht genauer beschrieben, wie mehrere unterschiedlichen Filter positioniert und wie sie gestaltet werden sollten.

Im Rahmen der Literaturrecherche wurden zu diesem Thema keine wissenschaftlichen Publikationen gefunden. In vielen Praktiker-Anleitungen und YouTube-Videos wird als Best Practice die Positionierung der Filter im Header eines Dashboards oder die Verwendung eines interaktiven Filterpanels empfohlen (ein Beispiel dazu ist in Abbildung 3 abgebildet). Ein interaktives Filterpanel kann je nach Bedarf ein- und ausgeblendet werden und alle notwendigen Filter können auf einen Blick übersichtlich präsentiert werden. Ein zentraler Nachteil könnte allerdings sein, dass während der Interpretation der Daten die aktivierten und deaktivierten Elemente nicht ständig sichtbar sind. Das könnte in der Analysephase zu Problemen, wie beispielsweise Verwechseln oder Vergessen der ausgewählten Optionen, führen. Klassische Filter hingegen sind oft statisch im Header oder an einer festen Stelle im Dashboard

positioniert und permanent sichtbar, was die Übersichtlichkeit beeinträchtigen kann. Allerdings kann es sein, dass bei vielen relevanten Filtermöglichkeiten nicht alle Filter entsprechend Platz finden.

Aufgrund mangelnder Literaturempfehlungen wird das Filterpanel im Quasi-Experiment mit klassischen Filtern im Header verglichen. Dazu werden folgende Hypothesen aufgestellt:

Tabelle 4: Hypothesen zur Gestaltung von Filtern im Dashboard

ID	Thema	Hypothesen
H5	Interaktion	Die Gestaltung der Filtermöglichkeit (Filter im Kopf vs. Filterpanel) hat keinen Einfluss auf a) die <i>Effektivität</i> , b) die <i>Effizienz</i> bei der Informationsverarbeitung.
H6	Komplexität	Die Komplexität (einfach vs. komplex) der Aufgabenstellung hat keinen Einfluss auf a) die <i>Effektivität</i> b) die <i>Effizienz</i> bei der Informationsverarbeitung.

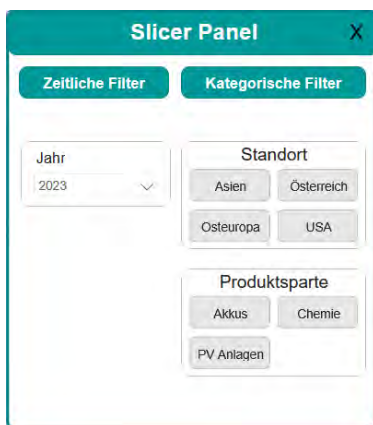


Abbildung 3: Möglichkeit zur Gestaltung von Filtern mittels Filterpanel

4 Beschreibung des Quasi-Experiments

In diesem Kapitel werden die Variablen und deren Manipulation sowie die Studienteilnehmer:innen und der Studienablauf beschreiben. Als unabhängige Variablen dienen je nach Aufgabe die Visualisierungsart bzw. Darstellungsform des Berichtselements sowie die Komplexität der Aufgabenstellung nach Zhu (Vgl. Zhu, 2023, S. 131–132). Die abhängigen Variablen umfassen die Effektivität und die Effizienz gemäß den Leitlinien Exzellenz im Reporting Design (Vgl. Eisl et al., 2018, S. 17–18).

4.1 Aufbau des Quasi Experiments

Das Quasi-Experiment wird mithilfe eines Within-Subject Designs durchgeführt. Das heißt, dass alle Teilnehmer:innen dasselbe Experiment mit denselben Themengebieten bearbeiten. (Vgl. Charness et al., 2012, S. 1–3) Dazu werden standardisierte Online-Fragebögen genutzt, die die Proband:innen gleichzeitig am Computer, also auf einem einheitlichen Bildschirm, ausfüllen. Für die Erstellung und für Teile der Auswertung wird das Tool „LimeSurvey“ genutzt, dass sich durch eine einfache Auswertung und hohe Benutzerfreundlichkeit auszeichnet. (LimeSurvey, 2024) Zudem erlaubt LimeSurvey, Aufgaben in randomisierter Reihenfolge ablaufen zu lassen. Insgesamt dauert das Experiment pro Durchführung durchschnittlich 30,8 Minuten.

Die Frage nach der optimalen Visualisierung wird durch zehn unterschiedliche Optionen getestet. Zur Frage nach der optimalen Interaktion werden vier mögliche Optionen verwendet. Diese sind untenstehend in der Tabelle zur

Veranschaulichung der unabhängigen Variablen aufgelistet. Der Effekt der Manipulation durch Visualisierung, Interaktion und Komplexität wird anhand zweier Variablen untersucht. Diese werden im Anschluss je nach Messniveau mittels unterschiedlicher statistischer Auswertungen untersucht. Während für metrische Variablen eine lineare Regression verwendet wird, wird für die nominale Variable eine logistische Regression herangezogen.

Als Kontrollvariablen werden Alter in Jahren, Erfahrung mit ESG (Likert Scale 1-5) sowie Erfahrung mit Power BI Dashboards (abgekürzt Erfahrung DB), ebenfalls gemessen mit einer 5-Punkte Likert Skala, herangezogen.

Tabelle 5: Manipulation der unabhängigen Variablen

	Thema	Manipulation
Optimale Visualisierung	Visualisierung Ziel- und Schwellenwert	Kombiniertes Diagramm mit schwarzer Linie Kombiniertes Diagramm mit grüner/roter Linie Inbar-Diagramm
	Komplexität	Niedrige Komplexität (Einfach – Identifikation eines Wertes) Hohe Komplexität (Komplex – Identifikation eines Trends/Zusammenhangs)
	Visualisierung hierarchische Informationen	Entscheidungsbaum Tabelle
	Komplexität	Niedrige Komplexität (Einfach – 3 Hierarchiestufen) Hohe Komplexität (Komplex – 4 Hierarchiestufen)
Optimale Interaktion	Filter	Filter im Header Filterpanel
	Komplexität	Niedrige Komplexität (Einfach: Filter muss zur Beantwortung 1x verändert werden) Hohe Komplexität (Komplex: Filter muss zur Beantwortung mehrmals geändert werden)

Tabelle 6: Definition und Messniveau der abhängigen Variablen

Variable	Beschreibung
Effektivität	Antwortrichtigkeit: Es gibt immer eine richtige Antwort. Richtige Antworten werden mit 1, falsche Antworten mit 0 kodiert. Damit können Fehlerquoten berechnet werden und wir arbeiten mit einer dichotomen bzw. nominal skalierten Variable.
Effizienz	Antwortgeschwindigkeit: Für die Fragen wird jeweils die Zeit dokumentiert und für einen Vergleich herangezogen. Auffällig kurze und auffällig lange Antwortzeiten wurden aus der Analyse ausgeschlossen (z-Score größer 3 oder kleiner -3). Zudem wurden für die Effizienz nur jene Antworten berücksichtigt, bei welchen die Teilnehmer:innen eine richtige Antwort identifizieren konnten. Die Variable ist metrisch.

4.2 Ablauf des Quasi-Experiments

Um Interviewer-Bias oder ähnliche Effekte zu vermeiden, erhalten die Studierenden ein einheitliches Briefing zum Experiment in Form einer voraufgezeichneten PowerPoint-Präsentation. Diese umfasst eine kurze Einführung in den EU Green Deal, den groben Aufbau der ESRS, eine Einführung in die Fallstudie, Details zum Experiment sowie

Erklärungen zu neuen Visualisierungsarten und Berichtselementen anhand von Beispielen. Dies hat sich in vergleichbaren Studien in der Vergangenheit als sehr hilfreich erweisen.

Anschließend wird den Proband:innen ein Link zur Teilnahme an der Studie geteilt. Dadurch kann jede Teilnehmer:in in einer selbst gewählten Geschwindigkeit durch das Experiment navigieren. Zu Beginn werden exemplarische Stimuli und Fragen präsentiert, um Teilnehmer:innen in das Experiment einzuführen und an den Aufbau zu gewöhnen. Initial wird auch Feedback zur Performance gegeben, um Sicherheit zu vermitteln. Nach der Übungsrunde wird auf den Start des Experiments nochmals explizit verwiesen. In der Datenerhebungsphase wird kein Hinweis auf Richtigkeit gegeben. Alle Teilnehmer arbeiten mit denselben Fragen und Stimuli, aber in unterschiedlicher Reihenfolge.

Die Proband:innen beantworten Fragen zu den jeweiligen Themen anhand von bereitgestellten Screenshots oder anhand von Links, welche zu einem interaktiven Beispiel eines Power-BI-Dashboards mit Nachhaltigkeitsinhalt führen. Aufgezeichnet wird die Antwortrichtigkeit (richtig oder falsch) sowie die benötigte Zeit zur Beantwortung der Frage.

4.3 Studienteilnehmer

Am Quasi-Experiment nehmen 42 Bachelorstudent:innen der FH Oberösterreich, Fakultät für Wirtschaft und Management, teil. Die Auswahl der Testpersonen erfolgt mittels Convenience Sampling aufgrund der Vorteile wie einfache Verfügbarkeit, geografische Nähe und hohe Rücklaufquoten. Die Teilnehmer:innen verfügen aufgrund ihrer Ausbildung über ein solides Grundverständnis für Zahlen und Reporting im Finanzbereich, was sie vergleichbar mit typischen externen Berichtsleser:innen macht und somit deren Meinung besser repräsentiert. Zudem bringen die berufsbegleitenden Studierenden bereits Arbeitserfahrung mit und können daher die internen Berichtsleser:innen repräsentieren (Vgl. Etikan, 2016, S. 1–3). Zusätzlich konnten die Studierenden im Zuge einer Lehrveranstaltung grundlegend auf die Inhalte und das Arbeiten mit den Nachhaltigkeitsdashboards vorbereitet werden.

Dennoch besteht ein Selection Bias, da die Studierenden alle in der Region Oberösterreich und Umgebung leben und eine ähnliche Ausbildung durchlaufen. Zudem gehören viele zur jungen Generation, insbesondere die Vollzeitstudent:innen sind in der Regel unter 25 Jahre alt. Obwohl sie nach einigen Kriterien nicht repräsentativ für die Gesamtheit der Berichtsleser:innen sind, ermöglicht das Experiment dennoch wichtige Einblicke in die relevanten Themen. (Vgl. Etikan, 2016, S. 1–3)

5 Auswertung

Zur besseren Einordnung der Ergebnisse und zur Schaffung eines Kontextes werden zunächst die demografischen Daten der Proband:innen analysiert. Das Experiment erreicht aufgrund der beschriebenen Auswahlstrategie eine Rücklaufquote von 100 %. Insgesamt werden 42 Bachelorstudierende getestet, wobei 24 weiblich und 18 männlich sind. Der Anteil an weiblichen Teilnehmenden überwiegt leicht, jedoch sind beide Gruppen insgesamt ausreichend vertreten.

Die Altersstruktur variiert stark je nach Zugehörigkeit zur Vollzeit- (VZ) oder berufsbegleitenden (BB) Gruppe. Die Altersspanne der Teilnehmenden liegt zwischen 20 und 49 Jahren, wobei das Gesamtdurchschnittsalter bei 24,6 Jahren liegt. Das Durchschnittsalter der VZ- und BB-Gruppen unterscheidet sich deutlich, wobei die VZ-Gruppe im Durchschnitt 21,3 Jahre alt ist, während die BB-Gruppe ein Durchschnittsalter von 28,2 Jahren aufweist. Die Altersgruppe der 20- bis 24-Jährigen ist dominant, was bei der späteren Analyse der Ergebnisse berücksichtigt werden sollte. Weitere Kontrollvariablen sind die allgemeine Vorerfahrung mit ESG-Themen ($\bar{x}$ 2,5; Skala 1-5) und die Erfahrung mit interaktiven Dashboards wie Power BI, SAC, usw. ($\bar{x}$ 2,1; Skala 1-5).

5.1 Darstellung von Ziel- und Schwellenwerten

Um das wichtige Thema der Ziel- und Schwellenwerte im Dashboard richtig darzustellen, werden drei Versionen zur Beurteilung in das Experiment eingebaut. Bei den erzielten Fehlerquoten in Tabelle 7 ist zunächst erkennbar, dass Inbar-Diagramme zur höchsten Anzahl richtiger Antworten führen. Zudem ist die erhöhte Komplexität durch die Manipulation bei allen drei Varianten feststellbar. Bei der niedrigen Komplexität bezieht sich die Aussage nur auf einzelne Jahre (geringere kognitive Last), während bei den Aussagen mit hoher Komplexität die Gesamtentwicklung der IST-Werte im Kontext der Vorgaben über alle Jahre hinweg bewertet werden muss. Die Anzahl an notwendigen aufeinanderfolgenden Schritten ist damit deutlich zu erkennen und spiegelt sich auch im Ergebnis wider.

Tabelle 7: Übersicht deskriptive Ergebnisse zu den Visualisierungsformen Ziel- und Schwellenwerte (Effektivität)

Manipulation Visua- lisierung	Antwortrichtigkeit in %	Manipulation Komplexität	Antwortrichtigkeit in %
Kombiniert (Säulen mit schwarzer Linie)	70,2%	Simple	78,0%
		Komplex	62,5%
Kombiniert (Säulen mit farbiger Linie)	69%	Simple	79,2%
		Komplex	58,9%
Inbar-Diagramm	83,3%	Simple	94,0%
		Komplex	72,6%

Um die Einflüsse zu modellieren, wurde eine logistische Regression berechnet. Die Grundlage zum Modell bildet Abbildung 4. Das Nagelkerke R^2 für die logistische Regression beläuft sich auf 0,126 und der Gesamtprozentsatz der Klassifizierungstabelle liegt bei 72,4%. Signifikant unterscheidet sich das Inbar-Diagramm vom kombinierten Diagramm (beide Farbvariationen), in der reinen Farbverwendung kann kein signifikanter Effekt festgestellt werden. Es zeigt sich allerdings, dass die Verwendung eines Inbar-Diagramms bei der Nutzer:innengruppe eine 2,4x bessere Antwort hervorruft ($\text{Exp}(B) = 0,421$ bzw. $0,448$), als die Verwendung des kombinierten Diagramms. Dies kann eventuell darin begründet werden, dass die Abweichung in einer richtig zu interpretierenden Farbe besser wirkt als eine Linie in Rot für die Schwelle unabhängig der Erreichung der Schwelle sowie, umgekehrt in Grün, für ein zu erreichendes Ziel. Da die Zielerreichung nicht durch die Farbe repräsentiert wird, müssen Teilnehmer das erst übersetzen und die Antwort kann nicht aus der Abbildung abgelesen werden. Das erhöht die inhärente kognitive Belastung und scheint damit zu schlechteren Ergebnissen zu führen. Auch der Effekt der Komplexität ist signifikant. Die einfache Aufgabe kann 2,7x besser beantwortet werden als die komplexe Antwort ($\text{Exp}(B) = 0,367$). Die Tabelle dazu ist dem Anhang zu entnehmen.

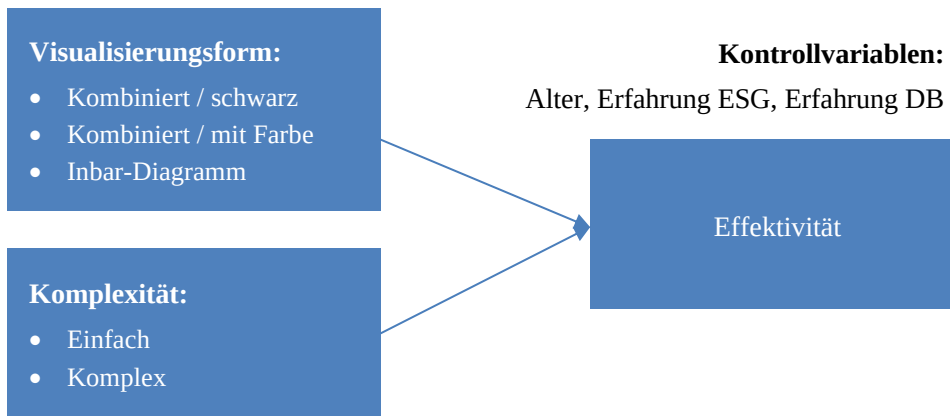


Abbildung 4: Übersicht Ergebnisse Inferenzstatistik zu den Visualisierungsformen Ziel- und Schwellenwerte (Effektivität)

Die Kontrollvariablen zeigen signifikante Effekte im Modell im Zusammenhang mit Alter und Erfahrung. Die beiden Variablen erhöhen die Wahrscheinlichkeit richtiger Antworten (Regressionskoeffizient Alter 0,06; Erfahrung ESG: 0,166) während eine hohe Erfahrung mit Power BI die Wahrscheinlichkeit reduziert (Regressionskoeffizient Power BI Erfahrung -0,226). Dies kann möglicherweise darauf zurückgeführt werden, dass Inbar-Diagramme nur sehr aufwendig in Power BI erstellt werden können und damit die Erfahrung und das Handling dieser Diagrammtypen nur in geringem Maße gegeben ist. Erfahrung ist einer der wichtigsten Einflussgrößen auf den kognitive Fit.

Als nächstes ist der Effekt dieser Darstellungsformen auf die Effizienz zu untersuchen. Bei diesen Aufgaben kann keine Unterscheidung nach Komplexität festgestellt werden, da die Zeit nur gesamt (also für einfache und komplexe Aufgaben gemeinsam) ausgezeichnet wurde.

Tabelle 8: Übersicht deskriptive Ergebnisse zu den Visualisierungsformen Ziel- und Schwellenwerte (Effizienz)

Manipulation	Visualisierung	Antwortzeit in Sekunden
Kombiniert / schwarz		32,9
Kombiniert / mit Farbe		46,1
Inbar-Diagramm		34,0

Im Hinblick auf die benötigte Zeit kann eine Verzögerung erkannt werden, sobald die Farben Rot und Grün für die Interpretation zur Verfügung stehen. Rot und Grün signalisieren aber wie bereits erwähnt nicht den Zielerreichungsgrad, sondern ob es sich um einen Ziel- oder einen Schwellenwert handelt. Die Übersetzung dieser Logik scheint zu einem höheren Zeitbedarf zu führen. Ob diese auch signifikant sind, wird mittels Regressionsanalyse überprüft.

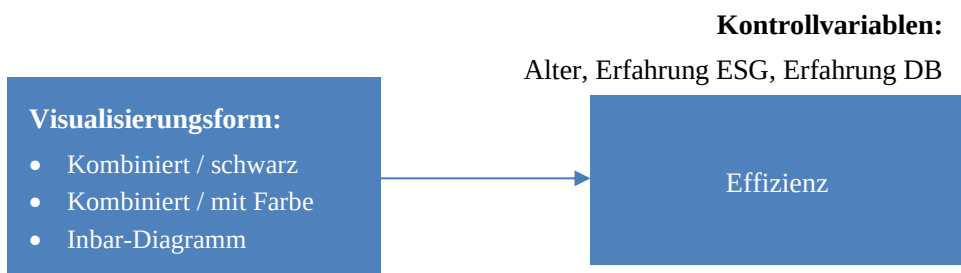


Abbildung 5: Übersicht Ergebnisse Inferenzstatistik zu den Visualisierungsformen Ziel- und Schwellenwerte (Effizienz)

Die Regressionsanalyse (für Tabelle siehe Anhang) untersucht den Effekt der Visualisierungsform sowie die Kontrollvariablen. Dazu müssen die Visualisierungsformen zuerst in Dummy Variablen übersetzt werden. Das R^2 , also der Grad der Erklärung durch die verwendeten Variablen ist mit 0,113 relativ gering. Lediglich der Effekt der Farbverwendung zeigt ein signifikantes Ergebnis (standardisierter Koeffizient $B = 0,300$, $p = 0,020$). Hier wird gezeigt, dass das Interpretieren der Farblogik vor allem bei wechselnder Grundlage (Schwellen- vs. Zielwerte) schwerfällt und damit die Effizienz signifikant negativ beeinflusst, also die benötigte Antwortzeit erhöht wird. Rot und Grün sollten damit für die Darstellung des Zielerreichungsgrades vorbehalten bleiben und nicht standardmäßig Ziel- und Schwellenwerte darstellen.

5.2 Darstellung von hierarchischen Zusammenhängen

Für die Darstellung von dimensionsübergreifenden Daten in einer Hierarchie werden die beiden Visualisierungstypen Analysebaum und Tabelle miteinander verglichen. Dazu wurde das Beispiel der Abfallart und Zuordnung im Test operationalisiert. Die Komplexität wurde bei diesen Aufgabestellungen mit der Anzahl an Hierarchiestufen und der Aufgabenschwierigkeit definiert. Bei einer niedrigen Schwierigkeit wurden die drei Hierarchiestufen aus Abbildung 2 verwendet und Single Choice Fragen gestellt. Bei der hohen Komplexität wurde zusätzlich die Hierarchiestufe „Standort“ hinzugefügt und die Proband:innen mussten Aussagen in der Form einer Multiple Choice Frage auf Richtigkeit überprüfen.

Tabelle 9: Übersicht deskriptive Ergebnisse Darstellung von Hierarchien (Effektivität)

Manipulation Visualisierung	Antwortrichtigkeit in %	Manipulation Komplexität	Antwortrichtigkeit in %
Entscheidungsbaum	85,7%	Einfach	88,1%
		Komplex	83,3%
Tabelle	94,5%	Einfach	95,2%
		Komplex	93,7%

Auffällig sind hierbei die niedrigen Fehlerquoten bei beiden Visualisierungstypen. Vor allem die Tabelle zeigt eine sehr hohe Effektivität, vermutlich da diese eine sehr geläufige Darstellungsform ist und damit auf hohe Erfahrungsniveaus der Berichtleser:innen trifft. Zwischen den beiden Komplexitätsstufen gibt es keine großen Unterschiede. Die Fehlerquoten deuten somit auf ein hohes Verständnis der Proband:innen für die beiden getesteten Visualisierungen, unabhängig der Komplexität der Aufgabenstellung, hin. Die Effekte werden wieder in einer logistischen Regression überprüft. Das Modell ist Abbildung 6 zu sehen.

Das Nagelkerke R^2 beläuft sich auf 0.142 und es der Gesamtprozentsatz der Klassifizierungstabelle liegt bei 89,3%. Der Unterschied zwischen den Komplexitätsstufen ist nicht signifikant, jener zwischen den beiden getesteten Visualisierungsoptionen schon. Die Tabelle führt im Modell zu 3,0x mehr richtigen Antworten im Vergleich zur Verwendung eines Entscheidungsbaumes (für die statistische Tabelle siehe Anhang).

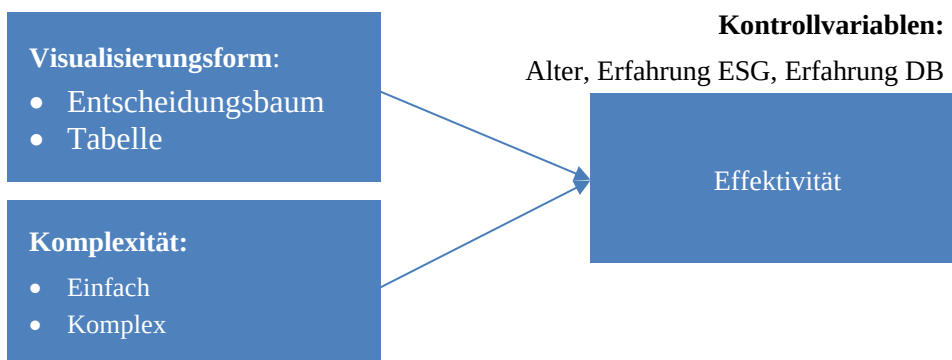


Abbildung 6: Übersicht Ergebnisse Inferenz zur Darstellung von Hierarchien (Effektivität)

In der nächsten Untersuchung wird die benötigte Zeit für richtige Antworten analysiert. Rein von den Mittelwerten kann darauf geschlossen werden, dass Entscheidungsbäume zu deutlich schnelleren Einsichten führen.

Tabelle 10: Übersicht deskriptive Ergebnisse Darstellung von Hierarchien (Effizienz)

Manipulation Visua- lisierung	Antwortzeit in Sekunden	Manipulation Kom- plexität	Antwortzeit in Sekunden
Entscheidungsbaum	41,2	Einfach	27,6
		Komplex	61,5
Tabelle	66,8	Einfach	44,2
		Komplex	90,7

Die ersten beiden Aufgaben mit niedriger Komplexität beanspruchten bei beiden Varianten wenig Zeit. Auffällig ist hingegen, dass die Bearbeitungszeit bei den Aufgaben mit hoher Komplexität sehr deutlich zugunsten des Analysebaums ausfällt. Vermutlich ist es einfacher, sich im Analysebaum mit vielen Hierarchieebenen zurechtzufinden, als bei einer Tabelle, die oft mit den +/- Symbolen auf- und zugeklappt werden muss, um einen Überblick zu behalten. Der Analysebaum kann hingegen durch einfache Klicks, schnell und intuitiv bedient werden. Um den Einfluss auf das Leseverhalten und damit einen Rückschluss auf eine möglicherweise Reduktion der kognitiven Belastung zu ziehen, wären weitere Untersuchungen, bspw. eine Eye-Tracking Untersuchung, hilfreich. Die statistische Prüfung mittels linearer Regression dieser Beobachtung erfolgt nachstehend.

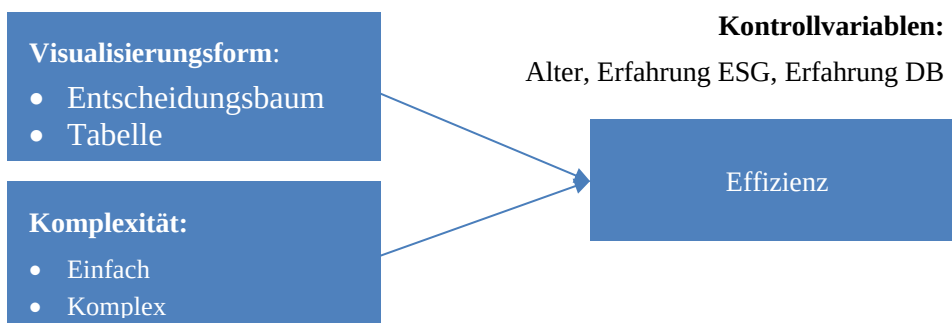


Abbildung 7: Übersicht Ergebnisse Inferenz zur Darstellung von Hierarchien (Effizienz)

Das R^2 liegt bei 0.625. Damit kann sehr viel von der beobachteten Varianz in der benötigten Zeit auf die in das Modell eingeschlossenen Variablen zurückgeführt werden. Beide Variablen, die Visualisierungsform sowie die Komplexität führen zu einer signifikanten Beeinflussung der Zeit. Während Entscheidungsbäume im Vergleich zu Tabellen zu einer Reduktion der benötigten Antwortzeit führen (standardisierter Koeffizient Beta = -0,365, $p = 0,000$) führt eine Erhöhung der Komplexität zu einer Erhöhung der Antwortzeit (standardisierter Koeffizient Beta = 0,652, $p = 0,000$). Die Kontrollvariablen Alter und Erfahrung mit ESG spielen keine Rolle. Lediglich die Erfahrung mit Dashboards in Power BI führt zu einer Reduktion der benötigten Zeit (standardisierter Koeffizient Beta = -0.149, $p = 0,008$). Entscheidungsbäume sind in dieser Umgebung schnell erstellt und ein beliebtes Visual. Daher lässt sich dieses Erkenntnis mit hoher Wahrscheinlichkeit erneut durch die Vorerfahrung und die Verbreitung der Visualisierungsform interaktiver Entscheidungsbaum erklären.

5.3 Interaktionseffekt für Nachhaltigkeitsdashboards

Um die Interaktionstechnik Filter zu bewerten, wurden abermals zwei Alternativen erstellt: In einer Version sind alle Filter im Kopfbereich des Dashboards ersichtlich und umzustellen, während in der nächsten Variante ein Filterpanel bei Bedarf aufgerufen und verwendet werden kann. Damit sind die Filter eventuell nicht immer sichtbar, der dafür zu reservierende Bereich am Bildschirm ist allerdings deutlich kleiner. Die Komplexität bezieht sich bei dieser Aufgabe auf die Häufigkeit, mit der die Filter bzw. das Filterpanel für dieselbe Aufgabestellung bedient werden müssen. Welche Effekte diese Design Optionen auf Effizienz und Effektivität bei unterschiedlicher Komplexität haben, wird in diesem Abschnitt untersucht.

Tabelle 11: Übersicht deskriptive Ergebnisse Interaktion Filter (Effektivität)

Manipulation Inter- aktion	Antwortrichtigkeit in %	Manipulation Kom- plexität	Antwortrichtigkeit in %
Filter im Header	68%	Einfach	76%
		Komplex	60%
Filterpanel	72,5%	Einfach	88%
		Komplex	57%

Die Auswertung zeigt, dass die Fehlerquote bei hoher Anzahl an Änderung (hohe Komplexität) der Filter in derselben Aufgabe deutlich höher ist. Die Proband:innen müssen bei dieser Aufgabe mehrere Werte miteinander vergleichen und sich die Zahlen im Kopf merken sowie genau lesen, was gefordert ist. Dies führt zu einem sehr starken Anstieg der inhärenten kognitive Belastung. Die möglichen Fehlerquellen sind deutlich größer als bei der niedrigen Komplexität, bei welcher nur ein Wert abgelesen werden muss. Dieser Effekt ist auch deutlich größer, wird ein Filterpanel anstelle eines Filterbereichs verwendet. Das Filterpanel muss dabei mehrmals geöffnet und geschlossen werden und die durch geänderte Filtereinstellungen verursachten Veränderung im Dashboard, sind schlechter überblickbar. Bei einfachen Anpassungen besticht klar das Filterpanel.

Das Nagelkerke R^2 beläuft sich auf 0.122 und der Gesamtprozentsatz der Klassifizierungstabelle liegt bei 70,2%. Der Unterschied zwischen Filter und Filterpanel im Hinblick auf Effektivität ist nicht signifikant. Lediglich der Faktor Komplexität weist einen interpretierbaren Effekt auf. Je höher die Komplexität, desto schlechter die Antwortrichtigkeit. Keine der getesteten Kontrollvariablen (Alter, Erfahrung mit ESG, Erfahrung mit Dashboard in Power BI) ist relevant. Die Tabelle zum den Koeffizienten kann dem Anhang entnommen werden.

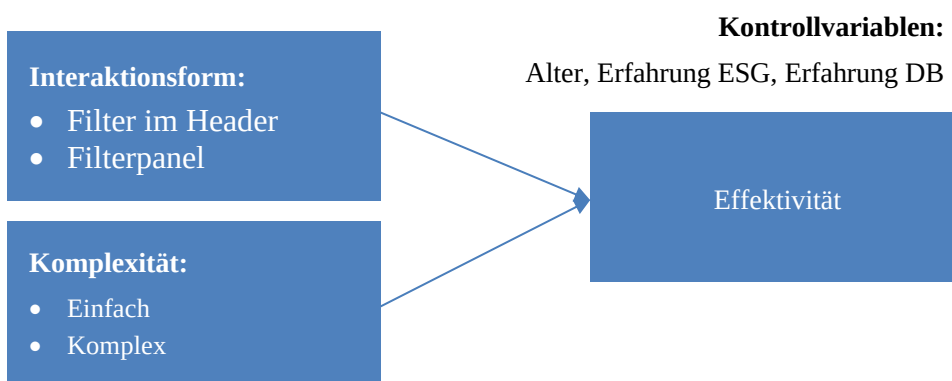


Abbildung 8: Übersicht Inferenzstatistik zur Darstellung der Interaktion Filter (Effektivität)

Ein anderes Bild zeigt sich bei der Analyse der Effizienz. Die Proband:innen benötigten mit dem Filterpanel durchschnittlich weniger Zeit für die Beantwortung.

Tabelle 12: Übersicht deskriptive Ergebnisse der Interaktion Filter (Effizienz)

Manipulation Interaktion	Antwortzeit in Sekunden	Manipulation Komplexität	Antwortzeit in Sekunden
Filter im Header	80,4	Einfach	53,6
		Komplex	114,7
Filterpanel	50,7	Einfach	37,9
		Komplex	70,5

Das R^2 des Modells liegt bei 0,517 und erlaubt damit etwas mehr als die Hälfte der Varianz in der Antwortzeit zu erklären. Der standardisierte Koeffizient Beta für Filter im Header liegt bei 0,352 und jener für Komplexität bei 0,596. Bezugnehmend auf die Effizienz liegt also das Filterpanel in beiden Komplexitätsausprägungen klar an erster Stelle. Ein Filterpanel ist damit schneller, da alle Filtereinstellungen auf einmal abgeändert werden können. Die Kontrollvariablen sind nicht signifikant.

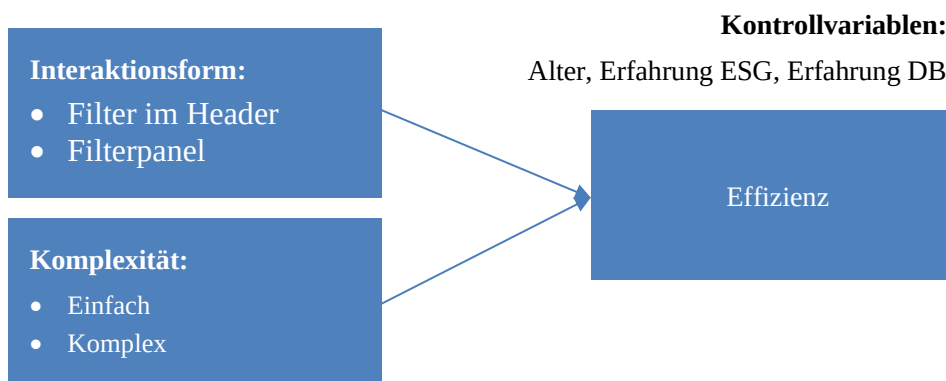


Abbildung 9: Übersicht Inferenzstatistik Darstellung der Interaktion Filter (Effizienz)

6 Diskussion und Fazit

Mit der Einführung der ESRS-Berichtspflichten sind viele Unternehmen im kommenden Jahr berichtspflichtig. Die Pflicht zur Nachhaltigkeitsberichterstattung ist für die Unternehmen mit einem hohen Aufwand und viel Unsicherheit verbuchen. Dies bestätigt auch die Green Controlling-Studie des Internationalen Controllervereins aus dem Jahr 2022, welche zusammenfasst, dass Unternehmen mit den Daten noch unzufrieden sind, unter anderem weil der Arbeitsaufwand hier sehr hoch und die Automatisierung noch gering ist (Kämmeler-Burrak u. a., 2022, S. 32). Erschwerend kommt hinzu, dass ESG-relevante Kennzahlen noch einen geringen Bekanntheitsgrad aufweisen und damit Lese- und Interpretationsaufwand für den Berichtleser überproportional höher ist als bei beispielsweise vergleichbaren und etablierten KPIs aus dem Finanzbereich (Reisinger, 2023, S. 37 ff.). Obwohl das vor CSRD freiwillige Reporting selten von Unternehmen genutzt wurde, lässt sich feststellen, dass gleichzeitig nicht nur die Umwelt, sondern auch das Unternehmen deutlich von diesem strategischen Ansatz profitieren konnte, sofern dieses richtig umgesetzt und strategisch verankert wurde (Dienes et al, 2026, S. 174).

Die Entwicklung von Maßnahmen, um die Nachhaltigkeitsziele intern und extern zu kommunizieren ist von großer Bedeutung. Dashboards können ein geeignetes Instrument dafür sein. Wie man mit Dashboards den Spezifika der Nachhaltigkeitszahlen begegnet, wurde in diesem vorliegenden Beitrag an geeigneten Stellen untersucht und bisherige Literatur ergänzt. Speziell haben wir uns mit der wichtigen Frage auseinandergesetzt, wie man Ziel- und Schwellenwerte sinnvoll in Dashboards integrieren bzw. wie man die Kommunikation mehrerer Auswertungsdimensionen an die Nutzer:innen optimieren kann. Zusätzlich, um das anfängliche Ausmaß an Informationen einschränken zu können, stellten wir uns der Frage nach dem optimalen Filterkonzept im Nachhaltigkeitsreporting. All

diese Fragen werden mithilfe der Cognitive Fit Theory bzw. der Cognitive Load Theory in Einklang gebracht und entsprechend diskutiert.

Darstellung von Ziel- und Schwellenwerten:

Unternehmen sollten im ökologischen Bereich gewisse Planwerte überschreiten (Recyclingquoten, Reduktion Emissionen, Anteil Energie erneuerbar) bzw. unterschreiten (CO₂ / Schadstoff Emissionen). Ähnlich dazu gibt im Finanzbereich die Darstellung von Erlösen und Kosten im Vergleich zum Budget, welche häufig durch Inbar-Diagramme visualisiert werden. Dieser Diagrammtyp benötigt jedoch zur richtigen und schnellen Interpretation ein gewisses Level an spezifischer Vorerfahrung. (Vgl. Eisl et al., 2018, S. 80–82) Als Alternative werden immer wieder kombinierte Säulen- und Liniendiagramme diskutiert, insbesondere für externe Berichtsempfänger:innen mit wenig Vorerfahrung.

Da man bei Nachhaltigkeitsdashboards von einer breiten Masse an Nutzer:innen ausgeht, ist womöglich die spezifische Erfahrung mit Inbar-Diagrammen nicht immer gegeben. Die Testgruppe hatte in diesem Zusammenhang auch keine nennenswerten Vorteile und einen eher unterdurchschnittlichen Vorkenntnisstand zum Thema Dashboards und Diagrammtypen. Für diesen Diagrammtyp wird das klassische Farbverständnis von Rot und Grün erweitert. Grün symbolisiert nicht nur positive Abweichungen, sondern auch die Zielwertlinie, während Rot negative Abweichungen und Schwellenwertlinien darstellt, die nicht überschritten werden sollen.

Die Ergebnisse zeigen, dass eine rote/grüne Linie im Vergleich zu einer einheitlichen schwarzen Linie signifikante Nachteile in der Antwortzeit mit sich bringt. Wenn Schwellen im Bericht verwendet werden, ist es ratsam, Begriffsdefinitionen und Farbkonzepte vorab ausreichend zu kommunizieren. Hier schien es, gerade beim Wechsel zwischen Ziel- und Schwellenwerten, zu einer längeren Bearbeitungszeit gekommen zu sein. Die zusätzliche Aufgabe Rot als Schwelle und Grün als Ziel zu identifizieren, war unüblich und damit gegen die gewohnte Erfahrung und führte damit zu zeitlichen Verzögerungen.

Bei Inbar Diagrammen ist die Farblogik der Abweichungsbalken an den Zielerreichungsgrad angepasst. Damit steht einer üblichen Interpretation nichts im Wege. In beiden getesteten Kategorien (Effizienz & Effektivität) ist deshalb das Inbar-Diagramm die bessere Wahl.

Tabelle 13: Übersicht Ergebnisse Hypothesentests zur Visualisierung von Ziel- & Schwellenwerten

#	Beschreibung	Ergebnis
H1	Die Visualisierungsform (kombiniertes Diagramm Schwarz vs. Kombiniertes Diagramm Farbe vs. Inbar-Diagramm) hat keinen Einfluss auf a) die <i>Effektivität</i> , b) die <i>Effizienz</i> bei der Informationsverarbeitung.	a) widerlegt : Empfehlung Inbar-Diagramm b) widerlegt : Vermeidung von farbigen Linien
H2	Die Komplexität (einfach vs. komplex) der Aufgabenstellung hat keinen Einfluss auf die a) <i>Effektivität</i> bei der Informationsverarbeitung.	a) widerlegt : Aufgaben, mit geringerer inhärenter Belastung erhöhen die Richtigkeit

Darstellung von hierarchischen Informationen:

Das Experiment befasst sich mit der Darstellung umfangreicher Hierarchien, die in den Themenbereichen Energie- oder Abfallmanagement häufig auftreten können. Hierfür wurden zwei Diagrammtypen miteinander verglichen: die

klassische Tabelle und der Entscheidungsbaum. Tabellen legen den Fokus auf präzise Zahlenwerte und Vergleichswerte, während Analysebäume die Hierarchieebenen übersichtlicher darstellen und zusätzliche Informationen über Tooltips bereitstellen können.

Die Ergebnisse des Experiments zeigen, dass die Entscheidungsbäume zu einer schnelleren Antwort führen. Der Analysebaum ermöglicht ein schnelles Ändern der Hierarchieebenen, was den Nutzer:innen erlaubt, die Daten nach ihren Bedürfnissen zu strukturieren. Allerdings schneiden Entscheidungsbäume im Zusammenhang mit der Qualität der Antworten (Richtigkeit) etwas schlechter ab. Die Manipulation der Komplexität führt auch in diesem Test zu einer Erhöhung der Antwortzeit sowie zu einer Reduktion der Antwortrichtigkeit.

Tabelle 14: Übersicht Ergebnisse Hypothesentests zur Visualisierung von hierarchischen Informationen

#	Beschreibung	Ergebnis
H3	Die Visualisierungsform (Entscheidungsbaum vs. Tabelle) hat keinen Einfluss auf a) die <i>Effektivität</i> , b) die <i>Effizienz</i> bei der Informationsverarbeitung.	a) widerlegt : die Tabelle führt zu mehr richtigen Antworten b) widerlegt : Entscheidungsbäume führen bei richtigen Antworten zu schnelleren Antworten
H4	Die Komplexität (einfach vs. komplex) der Aufgabenstellung hat keinen Einfluss auf a) die <i>Effektivität</i> b) die <i>Effizienz</i> bei der Informationsverarbeitung.	a) & b) widerlegt : Aufgaben, mit geringerer inhärenter Belastung, erhöhen die Antwortqualität

Darstellung von Filtern:

Bei den komplexeren Fragen zum Themenbereich Filtergestaltung mussten die Proband:innen die Filter mehrfach umstellen, sich die Werte merken und miteinander vergleichen. Daraus lässt sich eine hohe inhärenter Belastung durch die Aufgabe ableiten. Das wird im Experiment auch deutlich durch signifikante Verschlechterung der Testvariablen sichtbar. Ein Unterschied in der Gestaltung der Filter kann in der benötigten Antwortzeit, nicht aber in der Antwortrichtigkeit festgestellt werden. Filterpanels führen zu schnelleren Antworten, die Antwortrichtigkeit unterscheidet sich kaum.

Diese Gestaltung der Filter sowie die Verwendung von Diagrammtypen sollte im Unternehmen im Voraus geklärt und nach Möglichkeit vereinheitlicht werden, um die im Leitfaden für Exzellenz im Reporting Design erwähnten Vorteile für Schemen im Langzeitgedächtnis zu erzielen. (Vgl. Eisl et al., 2018, S. 28–31) Bei der Einführung von neuartigen Interaktionsmöglichkeiten und/oder Diagrammtypen sind zusätzlich eine entsprechende Schulung für die Funktionsweise des verwendeten Programms (z. B. Power BI) und des spezifischen Elements unerlässlich.

Tabelle 15: Übersicht Ergebnisse Hypothesentests zur Darstellung von Filtern

#	Beschreibung	Ergebnis
H5	Die Gestaltung der Filtermöglichkeit (Filter im Kopf vs. Filterpanel) hat keinen Einfluss auf a) die <i>Effektivität</i> , b) die <i>Effizienz</i> , bei der Informationsverarbeitung.	a) beibehalten b) widerlegt : Filterpanels führen zu schnelleren Einsichten
H6	Die Komplexität (einfach vs. komplex) der Aufgabenstellung hat keinen Einfluss auf a) die <i>Effektivität</i> b) die <i>Effizienz</i> bei der Informationsverarbeitung.	a) & b) widerlegt : Aufgaben, mit geringerer inhärenter Belastung, erhöhen die Antwortqualität

7 Ausblick und weiterer Forschungsbedarf

Welche Nachhaltigkeitskennzahlen sich für die Steuerung von Unternehmen in bestimmten Branchen in besonderem Maße eignen, wird sich in den nächsten Jahren herauskristallisieren. Mit Sicherheit werden die ersten nach ESRS veröffentlichten Berichte diesbezüglich genauestens analysiert. Auch insbesondere die Erweiterung der ESRS durch sektorspezifische Standards oder die Darstellung der Standards für KMUs könnten Gegenstand weiterer Forschung sein, sobald sie von der EU veröffentlicht werden.

Neben der Aufbereitung und richtigen Berechnung der Kennzahlen spielt die Datenbeschaffung und Qualität eine besonders große Rolle. Aktuell sind die Anforderungen für viele Unternehmen noch zu hoch und Potenziale der Automatisierung bleiben noch weitestgehend ungenutzt. Um die Prozesse zur Datenerfassung zu standardisieren und die Datenqualität zu gewährleisten, werden oftmals Datenerfassungsblätter (Data Capturing Sheets) verwendet. (Vgl. Sailer & Lösken, 2023, S. 10–12) Dieser Prozess könnte durch Automatisierungsmöglichkeiten wie Robotic Process Automation (RPA) oder Machine Learning (ML) beschleunigt und weniger fehleranfällig gestaltet werden. Der Einsatz von ML scheint für die Nachhaltigkeitsberichterstattung mit ihrer hohen Komplexität als ersehnter Ausweg. Wie genau die Unterstützung aussehen kann und ob bzw. wie dies in einem prüfungspflichtigen Bericht integriert werden kann, ist aktuell noch Gegenstand der Forschung bzw. der Entwicklung im Unternehmenskontext.

Technologische Einschränkungen (wie die Bildschirmgröße) wurden im Experiment nicht berücksichtigt. Es könnten daher spezielle Nachhaltigkeitsdashboards für die Nutzung auf dem Smartphone entwickelt und genauer untersucht werden. Kulturelle Unterschiede im Nachhaltigkeitsbereich könnten ebenfalls thematisiert werden, sowohl im europäischen Vergleich als auch im Vergleich zu Amerika oder Asien mit ihren jeweiligen (freiwilligen) Nachhaltigkeitsstandards. Darüber hinaus könnten in weiteren Experimenten gezielt verschiedene Gruppen, wie interne und externe Berichtsempfänger:innen, getestet werden, um so mögliche Unterschiede herauszuarbeiten. Eine Langzeitbetrachtung mit denselben Gruppen, bei der die Proband:innen mit wenig Vorerfahrung starten, könnte ebenfalls wertvolle Erkenntnisse liefern. Dabei könnte man untersuchen, inwieweit sich die Vorerfahrung mit den einzelnen Visualisierungsarten, Begrifflichkeiten und Interaktionskonzepten als signifikanter Unterschied herauskristallisiert. Das Thema Nachhaltigkeit wird sowohl die Praxis als auch die Forschung in den kommenden Jahren stark beschäftigen. Dieser Beitrag bietet einen Denkanstoß in diesem dynamischen und bedeutenden Themenfeld.

Zusätzlich gibt es neben den getesteten Visualisierungs- und Interaktionstechniken noch weitere Möglichkeiten. Unter anderem auch solche, welche noch nicht empirisch ausreichend untersucht wurden. So ist beispielsweise die Möglichkeit und die Gestaltung von Tooltips oder die Verwendung von Sunburst oder Treemap-Visualisierungen zur Visualisierung von hierarchischen Gliederungen in einem breiten und diversen Zielpublikum, wie jenem des Nachhaltigkeitsberichts, noch offen. Auch die empirische Erhebung an sich braucht mehr Varianz. Es ist nicht nur wertvoll zu wissen, bei welchen Formen der Visualisierung und Interaktion mehr oder weniger Zeitbedarf oder Fehler entstehen, empirische Erhebungen zur Anwendung (z.B. durch Analyse von veröffentlichten Nachhaltigkeitsberichten) oder Experimente zur direkten Erfassung der kognitiven Belastung (z.B. durch Eyetracking Experimente) wären für die Weiterentwicklung des Forschungsfeldes hilfreich.

Literaturverzeichnis

- Atkinson, R.C., & Shiffring, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation*. pp. 89- 195. New York: Academic Press.
- Baumüller, J., Eisl, C. & Leitner-Hanetseder, S. (2023). Neue KPI für die Unternehmenssteuerung durch CSRD und ESRS: Die erweiterte Nachhaltigkeits-Transparenz für Unternehmen bedingt eine Veränderung in diesen Steuerungssystemen. *REthinking:Finance*(4), 4–10.
- Brehmer, M., & Munzner, T. (2013). A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), 2376–2385.
- Charness, G., Gneezy, U. & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1–8.
- Delegierte Verordnung (EU) 2023/2772 der Kommission vom 31. Juli 2023 zur Ergänzung der Richtlinie 2013/34/EU des Europäischen Parlaments und des Rates durch Standards für die Nachhaltigkeitsberichterstattung, Amtsblatt der Europäischen Union (2023 & i.d.F.v. 22.12.2023).
- Dienes, D., Sassen, R., Fischer, J. (2016), What are the drivers of sustainability reporting? A systematic review, *Sustainability Accounting, Management and Policy Journal*, Vol. 7, No. 2 pp. 154 – 189.
- EFRAG. (2023). Draft EFRAG IG 3: Detailed ESRS datapoints implementation guidance. EFRAG.
- Eisl, C., Hofer, P., Losbichler, H. (2019). *Grundlagen der finanziellen Unternehmensführung*, Band IV: Controlling, 4. Überarbeitete Auflage, Linde Verlag, Wien.
- Eisl, C., Perkhofer, L., Hofer, P. & Losbichler, H. (Hrsg.). (2018). *Exzellenz im Reporting Design: Leitfaden für messbar bessere Berichte*. Haufe.
- Elkington, J. (1998), "ACCOUNTING FOR THE TRIPLE BOTTOM LINE", *Measuring Business Excellence*, Vol. 2 No. 3, pp. 18-22.
- Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1. <https://doi.org/10.11648/j.ajtas.20160501.11>
- Falschlunger, L., Lehner, O., und Treiblmaier, H., 2016, InfoVis: The Impact of Information Overload on Decision Making Outcome in High Complexity Settings. In: *Special Interest Group on Human-Computer Interaction, Proceedings of the 15th annual Pre-ICIS Workshop on HCI Research in MIS (1- 6, Paper 3)*. AIS Electronic Library: Association for Information Systems.
- Hamed et al., 2022, The impact of introducing new regulations on the quality of CSR reporting: Evidence from the UK, in: *Journal of International Accounting, Auditing and Taxation*, Vol. 46, SS. 100444-100462
- Heer, J. & Shneiderman, B. (2012). Interactive dynamics for visual analysis. *Communications of the ACM*, 55(4), 45–54.
- Hofer, P., Perkhofer, L. & Mayr, A. (2023). Interactive Big Data Visualizations: Potential for Management Reporting. In I. Keimer & U. Egle (Hrsg.), *The Digitalization of Management Accounting: Use Cases from Theory and Practice* (1st edition 2023, S. 143–169). Springer Fachmedien Wiesbaden.
- Hofinger, J. & Röhm, T. (2023). ESG-Controlling: von Deckungsbeitrag 1, 2, 3 zu Scope 1, 2, 3? *CFOaktuell*, 2023(3).
- Isil, O. & Hernke, M. (2012). The Triple Bottom Line: A Critical Review from a Transdisciplinary Perspective. In: *Business Strategy and the Environment*, Vol. 26., pp. 1235-1251
- Islam, M. & Jin, S. (2019). An Overview of Data Visualization. In *IEEE Conference (Hrsg.), 2019 International Conference on Information Science and Communications Technologies (ICISCT)* (S. 1–7). IEEE.
- Klovienè, L., Speziale, M., 2014, Sustainability Reporting as a Challenge for Performance Measurement: Literature Review, In: *Economics in Business*, Vol. 26, SS. 44-53
- LimeSurvey. (2024). LimeSurvey. <https://www.limesurvey.org/de-ch>

- Locke, E., und Latham, G., 2006, New Directions in Goal-Setting Theory, In: Current direction in psychological science, Vol. 15, No. 5, SS. 265-268
- Lösken und Sailer, 2023, ESG-Reporting in der Praxis: softwaregestützt und prüfungssicher, In: Rethinking Finance, August 2023, SS. 10-16
- Magdalena, R., Sensuse, D., Ruldeviyani, Y., Bernardo, C., 2019, Methods to Enhance the Utilization of Business Intelligence Dashboard by Integration of Evaluation and User Testing, In: Proceedings of the 3rd International Conference on Informatics and Computational Sciences (ICICoS), S. 1-6
- Moko, A., Victor-Ikoh, M., & Okardi, B. (2023). Information Overload: A Conceptual Model. In European Journal of Computer Science and Information Technology, 11(5), S. 19–29. <https://doi.org/10.37745/ejcsit.2013/vol11n51929>
- Myers, M. (2024, 18. April). Decomposition tree - Power BI. <https://learn.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-decomposition-tree>
- Ortner, R., Legény, R. & Schmutterer, B. (2022). ESG als strategisches Leit- und Steuerungssystem: Wie die nachhaltige Transformation gelingt. CFOaktuell, 2022(5).
- Pariante-Martinez, B., Gonzalez-Rodriguez, M., Fernandez-Lanvin, D. & Andres-Suarez, J. de (2016). Measuring the role of age in user performance during interaction with computers. Universal Access in the Information Society, 15(2), 237–247.
- Patloch, T. (2019). Dashboard-Erstellung – Tableau und Power BI im Vergleich: Ein Praxisvergleich der beiden führenden Dashboarding-Tools aus Fachanwendersicht. CFOaktuell(März), 69–72.
- Perkhofer, L., 2019, A cognitive load-theoretic framework for Information Visualization; In: Proceedings fo the Finance, Risk, and Accounting Perspectives, S. 9-25
- Perkhofer, L., Hofer, P., Walchshofer, C., Plank, T., Jetter, C. (2019). Interactive visualization of big data in the field of accounting A survey of current practice and potential barriers for adoption, In: Journal of Applied Accounting Research, Vol. 20, No. 4., S. 497-525
- Perkhofer, L., Walchshofer, C., Hofer, P. (2020). Does design matter when visualizing Big Data? An empirical study to investigate the effect of visualization type and interaction use. In: Journal of Management Control, 31. S. 55-95.
- Picard, N., Feiner, A., Brach, J., Gehrke, N. & Wallek, C. (2022). PWC Studie EU-Taxonomie 2022: Nachhaltigkeitsberichterstattung im Wandel. PWC.
- RICHTLINIE (EU) 2022/2464 DES EUROPÄISCHEN PARLAMENTS UND DES RATES zur Änderung der Verordnung (EU) Nr. 537/2014 und der Richtlinien 2004/109/EG, 2006/43/EG und 2013/34/EU hinsichtlich der Nachhaltigkeitsberichterstattung von Unternehmen (Corporate Sustainability Reporting Directive - CSRD), Amtsblatt der Europäischen Union (2022).
- Sailer, U. & Lösken, B. (2023). ESG-Reporting in der Praxis: softwaregestützt und prüfungssicher. REthinking:Finance, 2023(4), 10–16.
- Sattler, P. & Schentler, P. (2022). Von der Kür zur Pflicht: Wie Nachhaltigkeit in die Unternehmenssteuerung integriert werden kann (und zukünftig auch muss). CFOaktuell, 2022(2).
- Shneiderman, B. (1996). The eye have it: A task by data type taxonomy for information visualization. Proceedings of the IEEE Symposium on Visual Languages, 03-06 September 1996, Boulder, CO, USA, S. 336-343.
- Schönauer, K., Ratzinger, V. & Luggauer, M. (2022). Aktuelle Entwicklungen in der Nachhaltigkeitsberichterstattung: Überblick zu den derzeitigen und künftigen Anforderungen. CFOaktuell(3), 91–94.
- Speier, C. (2006) The influence of information presentation formats on complex task decision-making performance, International Journal of HumanComputer Studies, 64, 11, 1115-1131
- Stehle, A., Harms, S. G. & Sargl, M. (2023). Steuerung oder Compliance? Was Controller im Green Controlling tun. CFOaktuell, 2023(1), 32–34.
- Sweller, J. (1988). Cognitive load during problem solving: effects on learning. Cognitive Science, 12, 257- 285.

- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load . *Educational Psychology Review*, 22(2), 123- 138.
- Tichy, G., Fuhrmann, K., Aschauer, E., Benetik, C., Carius, M., Diener, M., Eberhard, A., Eckhardt, F., Harrer, C., Kerbl, G., Krall, K., Kühmayer, T., Lorenzen, T., Maiwald, J., Pamperl, E., Paul, F., Pichler, A., Regner, L., Regnery, M., . . . Winkelbauer, B. (Hrsg.). (2023). *Handbuch ESG-Berichterstattung* (1. Auflage 2023). Linde Verlag Ges.m.b.H.
- Tufte, E. (2001). *The Visual Display of Quantitative Information*. Second Edition, Graph Press LLC. Cheshire, Connecticut, USA.
- Velcu-Laitinen, O., Yigitbasioglu, 2021, The Use of Dashboards in Performance Management: Evidence from Sales Managers, In: *The International Journal of Digital Accounting Research*, Vol. 12, S. 39-58
- Vessey, I. (1991). Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature *Decision Sciences*, 22(2), 219–240. <https://doi.org/10.1111/j.1540-5915.1991.tb00344.x>
- Waniczek, M. (2021). Doing well by doing good – Nachhaltig(keit) steuern. *CFOaktuell*, 2021(März), 51–54.
- Yigitbasioglu, O. M., Velcu, O. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, 13(1), 41–59
- Zhu, Y. (2023). A Review of Complexity Metrics for Data Visualization. In *IEEE Conference (Hrsg.)*, 2023 27th International Conference Information Visualisation (IV) (S. 131–135). IEEE.

Anhang

Tabellen zur Gestaltung von Ziel- & Schwellenwerten

		Regressions- koeffizientB	Standardfehler	Wald	df	Sig.	Exp(B)
Schritt 1 ^a	Kombi_Schwarz			13,361	2	,001	
	Kombi_Farbe	-,873	,247	12,454	1	<,001	,418
	Inbar	-,811	,248	10,705	1	,001	,444
	Komplexität	1,013	,167	36,930	1	<,001	2,755
	Geschlecht	,535	,185	8,359	1	,004	1,708
	Alter	,082	,023	12,751	1	<,001	1,086
	ErfahrungESG	,130	,083	2,436	1	,119	1,139
	ErfahrungDB	-,186	,078	5,769	1	,016	,830
	Konstante	-,994	,768	1,672	1	,196	,370

a. In Schritt 1 eingegebene Variablen: Vergleichswert, Komplexität, Geschlecht, Alter, ErfahrungESG, ErfahrungDB.

		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten		
Modell		RegressionskoeffizientB	Std.-Fehler	Beta	T	Sig.
1	(Konstante)	31,946	14,739		2,168	,034
	Dummy_Inbar	,667	5,882	,014	,113	,910
	Dummy_Farbe	12,872	5,405	,300	2,381	,020
	Alter	-,137	,350	-,051	-,391	,697
	ErfahrungESG	-,293	2,398	-,016	-,122	,903
	ErfahrungDB	3,062	2,555	,140	1,198	,235

a. Abhängige Variable: Effizienz

Tabellen zur Gestaltung von Hierarchien

		Regressions- koeffizientB	Standardfehler	Wald	df	Sig.	Exp(B)
Schritt 1 ^a	Typ	1,115	,399	7,809	1	,005	3,049
	Komplexität	,385	,457	,712	1	,399	1,470
	Alter	,008	,035	,055	1	,814	1,008
	Erfahrung_ESG	,749	,215	12,158	1	<,001	2,115
	Erfahrung_DB	-,297	,172	2,974	1	,085	,743
	Konstante	,386	1,182	,107	1	,744	1,472

a. In Schritt 1 eingegebene Variablen: Typ, Komplexität, Alter, Erfahrung_ESG, Erfahrung_DB.

		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten		
Modell		Regressions- koeffizientB	Std.-Fehler	Beta	T	Sig.
1	(Konstante)	76,282	11,277		6,764	<,001
	Typ	22,778	3,330	,365	6,840	<,001
	Komplexität	-40,672	3,331	-,652	-12,210	<,001
	Alter	-,132	,303	-,026	-,436	,664
	Erfahrung_ESG	,539	1,663	,019	,324	,746
	Erfahrung_DB	-4,372	1,615	-,149	-2,707	,008

a. Abhängige Variable: Zeit

Tabellen zur Gestaltung von Filtern

		Regressions- koeffizientB	Standardfehler	Wald	df	Sig.	Exp(B)
Schritt 1 ^a	Filter	-,249	,354	,497	1	,481	,779
	Komplexität	-1,215	,365	11,064	1	<,001	,297
	Alter	,008	,034	,060	1	,806	1,009
	Erfahrung_ESG	,267	,181	2,174	1	,140	1,306
	Erfahrung_DB	-,188	,168	1,244	1	,265	,829
	Konstante	1,213	1,217	,994	1	,319	3,365

a. In Schritt 1 eingegebene Variablen: Filter, Komplexität, Alter, Erfahrung_ESG, Erfahrung_DB.

		Nicht standardisierte Koef- fizienten		Standardisierte Koeffizienten		
Modell		Regressions- koeffizientB	Std.-Fehler	Beta	T	Sig.
1	(Konstante)	6,348	17,340		,366	,715
	Filter	27,211	5,105	,352	5,330	<,001
	Komplexität	46,760	5,165	,596	9,054	<,001
	Alter	,760	,476	,119	1,596	,113
	Erfahrung_ESG	4,171	2,548	,120	1,637	,104
	Erfahrung_DB	-1,493	2,473	-,041	-,604	,547

a. Abhängige Variable: Zeit

Einsatz von Generativer KI in Controlling und Finance: Bedarf nach einer AI Governance?

Extended Abstract

Prof. Dr. rer. pol. Thomas Rautenstrauch

Ostschweizer Fachhochschule, Institut für Finance und Law, St.Gallen, E-Mail: thomas.rautenstrauch@ost.ch

Anja Lindenmann, BSc

Ostschweizer Fachhochschule, Institut für Finance und Law, St.Gallen, E-Mail: anja.lindenmann@ost.ch

Abstract

Der Beitrag befasst sich mit dem Einsatz von generativer KI in den Bereichen Controlling und Finance und präsentiert dazu die Ergebnisse einer Online-Befragung unter Fachleuten dieser Anwendungsdomäne. Deren Ergebnisse zeigen, dass generative KI in Controlling und Finance angekommen ist und weist darüber hinaus auf den Bedarf einer umfassenden AI Governance hin, um Risiken und Grenzen des KI-Einsatzes besser zu managen. Die Studie empfiehlt daher einen soliden, integrativen KI-Governance-Rahmen für eine nachhaltige und effektive Nutzung dieser Technologien, welcher KI-Risiken und -Leitlinien miteinander verbindet.

1 Einleitung

Die jüngste Entwicklung im Bereich der digitalen Transformation des Controllings wird maßgeblich durch die Fortschritte im Bereich der Generativen Künstlichen Intelligenz (kurz Generative KI) beeinflusst, einem Anwendungsgebiet des Maschinellen Lernens. Zu dieser gehören sogenannte Große Sprachmodelle (Large Language Models - LLMs), welche seit November 2022 mit der Einführung von Chat GPT (Generative Pre-trained Transformer) eine rasche Verbreitung in Wirtschaft und Gesellschaft erfahren haben (D’Onofrio, 2024). Derartige LLMs sind als spezifische Form der Generativen KI in der Lage, menschliche Sprache zu verarbeiten und zu erzeugen. Neben der Interaktion mit Menschen in natürlicher Sprache sind diese ebenso fähig, in grossen Datenmengen Muster zu erlernen, um dadurch neue Inhalte wie Texte, Bilder oder Musik zu generieren (Wecke, 2024). In Anbetracht der Leistungsfähigkeit von LLMs ist es nicht verwunderlich, dass diese transformative Technologie innerhalb kurzer Zeit in verschiedenen Sektoren und Branchen eingesetzt wird und eine rasch wachsende Anzahl von Einsatzbereichen und Anwendungsfällen identifiziert wurde, in denen die Generative KI vielen Unternehmen einen messbaren Geschäftswert liefert (Lamarre et al., 2024).

Zu den vorgenannten Einsatzbereichen, die von den Möglichkeiten der Generativen KI massgeblich beeinflusst werden, zählen auch das Controlling und Finance, welche das betriebswirtschaftliche Informationssystem sowohl für die internen Entscheidungsträger als auch externen Stakeholder darstellen und für die die generative KI in der Praxis eine Reihe von Aufgaben übernehmen kann (Leitner-Hanetseder et al., 2023; Grünbichler et al. 2023; Aiguo, 2022):

- Automatisierung und Beschleunigung von repetitiven Routinetätigkeiten im Zusammenhang mit Datenanalyseprozessen, wodurch Fachkräfte mehr Zeit für strategische Aufgaben haben.
- Integration von Geschäftsdaten mit Finanzdaten zur Erreichung einer verbesserten Entscheidungsfindung.
- Identifikation, Selektion und Integration relevanter Unternehmens-, Wettbewerbs- und Marktdaten in Entscheidungsprozesse des Managements, um dadurch Wettbewerbsvorteile für das Unternehmen zu schaffen.
- Kommentierung von Finanzberichten sowie die effiziente und zeitnahe Durchführung von Abweichungsanalysen.

In diesem Zusammenhang weisen verschiedene Autoren darauf hin, dass der erfolgreiche Einsatz generativer KI im Controlling nicht nur vom Wissen über die Anwendungsdomäne sowie über die Funktionsweise von LLM-Modellen abhängt, sondern ebenso von der Kenntnis der mit dem Einsatz dieser Technologien verbundenen Grenzen und Risiken (Leitner-Hanetseder, 2023; Schulze et al., 2023). Es kann als evident betrachtet werden, dass die Qualität der Datenbasis, mit der generative KI-Systeme trainiert werden, einen entscheidenden Einfluss auf die Qualität der erzeugten Ergebnisse hat. Ebenso sind bei der Beschaffung und dem Einsatz generativer KI ethische Aspekte, der Schutz geistigen Eigentums, Datenschutz, Haftungsfragen sowie relevante arbeitsrechtliche Bestimmungen zu berücksichtigen (Bitkom, 2024). In jüngster Zeit wird der Begriff der AI Governance verwendet, unter dem die Regeln, Richtlinien und Verfahren verstanden werden, mittels derer der Einsatz von KI ethisch, sicher und rechtskonform erfolgt.

Vor diesem Hintergrund ergeben sich folgende Fragen: Existieren für den Einsatz von generativer KI entsprechende Vorgaben, Richtlinien oder Weisungen in den Unternehmen und falls ja, auf welche Aspekte beziehen sich diese? Findet eine Sensibilisierung der Mitarbeitenden in denjenigen Unternehmen statt, in denen generative KI in Controlling und Finance eingesetzt wird? Wird den Anforderungen an eine umfassende AI Governance in den Unternehmen ausreichend Rechnung getragen?

2 Methodik

Die Untersuchung der oben genannten Fragen erfolgte nach einer umfassenden Literaturrecherche und -auswertung durch eine Online-Befragung, wobei auf die Bildung von Hypothesen zugunsten einer schnelleren Verfügbarkeit von Ergebnissen im Vergleich zu alternativen Datenerhebungsmethoden verzichtet wurde (Schützenmeister und Bussmann, 2009). Der zugehörige Online-Fragebogen wurde mittels Google Forms erstellt nach einem Pretest mit drei Fachleuten verfeinert, um Verständlichkeit und Relevanz sicherzustellen. Die Stichprobe wurde durch gezielte

Einladungen über LinkedIn gewonnen, wobei hierbei auf die als relevant angesehene maximale Bearbeitungszeit von fünf bis acht Minuten hingewiesen wurde. Die Auswahl der Stichprobe erfolgte durch aktive Einladung über die Social-Media-Plattform LinkedIn (linkedin.com). Dazu wurden die 543 Follower (Stand 28. Juni 2024) des Instituts für Finance und Law der OST – Ostschweizer Fachhochschule in einer LinkedIn-Nachricht zur Teilnahme eingeladen. In der Einladung wurde ein Link zur Umfrage auf die Website des Instituts veröffentlicht. Um die Angeschriebenen zur Teilnahme zu motivieren, wurde ihnen zugesichert, dass sie die Ergebnisse nach Abschluss der Untersuchung erhalten würden. Die Datenerhebung fand im Zeitraum vom 13. Mai bis zum 28. Juni statt. Insgesamt wurden 38 vollständig ausgefüllte Fragebögen erfasst, was einer Netto-Rücklaufquote von 7 % entspricht. Die Auswertung der 38 vollständig ausgefüllten Fragebögen erfolgte mittels deskriptiver Statistik, insbesondere durch Häufigkeitsverteilungen und Kreuztabellenanalysen, um zentrale Tendenzen und Zusammenhänge zu identifizieren. Hinsichtlich der Zusammensetzung der Stichprobe ist zu sagen, dass die Teilnehmenden an der Online-Befragung mehrheitlich (19 Teilnehmende) in grösseren Unternehmen mit mehr als 250 Mitarbeitenden beschäftigt sind und vorwiegend in Unternehmen der Branchen Finanzdienstleistungen, Industrie, Gross- und Detailhandel sowie die öffentliche Verwaltung beschäftigt sind.

3 Ergebnisse und Implikationen

Im Rahmen der nachfolgenden Datenauswertung wurden Methoden der deskriptiven Statistik angewendet. In der Folge gaben 23 von 38 Teilnehmenden (60,5 %) an, bereits generative KI-Tools in ihren Unternehmen zu nutzen, während 15 (39,5 %) dies bisher nicht tun.

Die von den Teilnehmenden am häufigsten verwendeten generativen KI-Tools sind: Die am häufigsten genannten generativen KI-Tools waren ChatGPT (n = 10), Microsoft Copilot (n = 5) und Eigene Entwicklungen (n = 4).

In Bezug auf die Existenz von Vorgaben und Richtlinien in ihren Unternehmen gaben 10 Befragte (26,3 %) an, dass in ihren Unternehmen Richtlinien oder Grundsätze für den Einsatz generativer KI bestehen. Bei 28 Befragten (73,7 %) ist dies nicht der Fall. Die von den Richtlinien abgedeckten Bereiche betreffen insbesondere die Datennutzung, die Datensicherheit sowie die Nutzung der Tools selbst.

In Bezug auf die Sensibilisierung der Mitarbeitenden in den Bereichen Controlling und Finance für den Einsatz generativer KI konnten drei zentrale Themen identifiziert werden: Die am häufigsten genannten Themen sind Cybersecurity/Datenschutz (19 Antworten), die korrekte Eingabe von Eingabeaufforderungen (11 Antworten) sowie Rechte an geistigem Eigentum (11 Nennungen).

Die Ergebnisse der vorliegenden Studie zeigen, dass die generative KI in Form verschiedener Applikationen im Anwendungsfeld Controlling und Finance in den Unternehmen deutlich angekommen ist. Allerdings zeigt sich, dass in fast drei von vier Unternehmen der Einsatz dieser Technologie nicht durch Vorgaben oder Richtlinien geregelt ist. Dies lässt den Bedarf nach einer klaren AI Governance erkennen, damit die Anwenderinnen und Anwender in Controlling und Finance sich der Grenzen und Risiken eines Einsatzes von generativer KI bewusster sind. Unternehmen sollten daher AI Governance-Rahmenwerke entwickeln, die spezifische Richtlinien für den Umgang mit generativen KI-Tools vorsehen (Sharma et al., 2020). Schliesslich ist eine wirksame Governance von entscheidender Bedeutung für den Umgang mit algorithmischen Verzerrungen, Datenschutzbedenken und potenziellem Missbrauch (Wirtz et al., 2022).

Es besteht jedoch die Möglichkeit, dass bei den Fachleuten in Controlling und Finance ein Verständnis der KI-Risiken und ihrer Auswirkungen fehlt. Für diesen Fall könnten beispielsweise Schulungen für Mitarbeitende zur sicheren und effektiven Nutzung von Tools wie ChatGPT oder Microsoft Copilot durchgeführt werden. Zudem ist es ratsam, klare Richtlinien zur Datensicherheit und zum Schutz geistigen Eigentums zu etablieren, um rechtliche und ethische Risiken zu minimieren.

4 Fazit

Im vorliegenden Extended Abstract erfolgt eine Untersuchung des aktuellen Stands der Anwendung Generativer KI in den Bereichen Controlling und Finance. Die Resultate der durchgeführten Online-Befragung legen dar, dass der Einsatz Generativer KI in Controlling und Finance bereits weit verbreitet ist, zumal die Potenziale dieser Technologie zur Effizienzsteigerung evident sind. Vor allem die Vorteile von LLMs in der Datenanalyse und Entscheidungsfindung sind erheblich, sodass davon auszugehen ist, dass die Zukunft des Controllings massgeblich durch eine verstärkte Nutzung von Technologien der generativen KI geprägt sein wird (Chen et al., 2023).

Unter Berücksichtigung der raschen Diffusion generativer KI legen die Ergebnisse der eigenen Online-Befragung nahe, dass der hohe transformative Einfluss dieser Technologie auf die Zukunft von Controlling und Finance nicht ohne ausreichende Beachtung der Herausforderungen in Bezug auf Datensicherheit, Regulierung und Ethik bleiben sollte. Es wird empfohlen, jede Nutzung generativer KI in den Domänen Controlling und Finance durch einen soliden Governance-Rahmen zu ordnen oder gar zu regulieren. Dadurch können ein effektives Risikomanagement, notwendige Transparenz und Rechenschaftspflicht gewährleistet werden. Dies ist erforderlich, um das volle Potenzial dieser Technologien nachhaltig ausschöpfen und sicherstellen zu können.

Literatur

- Aiguo, W.; Shengchuan, G.; Ruixue, L. (2022): Artificial intelligence technology enables the development of management Finance: The generation of Intelligent Accounting. In 2022 3rd Asia Service Sciences and Software Engineering Conference (ASSE' 22). Association for Computing Machinery, New York, NY, USA, 60–65. <https://doi.org/10.1145/3523181.3523190>
- Bitkom (2024): Generative KI im Unternehmen - Rechtliche Fragen zum Einsatz generativer Künstlicher Intelligenz im Unternehmen. <https://www.bitkom.org/Bitkom/Publikationen/Generative-KI-im-Unternehmen>
- D'Onofrio, S. (2024): Generative Künstliche Intelligenz – die neue Ära der kreativen Maschinen. HMD 61, 331–343 (2024). <https://doi.org/10.1365/s40702-024-01069-0>
- Chen, B., Wu, Z. & Zhao, R. (2023): 'From fiction to fact: The growing role of generative AI in business and accounting, Journal of Chinese Economic and Business Studies, vol. 21, no. 4, pp. 471-496. <https://doi.org/10.1080/14765284.2023.2245279>
- Grünbichler, R.; Sitter, A.; Fenzl, T. (2023): Den Weg für KI im Controlling ebnen. Controlling Management Review 67, 52–57. <https://doi.org/10.1007/s12176-023-1069-1>
- Lamarre, E.; Singla, A.; Sukharevsky, A.; Zimmel, R. (2024): A generative AI reset: Rewiring to turn potential into value in 2024. McKinsey Quarterly. March 4, 2024. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/a-generative-ai-reset-rewiring-to-turn-potential-into-value-in-2024>
- Leitner-Hanetseder, S., Losbichler, H., & Altendorfer, F. (2023): Generative KI im Controlling: Wie können Large Language Modelle das Controlling verändern? CFO aktuell, 17(5), 168-172. <http://www.lindedigital.at/doc-id/art-cfoaktuell-2023-05-168a>
- Schulze, M.; Brossmann, F.; Wallraff, B. (2023): KI in Controlling und Finance: Einsatzmöglichkeiten und Risiken. Controller Magazin 2023(06):60-65.
- Schützenmeister, F.; Bussmann, M. (2009). Online-Befragungen in der Wissenschaftsforschung. In: Jakob, N., Schoen, H., Zerback, T. (eds) Sozialforschung im Internet. VS Verlag für Sozialwissenschaften, Wiesbaden. https://doi.org/10.1007/978-3-531-91791-7_15
- Sharma, G.D. ; Yadav, A. ; Chopra, R. (2020) : Artificial intelligence and effective governance: A review, critique and research agenda. Sustainable Futures 2, 100004, <https://doi.org/10.1016/j.sfr.2019.100004>
- Wecke, B. (2024): Die Relevanz von Generativer Künstlicher Intelligenz. In: Wachstum durch den Einsatz von Generativer KI . Essentials. Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-44041-1_1
- Wirtz, B.W.; Weyerer, J.C.; Kehl, I. (2022): Governance of artificial intelligence : a risk and guideline-based integrative framework. Government Information Quarterly 39, 4, <https://doi.org/10.1016/j.giq.2022.101685>

Operative Planung und Budgetierung in Familienunternehmen: Eine empirische Analyse der Professionalisierung und Anwendung

Extended Abstract

Prof. Dr. Dr. habil. Robert Rieg

Hochschule Aalen, Aalen Institut für Unternehmensführung AAUF, D-73430 Aalen, E-Mail: robert.rieg@hs-aalen.de

Prof. Dr. habil. Patrick Ulrich

Hochschule Aalen, Aalen Institut für Unternehmensführung AAUF, D-73430 Aalen, E-Mail: patrick.ulrich@hs-aalen.de

Abstract

Diese Studie untersucht die operative Planung und Budgetierung in Familienunternehmen im Vergleich zu Nicht-Familienunternehmen. Während häufig davon ausgegangen wird, dass Familienunternehmen allgemein eine geringere Planungsaktivität aufweisen, da sie in geringerem Masse professionelle Führungsinstrumente einsetzen, zeigen die Ergebnisse, dass Familienunternehmen tendenziell zwar zentralisierter planen und weniger Forecasts durchführen, jedoch keine signifikanten Unterschiede in der Art der Planung und Budgetierung aufweisen. Dies deutet auf eine Anpassung der Managementpraktiken von Familienunternehmen an Nicht-Familienunternehmens hin. Die Studie trägt somit zur differenzierten Betrachtung der Professionalisierung in Familienunternehmen bei und zeigt die Notwendigkeit weiterer Forschung zur Optimierung dieser Prozesse auf.

1 Einführung

Familienunternehmen (FU) bilden das Rückgrat vieler Volkswirtschaften und zeichnen sich durch spezifische Merkmale und Herausforderungen aus. Ein zentrales Merkmal ist die oft geringere Formalität und Professionalisierung in der Unternehmensführung im Vergleich zu Nicht-Familienunternehmen (Hiebl & Mayrleitner, 2019). Dies betrifft insbesondere die operativen Planungs- und Budgetierungsprozesse, die für die Umsetzung von Strategien und den Erfolg eines Unternehmens von zentraler Bedeutung sind (Hamann et al., 2022). Trotz der Bedeutung von Familienunternehmen gibt es nur begrenzte empirische Forschung zur detaillierten Ausgestaltung der operativen Planung und Budgetierung in diesen Unternehmen. Die bestehende Literatur legt nahe, dass Familienunternehmen tendenziell weniger formalisierte und weniger professionelle Managementinstrumente nutzen, was zu einer eingeschränkten Nutzung des Controllings führt, das eher zur Dokumentation von Kosten als zur strategischen Planung verwendet wird (Duller et al., 2014; Hiebl et al., 2015; Hiebl & Mayrleitner, 2019). Des Weiteren kann man davon ausgehen, dass Entscheidungsprozesse auch zentralistischer ablaufen als in Nicht-Familienunternehmen (Wait & Nguyen, 2015).

2 Forschungslücke und Forschungsziele

Es gibt eine signifikante Forschungslücke in Bezug auf die spezifischen Merkmale und Unterschiede der Planungs- und Budgetierungssysteme zwischen Familien- und Nicht-Familienunternehmen. Diese Forschungslücke ist von besonderer Bedeutung, da widersprüchliche Theorien und Nachweise existieren, die eine umfassende und detaillierte Untersuchung erfordern. Die vorliegende Studie zielt darauf ab, diese Lücke zu schließen, indem sie empirische Daten nutzt, um die spezifischen Merkmale der operativen Planung und Budgetierung in Familienunternehmen zu analysieren. Dies umfasst Einblicke in die eingesetzten Planungs- und Budgetierungsinstrumente sowie die Häufigkeit und Art der durchgeführten Forecasts.

3 Literaturüberblick und Hypothesen

Planung und Budgetierung sind zentrale Managementinstrumente, die zur Steuerung und Kontrolle von Unternehmensaktivitäten genutzt werden. Zahlreiche empirische Studien haben ihre weit verbreitete Bedeutung bestätigt. Beispielsweise fanden Arnold und Artz (2019) heraus, dass die meisten Unternehmen ein einziges Budget sowohl für die operative Planung als auch für die Leistungsbewertung verwenden, was die zentrale Rolle dieser Instrumente unterstreicht. Familienunternehmen unterscheiden sich jedoch signifikant von Nicht-Familienunternehmen in ihren Planungs- und Budgetierungsprozessen. Familiäre Werte und eine langfristige Orientierung spielen oft eine größere Rolle (Vazquez & Rocha, 2016), was zu spezifischen Herausforderungen und Besonderheiten in der operativen Planung und Budgetierung führt.

Studien zeigen, dass Familienunternehmen tendenziell weniger formalisierte und weniger professionelle Planungs- und Budgetierungssysteme nutzen, was teilweise auf die enge Verflechtung von Familien- und Unternehmensinteressen zurückzuführen ist (Andric & Kammerlander, 2017; Hiebl, 2013). Dies kann zu einer geringeren Nutzung von Controlling-Instrumenten führen, die in erster Linie zur Dokumentation von Kosten anstatt zur strategischen Planung verwendet werden (Felden, 2013; Hiebl et al., 2013; Mitter, 2014; Ulrich, 2014). Ebenso zeigt sich eine stärker zentralisierte Entscheidungsfindung (Wait & Nguyen, 2015), die gegen Planungsmethoden spricht, die auf Partizipation setzen. Die Anwesenheit von externen Managern und die zunehmende Professionalisierung können jedoch zu einer stärkeren Strukturierung der Planungs- und Budgetierungsmethoden in Familienunternehmen führen.

Basierend auf der identifizierten Forschungslücke formuliert die Studie die grundlegende Vermutung, dass Familienunternehmen weniger formalisierte und professionelle Planungs- und Budgetierungssysteme im Vergleich zu Nicht-Familienunternehmen aufweisen. Dies drückt sich aus in mehreren Hypothesen:

Hypothese 1: in Unterschieden in der Ausgestaltung des Planungssystems,

Hypothese 2: in Unterschieden in der Art der Planung,

Hypothese 3: in Unterschieden in der Art der Budgetierung, und

Hypothese 4: in weniger Forecasts pro Jahr.

4 Methodik und Daten

Zur Beantwortung der Forschungsfrage wurde im Frühjahr 2021 eine Online-Umfrage durchgeführt, die an über 6.000 Mitglieder eines Berufsverbands von Controllern in Deutschland verschickt wurde. Insgesamt wurden 266 Antworten gesammelt, was einer Rücklaufquote von etwa 4 % entspricht. Die relevanten Umfrageelemente sind im Anhang 1 aufgeführt. Die Studie misst die operative Planung und Budgetierung durch eine Reihe von Items, die wie folgt gruppiert sind:

Planungssystem: Generelle Ausgestaltung der Zentralisierung, Planungsablauf, Verbindung strategischer und operativer Planung.

Art der Planung: Traditionell bis agile Planung.

Art der Budgetierung: Klassisch bis Zero-based Budgeting.

Häufigkeit der Forecasts pro Jahr.

Die Studie verwendet verschiedene statistische Methoden zur Analyse der Daten, einschließlich Chi-Quadrat-Tests und ordinaler Regressionen. Die Ergebnisse werden in Bezug auf abhängige Variablen (operative Planung und Budgetierung) und unabhängige Variablen (Familienunternehmensstatus) sowie Kontrollvariablen (z.B. KMU-Status, Fertigungs- und Dienstleistungssektor) präsentiert.

5 Wesentliche Ergebnisse und Diskussion

Die wesentlichen Ergebnisse der statistischen Hypothesentests sind:

- **Planungssystem:** Familienunternehmen weisen eine stärker zentralisierte Planung auf im Vergleich zu Nicht-Familienunternehmen. Es wurden jedoch keine signifikanten Unterschiede in der Integration von Volumen- und Finanzplanung festgestellt. Hypothese 1 wird teilweise gestützt.
- **Art der Planung:** Familienunternehmen zeigten keine signifikanten Unterschiede in der Art der Planung. Fertigungsunternehmen demonstrierten jedoch spezifische Planungstypen. Hypothese 2 wird abgelehnt.
- **Art der Budgetierung:** Es wurden keine signifikanten Unterschiede in der Art der Budgetierung zwischen Familien- und Nicht-Familienunternehmen festgestellt, obwohl Fertigungsunternehmen tendenziell bestimmte Budgetierungsmethoden anwenden. Hypothese 3 wird abgelehnt.
- **Häufigkeit der Forecasts:** Familienunternehmen führen weniger Forecasts durch als Nicht-Familienunternehmen. Hypothese 4 wird gestützt.

Die Ergebnisse zeigen, dass Familienunternehmen tendenziell zentralisierte Planungssysteme haben, aber keine signifikanten Unterschiede in den Planungs- und Budgetierungsmethoden im Vergleich zu Nicht-Familienunternehmen aufweisen.

Die Ergebnisse dieser empirischen Studie liefern wertvolle Einblicke in die operative Planung und Budgetierung von Familienunternehmen (FU) im Vergleich zu Nicht-Familienunternehmen (NFU). Entgegen der weit verbreiteten Annahme, dass Familienunternehmen generell eine geringere Professionalisierung aufweisen, zeigen unsere Daten ein differenzierteres Bild.

1. **Zentralisierte Planung:** Die Ergebnisse zeigen, dass Familienunternehmen im Vergleich zu Nicht-Familienunternehmen tendenziell eine stärker zentralisierte Planung aufweisen. Dies könnte darauf hindeuten, dass Familienunternehmen die zentrale Kontrolle und Steuerung bevorzugen, um Familieninteressen besser zu schützen und langfristige Ziele zu verfolgen. Diese zentrale Ausrichtung könnte als eine Form der Professionalisierung interpretiert werden, die spezifisch auf die Bedürfnisse und Werte von Familienunternehmen zugeschnitten ist.
2. **Art der Planung:** Es wurden keine signifikanten Unterschiede in der Art der Planung zwischen Familienunternehmen und Nicht-Familienunternehmen festgestellt. Dies widerspricht der Annahme, dass Familienunternehmen tendenziell weniger professionelle Planungsinstrumente einsetzen. Stattdessen zeigt sich, dass Familienunternehmen ähnliche Planungsprozesse nutzen, wie ihre nicht-familiären Pendants, was auf eine Angleichung der Methoden hinweist.
3. **Art der Budgetierung:** Auch in der Art der Budgetierung konnten keine signifikanten Unterschiede zwischen den beiden Unternehmensarten festgestellt werden. Dies legt nahe, dass Familienunternehmen ebenso wie Nicht-Familienunternehmen moderne Budgetierungsmethoden anwenden, was ein weiteres Indiz für eine höhere Professionalisierung ist als bisher angenommen.
4. **Häufigkeit der Forecasts:** Familienunternehmen führen weniger Forecasts pro Jahr durch als Nicht-Familienunternehmen. Diese geringere Häufigkeit könnte durch die langfristige Orientierung und die oft stabileren finanziellen Rahmenbedingungen in Familienunternehmen erklärt werden, die weniger häufige Anpassungen und Überprüfungen erfordern.

6 Schlussfolgerung

Die Studie zeigt, dass die pauschale Annahme einer geringeren Professionalisierung in Familienunternehmen nicht uneingeschränkt haltbar ist. Während Familienunternehmen tendenziell zentralisierter planen und weniger Forecasts durchführen, gibt es keine signifikanten Unterschiede in der Art der Planung und Budgetierung im Vergleich zu Nicht-Familienunternehmen. Diese Ergebnisse deuten darauf hin, dass Familienunternehmen ihre Managementpraktiken auf demselben Niveau wie Nicht-FU professionalisieren, um den Anforderungen einer zunehmend komplexen und dynamischen Wirtschaftsumwelt gerecht zu werden.

Diese Studie trägt somit wesentlich zum Verständnis der operativen Planung und Budgetierung in Familienunternehmen bei und zeigt, dass eine differenzierte Betrachtung notwendig ist, um die tatsächlichen Praktiken und Professionalisierungsgrade in diesen Unternehmen zu erfassen.

Literaturverzeichnis

- Andric, M. & Kammerlander, N. (2017). Motive zum Verzicht auf Controlling in Familienunternehmen – eine Mediatoranalyse. *Zeitschrift für KMU und Entrepreneurship*, 65(4), 223–251. <https://doi.org/10.3790/zfke.65.4.223>
- Arnold, M. & Artz, M. (2019). The use of a single budget or separate budgets for planning and performance evaluation. *Accounting, Organizations and Society*, 73, 50–67. <https://doi.org/10.1016/j.aos.2018.06.001>
- Duller, C., Feldbauer-Durstmüller, B. & Hiebl, M. R. (2014). Funktionen des Controllings in Familienunternehmen Die Informationsversorgungsfunktion wird weniger intensiv wahrgenommen als in Nicht-Familienunternehmen. *Controller Magazin*(1), 26–29. http://www.wiso-net.de/webcgi?START=A60&DOKV_DB=ZGEN&DOKV_NO=COWI201401012629&DOKV_HS=0&PP=1
- Felden, B. (2013). Controlling in Familienunternehmen: Ergebnisse einer empirischen Untersuchung. *Der Betrieb*, 66(11), 529–533.
- Hamann, P. M., Halw, O. & Günther, T. W. (2022). Meta-analysis of the corporate planning–organizational performance relationship: A research note. *Strategic Management Journal*, Artikel smj.3476. Vorab-Onlinepublikation. <https://doi.org/10.1002/smj.3476>
- Hiebl, M. R. (2013). Bean counter or strategist? Differences in the role of the CFO in family and non-family businesses. *Journal of Family Business Strategy*, 4(2), 147–161. <https://doi.org/10.1016/j.jfbs.2013.02.003>
- Hiebl, M. R., Duller, C., Feldbauer-Durstmüller, B. & Ulrich, P. (2015). Family Influence and Management Accounting Usage – Findings from Germany and Austria. *Schmalenbach Business Review*, 67, 368–404.
- Hiebl, M. R., Feldbauer-Durstmüller, B. & Duller, C. (2013). Die Organisation des Controllings in österreichischen und bayerischen Familienunternehmen. *Zeitschrift für KMU und Entrepreneurship*, 61(1-2), 83–114. <https://doi.org/10.3790/zfke.61.1-2.83>
- Hiebl, M. R. & Mayrleitner, B. (2019). Professionalization of management accounting in family firms: the impact of family members. *Review of Managerial Science*, 13(5), 1037–1068.
- Mitter, C. (2014). Controlling in Familienunternehmen. *Zeitschr. für KMU und Entrepreneurship*, 62(04), 345–352.
- Ulrich, P. (2014). "Controllen" Familienunternehmen anders als Nicht-Familienunternehmen? *Zeitschrift für Familienunternehmen und Stiftungen*(5), 171–176.
- Vazquez, P. & Rocha, H. O. (2016). On the Goals of Family Firms: A Review and Integration. *Academy of Management Proceedings*, 2016(1), 16731. <https://doi.org/10.5465/ambpp.2016.16731abstract>
- Wait, A. & Nguyen, B. T. (2015). Should I Stay or Should I Go? Participation and Decision-Making in Family Firms. *Journal of Institutional and Theoretical Economics*, 171(2), 215. <https://doi.org/10.1628/093245615X14273596658964>

Digitale Technologien in Planung, Budgetierung und Forecasting – es kommt (nicht) auf die Größe an?

Extended Abstract

Prof. Dr. habil. Patrick Ulrich

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: patrick.ulrich@hs-aalen.de

Prof. Dr. habil. Robert Rieg

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: robert.rieg@hs-aalen.de

Felix Stockert, M.Sc.

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: felix.stockert@hs-aalen.de

Abstract

Die Digitalisierung in Unternehmen stellt eine relevante Herausforderung dar. Bisherige Forschungen zeigen, dass grundsätzlich große Unternehmen besser in der Digitalisierung fortschreiten als kleine Unternehmen. Speziell im Controlling - das die strategische Ausrichtung von Unternehmen grundlegend prägt - kann die Digitalisierung eine massive Effizienzsteigerung bringen. In einer empirischen Erhebung mit einer Stichprobe von 231 Unternehmen in Deutschland, Österreich und Schweiz konnten wir zeigen, dass aktuelle digitale Tools, welche z.B. auf künstlicher Intelligenz basieren, erst geringfügig im Controlling implementiert wurden. Im Vergleich der kleinen und großen Unternehmen ist zu erkennen, dass beide im Controlling gleichauf beim Umsetzungsstand sind.

1 Einleitung und theoretische Grundlagen

Die Digitalisierung in der Finanzfunktion von Organisationen hat sich zu einer zentralen Herausforderung entwickelt. Speziell in den Controlling-Teilfunktionen Planung, Budgetierung und Forecasting bietet die Integration digitaler Tools massive Chancen eines Wettbewerbsvorteils. Durch digitale Tools kann die Planungsgenauigkeit gesteigert werden, eine höhere Effizienz erzielt werden und der strategische Nutzen der Controlling-Prozesse optimiert werden (Andreassen, 2020; Bhimani, 2020; Kappes & Leyk, 2018).

Dennoch befindet sich die Digitalisierung des Controllings von Unternehmen oft noch in einem sehr frühen Stadium oder ist lediglich durch einzelne Teilprojekte umgesetzt, ohne eine integrierte Digitalisierungsstrategie zu nutzen (Ivancic, 2019). Die Gründe hierfür sind vielfältig: Einführung und Training von Algorithmen zur Automatisierung oder Vorhersage benötigen umfangreiches Knowhow (Saidulu & Sasikala, 2017; Zhang et al., 2019). Oft scheitert es an der zeiteffizienten Handhabung der (bereits vorliegenden) Daten (Lopes & Ribeiro, 2017). Jedoch stellt nicht nur die Einführung digitaler Tools eine Herausforderung dar, sondern auch die laufende Nutzung und Wartung. Einerseits müssen technische Infrastrukturen möglicherweise optimiert werden (Hazelwood et al., 2018), andererseits Experten oder Keyuser ausgebildet werden (Boutaba et al., 2018). Schlussendlich ist auch der Change-Prozess nicht zu vernachlässigen, da es sich im Controlling um die Grundlage von weitreichenden Management-Entscheidungen handelt und Vertrauen für neue Tools aufgebaut werden muss, um bestehende ineffizientere Prozesse abzulösen (Baesens et al., 2016; Nunes & Jannach, 2020; Shafique et al., 2017). Defizite in der Handhabung digitaler Tools und kein hinreichendes Schulungssystem (Dillerup et al., 2020) stellen hierbei eine der Hauptherausforderungen dar.

Die existierende Forschung zeigt auf, dass die Unternehmensgröße eine relevante Korrelation mit dem Digitalisierungsstand von Unternehmen hat. So können große Unternehmen aufgrund deren umfangreicheren Ressourcen besser mit den dargestellten Herausforderungen umgehen und haben folglich einen höheren Digitalisierungsgrad als kleine und mittlere Unternehmen (KMU). So haben Großunternehmen mehr Fachkräfte in der Informations- und Kommunikationstechnologie beschäftigt. Außerdem nutzen momentan 35 % der Großunternehmen in Deutschland Verfahren der Künstlichen Intelligenz – bei KMU sind es lediglich 11 % (Statistisches Bundesamt, 2023; Institut für Mittelstandsforschung Bonn, 2023). Bislang in der Forschung wenig betrachtet ist, inwiefern sich die Unternehmensgröße auf den Digitalisierungsstand spezifisch im Controlling auswirkt. Anzunehmen ist, dass – entsprechend der generellen Digitalisierungsbemühungen von Unternehmen – auch in der Controllingfunktion größere Unternehmen besser vorankommen, da mehr Ressourcen vorhanden sind. Folglich stellen wir nachfolgende Hypothese auf:

H1: Der Digitalisierungsstand im Controlling hängt maßgeblich von der Unternehmensgröße ab.

2 Methodik

Zur Überprüfung der Hypothese haben wir eine empirische Primärdatenerhebung durchgeführt. Im Rahmen einer quantitativen Erhebung haben uns 231 Teilnehmende ihre Erfahrung zur Digitalisierung im Controlling mitgeteilt. Unsere Stichprobe bestand aus verschiedenen Organisationen in Deutschland, Österreich und Schweiz. Sowohl Kleinunternehmen als auch weltweit operierende Konzerne sind Teil der Stichprobe.

Um die Prüfung unserer Hypothese durchzuführen, haben wir unseren Rücklauf in zwei Gruppen kontrastiert: Einerseits kleine Unternehmen und andererseits große Unternehmen. In unserer Forschung haben wir uns entschieden, die Unternehmensgröße anhand der Mitarbeiterzahl zu messen. Die Schwelle haben wir bei einer Mitarbeiterzahl von 500 angesetzt, sodass die Stichprobe der kleinen Unternehmen eine Mitarbeiterzahl von 1-499 hatte und die Stichprobe der großen Unternehmen 500 und mehr Mitarbeitende.

3 Empirische Ergebnisse

Um ein umfangreicheres Bild zu erlangen, haben wir gesondert die Einsatzintensität digitaler Tools in Planung/Budgetierung und in Forecasting abgefragt. Da wir uns für eine standardisierte quantitative Erhebungsform entschieden haben, wurden folgende Antwortmöglichkeiten anhand bereits existierender Literatur entwickelt:

- Algorithmen-basierte Modelle & Simulationen
- Methoden der Künstlichen Intelligenz (inkl. Machine Learning & Deep Learning)
- Robotic Process Automation (RPA)
- Predictive Analytics
- Treibermodelle & Szenarioplanung
- Sonstige

Folgende Ergebnisse haben wir im Rahmen der Erhebung erhalten:

Einsatzintensität digitaler Tools in Planung/Budgetierung:

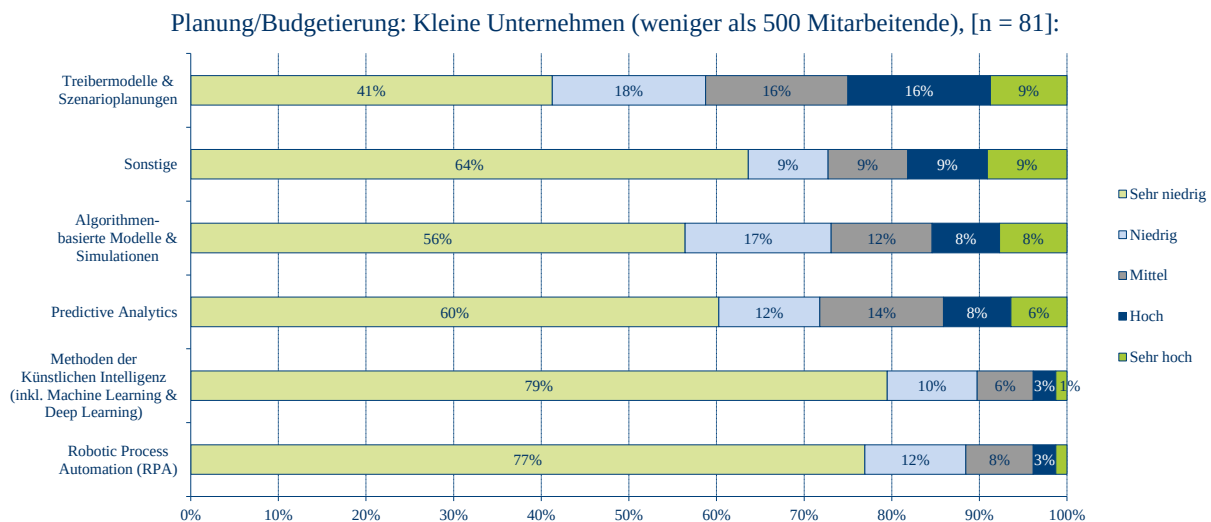


Abbildung 1: Einsatzintensität digitaler Tools in Planung/Budgetierung bei kleinen Unternehmen

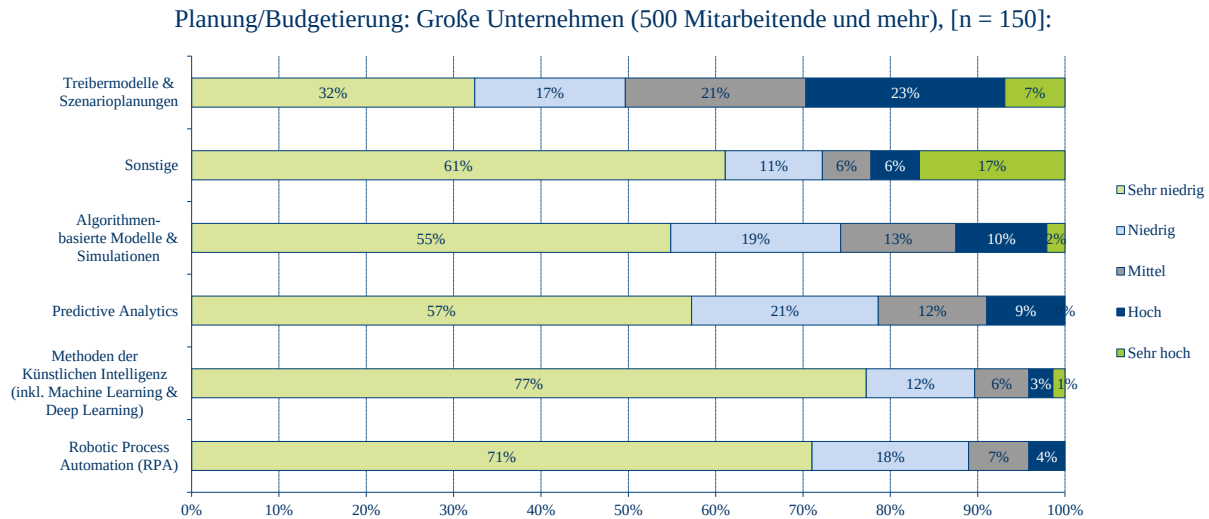


Abbildung 2: Einsatzintensität digitaler Tools in Planung/Budgetierung bei großen Unternehmen

Nachfolgend sind die summierten Prozentwerte für hohe und sehr hohe Einsatzintensität je Technologie gegenübergestellt. Der jeweils führende Wert hinsichtlich Unternehmensgröße ist in **grün**, der niedrigere Wert in **rot** markiert:

Tabelle 1: Einsatzintensität digitaler Tools in Planung/Budgetierung

Planung/Budgetierung	Kleine Unternehmen	Große Unternehmen
Treibermodelle & Szenarioplanung	25 %	30 %
Sonstige	18 %	23 %
Algorithmenbasierte Modelle & Simulationen	16 %	12 %
Predictive Analytics	14 %	9 %
Methoden der künstlichen Intelligenz (inkl. Machine Learning & Deep Learning)	4 %	4 %
Robotic Process Automation (RPA)	4 %	4 %

Es zeichnet sich ein ausgeglichenes Bild zwischen kleinen und großen Unternehmen bei der Einsatzintensität digitaler Tools in Planung und Budgetierung. Im Durchschnitt über alle Antwortmöglichkeiten hinweg weisen 13,5 % der kleinen Unternehmen die Nutzung moderner digitaler Tools auf, während 13,7 % der großen Unternehmen digitale Tools verwenden.

Einsatzintensität digitaler Tools im Forecasting:

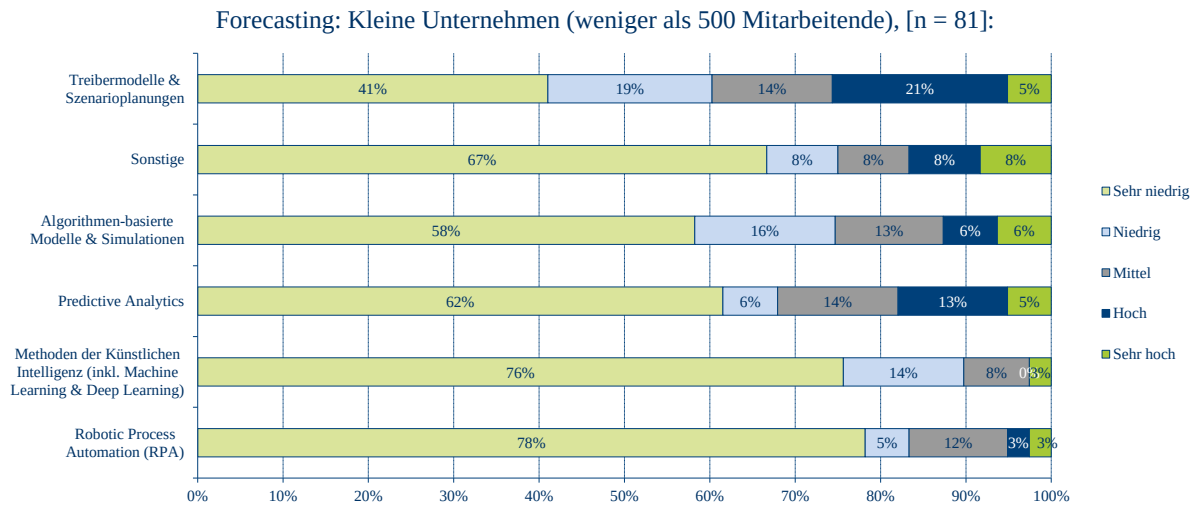


Abbildung 3: Einsatzintensität digitaler Tools im Forecasting bei kleinen Unternehmen

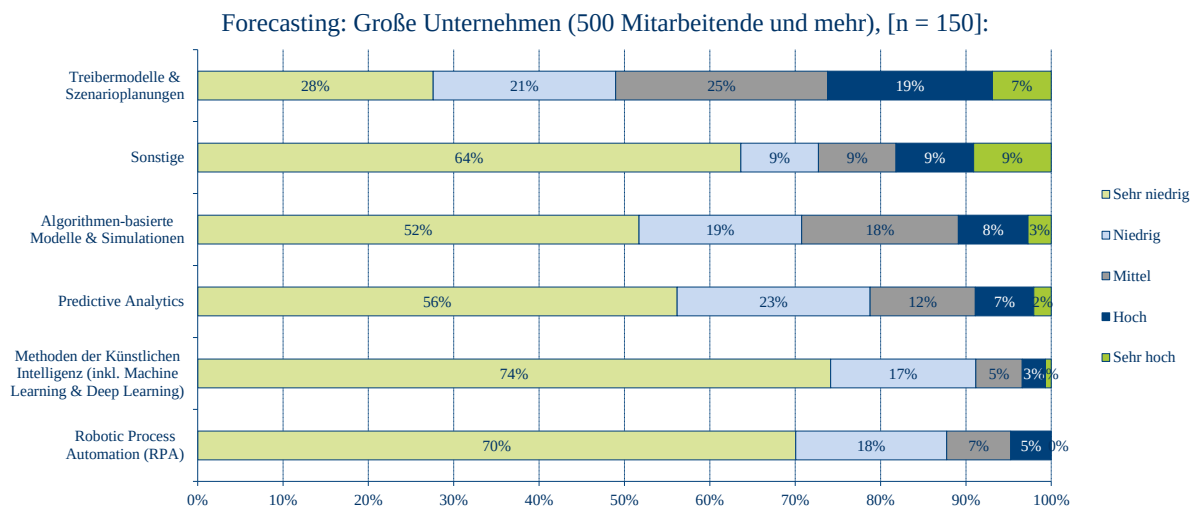


Abbildung 4: Einsatzintensität digitaler Tools im Forecasting bei großen Unternehmen

Nachfolgend sind die summierten Prozentwerte für hohe und sehr hohe Einsatzintensität je Technologie gegenübergestellt. Der jeweils führende Wert hinsichtlich Unternehmensgröße ist in **grün**, der niedrigere Wert in **rot** markiert:

Tabelle 2: Einsatzintensität digitaler Tools im Forecasting

Forecasting	Kleine Unternehmen	Große Unternehmen
Treibermodelle & Szenarioplanung	26 %	26 %
Sonstige	16 %	18 %
Algorithmenbasierte Modelle & Simulationen	12 %	11 %
Predictive Analytics	18 %	9 %
Methoden der künstlichen Intelligenz (inkl. Machine Learning & Deep Learning)	3 %	4 %
Robotic Process Automation (RPA)	6 %	5 %

Auch im Forecasting ist keine klare Tendenz ersichtlich, ob große oder kleine Unternehmen eine höhere Einsatzintensität moderner digitaler Tools aufweisen. Im Durchschnitt liegen kleine Unternehmen bei 13,5 % und große Unternehmen bei 12,2 %. Kleine Unternehmen sind somit sogar leicht führen.

4 Zusammenfassung und weitere Forschung

Die Einsatzintensität digitaler Tools ist sowohl in kleinen als auch in großen Unternehmen noch eher gering. Bei der Kontrastierung nach Unternehmensgröße ist keine klare Tendenz zu erkennen, ob große oder kleine Unternehmen eine höhere Einsatzintensität vorweisen. Dies führt dazu, dass unsere initiale Hypothese nicht verifiziert werden kann.

Es ist nun zu erforschen, warum kleine Unternehmen im Controlling genauso digital aufgestellt sind als große Unternehmen mit massiv höheren Ressourcen. Naheliegend ist, dass aufgrund der Vielzahl an heterogenen neuen Tools kleine Unternehmen diese einfacher implementieren und ausprobieren können als große Unternehmen, wo regulatorische Vorgaben die Einführung hemmen. In einer weiteren Forschung, die wir 2024 durchführen möchten, werden wir dieses Feld adressieren.

Literaturverzeichnis

- Andreassen, R.-I. (2020): Digital technology and changing roles: a management accountant's dream or nightmare? *Journal of Management Control* 31(3):209-238.
- Baesens, B.; Bapna, R.; Marsden, J. R.; Vanthienen, J.; Zhao, J. L. (2016): Transformational Issues of Big Data and Analytics in Networked Business. *MIS Quarterly* 40(4):807-818.
- Bhimani, A. (2020): Digital data and management accounting: why we need to rethink research methods. *Journal of Management Control* (31):9-23.
- Boutaba, R.; Salahuddin, M. A.; Limam, N.; Ayoubi, S.; Shahriar, N.; Estrada-Solano, F.; Caicedo, O. M. (2018): A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. *Journal of Internet Services and Applications* 9(16):1-99.
- Dillerup, R.; Witzemann, T.; Schröckhaas, B. (2020): Zehn Trends der Unternehmensplanung. *Controlling & Management Review* 64(3):46-54.
- Hazelwood, K.; Bird, S.; Brooks, D.; Chintala, S.; Diril, U.; Dzhulgakov, D.; Fawzy, M.; Jia, B.; Jia, Y.; Kalro, A.; Law, J.; Lee, K.; Lu, J.; Noordhuis, P.; Smelyanskiy, M.; Xiong, L.; Wang, X. (2018): Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective. 2018 IEEE International Symposium on High Performance Computer Architecture (HPCA):620-629.
- Institut für Mittelstandsforschung Bonn (2023): Digitalisierung der KMU im EU-Vergleich. <https://www.ifm-bonn.org/statistiken/mittelstand-im-einzelnen/digitalisierung-der-kmu-im-eu-vergleich>
- Ivancic, L.; Bosilj, V.; Spremic, M. (2019): Mastering the Digital Transformation Process: Business Practices and Lessons Learned. *Technology Innovation Management Review* 9(2):36-50.
- Nunes, I.; Jannach, D. (2020): A systematic review and taxonomy of explanations in decision support and recommender systems. *User Modeling and User-Adapted Interaction*, 27(3-5):393-444.
- Kappes, M.; Leyk, J. (2018): Digitale Planung. *Controlling* 30(6):4-12.
- Lopes, N.; Ribeiro, B. (2017): Novel Trends in Scaling Up Machine Learning Algorithms. *International Conference on Machine Learning and Applications*, 632-636.
- Statistisches Bundesamt (2023): Nutzung von IKT in Unternehmen. <https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/IKT-in-Unternehmen-IKT-Branche/IKT-U-Erhebung/info.html>
- Saidulu, D.; Sasikala, R. (2017): Machine learning and statistical approaches for big data: Issues, challenges and research directions. *International Journal of Applied Engineering Research* 12(21):11691-11699.
- Shafique, M.; Hafiz, R.; Javed, M. U.; Abbas, S.; Sekanina, L.; Vasicek, Z.; Mrazek, V. (2017): Adaptive and Energy-Efficient Architectures for Machine Learning: Challenges, Opportunities, and Research Roadmap. *IEEE Computer society annual symposium on VLSI (ISVLSI)*.

Ergebnisse einer Dokumentenanalyse zur Identifikation möglicher Einflüsse der Corporate Sustainability Reporting Directive auf das Controlling

Research Paper

Prof. Dr. habil. Patrick Ulrich

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: patrick.ulrich@hs-aalen.de

Jasmina Metzger, M.A.

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: jasmina.metzger@hs-aalen.de

Abstract

Jedes große Unternehmen, ob kapitalmarktorientiert oder nicht, muss sich aktuell mit der Implementierung der Corporate Sustainability Reporting Directive (CSRD) auseinandersetzen. Diese Richtlinie wird die Aufgaben der Unternehmen verändern, da künftig neben finanziellen auch nicht-finanzielle Informationen berichtet werden müssen. Dafür müssen Strukturen und Prozesse zur Ermittlung dieser Daten geschaffen werden, was auch Auswirkungen auf das Controlling hat. Zukünftig soll das Controlling nicht nur Finanzdaten, sondern auch nichtfinanzielle und Nachhaltigkeitsinformationen berücksichtigen.

Die Umsetzung der Richtlinie erweitert die Aufgaben des Controllings um die nicht-finanzielle Perspektive. Allerdings müssen dafür keine völlig neuen Strukturen geschaffen werden, da bestehende Prozesse, Instrumente, Strukturen und Verantwortlichkeiten beibehalten und durch die Integration von Nachhaltigkeitsaspekten ergänzt werden können. Das bedeutet, dass die einzelnen Elemente des Controllings, wie Funktionen, Aufgaben und Prozesse, erweitert werden und das Controlling neue Kompetenzen entwickeln muss, um weiterhin als Business Partner des Managements agieren zu können. Andernfalls müssten Teilorganisationen Parallelstrukturen aufbauen, was zu einem höheren Koordinationsaufwand führen und die Bedeutung des Controllings verringern würde.

1 Einleitung

Die Europäische Union strebt ehrgeizige Klimaschutzziele an. Infolgedessen wurde der Green Deal eingeführt, der die Klimaneutralität bis 2050 anstrebt und europäische Unternehmen an die Spitze der nachhaltigen Wirtschaft stellt (European Commission, 2019, 2023a). Es ist daher kaum überraschend, dass eines der Instrumente zur Verwirklichung der Ziele des Green Deal die nachhaltige Finanzierung ist. Durch die Integration von ESG-Faktoren (Umwelt-, Sozial- und Governance-Faktoren) unterstützt die nachhaltige Finanzierung das Wirtschaftswachstum und reduziert gleichzeitig die Umweltauswirkungen. Dementsprechend hat die Europäische Kommission in den letzten Jahren mehrere Richtlinien auf den Weg gebracht, um die Wirtschaft zu verändern (European Commission, 2023b).

Im Rahmen des Green Deal hat die Europäische Kommission auch die Überprüfung der nichtfinanziellen Berichterstattung, auch bekannt als "Nachhaltigkeitsberichterstattung", übernommen. Das Ergebnis ist die Corporate Sustainability Reporting Directive (CSRD), die neue europäische Richtlinie zur Nachhaltigkeitsberichterstattung (European Commission, 2022, S. 1–66). Die CSRD wird den Kreis der Unternehmen, die einen Nachhaltigkeitsbericht veröffentlichen müssen, erheblich erweitern. Sie spezifiziert und erweitert auch die verpflichtenden Inhalte, die in Nachhaltigkeitsberichten berichtet werden müssen (Lanfermann & Scheid, 2021, S. 1213–1214; Scheid & Reinke, 2021, S. 1355). Diese Ausweitung der Berichtspflichten bedeutet für die Unternehmen einerseits einen hohen Umsetzungsaufwand. Andererseits bieten sie auch einen Mehrwert für die Unternehmen, wenn die gewonnenen Informationen klug und vorausschauend genutzt werden (KPMG, 2023c, S. 16). Um die Anforderungen der CSRD zu erfüllen, sind die Unternehmen derzeit damit beschäftigt, Strukturen und Prozesse zu schaffen, um sie umzusetzen. Dies wird auch Auswirkungen auf einige der Unternehmensfunktionen haben. Wie schon bei der Integration der International Financial Reporting Standards (IFRS) wird die CSRD auch den Bereich Rechnungswesen und Controlling betreffen. Grund dafür ist zum einen die Beschaffung der notwendigen Daten, die zum Teil vom Controlling generiert werden müssen. Zum anderen sollen diese Informationen und damit auch das Reporting als Entscheidungsgrundlage dienen. Dies würde das Unternehmen nachhaltiger, erfolgreicher und gleichzeitig belastbarer machen (KPMG, 2023b, S. 16, 2023c, S. 16; Weißenberger, 2017, S. 472–492). Bislang ist allerdings noch nicht klar, in welchem Umfang und in welcher Form sich die verpflichtende Nachhaltigkeitsberichterstattung auf das Controlling auswirken wird und wie die Umsetzung der Regelungen durch das Controlling beeinflusst werden kann. Klar scheint bereits zu sein, dass die zunehmende Bedeutung von Nachhaltigkeitsaspekten und Nachhaltigkeitsrisiken eine Koordination der Masse an Aktivitäten im Bereich der Nachhaltigkeit im Unternehmen erfordert. Das Controlling sollte diese Aufgabe übernehmen. Daher wurde die folgende zentrale Forschungsfrage formuliert:

„Welchen Einfluss hat die CSRD auf die Elemente des Controllings“

Empirische Studien zeigen, dass die Umsetzung des Nachhaltigkeitscontrollings durch verschiedene Faktoren gebremst wird und die Integration von Nachhaltigkeitsaspekten in das Controlling noch ein Randthema ist (Lopatta et al., 2022, S. 1064). Dieser Artikel wird diese Aussage durch die Beantwortung der formulierten Forschungsfragen überprüfen.

2 Nachhaltigkeitsberichterstattung: Bedeutung der Corporate Sustainability Reporting Directive (CSRD) und der damit verbundenen European Sustainability Reporting Standards (ESRS)

Die CSRD verpflichtet Unternehmen ab dem Geschäftsjahr 2024 zur Erstellung einer nichtfinanziellen Erklärung, die oft als Nachhaltigkeitsbericht bezeichnet wird. Der Beginn der Berichtspflicht ist gestaffelt, so dass zunächst nur die Unternehmen berichten müssen, die bereits nach der aktuellen Richtlinie zur nichtfinanziellen Berichterstattung berichtspflichtig sind. Ein Jahr später müssen dann alle großen Unternehmen einen Nachhaltigkeitsbericht erstellen, bevor im Jahr 2027 auch kleine und mittlere Unternehmen von der CSRD betroffen sein werden. Diese nichtfinanzielle Erklärung muss immer in den Lagebericht des Unternehmens integriert sein und vom Wirtschaftsprüfer mit begrenzter Sicherheit geprüft werden (European Commission, 2022; KPMG, 2023a; Wunder, 2022,

S. 59). Im Rahmen der CSRD müssen u.a. Informationen über das Geschäftsmodell, die Risiken und deren Gegenmaßnahmen offengelegt werden (Lanfermann & Scheid, 2021, S. 1215; Müller & Reinke, 2022, S. 13; Müller et al., 2021, S. 1325). Ziel ist es zu klären, ob und inwieweit das Geschäftsmodell und die Strategie des Unternehmens die Interessen der Stakeholder und die Auswirkungen des Unternehmens auf Nachhaltigkeitsaspekte berücksichtigen. Gleichzeitig soll ermittelt werden, welche Anpassungen notwendig sind, um deren Interessen zu erfüllen. Zu diesem Zweck sollen die Unternehmen über ihre Nachhaltigkeitsorganisation, -strategie und die in diesem Bereich gesetzten Ziele berichten (Lanfermann & Scheid, 2021, S. 1215).

Die CSRD präzisiert auch die Definition der Wesentlichkeit und fordert die Berücksichtigung der doppelten Wesentlichkeit. Das bedeutet, dass in Zukunft neben der Betrachtung der Auswirkungen der eigenen Geschäftstätigkeit (Inside-Out-Perspektive) auch die Bedeutung der Themen für das Unternehmen selbst (Outside-In-Perspektive) zu berücksichtigen ist (European Commission, 2022; KPMG, 2023a). Die Einbeziehung einer externen Sicht auf das Unternehmen erhöht die Transparenz und Vergleichbarkeit des Berichts, während die interne Perspektive die Leistung des Unternehmens verbessern soll und u.a. die Nachhaltigkeitsstrategie und deren Management umfasst (Beusch, 2020, S. 34–35). Daraus ergibt sich ein Anpassungsbedarf in der Unternehmensführung, der sich vermutlich auch auf die Unternehmensfunktion des Controllings auswirken wird.

Um die nichtfinanzielle Berichterstattung in Europa künftig einheitlich und vergleichbar zu gestalten und die in der CSRD formulierten inhaltlichen Anforderungen zu konkretisieren, wurden europaweit einheitliche Berichtsstandards entwickelt und eingeführt: die European Sustainability Reporting Standards (ESRS). Wie in Abbildung 1 dargestellt, gliedern die ESRS die zu berichtenden Themen in Umwelt-, Sozial- und Governance-Themen sowie in übergreifende Anforderungen (Lanfermann & Scheid, 2021, S. 1213; Scheid & Reinke, 2021, S. 1355).

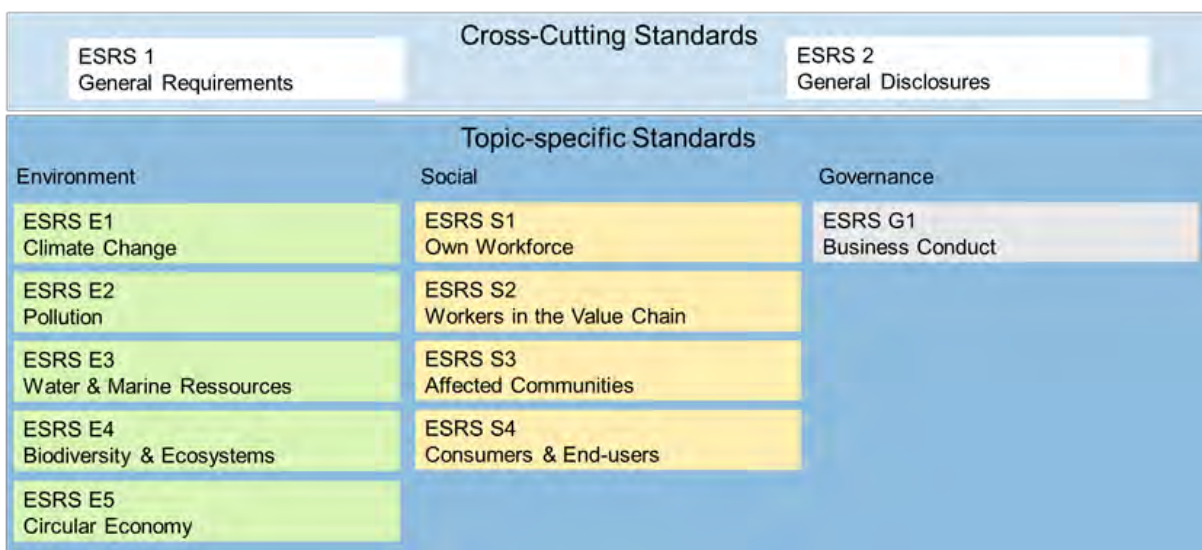


Abbildung 1: Überblick über die European Sustainability Reporting Standards (In Anlehnung an EFRAG, 2024)

Im Vergleich zu den bisherigen Themen, über die im Rahmen der alten Nachhaltigkeitsrichtlinie (Non-Financial Reporting Directive) berichtet werden musste, stellt der ESRS eine erhebliche Ausweitung der zu berichtenden Themen dar. Dazu gehören Offenlegungspflichten zu Risikomanagement, Strategien, Zielen und Maßnahmen sowie zu den finanziellen Auswirkungen einiger Themen. Der ESRS verlangt auch Informationen zu spezifischen Themen wie Energiemanagement, Ressourcenzuflüsse und -abflüsse, Informationen über die eigene Belegschaft und die Arbeitskräfte in der Wertschöpfungskette sowie über die Kunden (EFRAG, 2024).

3 Definition «Controlling» und Herausforderungen durch die Einführung der CSRD

3.1 Wertschöpfungsorientierte Controlling-Konzeption

Das zugrunde liegende Konzept des Controllings ist das wertschöpfungsorientierte Controllingkonzept nach Becker. Dabei dient die Wertschöpfung im Unternehmen als Ausgangspunkt (Becker et al., 2014, S. 57). Damit wird das Ziel verfolgt, die langfristige Existenz des Unternehmens zu sichern, was durch die Kontrolle der Effektivität und Effizienz des Handelns des Unternehmens erreicht wird (Petzold & Westerkamp, 2018, S. 32). Neben einer sehr eng gefassten Rationalität geht es bei der Wertschöpfung in Unternehmen auch um die Befriedigung von Stakeholder-Bedürfnissen, die sich in drei Kategorien unterteilen lassen: individuelle Bedürfnisse, soziale Bedürfnisse und ökologische Bedürfnisse. Dazu bedarf es einer spezialisierten Funktion, die sich auf den Wertschöpfungsprozess konzentriert und häufig in die Unternehmensorganisation integriert ist. Diese Aufgabe übernimmt das Controlling, das als "wertschöpfungsorientierte Managementfunktion" zu verstehen ist. Dementsprechend liegt der Fokus des Controllings hier auf der Sicherstellung der Erfüllung des Unternehmenszwecks (Becker, 2022, S. 7–8).

Das Controlling erfüllt die Lokomotion-Funktion als originäre Funktion. Das heißt, das Controlling stellt sicher, dass die Ausrichtung auf die Wertschöpfung bei allen Überlegungen, Entscheidungen und Handlungen des Managements berücksichtigt wird (Becker, 2022, S. 15; Becker et al., 2014, S. 57–58). Controlling ist somit ein integraler Bestandteil der Unternehmensführung, da die Lokomotion-Funktion auf dem Zweck des unternehmerischen Handelns basiert, der durch Wertschöpfung gekennzeichnet ist. Es umfasst demnach die Initiierung und Ausrichtung des unternehmerischen Handelns auf den Zweck der Wertschöpfung. Damit ist das Controlling ein wesentlicher und zentraler Bestandteil der Führungstätigkeit im Unternehmen (Becker, 2022, S. 14–16). Gleichzeitig bleiben aber die klassischen Funktionen des Controllings erhalten. Die Steuerungsfunktion der Lokomotion wird von der Unternehmensleitung wahrgenommen, die sich auch des Instruments des Controllings bedient. Das Controlling kann diese Funktion jedoch nur erfüllen, wenn es seine Koordinations- und Informationsfunktion wahrnimmt, denn diese bilden die Grundlage der Lokomotion und werden daher auch als Management-Support-Funktionen bezeichnet (Becker, 2022, S. 15; Becker et al., 2014, S. 57–58).

3.2 Herausforderungen bei der Umsetzung der Nachhaltigkeitsberichterstattung nach CSRD

Bereits 2011 hat sich der Internationale Arbeitskreis für Controlling (ICV) mit der Integration ökologischer Aspekte in das Controlling beschäftigt und in einer Studie festgestellt (International Group of Controlling, 2011) dass es bei den Kontrolleuren Wissensdefizite in Bezug auf Fragen der Nachhaltigkeit gibt (International Group of Controlling, 2011, S. 25). Dies wird auch von Möhrer und Vanini (2022) elf Jahre später bestätigt (Möhrer & Vanini, 2022, S. 27). Darüber hinaus ist die Kommunikation zwischen den Mitarbeitern im Umweltbereich und im Controlling unzureichend (International Group of Controlling, 2011, S. 25). Dies wirkt sich unter anderem negativ auf die Erfüllung der Koordinierungs- und Informationsfunktion der Kontrolleure aus, da diese aufgrund schlechter Kommunikation und mangelnder Kenntnisse nur schwer eine angemessene Informationsversorgung der Unternehmensleitung sicherstellen können. Das Gleiche gilt für die Koordinierung der Managementaktivitäten. Neben den Herausforderungen bietet die Integration von Nachhaltigkeitsaspekten in das Controlling aber auch eine große Chance, da die Controller das Management in Zukunft noch besser unterstützen und beraten können. Das bedeutet, dass sich Controller noch mehr in Richtung eines Business Partners für das Management entwickeln (Beusch, 2020, S. 35–36; International Group of Controlling, 2011, S. 21–25; Kämmler-Burrak & Bauer, 2022, S. 27; Keimer & Egle, 2023, S. 39; Sailer, 2020, S. 40–80; Schaltegger, 2016, S. 58).

Ein weiteres Problem im Zusammenhang mit der Informationsfunktion des Controllings ist das Problem der Datenerhebung bzw. Datenaufbereitung, für das Lösungen gefunden werden müssen (Colsman, 2016, S. 50–51). Ein besonderes Problem in diesem Zusammenhang ist, dass das Nachhaltigkeitscontrolling die Messmethoden teilweise erst entwickeln oder anpassen muss, um Daten erheben zu können. Dies erschwert den Zugang zum Thema und die Integration von Nachhaltigkeitsaspekten in das Controlling (Hartmann et al., 2016, S. 77; International Group of Controlling, 2011, S. 25; Kämmler-Burrak & Bauer, 2022, S. 27; Sailer & Biel, 2020, S. 89).

4 Methodik

Als Ausgangspunkt unserer Analyse dienten die beiden oben genannten Säulen: die Anforderungen an die Nachhaltigkeitsberichterstattung durch die CSRD und der Controlling-Ansatz. Auf dieser Grundlage wurde eine Dokumentenanalyse durchgeführt. In dieser Analyse haben wir die folgenden allgemeinen Suchbegriffe verwendet, um die relevante Literatur zu finden:

- Einflüsse der CSRD auf das Controlling
- Einflüsse der Nachhaltigkeitsberichterstattung auf das Controlling
- Corporate Sustainability Reporting Directive
- Konsequenzen der CSRD
- Controlling UND Nachhaltigkeitsberichterstattung

Da es sich um ein relativ neues Thema mit vergleichsweise wenig Literatur handelt, haben wir sowohl spezifische Suchbegriffe als auch allgemeine Suchbegriffe zu den Folgen der CSRD verwendet. Da es noch nicht allzu viele wissenschaftliche Publikationen zu diesem Thema gibt, enthalten die Ergebnisse sowohl wissenschaftliche Arbeiten als auch praxisorientierte Artikel und Beraterstudien.

Die ermittelte Literatur umfasst hauptsächlich deutschsprachige Artikel, da der Begriff "Controlling" im deutschsprachigen Raum am weitesten verbreitet ist. Ein weiterer Filter, den wir verwendet haben, ist die Relevanz der Artikel, da das Ziel der Analyse die Analyse der Controlling-Elemente war, um die notwendigen Anpassungen oder die Herausforderungen zu bestimmen, die bei der Umsetzung der CSRD bestehen.

Daher haben wir die Informationen aus der relevanten Literatur den einzelnen Steuerungselementen zugeordnet. Die Ergebnisse werden im nächsten Kapitel vorgestellt.

5 Ergebnisse

Da Nachhaltigkeitsaspekte immer in das bestehende Controlling-System integriert werden sollten (Endenich & Trapp, 2022), basieren die Ergebnisse auf dem bestehenden Controllingsystem und zeigen mögliche Einflussfaktoren auf diese Elemente auf. Die hier beschriebenen Einflussfaktoren und möglichen Veränderungen sind zusätzlich nochmal in Abbildung 2 zusammengefasst. Neben einer kurzen Beschreibung der jeweiligen Wirkungen durch die CSRD enthält die Abbildung auch eine kurze Beschreibung des jeweiligen Controlling-Elements, sodass Ausgangspunkt und Einflüsse auf einen Blick sichtbar werden.

Controlling-Philosophie

Integraler Bestandteil des Konzepts des wertschöpfungsorientierten Controllings ist die Wertschöpfungsorientierung und die damit verbundene Lokomotion (Becker et al., 2014, S. 57). Da die CSRD die Wertschöpfung in den Unternehmen voraussichtlich völlig verändern wird, wird sich dies auch auf Aufgaben, Prozesse, Methoden, Instrumente und Werkzeuge auswirken. Diese Auswirkungen sind in Abbildung 2 zusammengefasst. Der Grund dafür liegt zum einen in der CSRD selbst, aber auch in den Inhalten der ESRS, die die Umsetzung der CSRD beinhalten. So wird sich die Grundhaltung des Controllings an die neue Ausgangssituation anpassen müssen, da die Anwendung der ESRS auf Veränderungen im Unternehmensverhalten abzielt. So besteht beispielsweise eine Verpflichtung zur Analyse von Nachhaltigkeitsthemen wie Klimawandel und Menschenrechte sowie zur Transparenz der Zielvorgaben und der Messung der Zielerreichung. Positiv hervorzuheben ist jedoch, dass die korrekte Anwendung der CSRD zu einem besseren Verständnis der Auswirkungen von Nachhaltigkeitsfaktoren führt, was sich wiederum positiv

auf die Wertschöpfung und die finanziellen Ergebnisse auswirkt (PricewaterhouseCoopers, 2023b). Dies ist insbesondere dann der Fall, wenn durch die Berücksichtigung von Nachhaltigkeitsaspekten neue Erfolgspotenziale geschaffen werden können, die wiederum zu einem Mehrwert am Markt führen.

Diese Überlegungen führen bereits heute dazu, dass Unternehmen, die noch nicht zur Berichterstattung über Nachhaltigkeit verpflichtet sind, freiwillig einen Nachhaltigkeitsbericht veröffentlichen (Nienaber et al., 2023, S. 490). Dies zeigt, dass die neuen, umfassenden Berichtspflichten neben dem damit verbundenen Aufwand auch einen Mehrwert schaffen, sofern die Informationen im Entscheidungsprozess entsprechend berücksichtigt werden (KPMG, 2023c, S. 16).

Ziele

Hinsichtlich des Einflusses der CSRD auf die Ziele des Controllings konnten keine konkreten Informationen über Veränderungen durch die verpflichtende Nachhaltigkeitsberichterstattung ermittelt werden. Dementsprechend ist davon auszugehen, dass die Ziele gleichbleiben und die Sicherung des Unternehmenszwecks, auch im Hinblick auf Nachhaltigkeit, fortgesetzt wird.

Funktionen

Dies bedeutet aber auch, dass die Informationsfunktion des Controllings in Zukunft einen höheren Informationsbedarf abdecken muss. Dazu müssen innerhalb der Informationsfunktion neue Prozesse geschaffen werden, die es ermöglichen, die benötigten Daten zu erheben. Schließlich wird für die Entscheidungsfindung eine zuverlässige und glaubwürdige Datenbasis benötigt. Allerdings gibt es häufig das Problem der Datenverfügbarkeit, da oft nur Daten für das gesamte Jahr, wenn überhaupt, zur Verfügung stehen und unterjährig Daten fehlen (Oehler & Kämmler-Burrak, 2024, S. 29; PricewaterhouseCoopers, 2023b; Taubitz et al., 2023). Neben der Informationsfunktion wird auch die Koordinationsfunktion durch die CSRD beeinflusst werden, da zukünftig die zusätzlichen Aktivitäten der Unternehmensführung, die zur Umsetzung der zusätzlichen Anforderungen der Nachhaltigkeitsberichterstattung erforderlich sind, zusätzlich zu den bestehenden Aktivitäten koordiniert werden müssen. Da für die Umsetzung der CSRD und der Nachhaltigkeitsberichterstattung bisher häufig separate Nachhaltigkeitsabteilungen zuständig waren (PricewaterhouseCoopers, 2023a, S. 16–17), eine Koordinierung und klare Abgrenzung der Zuständigkeiten zwischen den beiden Unternehmensfunktionen erforderlich ist. Dieser Punkt wird im Bereich der Aufgabenträger wieder aufgegriffen. Hinsichtlich der Lokomotion-Funktion wurde bereits festgestellt, dass sich die Wertschöpfung der Unternehmen an die neuen Gegebenheiten anpassen muss, damit weiterhin Erfolgspotenziale geschaffen werden können und der Wertschöpfungskreislauf in Bewegung bleibt. Darüber hinaus wandelt CSRD auch die bisher lineare Wirtschaft in eine Kreislaufwirtschaft um, was zu Anpassungen des Wertschöpfungskreislaufs führt, da durch die Transformation der Wirtschaft die Recyclingfähigkeit von Produkten immer wichtiger wird, was bedeutet, dass ausrangierte Produkte wieder in den Wertschöpfungskreislauf zurückkehren sollten (Taubitz et al., 2023).

Aufgaben und Verantwortlichkeiten für die Aufgaben

Die Aufgaben des Controllings, die sich aus den Funktionen ableiten, werden dementsprechend auch von den Einflüssen auf die Funktionen beeinflusst. Wie bereits erwähnt, ergibt sich aus CSRD ein erhöhter Bedarf an Daten, die die gleiche Zuverlässigkeit und Glaubwürdigkeit aufweisen wie Finanzdaten. Gleichzeitig müssen für die Nachhaltigkeitsthemen Kennzahlen geschaffen werden, mit deren Hilfe ein Steuerungssystem aufgebaut werden kann (Oehler & Kämmler-Burrak, 2024, S. 29; PricewaterhouseCoopers, 2023b; Taubitz et al., 2023).

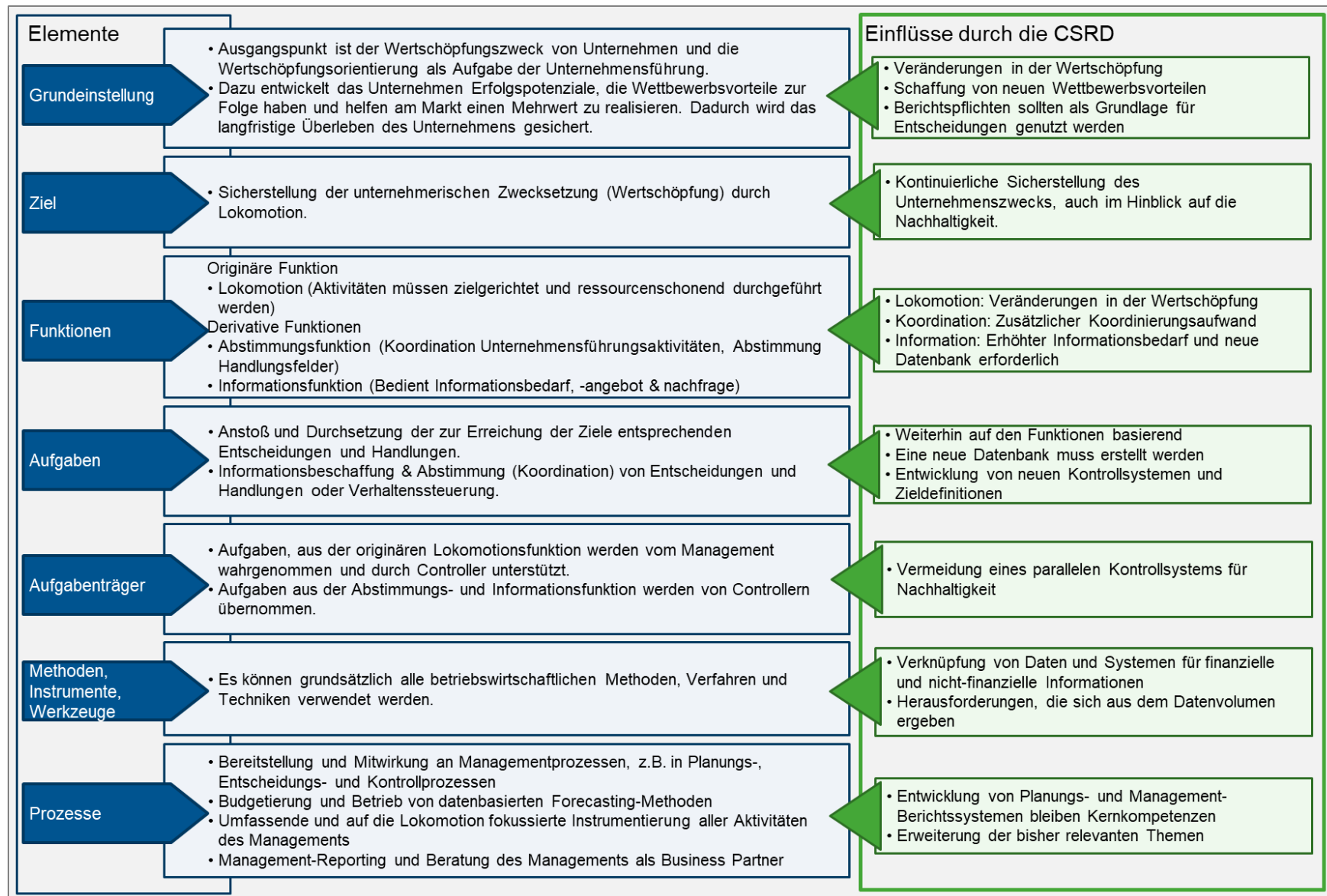


Abbildung 2: Überblick über die Einflüsse der CSRD auf die Controlling-Elemente

In diesem Zusammenhang ist auch zu betonen, dass die bereits erwähnte Erstellung von Kennzahlen einer der Schwerpunkte der Aufgaben des Controllings ist.

In diesem Zusammenhang ist auch darauf hinzuweisen, dass dies auch eine einheitliche Datenlandschaft voraussetzt, die ebenfalls vom Controlling geschaffen werden sollte. Dementsprechend werden auch hier die Aufgaben des Controllings aus dem Finanzbereich in den Bereich der Nachhaltigkeit verlagert (International Group of Controlling, 2023, S. 21).

Bei näherer Betrachtung wird jedoch deutlich, dass die Aufgaben, die sich aus CSRD für das Controlling ergeben, nicht völlig neu sind, denn Management und Informationsbeschaffung waren schon immer Aufgaben des Controllings. Neu ist die Ergänzung um die ökologische und soziale Säule der Nachhaltigkeit. Hinzu kommt die Aufgabe, ein neues Managementsystem für diese Themen zu etablieren, das von den Top-Level-Kennzahlen bis zu den operativen Kennzahlen sinnvoll und konsistent aufgebaut ist und sowohl die interne Steuerung als auch die externen Berichtspflichten abdeckt.

Der Aufbau eines parallelen Managementsystems für Nachhaltigkeit sollte vermieden werden. Das bedeutet, dass die Verantwortlichen für die Aufgaben dieselben sind wie im bisherigen Finanzmanagementsystem (Oehler & Kämmler-Burrak, 2024, S. 29), auch wenn Unternehmen derzeit oft eine eigene Nachhaltigkeitsabteilung haben, die für die Umsetzung der CSRD und die damit verbundene Nachhaltigkeitsberichterstattung zuständig ist (PricewaterhouseCoopers, 2023a, S. 16–17). Daher ist es wichtig, ein System zu etablieren, in dem die interne Managementsicht und die externe Nachhaltigkeitsberichterstattung abgedeckt und eng miteinander verzahnt sind, aber keine Parallelprozesse stattfinden (Oehler & Kämmler-Burrak, 2024, S. 29).

Das Management hat die Aufgabe, Nachhaltigkeit im Unternehmen zu verankern und die notwendigen Strukturen und Prozesse systematisch umzusetzen (International Group of Controlling, 2023, S. 21).

Instrumente

Ein Blick auf die im Controlling eingesetzten Instrumente zeigt, dass diese lediglich um ökologische und soziale Perspektiven erweitert werden müssen. Hinsichtlich der eingesetzten Instrumente besteht die Notwendigkeit, Daten und Systeme finanzieller und nicht-finanzieller Informationen zu verknüpfen (Oehler & Kämmler-Burrak, 2024, S. 30). Die zunehmende Datenmenge, die sich aus der Verknüpfung von finanziellen und nicht-finanziellen Informationen ergibt, stellt die Datenarchitekturen vor große Herausforderungen, da die wachsende Zahl von Nachhaltigkeitsinitiativen und Berichtsanforderungen den Bedarf an heterogenen und komplexen Datensätzen für die Informationsbereitstellung und ein erfolgreiches Management erhöhen. Um dieser Herausforderung gerecht zu werden, werden leistungsfähige, integrierte Informationssysteme benötigt, da die Analyse von finanziellen und nicht-finanziellen Daten miteinander verknüpft werden muss (Hoffmann, 2024, S. 22).

Die Balanced Scorecard bzw. ihre Weiterentwicklung zu einer nachhaltigen Balanced Scorecard ist ein mögliches Instrument, das auch für Nachhaltigkeitsfragen nützlich ist. Die Balanced Scorecard unterstützt die Abbildung von Vision und Leitbild des Unternehmens in der Unternehmensstrategie. Das bedeutet, dass die Balanced Scorecard zur Formulierung von Strategien genutzt wird. Sie konkretisiert auch die Maßnahmen und Parameter zur Umsetzung dieser Visionen (Arnold et al., 2003, S. 394; Baum et al., 2013, S. 415). Daraus lässt sich ableiten, dass es sich um ein umfassendes Controllingsystem zur Steuerung und Koordination handelt (Baum et al., 2013, S. 415; Buchholz, 2019, S. 247; Küpper et al., 2013, S. 500). Die Balanced Scorecard umfasst vier Perspektiven: Finanzen, Kunden, interne Geschäftsprozesse sowie Lernen und Entwicklung. Innerhalb jeder dieser Perspektiven sollten die Ziele, die zugehörigen Kennzahlen und Zielvorgaben sowie die Maßnahmen festgelegt werden (Kaplan & Norton, 1992, S. 72).

Da die Nachhaltigkeitsberichterstattung die Festlegung von Strategien, Maßnahmen und Zielen für jeden themenspezifischen Standard erfordert, kann es sinnvoll sein, eine Balanced Scorecard zu verwenden, um diese Strategien,

Maßnahmen und Ziele zu ermitteln und zu verfolgen. Zu diesem Zweck ist es ratsam, die traditionelle Balanced Scorecard um eine fünfte Perspektive zu ergänzen und eine Sustainable Balanced Scorecard zu verwenden. Dies wurde in der Literatur vor allem in den 2010er Jahren diskutiert, als das Konzept des Green Controlling entwickelt wurde. Derzeit wird die Balanced Scorecard eher weniger verwendet, zumindest hat die Anzahl der Beiträge in der Literatur über die Jahre abgenommen. Diese These wird im Rahmen der Literaturanalyse noch einmal überprüft werden. Dennoch ist die Sustainable Balanced Scorecard eine gute Möglichkeit, um Strategien, Maßnahmen und Ziele im Bereich der Nachhaltigkeit zu entwickeln und anschließend zu steuern. Der Einsatz der Sustainable Balanced Scorecard ist jedoch nicht unproblematisch, da die Maßnahmen zur Umsetzung der unternehmerischen Nachhaltigkeit z.B. kosteneffizient sein müssen, um tatsächlich umsetzbar zu sein. Dieser Beitrag zur ökonomischen Dimension wird häufig mit Kennzahlen wie Cashflow oder ROI gemessen. Der Erfolg einer Nachhaltigkeitsstrategie, insbesondere im sozialen und ökologischen Bereich, ist jedoch meist nicht finanzieller Natur. Dennoch stellen sie einen Maßstab für strategische Aktivitäten dar und dienen der Steigerung der unternehmerischen Nachhaltigkeit. Das bedeutet, dass diese Chancen, die nicht finanziell zu sein scheinen, dennoch erkannt werden müssen. Dennoch bleibt die Messbarkeit ein Problem, das es zu lösen gilt, denn die strategischen Erfolgspotenziale einer nachhaltigen Unternehmensausrichtung müssen in quantitative Zahlen übersetzt werden, um im Tagesgeschäft nachvollziehbar zu sein (Arnold et al., 2003, S. 394–395).

Die Herausforderung der Datenerhebung wurde bereits in den vorherigen Abschnitten angesprochen. Diese Frage der Datenmessung ist auch für die Investitionsprüfung relevant, da eine Entscheidungsgrundlage benötigt wird, um sich für oder gegen eine Investition unter Nachhaltigkeitsaspekten zu entscheiden (Schäffer, 2011, S. 82). Allerdings ist die Ermittlung der Informationen für die zu treffende Investitionsentscheidung auch heute noch oft problematisch, da es sich häufig um Prognosen handelt. Dennoch bietet die Investitionsrechnung eine gute Grundlage, um sich für die richtige Investitionsalternative zu entscheiden, da alle möglichen Alternativen anhand der zuvor definierten Teilziele verglichen und bewertet werden. Am Ende kann die beste Investitionsalternative ausgewählt werden (Poggensee, 2022, S. 31–32). Dass Investitionen auch in Zukunft in Bereichen mit Nachhaltigkeitsbezug relevant sind bzw. weiter an Relevanz gewinnen werden, liegt nicht zuletzt an den Anforderungen der CSRD. So lassen sich beispielsweise die Treibhausgasemissionen in den meisten Fällen nicht reduzieren, ohne in neue Anlagen, Produktionsprozesse oder Photovoltaikanlagen zur Energiegewinnung zu investieren. Ähnlich verhält es sich im Bereich der Kreislaufwirtschaft, wo Unternehmen ebenfalls in Forschung und Entwicklung oder neue Produktionsverfahren investieren müssen. Diese und andere Investitionsprojekte müssen zwei Anforderungen erfüllen: Zum einen müssen sie zur Verbesserung der Nachhaltigkeitsleistung des Unternehmens beitragen, zum anderen müssen sie auch kosteneffizient sein. Ein Beispiel soll das Ganze noch einmal verdeutlichen, denn mit der Investitionsrechnung kann beispielsweise die Energieeffizienz einer Anlage dargestellt werden, sofern die eingesparte Energie quantifiziert werden kann. Auf diese Weise ist es möglich, die Einsparungen in Form von reduzierten Zahlungen in die Berechnung einzubeziehen. Damit ist es auch möglich, das Thema Nachhaltigkeit in die Investitionsrechnung einzubeziehen. (Endenich & Trapp, 2022, S. 188–189). Da es jedoch immer schwierig ist, qualitative Auswirkungen zu quantifizieren, ist es sinnvoll, die quantitative Analyse durch eine Nutzenanalyse zu ergänzen. Die Nutzwertanalyse ist eine Technik innerhalb der Investitionsrechnung und ermöglicht eine mehrdimensionale Betrachtung der Alternativen. Sie ist daher ein mögliches Instrument, um Investitionsalternativen im Bereich der Nachhaltigkeit zu bewerten (Endenich & Trapp, 2022, S. 188–189; Poggensee, 2022, S. 187).

Das letzte Instrument, das in diesem Beitrag betrachtet werden soll, sind Kennzahlen und Kennzahlensysteme. Sie sind ein etabliertes Instrument im Controlling, mit dem kurz und prägnant über wirtschaftliche Themen informiert werden kann (Reichmann et al., 2017, S. 38–39; Weber & Schäffer, 2022, S. 179). Die wichtigsten Elemente sind der Informationscharakter, die Quantifizierbarkeit und die spezifische Form der Information. Die Aussagekraft einer einzelnen Kennzahl ist jedoch begrenzt, da die Qualität der informativen Eigenschaften von der Gestaltung des zugrundeliegenden Informationssystems abhängt, denn Kennzahlen können keine Mängel in der Basisinformation beheben. Außerdem liefern Kennzahlen nur quantitative Informationen. Sobald zusätzliche qualitative Informationen benötigt werden, ist eine Kombination aus quantitativen und qualitativen Informationen zu empfehlen. Aus diesem Grund sind Kennzahlensysteme entwickelt worden. Darunter versteht man einen organisierten Satz von

Kennzahlen, die sich gegenseitig erklären oder ergänzen und miteinander in Beziehung stehen. Dadurch können sie vollständige Informationen über eine Situation als Ganzes liefern (Horváth et al., 2020, S. 307–309; Reichmann et al., 2017, S. 41; Weber & Schäffer, 2022, S. 197).

Da in der CSRD und dem dazugehörigen ESRS eine Vielzahl quantitativer Informationen in Form von Kennzahlen berichtet werden müssen und diese auch zur Steuerung des Unternehmens herangezogen werden, liegt es nahe, dass hier auch Nachhaltigkeitsaspekte integriert werden müssen. Mit anderen Worten: Die für die Steuerung der Nachhaltigkeit erforderlichen Kennzahlen müssen ermittelt und dann erhoben werden. Damit werden neue Kennzahlensysteme geschaffen.¹

Prozesse

Die CSRD verlangt die Offenlegung der Verbindung zwischen der Nachhaltigkeitsagenda und der Geschäftsstrategie. Zudem werden Informationen darüber gefordert, wie die Planung zur Verbesserung der Nachhaltigkeitsleistung aussieht. Da das Controlling in die Strategieentwicklung und -umsetzung eingebunden ist und sich daher künftig auch mit Nachhaltigkeitsaspekten befassen muss, führt dies zu einer Ausweitung der bisher relevanten Themen, da die CSRD auch eine Erläuterung verlangt, wie die Nachhaltigkeitsleistung gesteuert wird. Aus diesem Grund müssen bestehende Prozesse an die neuen Berichtspflichten angepasst werden (PricewaterhouseCoopers, 2023b). Eine der Kernkompetenzen des Controllings ist die Entwicklung und der Betrieb von Systemen zur Planung und Managementberichterstattung. Dabei ist es unerheblich, ob sie sich auf finanzielle oder nicht-finanzielle Informationen beziehen. Diese Aufgabe und die damit verbundenen Prozesse sollten daher beim Controlling verbleiben (Wiltinger, 2024, S. 35).

Die Prozesse der strategischen Planung, des Investitionscontrollings, des Management-Reportings und des Datenmanagements stehen im Fokus der neuen Aufgaben des Controllings und sind auch am stärksten von der CSRD beeinflusst (International Group of Controlling, 2023, S. 21).

Da die Anpassung der Controlling-Prozesse in der Regel schrittweise erfolgt und vom Reifegrad des Unternehmens im Bereich der Nachhaltigkeit abhängt, werden die Anpassungen in Teilprozessen vorgenommen. Einer der ersten betroffenen Prozesse ist die strategische Planung. Hier gilt es, zusätzliche Stakeholder zu berücksichtigen und nachhaltigkeitsrelevante Aspekte zu integrieren. Darüber hinaus gewinnt die Betrachtung von Szenarien und langfristigen Betrachtungszeiträumen an Bedeutung. Ein Schwerpunkt liegt dabei auf der Übersetzung der Ziele in konkrete Maßnahmen und Metriken. Dazu muss im Vorfeld das Verhältnis zwischen ökologischen, ökonomischen und sozialen Zielen im Unternehmen geklärt werden. Es muss festgelegt werden, ob diese Ziele gleichrangig sind oder ob soziale und ökologische Ziele den finanziellen Zielen untergeordnet sind oder nicht.

Die realistischste Form ist wohl immer noch die Dominanz der finanziellen Ziele, um die Wettbewerbsfähigkeit des Unternehmens zu sichern. Diesen können dann ökologische und soziale Ziele untergeordnet werden. Diese können als Nebenbedingungen zu den Mindeststandards formuliert werden oder alle drei Zielkategorien können gewichtet werden. Die Mindeststandards können aus regulatorischen Vorgaben oder der Nachhaltigkeitsstrategie abgeleitet werden, zum Beispiel (International Group of Controlling, 2023, S. 21–52).

Der Prozess, der in der Regel als nächstes angepasst wird, ist das Investitionscontrolling. Hier muss z.B. der Prozess um die nicht-finanzielle Perspektive erweitert werden, ähnlich dem Instrument der Investitionsplanung, damit zukünftig auch nachhaltigkeitsbezogene Bewertungen im Rahmen von Investitionsentscheidungen vorgenommen werden können (International Group of Controlling, 2023, S. 21–23).

Der Management-Reporting-Prozess wird auch inhaltlich erweitert, da hier künftig auch nicht-finanzielle Informationen aufgenommen werden sollen.

¹ Eigene Überlegungen auf der Grundlage der oben genannten Informationen

Eine qualitativ hochwertige Berichterstattung über nicht-finanzielle Informationen schafft Vertrauen und Glaubwürdigkeit nach innen und außen. Dies setzt aber, wie bereits mehrfach erwähnt, eine vernünftige Datenbasis und aussagekräftige Kennzahlen voraus, damit die Daten auch in Zukunft die Grundlage für Managemententscheidungen bilden können. Es ist daher nicht verwunderlich, dass auch der Datenmanagementprozess durch die CSRD verändert wird und angepasst werden muss. Alle Teilprozesse sind von der Umsetzung der CSRD betroffen: die kontinuierliche Überprüfung der Qualität des Controllings, die ständige Weiterentwicklung und das Zusammenspiel der verschiedenen Controller-Ebenen. Darüber hinaus sind neue Richtlinien und Standards zu entwickeln und zu dokumentieren sowie das vorhandene Wissen über bestehende und neue Instrumente zu erweitern. Nicht zuletzt gilt es, die Überprüfung des möglichen Outsourcings von Controllingprozessen anzupassen (International Group of Controlling, 2023, S. 21–62).

Der letzte Prozess, der in diesem Beitrag näher betrachtet werden soll, ist das Risikocontrolling. Da die CSRD für fast alle themenspezifischen Offenlegungsanforderungen die Offenlegung von Chancen und Risiken sowie die zusätzliche Offenlegung der Ermittlung von Risiken und Chancen vorschreibt, ist es nicht verwunderlich, dass auch dieser Prozess von Veränderungen durch die Umsetzung der CSRD betroffen ist. Auch hier müssen die Datengrundlagen um Nachhaltigkeitsthemen erweitert werden. Die Studie der International Group of Controlling (2023) zeigt hier einen enormen Nachholbedarf bei der Integration von nicht-finanziellen Risiken in das Risikomanagement (International Group of Controlling, 2023, S. 21–29).

Zusammenfassend lässt sich sagen, dass ein Großteil der Controlling-Prozesse durch die Implementierung von CSRD im Unternehmen angepasst und erweitert werden muss, um weiterhin ein kompetentes und gut funktionierendes Controlling im Unternehmen zu haben.

6 Diskussion

Die Ergebnisse zeigen, dass die Ausweitung der Regulierung im Bereich der Nachhaltigkeitsberichterstattung nicht nur Auswirkungen auf die offensichtlich betroffenen Funktionen wie das Umweltmanagement, das Personalwesen oder die Produktion eines Unternehmens haben wird. Vielmehr wird auch die Finanzfunktion im Unternehmen in die Umsetzung der Regelungen einbezogen werden, da der Beitrag deutlich macht, dass sich auch das Controlling mit diesem Thema auseinandersetzen muss. Andernfalls sind die Unternehmen gezwungen, z.B. eine Parallelstruktur zu schaffen, die die Informationsversorgung des Managements, die Koordination aller relevanten Aktivitäten oder auch die Sicherstellung der Wertschöpfungsorientierung bei allen Entscheidungen übernimmt. Dies kann kaum im Interesse des Controllings sein. Auch das Unternehmen sollte an dieser Lösung nicht interessiert sein, da sie zusätzlichen Verwaltungsaufwand schafft und zusätzliche Ressourcen erfordert. Stattdessen sollte das Ziel sein, nicht-finanzielle Themen in die bestehende Controllingfunktion gemäß der CSRD zu integrieren. Das schont Ressourcen, nutzt vorhandene Kompetenzen und reduziert zudem Schnittstellen. Zudem würden so finanzielle und nicht-finanzielle Themen innerhalb des Unternehmens angeglichen, statt eine Parallelstruktur zu schaffen.

Die Ergebnisse zeigen auch, dass das Controlling bereits über die notwendigen Prozesse und Instrumente verfügt, um Nachhaltigkeitsaspekte oder nicht-finanzielle Themen im Allgemeinen in die Unternehmensführung und Entscheidungsfindung zu integrieren.

Die Ergebnisse zeigen aber auch, dass das Controlling zwar über die notwendigen Methoden, Prozesse und Strukturen verfügt, es ihm aber noch an der notwendigen Expertise im Bereich Nachhaltigkeit mangelt. Hier muss das Controlling definitiv nachbessern, wenn es seine Position als Geschäftspartner behalten will.

Natürlich birgt die Studie selbst auch einige Einschränkungen. So wurde beispielsweise nicht zwischen verschiedenen Eigentümerstrukturen oder Unternehmensgrößen unterschieden. Außerdem wurden nur die Einflüsse der Corporate Sustainability Reporting Directive berücksichtigt. Andere Regelungen, wie die EU-Taxonomie oder das Lieferkettengesetz, wurden nicht einbezogen. Darüber hinaus wurde das Controlling nur oberflächlich betrachtet, was den Weg für weitere Forschungen in diesem Bereich ebnen sollte. Weitere Forschungsfelder können sein:

- Nähere Betrachtung eines spezifischen Controlling-Elements (z.B. Funktionen oder Aufgaben)
- Einflüsse der EU-Taxonomie auf das Controlling

7 Fazit

Werfen wir abschließend noch einmal einen Blick auf die eingangs formulierte Forschungsfrage (*"Welchen Einfluss hat die CSRD auf die Elemente des Controllings"*).

Die Ergebnisse zeigen deutlich, dass es Einflüsse durch die CSRD auf das Controlling gibt. Sie zeigen aber auch, dass diese Einflüsse bisher noch wenig untersucht wurden. Mit Ausnahme der Instrumente und der Prozesse des Controllings wurden die Auswirkungen der CSRD bisher nicht untersucht. Die festgestellten Ergebnisse basieren daher auf den Folgen der Umsetzung der CSRD im Unternehmen im Allgemeinen, die auf die theoretischen Anforderungen an das Controlling übertragen wurden. Dementsprechend wurde in diesem Beitrag auch eine eher praxisorientierte Perspektive gewählt, um die Wirkungen der CSRD auf das Controlling aufzuzeigen, die aber in allen Elementen des Controllings zu beobachten sind.

Abschließend bleibt festzustellen, dass es vielfältige Einflüsse und Veränderungen im Controlling durch die CSRD geben wird, jedoch sind weitere Untersuchungen notwendig, um diese Veränderungen genauer zu identifizieren und Empfehlungen hinsichtlich der Implementierung der Anforderungen der Nachhaltigkeitsberichterstattung in das Controlling geben zu können.

Literaturverzeichnis

- Arnold, W., Freimann, J. & Kurz, R. (2003). Sustainable Balanced Scorecard (SBS): Integration von Nachhaltigkeitsaspekten in das BSC-Konzept. *Zeitschrift für Controlling & Management (ZfCM)*, 47(6), 391–400.
- Baum, H.-G., Coenenberg, A. G. & Günther, T. (2013). *Strategisches Controlling* (5. Auflage).
- Becker, W. (2022). Wertschöpfungsorientiertes Controlling in Zeiten der Digitalisierung. In W. Becker & P. Ulrich (Hrsg.), *Springer NachschlageWissen. Handbuch Controlling* (2. Auflage, S. 7–27). Springer Gabler.
- Becker, W., Baltzer, B. & Ulrich, P. (2014). *Wertschöpfungsorientiertes Controlling: Konzeption und Umsetzung*.
- Beusch, P. (2020). Management accounting and control for sustainability. In G. Rimmel (Hrsg.), *Accounting for Sustainability* (S. 30–55).
- Buchholz, L. (2019). *Strategisches Controlling: Grundlagen - Instrumente - Konzepte* (3. Auflage).
- Colsmann, B. (2016). *Nachhaltigkeitscontrolling: Strategien, Ziele, Umsetzung* (2. Aufl. 2016). Springer Gabler. <https://doi.org/10.1007/978-3-658-09437-9>
- EFRAG. (2024). The first set of ESRS - the journey from PTF to delegated act (adopted on 31 July 2023). <https://www.efrag.org/lab6>
- Endenich, C. & Trapp, R. (2022). Nachhaltigkeitscontrolling in Klein- und Mittelunternehmen. In B. Feldbauer-Durstmüller & S. Mayr (Hrsg.), *Controlling – Aktuelle Entwicklungen und Herausforderungen* (S. 185–200). Springer Fachmedien Wiesbaden.
- European Commission. (2019). What is the European Green Deal?
- European Commission. (2022). RICHTLINIE (EU) 2022/2464 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 14. Dezember 2022: zur Änderung der Verordnung (EU) Nr. 537/2014 und der Richtlinien 2004/109/EG, 2006/43/EG und 2013/34/EU hinsichtlich der Nachhaltigkeitsberichterstattung von Unternehmen. <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32022L2464>
- European Commission. (2023a). The European Green Deal. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en
- European Commission. (2023b). Overview of sustainable finance. https://finance.ec.europa.eu/sustainable-finance/overview-sustainable-finance_en
- Hartmann, F., Maas, K. & Perego, P. (2016). Den Wald vor lauter Bäumen nicht sehen: Controller auf der Suche nach Nachhaltigkeit. In E. Günther & K.-H. Steinke (Hrsg.), *Management-Reihe Corporate Social Responsibility. CSR und Controlling: Unternehmerische Verantwortung als Gestaltungsaufgabe des Controlling* (S. 71–82). Springer-Gabler.
- Hoffmann, K. (2024). Von Daten zu Taten: Die Schlüsselrolle von Daten für ein erfolgreiches Nachhaltigkeitsreporting. *Controller Magazin*(1), 22–23.
- Horváth, P., Gleich, R. & Seiter, M. (2020). *Controlling* (14. Auflage).
- International Group of Controlling. (2011). *Green Controlling - eine (neue) Herausforderung für den Controller? Relevanz und Herausforderungen der Integration ökologischer Aspekte in das Controlling aus Sicht der Controllingpraxis*.
- International Group of Controlling. (2023). *Controlling & Nachhaltigkeit: Die Rolle des Controllings im Transformationsprozess zur nachhaltigen Unternehmenssteuerung: Dream-Car-Bericht der Ideenwerkstatt im ICV 2022/2023*.
- Kämmler-Burrak, A. & Bauer, R. (2022). Nachhaltigkeit wird Standardaufgabe im Controlling: Vom «Why» zum «How». *Controller Magazin*(01), 22–27.
- Kaplan, R. S. & Norton, D. P. (1992). The Balanced Scorecard: Measures That Drive Performance. *Harvard Business Review*(Januar-Februar), 71–79.
- Keimer, I. & Egle, U. (2023). Nachhaltigkeit im Controlling: Mit Growth Mindset zu neuen Kompetenzen. *Controller Magazin*(2023), 36–41.

- KPMG. (2023a). Corporate Sustainability Reporting Directive (CSRD). <https://kpmg.com/de/de/home/themen/uebersicht/esg/corporate-sustainability-reporting-directive.html>
- KPMG. (2023b). Nachhaltig steuern - Studie zum ESG Management und Steering.
- KPMG. (2023c). Nachhaltigkeitsberichterstattung: Bestandsaufnahme und Trends.
- Küpper, H.-U., Friedl, G., Hofmann, C., Hofmann, Y. & Pedell, B. (2013). Controlling: Konzeption, Aufgaben, Instrumente (6. Auflage).
- Lanfermann, G. & Scheid, O. (2021). Vorschlag der EU-Kommission zur Corporate Sustainability Reporting Directive (CSRD). *Der Betrieb*(23/ 2021), 1213–1218.
- Lopatta, K., Buchholz, F. & Kornblum, W. (2022). Controlling und Corporate Sustainability: eine exemplarische Darstellung anhand der Kreditwirtschaft. In W. Becker & P. Ulrich (Hrsg.), *Springer NachschlageWissen. Handbuch Controlling* (2. Auflage, S. 1057–1076). Springer Gabler.
- Möhrer, M. & Vanini, U. (2022). Entwicklungen im Nachhaltigkeitscontrolling vor dem Hintergrund der EU-Taxonomie für Sustainable Finance. *Rethinking Finance*(04), 25–27.
- Müller, S. & Reinke, J. (2022). Regulatorische Vorgaben für eine Nachhaltigkeitsberichterstattung: Deutliche qualitative und qualitative Ausweitung von kurzfristig zu erwarten. *Controller Magazin*(1/2022), 10–15.
- Müller, S., Scheid, O. & Baumüller, J. (2021). Kommissionsvorschlag zur Corporate Sustainability Reporting Directive: von der nichtfinanziellen Berichterstattung zur Nachhaltigkeitsberichterstattung. *Betriebs-Berater*(22.2021), 1323–1327.
- Nienaber, M., Nienhaus, M. & Stanka, P. (2023). Nachhaltigkeitsberichterstattung von nichtkapitalmarktorientierten Unternehmen – Aktuelle Erkenntnisse aus Theorie und Praxis. *Der Konzern*(12), 487–494.
- Oehler, K. & Kämmler-Burak, A. (2024). Systemgestütztes Controlling zur Steuerung der Nachhaltigkeit. *Controller Magazin*(1), 28–33.
- Petzold, J. & Westerkamp, M. (2018). Informationssysteme im wertorientierten Controlling: Grundlagen - Aufbau - Anforderungen - Integration - Anwendungen.
- Poggensee, K. (2022). *Investitionsrechnung* (4. Auflage). Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-37226-2>
- PricewaterhouseCoopers. (2023a). Corporate Sustainability Reporting Directive (CSRD) 2023 – eine Analyse: Wie weit Unternehmen aus Deutschland, Österreich, der Schweiz und den Niederlanden mit der Umsetzung sind.
- PricewaterhouseCoopers. (2023b). Die Corporate Sustainability Reporting Directive (CSRD) und ihre Folgen. <https://www.pwc.de/de/nachhaltigkeit/die-corporate-sustainability-reporting-directive-csrd-und-ihre-folgen.html#veraenderungen-fuer-unternehmen>
- Reichmann, T., Kißler, M. & Baumöl, U. (2017). Controlling mit Kennzahlen: Die systemgestützte Controlling-Konzeption (9. Auflage).
- Sailer, U. (2020). Nachhaltigkeitscontrolling. 3. Auflage.
- Sailer, U. & Biel, A. (2020). Nachhaltigkeit ein Controlling-Thema? Interview mit Prof. Dr. Ulrich Sailer, Studiendekan Master Controlling, Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen. *Controller Magazin*(06), 86–90.
- Schäffer, U. (2011). Nachhaltigkeit und Controlling: Drei Herausforderungen für die Unternehmenspraxis. *Controller Magazin*(November/Dezember), 81–84.
- Schaltegger, S. (2016). CSR, Nachhaltigkeit und Controlling - Zwischen Praxislücke und Forschungskonzepten. In E. Günther & K.-H. Steinke (Hrsg.), *Management-Reihe Corporate Social Responsibility. CSR und Controlling: Unternehmerische Verantwortung als Gestaltungsaufgabe des Controlling* (S. 55–70). Springer-Gabler.
- Scheid, O. & Reinke, J. (2021). Unternehmerische Nachhaltigkeitsberichterstattung in der deutschen Sustainable-Finance-Strategie. *Der Betrieb*(25/ 2021), 1355.

- Taubitz, P., Csernalabics, P., Hosp, M. & Abraham, K. (2023). Neue Berichtspflichten: Warum Nachhaltigkeit keinen Aufschub duldet. https://www.ey.com/de_de/sustainability/die-csrd-nachholbedarf-bei-der-nachhaltigkeit#accordion-content-445439438-4
- Weber, J. & Schäffer, U. (2022). Einführung in das Controlling (17. Auflage).
- Weissenberger, B. E. (2017). Controller und IFRS: Konsequenzen der IFRS-Berichterstattung für die Controlleraufgaben. In W. Funk & J. Rossmannith (Hrsg.), Internationale Rechnungslegung und internationales Controlling: Herausforderungen - Handlungsfelder - Erfolgspotenziale (3., überarbeitete Auflage, S. 472–492). Springer Gabler.
- Wiltner, K. (2024). Nachhaltigkeitscontrolling – Herausforderungen für mittelständische Unternehmen. Controller Magazin(1), 34–35.
- Wunder, T. (2022). Strategieverankerung der CSRD: Von ESG-Compliance zu Wettbewerbsvorteilen. Rethinking Finance(6), 57–64.

2

Accounting



Wertschöpfungsverständnis in der Corporate Governance

Extended Abstract

Prof. Dr. Björn Baltzer

Technische Hochschule Würzburg-Schweinfurt, THWS Business School, E-Mail: bjoern.baltzer@thws.de

Prof. Dr. habil. Patrick Ulrich

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: patrick.ulrich@hs-aalen.de

Abstract

Der Begriff Wertschöpfung ist für die Corporate Governance in Deutschland von herausragender Bedeutung, da er im Deutschen Corporate Governance Codex (DCGC) als Handlungsmaxime für Vorstand und Aufsichtsrat genannt wird. Gleichzeitig gibt es weder im DCGK noch in den darauf verweisenden Gesetzen eine Erläuterung, was unter Wertschöpfung zu verstehen ist. Innerhalb der Betriebswirtschaftslehre wird der Wertschöpfungsbegriff wiederum in zahlreichen unterschiedlichen Kontexten verwendet. Zur Klärung des Begriffsverständnisses von Wertschöpfung in der Unternehmenspraxis wird daher momentan eine empirische Untersuchung durchgeführt.

1 Einordnung des Themas

Als Corporate Governance (CG) sollen im Folgenden diejenigen Regelungen bezeichnet werden, welche die Leitung und Überwachung von Unternehmen (Innenverhältnis) sowie die Beziehungen von Unternehmen mit ihren Stakeholdern (Außenverhältnis) betreffen (Ulrich 2011, S. 56-65). Die vielfältigen CG-Regelungen weisen hierbei unterschiedliche Verbindlichkeitsgrade auf und finden sich u.a. in Gesetzen oder Unternehmenssatzungen.

Eines der wichtigsten Dokumente zur CG in Deutschland ist der Deutsche Corporate Governance Codex (DCGK). Dieser wurde von einer von der Bundesregierung eingesetzten Expertenkommission entwickelt, erstmals 2002 veröffentlicht und seitdem etwa alle ein bis zwei Jahre überarbeitet - die aktuelle Fassung stammt aus dem Jahr 2022. Der DCGK fasst einerseits Regelungen zusammen, die bereits an anderer Stelle gesetzlich niedergeschrieben und somit verbindlich zu befolgen sind. Darüber hinaus enthält der DCGK auch anerkannte Empfehlungen, die anhand des Begriffs ‚soll‘ identifizierbar sind. Desweiteren finden sich im DCGK auch Anregungen, deren Kennzeichnung durch den Begriff ‚sollte‘ erfolgt. Kapitalmarktorientierte Unternehmen müssen gemäß §161 Aktiengesetz (AktG) jedes Jahr erklären, inwieweit den Empfehlungen des DCGK gefolgt wird. Falls Empfehlungen nicht entsprochen wird, so ist dies angemessen zu begründen („comply or explain-Prinzip“). Diese Entsprechenserklärung ist gemäß §289f bzw. §315d Handelsgesetzbuch (HGB) in den Lagebericht des Unternehmens bzw. der Unternehmensgruppe aufzunehmen. Über diesen relativ engen Anwenderbereich hinaus ist jedoch festzuhalten, dass der DCGK eine breite Ausstrahlungswirkung auch auf andere Unternehmen aufweist und als Best Practice für gute und verantwortungsvolle Unternehmensführung angesehen werden kann.

Der Begriff ‚Wertschöpfung‘ nimmt innerhalb des DCGK eine prominente Stellung ein. In der den Einzelregelungen vorangestellten Präambel werden das oberste Leitungsgremium und das Aufsichtsgremium dazu verpflichtet, stets und ausschließlich im Interesse des Unternehmens zu handeln. Das Unternehmensinteresse wird hierbei definiert als Existenzsicherung durch „nachhaltige Wertschöpfung“.¹ Alle Einzelregelungen des DCGK spezifizieren somit diese Generalnorm. Der Begriff Wertschöpfung hatte dabei nicht bereits von Anfang an diesen herausragenden Stellenwert. In den ersten Versionen des DCGK von 2002 bis einschließlich 2008 findet sich kein entsprechender Passus in der Präambel, und der Begriff Wertschöpfung wurde nicht verwendet. Stattdessen war der Vorstand gemäß der damaligen Einzelregelung 4.1.1 zu einer „Steigerung des nachhaltigen Unternehmenswertes verpflichtet“. Der Austausch des Begriffs ‚Wertsteigerung‘ durch den Begriff ‚Wertschöpfung‘ in der damaligen Einzelregelung 4.1.1 sowie die Betonung der Wertschöpfung durch die Ergänzung der Präambel erfolgte erst in der DCGK-Version des Jahres 2009. Diese Änderung kam einem Paradigmenwechsel gleich: War der Vorstand bis dahin vorrangig den Interessen der Eigentümer und damit der Steigerung des Shareholder Value verpflichtet, so sind Vorstand wie auch Aufsichtsrat seitdem den Interessen aller wesentlichen Anspruchsgruppen des Unternehmens und damit der Generierung eines Stakeholder Value verpflichtet. Dieser Paradigmenwechsel wurde seinerzeit in der Literatur durchaus zur Kenntnis genommen (Hecker 2009; Kann & Keiluweit 2009; Fabisch 2010), allerdings insgesamt relativ wenig beachtet (Lingnau & Kreklow 2011).

2 Begründung des Forschungsbedarfs

In ihren Entsprechenserklärungen nehmen die berichtspflichtigen Unternehmen oftmals direkten Bezug zur Maxime der nachhaltigen Wertschöpfung gemäß der Präambel des DCGK, wie das Beispiel von adidas zeigt:

„Der Vorstand leitet das Unternehmen in eigener Verantwortung mit dem Ziel nachhaltiger Wertschöpfung und im Unternehmensinteresse, entwickelt die strategische Ausrichtung des Unternehmens, stimmt diese mit dem Aufsichtsrat ab und sorgt für deren Umsetzung.“²

¹ Der Wortlaut in der aktuellen Version des DCGK ist: „Der Kodex verdeutlicht die Verpflichtung von Vorstand und Aufsichtsrat, im Einklang mit den Prinzipien der sozialen Marktwirtschaft unter Berücksichtigung der Belange der Aktionäre, der Belegschaft und der sonstigen mit dem Unternehmen verbundenen Gruppen (Stakeholder) für den Bestand des Unternehmens und seine nachhaltige Wertschöpfung zu sorgen (Unternehmensinteresse).“

² Quelle: adidas Geschäftsbericht 2023, S. 37.

Problematisch ist hierbei, dass weder der Begriff ‚nachhaltig‘ noch der Begriff ‚Wertschöpfung‘ in den einschlägigen Dokumenten definiert oder zumindest erläutert werden. Der hier besonders interessierende Begriff ‚Wertschöpfung‘ findet sich außer in der Präambel an keiner weiteren Stelle im DCGK, und ebenso wenig im AktG oder im HGB. Wertschöpfung ist somit als unbestimmter Rechtsbegriff anzusehen.

Diese Feststellung wäre unproblematisch, sofern es anderweitig, also insb. innerhalb der Betriebswirtschaftslehre, ein eindeutiges Verständnis von Wertschöpfung gäbe. Dies ist allerdings nicht der Fall, was allein daran ersichtlich ist, dass Küpper (2007, S. 1) die gesamte „Betriebswirtschaftslehre als Wertschöpfungstheorie“ bezeichnet. Betrachtet man die betriebswirtschaftlichen Teildisziplinen, so stößt man auf eine vielfältige Verwendung des Wertschöpfungsbegriffs, so z. B.:³

- Produktionswirtschaft: Wertschöpfung als Grundlage der Produktionstiefe
- Steuerlehre: Wertschöpfung als Grundlage der Umsatzsteuer
- Strategisches Management: Wertschöpfung als Aspekt von Geschäftsmodellen sowie Wertschöpfungsketten
- Supply Chain Management: Wertschöpfungsnetzwerke
- Controlling: Wertschöpfungsorientierte Controlling-Konzeption⁴
- Rechnungswesen: Wertschöpfungsrechnung

Auch innerhalb dieser Teildisziplinen wird der Wertschöpfungsbegriff nicht einheitlich verwendet. So gibt es selbst im Rechnungswesen unterschiedliche Vorgehensweisen bei der Erstellung von Wertschöpfungsrechnungen (Haller, Staden & Landis 2018).⁵ Somit ergibt sich insgesamt die Situation, dass Unternehmen zwar einerseits eine Erklärung abgeben, dass der Maxime der nachhaltigen Wertschöpfung gefolgt werde – andererseits aber unklar ist, was konkret mit Wertschöpfung gemeint ist.

3 Beschreibung des Forschungsvorhabens

Angesichts der hohen Bedeutung des Wertschöpfungsbegriffs für die Corporate Governance auf der einen Seite, und der fehlenden Klarheit hinsichtlich des Begriffs Wertschöpfung auf der anderen Seite, führen die Autoren derzeit eine empirische Querschnittsuntersuchung zum Thema Wertschöpfung durch. Damit wird die zentrale Empfehlung von Haller, Staden & Landis (2018, S. 778) umgesetzt: „First of all, academia has to (re)investigate more intensively the concept of value added“.

In der Studie werden Unternehmen im deutschsprachigen Raum mittels einer online-Fragebogens zu verschiedenen Aspekten der Wertschöpfung befragt, u.a. zur Relevanz des Wertschöpfungsbegriffs im Unternehmen sowie zum konkreten Verständnis von Wertschöpfung. Aufgrund der Ausstrahlungswirkung des DCGK wird hierbei keine Beschränkung auf kapitalmarktorientierte Unternehmen vorgenommen.

³ Siehe hierzu bspw. die zahlreichen Beiträge in Ulrich & Baltzer (2019).

⁴ Siehe Becker, Baltzer & Ulrich 2014.

⁵ Hierbei kann insb. ein kontinentaleuropäisches von einem angelsächsischen Modell unterschieden werden.

Literaturverzeichnis

- Becker, W.; Baltzer, B.; Ulrich, P. (2014): Wertschöpfungsorientiertes Controlling. Kohlhammer, Stuttgart.
- Fabisch, M. (2010): Neuausrichtung des DCGK mit Schwerpunktsetzung auf Nachhaltigkeitsmanagement. Zeitschrift für Corporate Governance, 7(3):119-122.
- Haller, A.; Staden, C.J. van; Landis, C. (2018): Value Added as part of Sustainability Reporting: Reporting on Distributional Fairness or Obfuscation?. Journal of Business Ethics, 152:763-781.
- Hecker, A. (2009): Die aktuellen Änderungen des Deutschen Corporate Governance Kodex im Überblick. Betriebs-Berater, 58(32):1654-1659.
- Kann, A. van; Keilueit, A. (2009): Die aktuellen Änderungen des Deutschen Corporate Governance Kodex. Der Betrieb, 56(50):2699-2702.
- Küpper, H.-U. (2007): Betriebswirtschaftslehre als Wertschöpfungstheorie – Perspektiven für die Entwicklung einer mehrdimensionalen Theorie der Unternehmung. Zeitschrift für betriebswirtschaftliche Forschung, 56(Sonderheft): 1-26.
- Lingnau, V.; Kreklow, K. (2011): Ausrichtung der Unternehmensführung auf nachhaltige Wertschöpfung. Zeitschrift für Corporate Governance, 6(4):192-197.
- Ulrich, P. (2011): Corporate Governance in mittelständischen Familienunternehmen. Gabler, Wiesbaden.
- Ulrich, P.; Baltzer, B. (2019): Wertschöpfung in der Betriebswirtschaftslehre. Springer Gabler, Wiesbaden.

Die Bilanzierung von Crypto Assets nach verschiedenen Rechnungslegungsnormen – Vergleich und mögliche Vorschläge zur Anpassung

Extended Abstract

Dr. Wladislav Gawenko

Technische Universität Chemnitz, Fakultät für Wirtschaftswissenschaften, Chemnitz, E-Mail: wladislav.gawenko@wiwi.tu-chemnitz.de

Abstract

Crypto Assets gewinnen immer mehr an Bedeutung. Damit steigt auch das Interesse, wie diese bilanziell erfasst werden können. Die Veröffentlichung des Accounting Standards Update zur Bilanzierung von Crypto Assets vom FASB im Dezember 2023 verdeutlicht dies. Im Rahmen des Vortrags soll die aktuelle Behandlung von Crypto Assets nach HGB, OR, IFRS und US-GAAP dargestellt und Lösungsvorschläge zur Anpassung von Bilanzierungsregeln diskutiert werden. U.a. wird dabei auf einen Beitrag, der derzeit in Arbeit ist und Ende des Jahres veröffentlicht wird, eingegangen.

Extended Abstract

Im Rahmen der diesjährigen CARF-Konferenz soll eine Präsentation zu einem für Ende dieses Jahres geplanten Beitrag zur Bilanzierung von Crypto Assets nach verschiedenen Rechnungslegungsnormen gehalten werden. Dabei werden aktuelle Regelungen ausgewählter Rechnungslegungsnormen dargestellt, miteinander verglichen und mögliche Anpassungsvorschläge diskutiert. Als Rechnungslegungsnormen werden dabei das HGB, Obligationenrecht, die IFRS und US-GAAP betrachtet.

In den letzten Jahren wurden in der Wissenschaft und Praxis einige Beiträge und Hilfestellungen publiziert, die sich mit Bilanzierungs- und Bewertungsfragen von Crypto Assets beschäftigen (vgl. z.B. Gerlach & Oser, 2018, S. 1541; Kirsch & Wieding, 2018, S. 2731 ff.). Dabei werden auch mehrere Rechnungslegungsnormen betrachtet und miteinander verglichen (vgl. z.B. Behringer/Passardi/Noto, 2021, S. 27). Allerdings beschäftigen sich die bisherigen Beiträge nur mit einem Teilbereich von Crypto Assets und damit gibt es noch kein (zusammenfassendes) Werk, welches den Großteil von Crypto Assets nach verschiedenen Rechnungslegungsnormen betrachtet. So lassen sich Crypto Assets in verschiedene Oberkategorien einteilen, die noch in weitere Unterkategorien differenziert werden können (Abb. 1) (vgl. z.B. EFRAG, 2020, S. 56; Gawenko/Hinz, 2023, S. 166 f.).

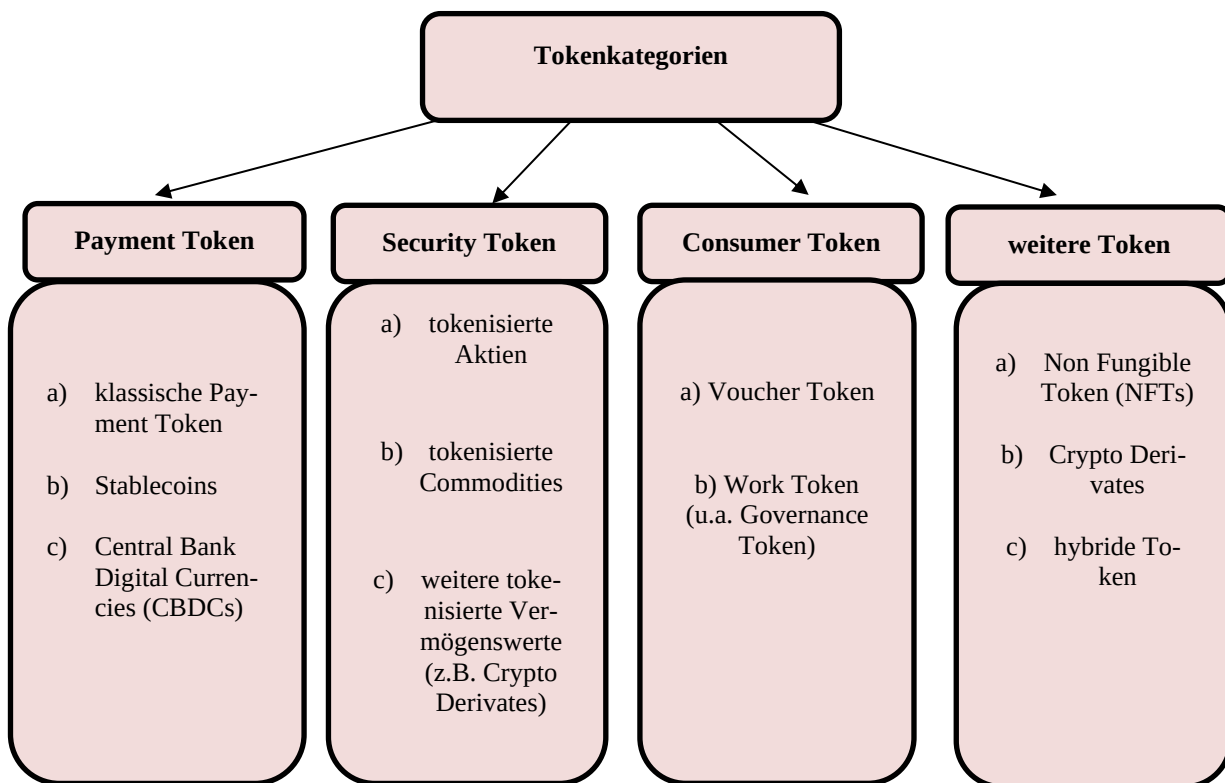


Abbildung 1: Tokenkategorien

Payment Token sind bisher im Schrifttum am häufigsten betrachtet worden. Klassische Payment Token sind Zahlungsmittel in einem konkreten Blockchain-Netzwerk. Ein Beispiel ist Bitcoin, welches als Zahlungsmittel in der Bitcoin-Blockchain fungiert. Außerhalb des Blockchain-Netzwerks sind klassische Payment Token i.d.R. nicht als gesetzliche Nominalgeldwährungen zugelassen (EFRAG, 2020, S. 123). Stablecoins sind auf der Kryptographie basierte Zahlungsmittel, die an eine übliche FIAT-Währung, wie z.B. dem US-Dollar gebunden sind (z.B. Tether USD). Mit diesen soll Wertstabilität geschaffen werden und es ist möglich, andere Kryptowährungen oder Crypto Assets auf zentralen oder dezentralen Börsen zu erwerben. CBDCs sind von einer Zentralbank herausgegebene digitale Zahlungsmittel und dürften damit (anders als klassische Payment Token und Stablecoins) als gesetzliche Nominalgeldwährung zählen (vgl. Gawenko/Hinz, 2023, S. 166 f. m.w.N.).

Security Token sind Rechte an bestimmten Werten, wie z.B. Aktien oder auch anderen Crypto Assets. Consumer Token dienen dazu, in einem Programm zum bestimmten Zweck genutzt oder „konsumiert“ zu werden. Eine übliche Form sind sog. Work Token, die dazu dienen, eine bestimmte Anwendung „am Laufen“ zu halten und/oder bestimmte Funktionen in einer dezentralen Plattform ausführen zu können. Z.B. können in einer Metaverse-Plattform zahlreiche digitale Gegenstände erst durch bestimmte Work Token repariert werden usw. Zu den weiteren Token gehören z.B. NFTs, die einzigartige, digitale Kunststücke, Figuren usw. sind und Zugang zu einer Metaverse-Plattform gewähren können (aber nicht zwingend müssen). Hybride Token sind Token, die die Funktionen mehrerer Token kombinieren und demnach keiner Kategorie zugeordnet werden können (vgl. Gawenko/Hinz, 2023, S. 166 f. m.w.N.).

Wie bereits erwähnt, wurde bisher primär auf die Bilanzierung von Payment Token bzw. Kryptowährungen im Schrifttum eingegangen. Dabei wird i.d.R. eine Bilanzierung von Kryptowährungen als immaterielle Vermögenswerte, im Vorratsvermögen, Finanzinstrumente und/oder Zahlungsmittel und Zahlungsmitteläquivalenten diskutiert (Tab. 1).

Tabelle 1: Bilanzierung Kryptowährungen -Vergleich

	HGB	Obligationenrecht	IFRS	US-GAAP
klassische payment token	Immaterielle Vermögensgegenstände, Vorräte, sonstige Vermögensgegenstände	Immaterielle Vermögenswerte, Wertschriften (Finanzinstrumente), Vorräte	intangibles, inventory	intangibles

Zu erwähnen ist, dass es nach HGB, Obligationenrecht (und SWISS GAAP FER) sowie IFRS noch keine gesetzlichen Regelungen gibt, die sich speziell auf Crypto Assets beziehen. Allerdings wurden einige Stellungnahmen z.B. seitens der ExpertSuisse, IFRS IC und EFRAG veröffentlicht, die die Bilanzierung und Bewertung von Kryptowährungen diskutieren (vgl. ExpertSuisse, 2017; IFRS IC, 2019; EFRAG, 2020). Lediglich in den USA hat das FASB im Dezember 2023 einen Standard zur Bilanzierung von Crypto Assets publiziert (FASB, 2023) und nimmt damit eine gewisse Vorreiterrolle ein.

Kryptowährungen werden in den bisherigen Stellungnahmen und Standards nicht als liquide Mittel gesehen, da sie keine gesetzlich anerkannten Nominalgeldwährungen sind. Des Weiteren wird nach HGB, Obligationenrecht und IFRS eine Bilanzierung als immaterielle Vermögenswerte und im Vorratsvermögen gesehen (Behringer/Passardi/Noto, 2021, S. 27). Lediglich US-GAAP erlaubt in dem neuen Standard 350-60 eine Bilanzierung ausschließlich als immaterieller Vermögenswert (unabhängig davon, ob die Kryptowährungen erworben oder geschürft werden). Interessant ist zudem, dass in der Schweiz zusätzlich eine Bilanzierung als Finanzinstrumente gesehen wird, was in den anderen Ländern nicht der Fall ist.

Ebenso ergeben sich bei der Bewertung einige Unterschiede. Nach HGB gilt bei der Folgebewertung entweder das gemilderte oder strenge Niederstwertprinzip. Nach Obligationenrecht besteht ein Wahlrecht, die Folgebewertung mit den fortgeführten Anschaffungs- oder Herstellungskosten (AHK) oder mit den Marktpreisen durchzuführen. Nach IFRS wird ermöglicht, auf das cost model oder das revaluation model zurückzugreifen, wenn die Kryptowährungen als intangibles erfasst werden. Bei inventories kann entweder der niedrigere Wert aus AHK und dem Nettoveräußerungswert oder gem. IAS 2.3 (b) der fair value abzgl. Veräußerungskosten verwendet werden (Kirsch & Wieding, 2018, S. 2731 ff.; Behringer/Passardi/Noto, 2021, S. 42 f.). Nach 360-50 US-GAAP ist bei der Folgebewertung auf den fair value abzustellen mit erfolgswirksamer Erfassung der Werterhöhungen und -minderungen (FASB, 2023, S. 10).

Im Rahmen des geplanten Beitrags wird zudem die Bilanzierung bisher im Schrifttum kaum oder zum Teil betrachteten Crypto Assets dargestellt und Lösungsvorschläge zur Bilanzierung diskutiert. In der nachfolgenden Tabelle sind diese Bilanzierungsmöglichkeiten abgebildet. Darüber hinaus wird auch dargestellt, wie bestimmte DeFi-Funktionen (z.B. Staking und Liquidity Mining) erfasst werden können. Dies soll im Vortrag ebenfalls diskutiert werden.

Tabelle 2: Bilanzierung Crypto Assets - Lösungsvorschläge

	HGB	Obligationenrecht	IFRS	US-GAAP
CBDCs, Stablecoins	Liquide Mittel, Finanzinstrumente	Liquide Mittel, Finanzinstrumente	Liquide Mittel, Finanzinstrumente	Liquide Mittel, Finanzinstrumente
Security Token	Finanzinstrumente	Finanzinstrumente	Finanzinstrumente	Finanzinstrumente
Consumer Token	Immaterielle Vermögensgegenstände, Vorräte, sonstige Vermögensgegenstände, Finanzinstrumente	Immaterielle Vermögenswerte, Finanzinstrumente, Vorräte	intangibles, inventory, Finanzinstrumente	intangibles, Finanzinstrumente
NFTs	Immaterielle Vermögensgegenstände, Vorräte, sonstige Vermögensgegenstände	Immaterielle Vermögenswerte, Vorräte	intangibles, inventory	intangibles

Literaturverzeichnis

- Behringer, S.; Passardi, M.; Noto, D. (2021): Kryptowährungen im Rechnungswesen. Springer, Wiesbaden.
- European Financial Reporting Advisory Group (2020): Accounting for Crypto-Assets (Liabilities). <https://www.efrag.org/Assets/Download?assetUrl=/sites/webpublishing/SiteAssets/EFRAG%2520Discussion%2520Paper-Accounting%2520for%2520Crypto-Assets%2520%28Liabilities%29-%2520July%25202020.pdf>. Abgerufen am 10.04.2024.
- ExpertSuisse (2017): Behandlung von Bitcoin nach OR-Rechnungslegungsrecht. In: Ausgewählte Fragen und Antworten zum neuen Rechnungslegungsrecht. EXPERTsuisse.
- Financial Accounting Standards Board (2023): Accounting Standards Update. Intangibles-Goodwill and Other-Crypto Assets (Subtopic 350-60). FASB.
- Gawenko, W.; Hinz, M. (2023): Crypto Assets – Anwendung des Hedge Accounting nach IFRS. IRZ, 18: 165-171.
- Gerlach, I.; Oser, P. (2018): Ausgewählte Aspekte zur handelsrechtlichen Bilanzierung von Kryptowährungen. Der Betrieb, 65: 1541–1547.
- Kirsch, H.J.; Wieding, F. (2017): Bestandsbilanzierung von Bitcoin im IFRS-Kontext. IRZ, 12: 115–120.
- IFRS Interpretations committee (2019): Holdings of cryptocurrencies. <https://www.ifrs.org/content/dam/ifrs/supporting-implementation/agenda-decisions/2019/holdings-of-cryptocurrencies-june-2019.pdf>. Abgerufen am 10.04.2024.

Evaluating the Impact of AI Mention in Key Audit Matters on Stakeholder Perceptions: An Experimental Study on Decision Usefulness and Trust

Extended Abstract

Prof. Dr. Theresa Herrmann

Technische Hochschule Ulm, Institut für Management und Unternehmertum, Ulm, E-Mail: theresa.herrmann@thu.de

Abstract

Broadly, the integration of artificial intelligence in audit procedures is often associated with benefits such as higher accuracy and efficiency of the audit as well as higher reliability of the auditor's opinion. This research aims to investigate if these benefits hold, when stakeholders' technology acceptance varies. Drawing onto prior research, which focuses on artificial intelligence in basic audit procedures, this project aims to investigate the perception of Key Audit Matters mentioning the use of artificial intelligence to address complex audit issues. In line with the Technology Acceptance Model this experimental research project predicts that stakeholders with higher (lower) technology acceptance value the use and mention of artificial intelligence in Key Audit Matters positively (negatively). This will then lead to higher (lower) perceived usefulness of the audit report as well as higher (lower) trust in the auditor.

1 Introduction

In recent years, artificial intelligence (AI) was integrated into audit processes to help with accuracy and efficiency. Current research finds that the integration of AI leads to higher reliability in the audit process (Fedyk et al., 2022). While in the early stages AI was only used for rather simple audit procedures (i.e. analyzing large data sets), technology can now support the audit of complex audit issues (i.e. impairment testing level 3 fair values) (Föhr et al., 2024). Such complex audit issues for public interest entities (PIEs) are enclosed in Key Audit Matters (KAM) as part of the audit report. While admittedly practical use of task-specific models for complex audit tasks is scarce at present, the future application is beyond doubt (Cao et al., 2024; Föhr et al., 2024). While there is research on the effects of AI in terms of the audit procedures, it remains unclear how financial statement users perceive the use of AI in audit procedures. This research predicts that stakeholders perceive the use of AI in audit procedures differently, depending on their level of technology acceptance in accordance with the Technology Acceptance Model (TAM). Specifically, stakeholders with higher levels of technology acceptance value the use of AI in audit processes more, leading to higher decision usefulness and trust. For stakeholders with lower levels of technology acceptance, who find AI difficult to understand or unreliable, the use of AI in audit processes will have the opposite effect.

2 Theoretical Background and Hypotheses

KAM provide detailed information on significant issues during the audit (Gold et al., 2012). Thus, they can enhance the overall decision usefulness of the audit report, which leads to perceived increased transparency and ultimately to increased user confidence (Gold et al., 2020; Sirois et al., 2018). Consequently, KAM contribute to narrow the audit expectation gap by providing more insights into the audit process as well as the critical judgement of the auditor (Elmarzouky et al., 2024; Gold et al., 2012). Further, KAM make critical audit issues salient for the financial statement users (i.e. stakeholders) (Maines & McDaniel, 2000).

The integration of Artificial Intelligence (AI) in audit processes represents a significant evolution in the field of auditing, promising enhanced accuracy, efficiency, and insight (Azimi & Agrawal, 2021; Föhr et al., 2024; Seidenstein et al., 2024; Wei et al., 2024). While research provides evidence that AI improves audit quality also for complex audit tasks (Fedyk et al., 2022) this research aims to shed more light onto the perception of AI in audit processes by stakeholders. To address this aim the author incorporates aspects from agency and signaling theory as well as the TAM in this research to contextualize the investigation of AI audit procedures in the audit process and its impact on perceived decision usefulness as well as perceived trust in the technology and the auditor.

Agency theory highlights the relationship between principals (stakeholders) and agents (auditors) driven by trust and accountability. When auditors use AI and mention the use in KAM they contribute to reducing information asymmetry and enhancing audit quality. By doing so, they can strengthen the principal-agent relationship, as they use the best available tools to act in the principals' interest (Jensen & Meckling, 1976). To reduce information asymmetry further, auditors enhance stakeholders trust by signaling the use of advanced audit technologies. Thus, the use of AI can serve as a positive signal that the auditor conducts a thorough and reliable audit (Spence, 1973). Yet, the effectiveness of this signal depends on the stakeholders' perception of using AI for the audit of a significant issue.

The TAM suggests that both perceived usefulness and ease of use are critical factors that influence stakeholders' acceptance of new technologies (i.e. AI) (Davis, 1989; Inayatulloh et al., 2021). Consequently, mentioning the use of AI to address a KAM may enhance stakeholders' confidence in the audit if they have a high acceptance of new technologies. However, if stakeholders perceive AI as difficult to understand or unreliable, and thus have a rather low acceptance of new technologies, it may negatively affect their trust in the audit (Altiero et al., 2024; van Liempd et al., 2024).

This leads to the following hypotheses:

- H1a Mentioning AI in KAM increases (decreases) the perceived usefulness of the audit report for stakeholders with higher (lower) technology acceptance.
- H1b Mentioning AI in KAM positively (negatively) influences stakeholders' perceptions of the ease of understanding and using the audit report for stakeholders with higher (lower) technology acceptance.
- H2 Mentioning AI in KAM enhances (diminishes) the perceived decision usefulness of the audit report for stakeholders with higher (lower) technology acceptance.
- H3 Mentioning AI in KAM increases (decreases) stakeholders' trust in the audit report for stakeholders with higher (lower) technology acceptance.

Further, an interaction effect between perceived usefulness and perceived ease of use on stakeholders' trust and decision usefulness is anticipated.

- H4 High (low) perceived usefulness combined with high (low) perceived ease of use will lead to the greatest increase (decrease) in trust and decision usefulness.

3 Methodology

This study uses a 2x2 between-subjects experimental design to investigate the impact of mentioning AI in Key Audit Matters (KAM) for different stakeholder groups (high vs. low technology acceptance) on decision usefulness and trust in audit opinions. The first independent variable Technology Acceptance Level is manipulated at two levels: TAM_high vs. TAM_low. Students currently enrolled in the degree for Business informatics proxy for stakeholders with TAM_high and students currently obtaining a degree in social sciences proxy for stakeholders with TAM_low. The second independent variable KAM is again manipulated at two levels: AI vs. No_AI. Participants of both groups will be randomly assigned to one of two groups: one receiving an audit report with KAM mentioning AI for auditing a complex task (i.e. Level 3 fair value estimates), and the other receiving an audit report without AI use and mention. The dependent variables are Decision Usefulness and Trust. Decision Usefulness is the extent to which the audit report aids stakeholders in making informed decisions. This is measured on a Likert-scale using items like "The information provided in the audit report is useful for making investment decisions" and "The audit report provides relevant information for assessing the company's financial health." Trust is the degree of confidence stakeholders have in the audit opinion, measured using items like "The mention of AI in the KAM increases my trust in the audit opinion" and "I believe the audit report is more reliable with AI mentioned in the KAM."

4 Outlook

This study investigates how the use and mention of AI affects decision usefulness of the audit report as well as the trust in the auditor. The project is the first to evaluate the perceived usefulness and trust of AI audit mentions and distinguish between different stakeholder groups. Thus, this research could be relevant for auditors, standard setters as well as practitioners.

References

- Altiero, E. C., Baudot, L., & Hazgui, M. (2024). Technology and Evidence in Non-Big 4 Assurance Engagements: Insights from the COVID-19 Pandemic. *Auditing: A Journal of Practice & Theory*, 43(2), 1–18. <https://doi.org/10.2308/AJPT-2022-068>
- Azimi, M., & Agrawal, A. (2021). Is Positive Sentiment in Corporate Annual Reports Informative? Evidence from Deep Learning. *The Review of Asset Pricing Studies*, 11(4), 762–805. <https://doi.org/10.1093/rapstu/raab005>
- Cao, J., Kristanto, A. B., & Gu, Z. (2024). Evolution of Research Streams and Future Research Directions in Accounting Education: Quantitative Systematic Literature Review. *Issues in Accounting Education*, 1–35. <https://doi.org/10.2308/ISSUES-2023-094>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319. <https://doi.org/10.2307/249008>
- Elmarzouky, M., Hussainey, K., & Abdelfattah, T. (2024). Key audit matters: a systematic review. *International Journal of Accounting, Auditing and Performance Evaluation*, 20(3/4), Article 138473, 319–341. <https://doi.org/10.1504/IJAAPE.2024.138473>
- Fedyk, A., Hodson, J., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies*, 27(3), 938–985. <https://doi.org/10.1007/s11142-022-09697-x>
- Föhr, T. L., Schreyer, M., Moffitt, K., & Marten, K.-U. (2024). Deep Learning Meets Risk-Based Auditing: A Holistic Framework for Leveraging Foundation and Task-Specific Models in Audit Procedures: Working Paper (SSRN).
- Gold, A., Gronewold, U., & Pott, C. (2012). The ISA 700 Auditor's Report and the Audit Expectation Gap – Do Explanations Matter? *International Journal of Auditing*, 16(3), 286–307. <https://doi.org/10.1111/j.1099-1123.2012.00452.x>
- Gold, A., Heilmann, M., Pott, C., & Rematzki, J. (2020). Do key audit matters impact financial reporting behavior? *International Journal of Auditing*, 24(2), 232–244. <https://doi.org/10.1111/ijau.12190>
- Inayatulloh, Kumala, D., Mangruwa, R. D., & Dewi, E. P. (2021, September 18–19). Technology Acceptance Model for Adopting E-Accounting Information System based on open source for SMEs. In 2021 International Seminar on Application for Technology of Information and Communication (iSemantic) (pp. 263–267). IEEE. <https://doi.org/10.1109/iSemantic52711.2021.9573203>
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)
- Maines, L. A., & McDaniel, L. S. (2000). Effects of Comprehensive-Income Characteristics on Nonprofessional Investors' Judgments: The Role of Financial-Statement Presentation Format. *The Accounting Review*, 75(2), 179–207. <https://doi.org/10.2308/accr.2000.75.2.179>
- Seidenstein, T., Marten, K.-U., Donaldson, G., Föhr, T. L., Reichelt, V., & Jakoby, L. B. (2024). Innovation in Audit and Assurance: A Global Study of Disruptive Technologies. *Journal of Emerging Technologies in Accounting*, 21(1), 129–146. <https://doi.org/10.2308/JETA-2022-026>
- Sirois, L.-P., Bédard, J., & Bera, P. (2018). The Informational Value of Key Audit Matters in the Auditor's Report: Evidence from an Eye-Tracking Study. *Accounting Horizons*, 32(2), 141–162. <https://doi.org/10.2308/acch-52047>
- Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355. <https://doi.org/10.2307/1882010>
- van Liempd, D., Theis, J. C., & Sutton, S. G. (2024). The Risk of Technology Dominance in Using Digital Decision Aids in Assurance Engagements—Evidence from a Survey among Danish Auditors. *Accounting Horizons*, 1–20. <https://doi.org/10.2308/HORIZONS-2023-070>

Wei, D., Cho, S., Vasarhelyi, M. A., & Te-Wierik, L. (2024). Outlier Detection in Auditing: Integrating Unsupervised Learning within a Multilevel Framework for General Ledger Analysis. *Journal of Information Systems*, 1–20. <https://doi.org/10.2308/ISYS-2022-026>

Digitalisierung als Antwort auf den Fachkräftemangel? Sich ändernde Prozesse und Anforderungen in Rechnungslegung und Wirtschaftsprüfung und ihre Auswirkungen auf die Hochschulausbildung

Research Paper

Prof. Dr. Tim Kampe

Hochschule Bielefeld – University of Applied Sciences and Arts, Fachbereich Wirtschaft, Bielefeld, E-Mail: tim.kampe@hsbi.de

Abstract

Die fortschreitende Digitalisierung in Unternehmen führt dazu, dass Routinetätigkeiten zunehmend automatisiert werden und der Bedarf an menschlicher Arbeitskraft sinkt. Dies zeigt sich insbesondere auch in der Rechnungslegung von Unternehmen und in gewissem Umfang bei der Prüfung von Jahresabschlüssen. Der Purchase-to-Pay-Prozess bzw. der Teilprozess Invoice-to-Pay dient im vorliegenden Beitrag⁶ als anschauliches Beispiel für die erheblichen Rationalisierungspotentiale, die sich aus der Digitalisierung ergeben. Dieses Einsparpotential kann auf der einen Seite dazu beitragen, dem Fachkräftemangel zu begegnen. Auf der anderen Seite ergeben sich für Wirtschaftsprüfungsgesellschaften auch neue Chancen, weitere innovative Dienstleistungen anzubieten. Der Wegfall von Routineaufgaben und das Entstehen neuer Möglichkeiten durch die Digitalisierung führen jedoch auch zu einem Anstieg der qualitativen Anforderungen an (zukünftige) Mitarbeitende im Rechnungswesen und in der Wirtschaftsprüfung. Daher ist es erforderlich, dass die Hochschulen ihre Lehre entsprechend anpassen, wie ebenfalls im Beitrag diskutiert wird.

⁶ Eine frühere Version dieses Beitrags wurde am 17.05.2024 auf der 3rd Bielefeld International Conference on Applied Business (BiCAB) on “Skills Shortage- Approaches and Solutions from Universities, Companies and Politics” präsentiert.

1 Einleitung und Grundlagen

Viele Jahre lang galt der Beruf des Wirtschaftsprüfers als attraktives Beschäftigungsfeld für Absolventinnen und Absolventen von betriebswirtschaftlichen Studiengängen. Die großen (Big4) Wirtschaftsprüfungsgesellschaften, darunter Deloitte, PricewaterhouseCoopers (PWC), KPMG und Ernst & Young (EY), gehörten in Absolventenbefragungen lange Zeit zu den attraktivsten Arbeitgebern (Kampe, 2011, S. 96). Der Beruf des Wirtschaftsprüfers ist in Deutschland, Österreich und der Schweiz, wie in den meisten Industrienationen, als freier Beruf ausgestaltet; der Nachweis der persönlichen und fachlichen Eignung muss durch das Bestehen eines (Staats-) Examens erbracht werden. Im Gegenzug haben Wirtschaftsprüferinnen und Wirtschaftsprüfer dann die alleinige Befugnis (Vorbehaltsaufgabe) zur Durchführung von gesetzlichen Prüfungen. Dabei handelt es sich um die Bestätigung der Zuverlässigkeit und Richtigkeit der Angaben, die Unternehmen in ihrer Buchführung, ihrem Jahres- oder Konzernabschluss und ggf. im Lagebericht oder Konzernlagebericht dokumentieren. Das Berufsbild umfasst aber auch andere Tätigkeiten, wie z.B. betriebswirtschaftliche Prüfungen, Steuerberatung, Unternehmensberatung, Gutachten und treuhänderische Tätigkeiten (Kampe, 2011, S. 17-24).

Allerdings ist das Interesse am Berufsfeld Wirtschaftsprüfung bei Absolventinnen und Absolventen in den letzten 20 Jahren stark zurückgegangen, was nicht allein mit der demografischen Entwicklung zu erklären ist. Vielmehr haben auch das Image und der Ruf des Berufsstandes durch die zahlreichen Bilanzskandale der Vergangenheit erheblich gelitten (Kampe, 2011, S. 97). Eine Tätigkeit in der Rechnungslegung oder Wirtschaftsprüfung wird von Studierenden zudem als eher langweilig und eintönig empfunden (Hossenfelder, 2020, S. 241; Mujkanovic, 2014, S. 25). Derzeit suchen die Wirtschaftsprüfungsgesellschaften händierend nach jungen Talenten, und Stellen im Rechnungswesen bleiben lange unbesetzt (Bleiber, 2023). So hat sich beispielsweise in Deutschland die Zahl der offenen Stellen im Rechnungswesen seit 2012 fast verdoppelt (IAB, 2024a).

Die Digitalisierung bzw. digitale Transformation bringt nun nicht nur erhebliche Veränderungen im allgemeinen (Berufs-) Leben mit sich, sondern birgt durch den Einsatz von künstlicher Intelligenz und Automatisierung auch erhebliche Einsparpotentiale beim Personaleinsatz in Unternehmen. Dies wird auf der einen Seite absehbar zu mehr Effizienz und Produktivität führen. Auf der anderen Seite wird mit erheblichen Auswirkungen auf die gesamte Belegschaft in Unternehmen gerechnet, was auch in der Notwendigkeit einer kooperativen Zusammenarbeit zur Gestaltung der digitalen Transformation resultiert (Boritz & Stratopoulos 2023; IAB, 2024a). Nach Bravidor & Lösse (2018, S. 784) können drei verschiedene Dimensionen bzw. Stufen der Digitalisierung unterschieden werden:

- „Digitization“ oder Digitalisierung im engeren Sinne beschreibt zunächst lediglich den Übergang von der analogen Erfassung und Verarbeitung von Daten zur digitalen. Diese grundlegende, technische Form der Digitalisierung ist damit die Voraussetzung für die folgende Stufe.
- „Digitalization“ oder Digitalisierung im weiteren Sinne bezieht sich auf die Nutzung digitaler Technologien, um einen höheren Automatisierungsgrad zu erreichen und Geschäftsmodelle anzupassen.
- „Digital Transformation“ geht über die Unternehmensebene hinaus und beschreibt die Auswirkungen auch auf die gesellschaftliche und private Ebene. Digitale Transformation ist ein kontinuierlicher Prozess, der die Veränderung durch Informationstechnologien auf mikro- und makroökonomischer Ebene umfasst.

Im Folgenden werden – nach einer kurzen Darstellung der Methodik und Vorgehensweise (2) – die Auswirkungen der Digitalisierung auf das Rechnungswesen in Unternehmen und die Abschluss- bzw. Wirtschaftsprüfung eingehender untersucht (3). Die sich verändernden Prozesse und Anforderungen werden ebenso diskutiert wie die sich daraus ergebenden Auswirkungen auf die erforderlichen Fähigkeiten und Kompetenzen der Mitarbeitenden in diesen Bereichen. Darüber hinaus werden abschließend die Auswirkungen auf die Hochschulausbildung und die Rekrutierungsstrategien von Unternehmen diskutiert (4).

2 Methodik und Vorgehensweise

Der Beitrag verfolgt grundsätzlich einen Mixed-Methods-Ansatz (Döring, 2023, S. 26-28). Der Grundlagenteil basiert auf einer umfangreichen Literaturanalyse (Pargmann et al., 2023), der Stand und die Perspektiven der Digitalisierung im Rechnungswesen werden unter Bezugnahme auf bestehende Studien (Kreher et al., 2024; Prümm et al., 2023; Demelius et al., 2023) dargestellt.

Die Wirtschaftsprüfungsgesellschaft KPMG führt seit 2017 eine regelmäßige Online-Befragung zum Stand der Digitalisierung im Rechnungswesen durch (Kreher et al., 2017). Hierfür wurden zuletzt mehr als 2.500 Unternehmen in Deutschland, Österreich und der Schweiz befragt (Kreher et al., 2022, S. 9). Zwischen 350 Unternehmen (Kreher et al., 2021, S. 5) und zuletzt 232 Unternehmen (Kreher et al., 2024, S. 9) nahmen regelmäßig teil. Die Unternehmen kamen aus allen Größenklassen und Branchen (Kreher et al., 2024, S. 10). Die Wirtschaftsprüfungsgesellschaft PWC führt bereits im achten Jahr eine ähnlich konzipierte Studie mit rund 100 teilnehmenden Unternehmen durch (Prümm et al., 2023, S. 3). Die Wirtschaftsprüfungsgesellschaft Deloitte Österreich hat zuletzt im Jahr 2023 unter 268 österreichischen Unternehmen aller Größen und Branchen eine Studie zur Automatisierung im Rechnungswesen und deren Auswirkungen auf die Kompetenzen der Mitarbeiter im Rechnungswesen durchgeführt (Demelius et al., 2023, S. 8). Dies war die vierte Wiederholung einer Studie, die erstmals im Jahr 2016 durchgeführt wurde (Vlk et al., 2016).

Anhand von synthetischen und persönlichen Erfahrungen werden die Veränderungen durch die Digitalisierung am Beispiel des Purchase-to-Pay- bzw. Invoice-to-Pay-Prozesses vor und nach Digitalisierung dargestellt (Diehm & Benzinger, 2018). Die Auswirkungen auf die Wirtschaftsprüfung basieren zum einen auf vorhandener Literatur (Feld & Pöhlmann, 2017; Langhein et al., 2018; Bruckner, 2019; Downar & Fischer, 2019) und zum anderen auf zwölf Experteninterviews mit Wirtschaftsprüferinnen und Wirtschaftsprüfern aus der Region Ostwestfalen-Lippe in Deutschland (Döring, 2023, S. 371-372; Strübing, 2018, S. 106-108). Diese Interviews wurden aufgezeichnet und transkribiert (Strübing, 2018, S. 117-119). Auf Basis der Literatur und der Interviews wurden Kategorien gebildet und die Transkriptionen entsprechend kodiert (Mayring, 2022, 49-103). Ziel war es, zum einen einen aktuellen und explorativen Zugang zum Feld zu erhalten und zum anderen die theoretischen Ausführungen mit konkreten Aussagen illustrieren zu können.

Die Auswirkungen auf und die Umsetzung in der Hochschullehre basieren auf dem aktuellen Diskussionsstand innerhalb der Fachgruppe Controlling, Finanz- und Rechnungswesen der Hochschule Bielefeld und werden durch Literatur (Kesten, 2019; Vanini, 2019; Weißenberger et al., 2019) flankiert.

3 Ergebnisse

Die Digitalisierung ist an sich kein neues Phänomen in Rechnungswesen und Wirtschaftsprüfung, jedoch hat die Dynamik und Durchdringung dieser Bereiche deutlich zugenommen (Demelius et al., 2023, S. 10; Prümm et al., 2023, S. 9). Dadurch ergibt sich die Möglichkeit, insbesondere repetitive Tätigkeiten stärker zu automatisieren, Mitarbeiter zu entlasten und Kapazitäten freizusetzen (Kreher et al., 2024, S. 29; Prümm et al., 2023, S. 6). Im Folgenden wird zunächst allgemein der Stand der Digitalisierung im Rechnungswesen dargestellt (3.1). Die konkreten Potentiale der Digitalisierung werden anhand des Invoice-to-Pay-Prozesses vor (3.2) und nach Digitalisierung (3.3) dargestellt. Aus der Digitalisierung des Rechnungswesens ergeben sich Konsequenzen für die Abschlussprüfung aber auch darüber hinaus ist die Wirtschaftsprüfung durch die digitale Transformation betroffen (3.4).

3.1 Stand der Digitalisierung im Rechnungswesen

Kreher et al. (2021; 2022; 2024) können in ihrer regelmäßigen Befragung deutscher Unternehmen zeigen, dass die Digitalisierung im Rechnungswesen, insbesondere in den Kernprozessen, voranschreitet. Die Basissysteme werden zunehmend homogenisiert, die Stammdatenqualität verbessert und papierlose End-to-End-Prozesse etabliert. Mehr als die Hälfte der befragten Unternehmen hat dies bereits flächendeckend oder zumindest in Teilbereichen umgesetzt, während die Mehrheit der übrigen Unternehmen in der Umsetzung ist oder dies plant (Kreher et al, 2024, S. 14). Zusammengefasst ist der Stand der Digitalisierung im Rechnungswesen in Abbildung 1 wiedergegeben.



Abbildung 1. Stand der Digitalisierung im Rechnungswesen (eigene Abbildung).

Im Rechnungswesen der befragten Unternehmen hat sich die Cloud-Technologie als führender Trend etabliert: 22 % der befragten Unternehmen nutzen sie flächendeckend, 37 % in ersten Pilotprojekten (Kreher et al., 2024, S. 17). Weitere 12 % planen konkret die Einführung und 14 % diskutieren die Einführung. Lediglich 14 % der befragten Unternehmen planen dies nicht, sehen darin keinen Mehrwert oder haben keine Angaben gemacht (Kreher et al., 2024, S. 17). Blockchain oder andere neuere, spezielle Trends spielen dagegen kaum eine Rolle (Kreher et al., 2024, S. 18; Prümm et al. 2023, S. 23). Prümm et al. (2023, S. 23) zeigen in ihrer Studie, dass der Einsatz von maschinellem Lernen und künstlicher Intelligenz mit 21 % bzw. 8 % der befragten Unternehmen zuletzt deutlich zugenommen hat. Die operativen Geschäftsprozesse werden weiterhin (vollständig) digitalisiert. Dazu werden die Prozesse (innerhalb des Unternehmens) standardisiert, bevor sie in den ERP-Systemen digital abgebildet werden. Dies erfordert Anpassungen in der Aufbau- und Ablauforganisation, wie z. B. die Neuordnung von Verantwortlichkeiten (Bruckner, 2019, S. 237; Kreher et al., 2021, S. 7).

Eine große Herausforderung liegt jedoch im Bereich der nicht-finanziellen Berichterstattung (Kreher et al, 2024, S. 39). Die fehlende Standardisierung der nicht-finanziellen Daten führt häufig zu einer mangelnden Qualität der Berichterstattung. Die regulatorischen Anforderungen in diesem Bereich sind unsicher bzw. verschärfen sich kontinuierlich, z.B. im Hinblick auf die Nachhaltigkeitsberichterstattung. Die Prozessabbildung ist hier unsicher.

Digitale Rechnungen ersetzen zunehmend Papierrechnungen (Kreher et al, 2024, S. 14). Die Digitalisierung des Purchase-to-Pay-Prozesses führt letztlich zur Papierlosigkeit. Wo Rechnungen noch nicht digital sind, werden Scan- und OCR-Lösungen eingesetzt. Der Prozess wird durch digitale Einkaufs- und Beschaffungsplattformen unterstützt (Kreher et al, 2021, S. 7). Die Auswirkungen der Digitalisierung und der Digitalisierung werden daher im Folgenden am Beispiel des Purchase-to-Pay-Prozesses näher untersucht.

3.2 Purchase-to-Pay bzw. Invoice-to-Pay Prozess ohne Digitalisierung

Der Purchase-to-Pay Prozess umfasst in der Regel den gesamten operativen Prozess von der Bedarfsmeldung bis zur Bezahlung des Einkaufs. Während der erste Teil des Prozesses (Bedarfsmeldung, Prüfung/Genehmigung, Bestellung, Wareneingang) in der Regel in der Einkaufsabteilung stattfindet und in deren Zuständigkeit fällt, ist für den zweiten Teil des Prozesses (Rechnungsbearbeitung) in der Regel das Rechnungswesen zuständig. Aus diesem Grund und um den Prozess nicht zu komplex werden zu lassen, liegt der Schwerpunkt auf dem zweiten Teilprozess des Gesamtprozesses, dem Invoice-to-Pay Prozess.

Der Prozess vom Rechnungseingang bis zur Zahlung kann grundsätzlich in sieben Schritte unterteilt werden. Dies ist in Abbildung 2 zusammengefasst. Eingehende Rechnungen (Schritt 1) können über verschiedene Eingangskanäle empfangen werden. Prinzipiell sind Post, Fax und elektronische Rechnungen möglich. Für letztere gibt es sowohl standardisierte als auch halbstandardisierte oder nicht standardisierte Kanäle, die von E-Mail, PDF, XML und EDI bis hin zu ZUGFeRD reichen können. ZUGFeRD ist ein branchenübergreifendes Datenformat für den elektronischen Rechnungsdatenaustausch, das vom Forum elektronische Rechnung Deutschland (FeRD) - mit Unterstützung des Bundesministeriums für Wirtschaft und Energie - entwickelt wurde. (FeRD, 2024). Paradoxerweise bedeutet die Vielfalt der Formate auch, dass Unternehmen eine elektronisch empfangene Rechnung häufig zur Weiterverarbeitung ausdrucken und einen standardisierten, papierbasierten Rechnungsbearbeitungsprozess beibehalten (Diehm & Benzinger, 2018, S. 841; Schmidt, 2020, S. 249).

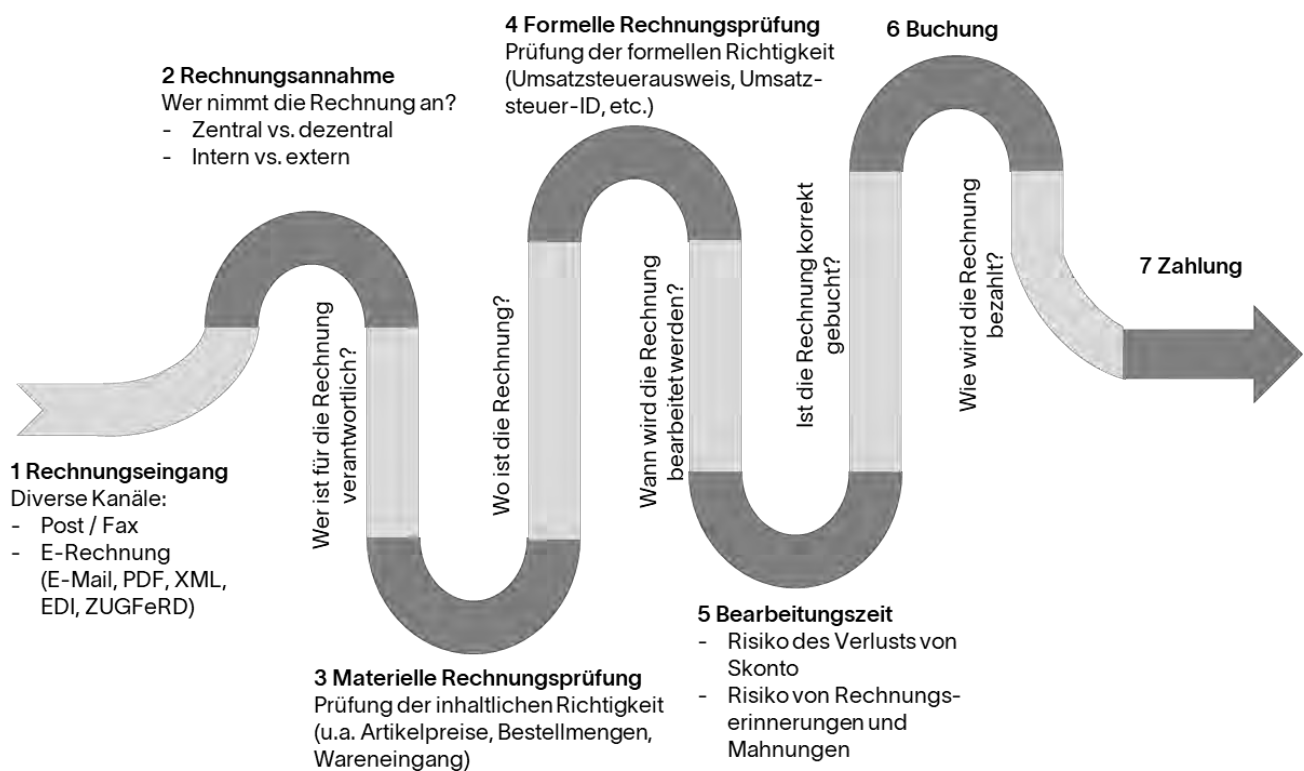


Abbildung 2. Rechnungsbearbeitung ohne Digitalisierung (modifiziert aus Diehm & Benzinger, 2018, S. 842)

Die Rechnungsannahme (Schritt 2) kann ebenfalls auf unterschiedliche Weise durchgeführt werden. Sie kann zentral im Unternehmen erfolgen (Poststelle o.ä.), aber auch dezentral in den Bereichen oder Abteilungen. Neben der internen Rechnungsannahme ist auch eine externe Annahme, z.B. durch einen Dienstleister, denkbar. Unabhängig von der Form der Annahme stellt sich die Frage: Wer ist für die Rechnung verantwortlich?

Mit der Verantwortung für die Rechnung verbunden ist die inhaltliche (materielle) Prüfung der Rechnung (Schritt 3). Das bedeutet, dass die wesentlichen Parameter der Rechnung (Bestellmenge, Preise, Bestellkonditionen) mit den vereinbarten Konditionen abgeglichen werden und vor allem der Wareneingang bestätigt wird. Da dies in der Regel dezentral durch die jeweiligen Auftraggeber erfolgt, stellt sich bei papierbasierten Abläufen oft die Frage: Wo ist die Rechnung?

Die formelle Prüfung (Schritt 4) der Rechnung wird in der Regel von der Buchhaltung, d.h. in der Abteilung Rechnungswesen durchgeführt. Formelle Richtigkeit bedeutet insbesondere die Einhaltung der steuerrechtlichen Anforderungen an eine Rechnung (korrekter Umsatzsteuerausweis, Verwendung einer Umsatzsteuer-Identifikationsnummer usw.), damit die Vorsteuer zutreffend gezogen werden kann. Der interne Postfluss zwischen den Abteilungen und die notwendigen Abstimmungen, insbesondere bei Unstimmigkeiten, können zu Verzögerungen führen, was die Frage aufwirft: Wann wird die Rechnung bearbeitet werden?

Der Zeitpunkt der Rechnungsbearbeitung bzw. die Bearbeitungszeit (Schritt 5) ist insofern wichtig, als Verzögerungen zu (finanziellen) Nachteilen führen können. Dies ist zum einen der mögliche Verlust von Skonti, wenn die entsprechenden Skontofristen nicht eingehalten werden können, und zum anderen das Risiko von Mahnungen (Reputations- bzw. Bonitätsverlust und ggf. Mahngebühren). Die zentrale Frage lautet: Ist die Rechnung korrekt gebucht?

Die eigentliche Buchung der Rechnung (Schritt 6) sorgt dafür, dass die entsprechenden Daten im Buchungssystem erfasst werden. Die Frage bleibt: Wie wird die Rechnung bezahlt? Die Zahlung (Schritt 7) erfolgt dann im Rahmen eines so genannten "Zahlungslaufs" und führt so schließlich zur Begleichung der Schuld.

Insgesamt lässt sich festhalten, dass der analoge Prozess zahlreiche Einzeltätigkeiten, Prozessbeteiligte, Schnittstellen zwischen Abteilungen und Personen und damit potenzielle Fehlerquellen aufweist. Insgesamt bindet der papierbasierte Prozess Mitarbeiterressourcen für Routineaufgaben und birgt das Risiko von Erfassungsfehlern (Diehm & Benzinger, 2018, S. 842; Schmidt, 2020, S. 249).

3.3 Invoice-to-pay Prozess nach der Digitalisierung

Die Kosten für die Bearbeitung einer Papierrechnung liegen bei den meisten Unternehmen schätzungsweise zwischen 5 und 20 Euro (Diehm & Benzinger, 2018, S. 842). Diehm und Benzinger (2018, S. 842) gehen davon aus, dass die digitale Rechnungsverarbeitung in Kombination mit der Optimierung vor- und nachgelagerter Prozesse insgesamt Kosteneinsparungen von bis zu 80 % pro Vorgang erzielen kann. Diese Kosteneinsparungen gelten jedoch nicht für alle rechnungslegungsrelevanten Prozesse, wie Kreher et al. (2024, S. 29) in ihrer Studie zeigen: für die Mehrheit der Befragten sind die Kosten für die Erstellung des Jahresabschlusses mit zunehmender Digitalisierung eher nicht gesunken, teilweise sogar gestiegen. Andererseits hat die Digitalisierung bei rund der Hälfte der befragten Unternehmen spürbar zur Beschleunigung der Prozesse bei der Abschlusserstellung und -prüfung beigetragen (Kreher et al., 2024, S. 29).



Abbildung 3. Digitale Rechnungsverarbeitung (modifiziert aus Diehm & Benzinger, 2018, S. 843)

Der typische digitale Invoice-to-Pay Prozess kann in acht Schritte unterteilt werden, wie in Abbildung 3 dargestellt (Diehm & Benzinger, 2018, S. 843). Eingangsrechnungen (Schritt 1) stellen aufgrund der vielen Kanäle, Formate und Dokumententypen eine besondere Herausforderung für Unternehmen dar (Diehm & Benzinger, 2018, S. 843; Reuschenbach et al, 2020, S. 145). Papierdokumente oder Dokumente in einem unstrukturierten Dateiformat (PDF, TIF, JPEG), die in vielen Unternehmen immer noch den Großteil der Eingangsrechnungen ausmachen, müssen mit Hilfe von OCR-Systemen (Optical Character Recognition) in ein lesbares und verarbeitbares Format umgewandelt werden (Schritt 2) (Reuschenbach et al, 2020, S. 145). Da OCR-Systeme in der Regel eine Erkennungsrate von 70-80% haben und somit keine vollständige Automatisierung ermöglichen, sind sie lediglich ein Zwischenschritt auf dem Weg zu einem vollständig digitalisierten Prozess (Diehm & Benzinger, 2018, S. 843). Im Kontext der Dokumenterkennung werden daher auch erhebliche Einsatzmöglichkeiten für KI-basierte Lösungen gesehen. In der Studie von Kreher et al. (2024, S. 15) arbeiten 80 % der befragten Unternehmen bereits mit OCR-Texterkennung oder sind gerade dabei, diese zu implementieren.

Ziel der digitalisierenden Unternehmen sollte es jedoch sein, alle Lieferanten dazu zu bewegen, Rechnungen in einem strukturierten Datenformat (EDI, XML) oder einem hybriden Datenformat (PDF/A, ZUGFeRD) auszustellen. Während die Daten in einem strukturierten Format direkt lesbar sind, kombinieren hybride Formate die visuelle Darstellung der Rechnung in einem unstrukturierten Format mit einem XML-Anhang und können daher auch automatisch verarbeitet werden. Alternativ besteht die Möglichkeit, einen externen E-Invoicing-Dienstleister zu beauftragen, der die unstrukturierten Daten verarbeitet und dem beauftragenden Unternehmen in einem standardisierten, elektronischen Format zur Verfügung stellt (Diehm & Benzinger, 2018, S. 844).

Die Dokumentenklassifizierung (Schritt 3) erfolgt nach dem Auslesen oder Übertragen der Dokumentendaten. Je nach Belegart wird in der Regel sofort die Vollständigkeit des Beleges, z.B. hinsichtlich der steuerlichen Mindestangaben, geprüft und die Stammdaten mit den im ERP-System hinterlegten Daten abgeglichen (Reuschenbach et al, 2020, S. 146). Fehlerhafte Dokumente können so frühzeitig erkannt und abgelehnt werden (Diehm & Benzinger, 2018, S. 844). Die weitere Dokumentenvalidierung (Schritt 4) kann dann je nach Auftragsart und Systemreife mehr oder weniger automatisiert erfolgen. Der automatisierte Abgleich von Mengen und Preisen mit der Bestellung wird als 2-Way Match bezeichnet. Wird dieser durch tatsächliche Liefer- und Leistungsmengen auf der Basis eines Wareneingangsbelegs ergänzt, handelt es sich um einen 3-Way Match. Die Ergänzung um abgenommene Mengen aus einem Qualitätssicherungsprozess wird als 4-Way Match bezeichnet (Diehm & Benzinger, 2018, S. 845). Rechnungen ohne Bestellung oder mit Abweichungen werden einem manuellen Prüf- und Freigabeprozess zugeführt.

Die Rechnungsprüfung (Schritt 5) und die Rechnungsfreigabe (Schritt 6) können bei validierten Dokumenten ohne Unstimmigkeiten vollständig automatisiert werden (Reuschenbach et al, 2020, S. 146). Darüber hinaus sind verschiedenste Szenarien denkbar: manuelle Prüfung durch eine Mitarbeiterin oder einen Mitarbeiter, Prüfung durch mehrere Mitarbeitende (Mehraugenprinzip) in sequentieller oder paralleler Reihenfolge. Im Gegensatz zum analogen Workflow werden die Rechnungen im automatisierten Prozess alle in einem zentralen, elektronischen Rechnungseingangsbuch erfasst, was zu einer klaren Transparenz über den Standort und den Bearbeitungsstatus der einzelnen Dokumente führt (Diehm & Benzinger, 2018, S. 845).

Buchung (Schritt 7) und Zahlung (Schritt 8) können je nach Belegart und Unternehmensvorgaben zumindest für Standardfälle und unkritische Belege automatisiert werden; alle anderen Belege werden schließlich in einem digitalen Workflow geprüft und gebucht (Diehm & Benzinger, 2018, S. 846).

Der Invoice-to-Pay Prozess hat wie dargestellt ein großes Potenzial, manuelle Arbeit durch Automatisierung zu ersetzen (Reuschenbach et al, 2020, S. 147). Auch das Institut für Arbeitsmarkt- und Berufsforschung (IAB), das Forschungsinstitut der deutschen Bundesagentur für Arbeit, kommt zu dem Ergebnis, dass 100 % der Aufgaben einer Buchhalterin bzw. eines Buchhalters automatisiert werden können (IAB, 2024a; Dengler & Matthes, 2018). Allerdings ist der Automatisierungsgrad in vielen Betrieben derzeit noch zu gering, um ganze Arbeitsplätze abzuschaffen (Demelius et al, 2023, S. 29; Dengler & Matthes, 2018, S. 8).

3.4 Auswirkungen und Stand der Digitalisierung in der Wirtschaftsprüfung

Aus der Digitalisierung der Mandanten ergeben sich auch direkte Digitalisierungspotentiale für die Abschlussprüfung bzw. Wirtschaftsprüferinnen und Wirtschaftsprüfer sowie Wirtschaftsprüfungsgesellschaften (Bruckner, 2019, S. 235). Die Potentiale gehen jedoch noch weit über die Abschlussprüfung hinaus. Zum einen können weitere Potentiale durch die Digitalisierung der eigenen Praxis gehoben werden. Andererseits eröffnet die Digitalisierung auch Chancen für neue Dienstleistungen, die den Mandanten angeboten werden können (Feld & Pöhlmann, 2017).

Die Digitalisierung der Mandanten und die Digitalisierung der Prüfung führen unmittelbar zu einer Reduzierung des Prüfungsaufwands (Downar & Fischer, 2019, S. 762). Bei automatisierten Prozessen ist die Fehleranfälligkeit im Vergleich zur manuellen Erfassung sehr gering, so dass weniger Stichproben und ein geringerer Stichprobenumfang für die Prüfung ausreichend sind (Marten & Harder, 2019, S. 767). Mit modernen Analysetools (Big Data) können große Datenmengen in kurzer Zeit analysiert werden.

Generell ist eine stärkere Integration von datenanalytischen Verfahren in den risikoorientierten Prüfungsansatz zu beobachten (Bruckner, 2019, S. 242). Insgesamt führt dies zu zwei zentralen Trends in der Abschlussprüfung: Erstens wird es zunehmend möglich und üblich, die Grundgesamtheit statt Stichproben zu prüfen (Downar & Fischer, 2019, S. 763). Zum anderen ermöglicht die digitale Anbindung des Prüfers an die Systeme des Mandanten eine zeitliche Entzerrung der Prüfung durch eine kontinuierliche Prüfung („continuous auditing“) der Daten des Mandanten in Echtzeit (Blum, 2020, S. 277; Downar & Fischer, 2019, S. 762). Andererseits ist jedoch zu erwarten, dass vor dem Hintergrund des gestiegenen Datenvolumens der Aufwand für eine sogenannte Erstprüfung, d.h. die erstmalige Durchführung der Jahres- oder Konzernabschlussprüfung für ein neues Mandat, deutlich zunehmen wird (Bruckner, 2019, S. 241).

Die Digitalisierung in der eigenen Praxis ermöglicht den Prüfern weitere Effizienzgewinne. So liegen Arbeitspapiere in der Regel nur noch in digitaler Form vor und der Einsatz mobiler Geräte ist in der Praxis inzwischen die Regel. Die verstärkte Nutzung von Cloud-Diensten ermöglicht den weltweiten Zugriff auf Daten. Dies führt zu einer Reihe interessanter Entwicklungen, auch im Hinblick auf die Belegschaft der Prüfungsgesellschaften:

- Zum einen ist die Schaffung von zentralen oder teilweise ausgelagerten Shared Service Centern zu beobachten (Langhein et al, 2018, 420-421). Diese übernehmen bestimmte Dienstleistungen zentral und weltweit, z. B. die Durchführung und Auswertung von rein digitalen Saldenbestätigungsaktionen.
- Auf der anderen Seite ist eine zunehmende Spezialisierung des Prüfungspersonals zu beobachten, insbesondere im Hinblick auf IT-Systemprüfungen. Diese können aufgrund der zunehmenden Komplexität in der Regel nicht mehr vom "normalen" Prüfungspersonal durchgeführt werden. Dementsprechend ist - insbesondere bei großen Wirtschaftsprüfungsgesellschaften - zu beobachten, dass sich die Rekrutierungsstrategien ändern (Weissenberger et al., 2019, S. 1121). Während in der Vergangenheit vor allem Studierende der Betriebswirtschaftslehre rekrutiert wurden, setzen die Unternehmen inzwischen auch stark auf IT- und techniknahe Disziplinen wie Informatik oder (Wirtschafts-) Ingenieurwesen (Downar & Fischer, 2019, S. 772; Hossenfelder, 2020, S. 245; Marten & Harder, 2019, S. 763).

Die Digitalisierung wird sich auch weiter auf die Personalgewinnung von Wirtschaftsprüfungsgesellschaften auswirken. In der Vergangenheit war die Wirtschaftsprüfung durch ein hohes Maß an Zeitverdichtung während der sogenannten "Busy Season" gekennzeichnet. Prüfungsteams mussten über mehrere Wochen hinweg tagelang beim Mandanten vor Ort arbeiten. Dies war teilweise mit erheblichen Überstunden und Reisen verbunden, insbesondere bei den großen (Big 4) Prüfungsgesellschaften (Mujkanovic, 2014, S. 23). Dank der ortsunabhängigen Verfügbarkeit von Daten können Prüfungen heute vom Büro oder Homeoffice aus durchgeführt werden. Die Anwesenheit vor Ort ist zunehmend die Ausnahme und nicht mehr die Regel. Es wird auch erwartet, dass die Zeit, die für die kontinuierliche Prüfung (continuous auditing) aufgewendet wird, entzerrt oder über das Jahr verteilt wird (Blum, 2020; S. 277; Downar & Fischer, 2019, S. 772). Dies wird möglicherweise auch positive Auswirkungen auf das Image des Berufsstandes bei Studierenden haben.

Künstliche Intelligenz (KI) und maschinelles Lernen werden in der Wirtschaftsprüfungsbranche bereits punktuell eingesetzt, haben sich aber noch nicht flächendeckend etablieren können (Boritz & Stratopoulos, 2023, S. 2; Gierbl et al, 2020, S. 613; Odenthal & Odenthal, 2023, S. 358). Die Einsatzmöglichkeiten generativer KI können sich über den gesamten Prüfungsprozess erstrecken: von der Recherche, Wissensbasis und Informationssammlung für die Prüfungsstrategie und Prüfungsplanung bis hin zur Unterstützung bei Risikoanalysen, der Erstellung und Bearbeitung von Prüfungsunterlagen, Datenanalysen und der Arbeit mit Texten und Prüfungsberichten (Odenthal & Odenthal, 2023, S. 362). Die inhärenten Schwächen der KI (einfache, oft oberflächliche Antworten, Auslassungen, falsche Behauptungen, fehlerhafte Schlussfolgerungen etc.) erfordern jedoch einen geübten und kritischen Umgang mit diesen Systemen (Berge, 2023, S. 609; Föhr et al, 2023, S. 1691; Odenthal & Odenthal, 2023, S. 367). Es ist zu beobachten, dass es in diesem Bereich eine große Diskrepanz zwischen den Big-4-Unternehmen und den kleinen und mittleren Prüfungsgesellschaften gibt. Während die Big 4 erhebliche finanzielle Mittel investieren, formuliert es ein Interviewpartner der Expertengespräche, ein Partner einer mittelständischen Wirtschaftsprüfungsgesellschaft in Bielefeld, so: "Wir haben das ein bisschen auf dem Radar. Aber künstliche Intelligenz ist für uns intern noch kein Thema und unsere Mandanten sind auch noch nicht so weit. Ja, und jetzt warten wir einfach mal ab [...]".

Dabei ist zu beachten, dass viele Prüfungsschritte Ermessensentscheidungen des Prüfers erfordern, die nicht digitalisiert oder automatisiert werden können (Bruckner, 2019, S. 234). Auch das IAB schätzt den Automatisierungsgrad in der Wirtschaftsprüfung mit „nur“ 57 % deutlich niedriger ein als die Automatisierung des Rechnungswesens (IAB, 2024b). Bislang sind auch keine unmittelbaren Arbeitsplatzverluste zu verzeichnen (Bruckner, 2019, S. 243; Warning et al, 2020, S. 408). Eine befragte Person im Rahmen der Experteninterviews kommentierte: „Die Digitalisierung ist eine Hilfe. Aber sie verhindert nicht, dass ich mich weiterhin selbst um das Ergebnis kümmern muss.“ Insgesamt sind die Auswirkungen der Digitalisierung auf die Wirtschaftsprüfung in Abbildung 4 zusammengefasst.



Abbildung 4. Auswirkungen der Digitalisierung auf die Wirtschaftsprüfung (eigene Abbildung).

Die Digitalisierung eröffnet auch Chancen für neue Dienstleistungen (Hossenfelder, 2020, S. 241). Dies betrifft zunächst die Beratung im Kontext der Digitalisierung selbst und umfasst Unterstützungsleistungen bei der Anpassung der Systeme und internen Prozesse der Mandanten (Downar & Fischer, 2019, S. 768). Dienstleistungen rund um das Thema Datenschutz und Cybersicherheit werden immer relevanter und passen gut zu den Kompetenzen und dem Image der Wirtschaftsprüfer (Bruckner, 2019, S. 244). Die Daten und Ergebnisse der Datenanalyse im Rahmen des Audits könnten für die Unternehmensberatung genutzt werden (Schwachstellen, Visualisierungen, Benchmarking). (Spezial-)Audits könnten sich auch auf die Ordnungsmäßigkeit, Sicherheit und Compliance von IT-Systemen konzentrieren. Wenn Kunden im Rahmen der Digitalisierung Prozesse auslagern, können in diesem Zusammenhang

Beratungsleistungen angeboten werden, die auch Prüfungen des internen Kontrollsystems bei Dienstleistungsunternehmen umfassen (Downar & Fischer, 2019, S. 768). Allerdings sind in diesem Zusammenhang auch die besonderen berufsrechtlichen Anforderungen an Wirtschaftsprüfer zu berücksichtigen, die eine gleichzeitige Prüfung und Beratung desselben Mandanten einschränken bzw. bei kapitalmarktorientierten Unternehmen (sog. PIEs - Public Interest Entities) innerhalb der EU fast vollständig verbieten (Hayn et al., 2022).

Insgesamt lässt sich feststellen, dass die Digitalisierung zahlreiche Chancen für Wirtschaftsprüfungsgesellschaften bietet. Sie verändert aber auch die Anforderungen an das eingesetzte Personal im Hinblick auf die für die Jahresabschlussprüfung erforderlichen Fähigkeiten und Kompetenzen (Bruckner, 2019, S. 243; Hossenfelder, 2020, S. 243; Marten & Harder, 2019, S. 767). Oder wie es ein Befragter, Partner einer mittelständischen Wirtschaftsprüfungsgesellschaft, formuliert: „Die Digitalisierung ist vor allem ein mentales Thema. Man muss sich hier umstellen und einfach lernen und mit den alten Gewohnheiten brechen, sagen wir es mal so.“

4 Diskussion: Auswirkungen auf die Hochschullehre

Routinetätigkeiten in der Buchführung und Rechnungsprüfung wurden oder werden automatisiert. Dies führt dazu, dass die Quantität der benötigten Arbeitskräfte abnimmt, während die Qualität der Tätigkeiten zunimmt. Die Digitalisierung scheint daher eine mögliche und notwendige Strategie für Unternehmen im Kontext des Fachkräftemangels zu sein. Das bedeutet aber auch, dass die Anforderungen an die Fähigkeiten der dort (zukünftig) Tätigen steigen (Demelius et al, 2023, S. 28; Vanini, 2019, S. 56).

Mit dem Wandel der Praxis werden auch die Hochschulen ihre Ausbildung stärker auf die Vermittlung digitaler Kompetenzen ausrichten müssen (Bruckner, 2019, S. 243; Vanini, 2019, S. 59). Darauf soll schließlich der Fokus gelegt werden. Die für die Digitalisierung erforderlichen Kompetenzen können dabei in Basiskompetenzen, spezifische Kompetenzen und Metakompetenzen differenziert werden, wie es in Abbildung 5 dargestellt ist.

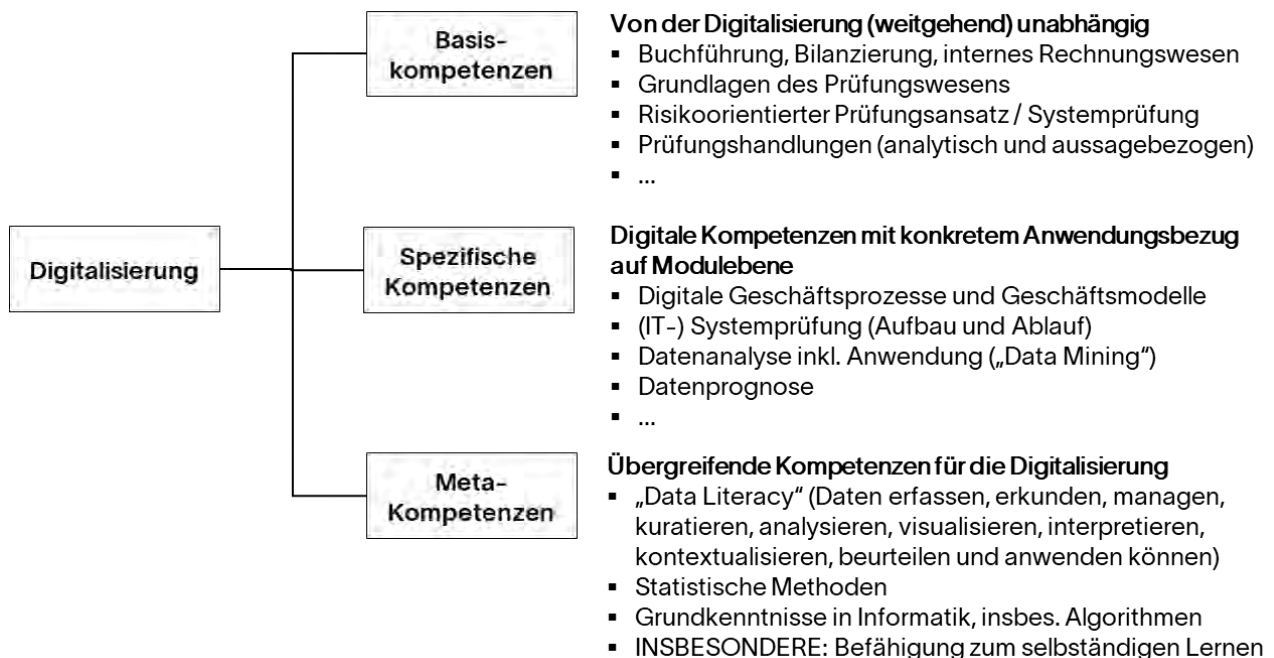


Abbildung 5. Im Rahmen der Digitalisierung zu vermittelnde Kompetenzen (eigene Abbildung).

Grundfertigkeiten bzw. Basiskompetenzen sind weitgehend unabhängig von der Digitalisierung, da sie die wesentlichen grundlegenden Fertigkeiten der Disziplinen bilden (Kesten, 2019, S. 47). Dazu gehören Grundkenntnisse der Buchführung, des Rechnungswesens und des Controllings. In der Abschlussprüfung gehören dazu auch die Grund-

lagen der Wirtschaftsprüfung, des risikoorientierten Prüfungsansatzes und Grundkenntnisse der Durchführung analytischer und aussagebezogener Prüfungshandlungen. Diese Kompetenzen und Fertigkeiten bilden de facto die Grundlage der beruflichen Tätigkeit. Sie müssen auch in Zukunft, wenn auch möglicherweise in geringerer Intensität, vermittelt werden, um mit dem Grundverständnis mögliche Fehler im digitalen Prozess erkennen zu können (Klein & Küst, 2020, S. 94)

Spezifische (digitale) Kompetenzen zielen auf konkrete Anwendungen im Kontext der Digitalisierung ab. Diese können an der Hochschule in eigenen Modulen vermittelt werden. Dazu gehören insbesondere die Auseinandersetzung mit digitalen Geschäftsprozessen und Geschäftsmodellen, der Aufbau und Ablauf von (IT-) Systemtests sowie die Datenanalyse (Data Mining) und -prognose (Predictive Analytics) (Leitner-Hanetseder et al, 2022, 199; Weißenberger et al., 2019, S. 1124). Es ist jedoch wichtig, diese Module aufeinander zu beziehen und die Studierenden nicht zu IT- und Datenanalyse-Profis auszubilden, sondern sie darin zu schulen, die Anwendung von Methoden und Techniken in der beruflichen Praxis zu reflektieren (Bravidor et al, 2020, S. 293).

Metakompetenzen befassen sich mit den übergreifenden Kompetenzen für die Digitalisierung. Data Literacy bezieht sich auf die Fähigkeit, Daten zu sammeln, zu erkunden, zu verwalten, zu kuratieren, zu analysieren, zu visualisieren und zu interpretieren (Stifterverband, 2024). Dies erfordert auch statistische Methoden und Grundkenntnisse der Informatik, insbesondere Algorithmen (Marten et al, 2017, S. 1239; Kesten, 2019, S. 47; Klein & Küst, 2020, S. 94). Das Wichtigste ist jedoch die nachhaltige Fähigkeit zum selbstständigen Lernen (Kesten, 2019, S. 49).

Zusammenfassend lässt sich sagen, dass die Digitalisierung kein neues Phänomen ist - neu sind vielmehr die Dynamik und die Durchdringung aller Lebensbereiche. Für die Arbeitswelt im Allgemeinen und Rechnungswesen bzw. Wirtschaftsprüfung im Besonderen bringt die fortschreitende Digitalisierung zahlreiche Herausforderungen, aber auch Chancen. Die Hochschulen müssen auf die Digitalisierung nicht nur reagieren, sondern sie aktiv gestalten.

Die Digitalisierung ist kein Selbstzweck - unabhängig von der fortschreitenden Durchdringung bleiben bestimmte Grundkompetenzen, die an den Hochschulen vermittelt werden müssen. Zudem sollte (auch) in den Vertiefungsmodulen ein Fokus auf die Digitalisierung gelegt werden (sog. "Hybridwissen" oder Schnittstellenwissen). Bestimmte Metakompetenzen (insbesondere in Datenanalyse, Statistik, Informatik) können jedoch nur modul- bzw. fächerübergreifend vermittelt werden. Die Weiterentwicklung erfordert eine verstärkte Kooperation innerhalb und außerhalb der Hochschule, da die Kompetenzen in der Praxis des Rechnungswesens und der Wirtschaftsprüfung dringend benötigt werden.

Unternehmen sind zunehmend auch bereit, offene Stellen mit fachfremden Bewerbern zu besetzen, wie Demelius et al. (2023, S. 32) in ihrer Studie in österreichischen Unternehmen beschreiben. Bei diesem sogenannten Skill-based Recruiting wird auf relevante Vorerfahrungen im Rechnungswesen verzichtet und damit der Talentpool vergrößert. Stattdessen werden relevante Kenntnisse und Fähigkeiten aus verwandten und ähnlich komplexen Tätigkeitsfeldern in den Fokus gestellt und gezielt angesprochen. Demelius et al. (2023, S. 33) führen als Beispiel die Besetzung einer Stelle im Bereich der internen Lohnbuchhaltung an. Da es hier an Fachkräften mangelt, sucht das Unternehmen nach Kompetenzen in ähnlichen Berufen. Genauigkeit, Zahlenverständnis, Zuverlässigkeit und Serviceorientierung sind relevante Fähigkeiten, über die auch Reisebüromitarbeiterinnen und -mitarbeiter verfügen. Da viel Personal in diesem Bereich durch die Digitalisierung freigesetzt wurde, stellt das Unternehmen gezielt solche Personen ein und bildet sie on the job aus.

Auch die Wirtschaftsprüfung rekrutiert zunehmend Absolventen, die über keine oder nur geringe einschlägige Fachkenntnisse verfügen. Eine umfassende, interdisziplinäre Einarbeitung und praktische Ausbildung gleicht das fehlende Wissen aus (Hossenfelder, 2020, S. 245).

Literaturverzeichnis

- Berge, T. (2023): IT and more: Generative Künstliche Intelligenz. Nutzungsmöglichkeiten für die Wirtschaftsprüfung. WPg - Die Wirtschaftsprüfung, 76(11): 607-609.
- Bleiber, R. (2023): Fachkräftemangel im Rechnungswesen. https://www.haufe.de/finance/steuern-finanzen/fachkraeftemangel-im-rechnungswesen_190_606244.html. Abgerufen am 30.06.2024.
- Blum, A. (2020): Continuous Audit. In: Preus, P. (Hrsg.), Accounting und Taxation 4.0: Digitalisierung und Automatisierung im Rechnungswesen, Schäffer-Poeschel, Stuttgart: 265-278.
- Boritz, J. E.; Stratopoulos, T. C. (2023): AI and the Accounting Profession: Views from Industry and Academia. Journal of Information Systems, 37(3): 1-9.
- Bravidor, M.; Förster, G.; Weißenberger, B. E. (2020): Berufsstand 4.0: Wirtschaftsprüfer und Steuerberater zwischen Fachexpertise, IT und Datenanalyse. Ansätze für die Ausbildung künftiger Wirtschaftsprüfer und Steuerberater an Hochschulen und in der Praxis. WPg - Die Wirtschaftsprüfung, 73(5): 287-293.
- Bravidor, M.; Lösse, L. J. (2018): Digitalisierung - ein Wieselwort. Einordnung von Begriff, Facetten und Technologie in die Bedeutung für Rechnungswesen und Besteuerung. StuB - Unternehmenssteuern und Bilanzen, 20(21): 783-787.
- Bruckner, A. (2019): Abschlussprüfung in der digitalen Welt - Prüfungsprozesse und Geschäftsmodell. In: Fink, C.; Kunath, O. (Hrsg.). Digitale Transformation im Finanz- und Rechnungswesen, Schäffer-Poeschel, Stuttgart: 235-249.
- Demelius, K.; Viveros, N.; Herrmann, M.; Wolfsberger, J. (2023): Automatisierung und Digitalisierung im Rechnungswesen 2023. Eine Studie von Deloitte Österreich. Deloitte, Wien. <https://www2.deloitte.com/content/dam/Deloitte/at/Documents/presse/at-digitalisierung-im-rechnungswesen-2023.pdf>. Abgerufen am 30.06.2024.
- Dengler, K.; Matthes, B. (2018): Substituierbarkeitspotenziale von Berufen: Wenige Berufsbilder halten mit der Digitalisierung Schritt. IAB-Kurzbericht 4/2018, Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur für Arbeit, Nürnberg.
- Diehm, J.; Benzinger, L. (2018). Digital Finance: Digitale Rechnungsverarbeitung und Workflows als Basis für ein Rechnungswesen 4.0. Der Betrieb, 74(15): 841-847.
- Döring, N. (2023): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 6., vollst. überarb. u. erw. Aufl., Springer, Berlin.
- Downar, B.; Fischer, D. (2019): Wirtschaftsprüfung im Zeitalter der Digitalisierung. In Obermaier, R. (Hrsg.), Handbuch Industrie 4.0 und Digitale Transformation -Betriebswirtschaftliche, technische und rechtliche Herausforderungen. Gabler, Wiesbaden: 753-779.
- Feld, K.-P.; Pöhlmann, A. (2017): Digitalisierung. Eine Bestandsaufnahme für den Wirtschaftsprüfer. IDW Life, 04/2017, 356-362.
- FeRD (2024). Was ist ZUGFeRD? <https://www.ferd-net.de/standards/was-ist-zugferd/was-ist-zugferd.html>. Abgerufen am 30.06.2024.
- Föhr, T. L.; Marten, K.-U.; Schreyer, M. (2023): Generative Künstliche Intelligenz und risikoorientierter Prüfungsansatz. Der Betrieb, 76(30): 1681-1693.
- Gierbl, A.; Schreyer, M.; Leibfried, P.; Borth, D. (2020): Künstliche Intelligenz in der Prüfungspraxis. Eine Bestandsaufnahme aktueller Einsatzmöglichkeiten und Herausforderungen. Expert Focus, 94(9): 612-617
- Hayhn, S.; Kelm, D.; Schmitz-Herkendell, A. (2022): Zulässigkeit von Nichtprüfungsleistungen des Abschlussprüfers nach dem FISG. WPg - Die Wirtschaftsprüfung, 75(07), 374-384.
- Hossenfelder, J. (2020): Audit 2.0 - Der Wirtschaftsprüfer der Zukunft. In: Preus, P. (Hrsg.), Accounting und Taxation 4.0. Digitalisierung und Automatisierung im Rechnungswesen, Schäffer-Poeschel, Stuttgart: 241-246.
- IAB (2024a): Werden digitale Technologien Ihren Beruf verändern? Automatisierbarkeit im Beruf Buchhalter/in. <https://job-futuromat.iab.de/>. Abgerufen am 30.06.2024.

- IAB (2024b): Werden digitale Technologien Ihren Job verändern? Automatisierbarkeit im Beruf Wirtschaftsprüfer/in. <https://job-futuroomat.iab.de/>. Abgerufen am 30.06.2024.
- Kampe, T. (2011): Führung von Wirtschaftsprüfungsgesellschaften: Intellektuelles Kapital als Rahmenkonzept des strategischen Managements in Professional Service Firms. Gabler: Wiesbaden.
- Kesten, R. (2019): Digitalisierung in Rechnungswesen und Controlling und ihre Folgen für die Hochschullehre: Chancen und Herausforderungen. *Controller Magazin*, 44(6): 44-49.
- Klein, J.; Küst, C. (2020): Wie die Digitalisierung im Rechnungswesen die Aufgaben und Anforderungen an die Mitarbeiter/-innen verändert. In: Berding, F.; Jahncke, H.; Slopinski, A. (Hrsg.), *Moderner Rechnungswesenunterricht 2020*. Springer VS, Wiesbaden: 83-97.
- Kreher, M.; Eichner, V.; Sellhorn, T.; Hess, T.; Köster, A.; Weiß, K. (2021): Digitalisierung im Rechnungswesen 2021. Studie. KPMG, Berlin. <https://hub.kpmg.de/studie-de-digitalisierung-im-rechnungswesen-2021>. Abgerufen am 30.06.2024.
- Kreher, M.; Sellhorn, T.; Hess, T. (2017): Digitalisierung im Rechnungswesen: Studie zum Status quo deutscher Unternehmen. Berlin, KPMG.
- Kreher, M.; Winkler, M.; Sellhorn, T.; Hess, T.; Köster, A. (2022): Digitalisierung im Rechnungswesen: Ausgabe 2022/2023. Studie. KPMG, Berlin. <https://hub.kpmg.de/digitalisation-in-accounting-2022>. Abgerufen am 30.06.2024.
- Kreher, M.; Winkler, M.; Sellhorn, T.; Hess, T.; Sehn, V.; Meythaler, A. (2024): Digitalisierung im Rechnungswesen. Studienausgabe 2023/2024. KPMG, Berlin. <https://hub.kpmg.de/de/digitalisierung-im-rechnungswesen>. Abgerufen am 30.06.2024.
- Marten, K.-U.; Czupalla, K.; Harder, R. (2017): Digitalisierung in der Wirtschaftsprüfung und in der Internen Revision. Herausforderungen für die Aus- und Weiterbildung. *WPg - Die Wirtschaftsprüfung*, 70(21): 1233-1241.
- Marten, K.-U.; Harder, R. (2019): Digitalisierung in der Abschlussprüfung. Experteninterviews auf explorativer Basis. *WPg - Die Wirtschaftsprüfung*, 72(14): 761-769.
- Langhein, J.; Kiesow, A.; Strobel, C.; Thomas, O. (2018): Digitale Wirtschaftsprüfung - Make or Buy? *HMD Praxis der Wirtschaftsinformatik*, 55(5): 412-426.
- Leitner-Hanetseder, S.; Eisl, C.; Knoll, C.; Koislgruber, M. (2022): Neue Berufsfelder im Rechnungswesen: Die Stellenprofile des Business Data Analyst und Business Data Scientist. *zfo - Zeitschrift Führung und Organisation*, 91(3): 198-203.
- Mayring, P. (2022): *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 13., überarb., Aufl., Beltz, Weinheim.
- Mujanovic, R. (2014): Warum merkt nur keiner, wie toll der WP-Beruf ist? Weshalb eine Imagekampagne allein nicht ausreichend ist. *WP-Praxis*, 3(1): 23-25.
- Odenthal, R.; Odenthal, K. (2023): Mehr als ein neues Werkzeug. Generative KI-Techniken (ChatGPT) in der Prüfung. *WP Praxis*, 12(12): 358-367.
- Pargmann, J.; Riebenbauer, E.; Flick-Holtsch, D.; Berding, F. (2023): Digitalisation in accounting: a systematic literature review of activities and implications for competences. *Empirical Research in Vocational Education and Training*, 15(1).
- Prümm, D.; Loitz, R.; Dittmar, H.-P. (2023): Digitalisierung im Finanz- und Rechnungswesen 2023 - und was sie für die Zukunft der Abschlussprüfung bedeutet: Befragungen zum Status quo und zur digitalen Weiterentwicklung. PricewaterhouseCoopers, Frankfurt am Main. <https://pages.pwc.de/digitalisierung-im-finanz-und-rewe-2023>. Abgerufen am 30.06.2024.
- Reuschenbach, D.; Beuckes, T.; Isensee, J. (2020): Intelligente Prozessautomatisierung im Accounting. In: Preus, P. (Hrsg.), *Accounting und Taxation 4.0. Digitalisierung und Automatisierung im Rechnungswesen*, Schäffer-Poeschel, Stuttgart: 131-150.

- Schmidt, A. (2020): Digitalisierung im Rechnungswesen - Chancen, Herausforderungen. In: Preus, P. (Hrsg.), Rechnungswesen und Steuern 4.0. Digitalisierung und Automatisierung im Rechnungswesen, Schäffer-Poeschel, Stuttgart: 247-264.
- Stifterverband (2024): Data Literacy Education. <https://www.stifterverband.org/data-literacy-education>. Abgerufen am 30.06.2024.
- Strübing, J. (2018): Qualitative Sozialforschung: Eine komprimierte Einführung. 2., überarb. u. erw. Aufl., De Gruyter Oldenbourg, Berlin und Boston.
- Vanini, U. (2019): Wie viel "Digitalisierung" lernt man an Hochschulen? Controlling & Management Review, 63(6): 58-63.
- Vlk, G.; Demelius, K.; Karigl, B. (2016): Automatisierung und Digitalisierung im Rechnungswesen. Eine Studie von Deloitte Österreich. Deloitte, Wien. <https://www2.deloitte.com/content/dam/Deloitte/at/Documents/about-deloitte/at-studie-automatisierung-und-digitalisierung-im-rechnungswesen.pdf>. Abgerufen am 30.06.2024.
- Warning, A.; Sellhorn, T.; Kummer, J.-P. (2020): Digitalisierung und Beschäftigung: Empirische Befunde für die Rechts- und Steuerberatung sowie Wirtschaftsprüfung. BFuP - Betriebswirtschaftliche Forschung und Praxis, 72(4): 391-412.
- Weissenberger, B. E.; Förster, G.; Wesser, M. B.; Bravidor, M. (2019): Wohin führt die Digitalisierung? Auswirkungen auf Wirtschaftsprüfung, Steuerberatung, Finanzfunktion und Hochschullehre. WPg - Die Wirtschaftsprüfung, 72(20): 1118-1124.

Berichterstattung über immaterielle Ressourcen – Umfang und Grenzen der neuen Berichtspflicht

Research Paper

Prof. Dr. Robert Nothhelfer

Hochschule Pforzheim, Pforzheim, E-Mail: robert.nothhelfer@hs-pforzheim.de

Abstract

Die Bedeutung immaterieller Ressourcen in der Wirtschaft hat in den letzten Jahrzehnten kontinuierlich zugenommen. 2023 schätzte BrandFinance ihren globalen Wert auf 61,9 Billionen USD, wovon 45,7 Billionen USD nicht bilanziert werden dürfen. Diese immateriellen Ressourcen, wie Marken, Kundenlisten oder Humankapital, erfüllen oft nicht die Kriterien für die Bilanzierung nach HGB und IFRS. Die EU hat im Rahmen der Corporate Social Responsibility Directive (CSRD) eine neue qualitative Berichterstattung über immaterielle Ressourcen im Lagebericht eingeführt, die bestimmte große Kapitalgesellschaften ab 2024 verpflichtend einhalten müssen. Die Berichterstattung soll wesentliche immaterielle Ressourcen identifizieren und deren Bedeutung für das Geschäftsmodell und die Wertschöpfung erläutern, wobei keine spezifische Methodik vorgegeben wird. Der Beitrag diskutiert die neuen gesetzlichen Anforderungen und deren Anwendung. Als mögliche Probleme werden die Beschränkung der Berichtspflicht auf den vorhandenen Bestand immaterieller Ressourcen sowie der Zielkonflikt zwischen Transparenz und Schutz von Wettbewerbsvorteilen erkannt.

1 Einführung

Schon seit einiger Zeit nimmt die Bedeutung immaterieller Ressourcen in der Wirtschaft kontinuierlich zu (Zambon et al., 2020, S. 27 ff.; von Keitz & Schwedler, 2023, S. 26 ff.). BrandFinance schätzt den Wert der globalen immateriellen Ressourcen von Unternehmen für das Jahr 2023 auf 61,9 Billionen USD, von denen 45,7 Billionen USD nicht bilanziert werden dürfen. Die Marktwert-Buchwert-Lücke beträgt bei einem geschätzten Gesamtwert des Nettovermögens von 124,6 Billionen USD etwa 37 % (BrandFinance, 2023, S. 16). Dies bedeutet, dass ein erheblicher Teil der Unternehmenswerte durch das bilanzierte Eigenkapital nicht erklärt wird. Der verpflichtenden Unternehmensberichterstattung fehlen somit wesentliche Informationen zum Unternehmenswert und zur Wertentwicklung sowie zur Beurteilung zukünftiger Chancen und Risiken (Haller & Fischer, 2023, S. 79).

Immaterielle Ressourcen erfüllen aufgrund ihrer Eigenschaften wie z. B. mangelnde Identifizierbarkeit und Eigentumsrechte, unbestimmbare wirtschaftliche Vorteile oder unklare Nutzungsdauer häufig nicht die Kriterien zum Ansatz als Vermögensgegenstand und dürfen dann sowohl nach HGB als auch nach IFRS nicht angesetzt werden (Haller & Fischer, 2023, S. 79; Coenenberg u. a., 2021, S. 84 und S. 186 f.). Während erworbene immaterielle Vermögensgegenstände angesetzt werden müssen, besteht bei selbst geschaffenen immateriellen Vermögensgegenständen nach HGB ein Ansatzwahlrecht (§ 248 Abs. 2 Satz 1) und nach IFRS eine Ansatzpflicht unter weiteren strengen Bedingungen (IAS 38.51ff.). Für selbst erstellte Marken, Kundenlisten und ähnliche Vermögensgegenstände besteht sowohl nach § 248 Abs. 2 Satz 2 als auch nach IAS 38.63 ein Ansatzverbot.

Aufgrund der zunehmenden Bedeutung immaterieller Ressourcen seit den 1980er Jahren und der restriktiven Rechnungslegungsvorschriften wurden vielfältige Initiativen und Ansätze zur erweiterten freiwilligen Berichterstattung über immaterielle Ressourcen entwickelt.¹ Parallel zu diesen Diskussionen entstand die zunächst ebenfalls freiwillige Nachhaltigkeitsberichterstattung, die teilweise eine Berichterstattung über immaterielle Ressourcen erfordert. In diesem Kontext hat die EU im Rahmen der Corporate Social Responsibility Directive (CSRD) beschlossen, eine neue qualitative Berichterstattung über immaterielle Ressourcen im Lagebericht zu verankern.² Dabei soll über die wichtigsten immateriellen Ressourcen berichtet werden, die für das Geschäftsmodell grundlegend sind und eine Wertschöpfungsquelle für das Unternehmen darstellen. Es wird jedoch keine Methodik vorgegeben, wie diese Ressourcen zu ermitteln oder deren Wertschöpfungsqualität zu beurteilen ist (Europäische Union, 2022, Art. 1 Nr. 3). Zur Umsetzung in nationales Recht liegt bisher der Referentenentwurf des CSRD-Umsetzungsgesetzes (CSRD-UmsG) vor (Bundesjustizministerium, 2024).

Dieser Beitrag stellt zunächst die neue Berichtspflicht dar und gibt erste Ansätze für die vorzunehmende Wesentlichkeitsprüfung, bei der zu beurteilen ist, welche immateriellen Ressourcen grundlegende Werttreiber im Unternehmen sind. Danach werden die Grenzen der Berichterstattung im Hinblick auf das Berichtsziel aufgezeigt.

2 Neue Berichterstattung über immaterielle Ressourcen

2.1 Stand der Gesetzgebung

Die CSRD wurde am 14. Dezember 2022 verabschiedet und wäre bis zum 6. Juli 2024 in nationales Recht umzusetzen gewesen (Europäische Union, 2022, Art. 5). In Deutschland liegt seit dem 23. Juli 2024 ein Regierungsentwurf zur Umsetzung vor (Bundesjustizministerium, 2024); die Beschlussfassung im Bundestag ist derzeit noch offen. Trotz der verspäteten Umsetzung sieht der Regierungsentwurf bislang keine veränderten Anwendungszeitpunkte vor.

2.2 Anwendungsbereich

Die Berichterstattung über immaterielle Ressourcen ist nach Art. 19 Abs. 1 der Richtlinie 2013/34/EU in der Fassung der CSRD verpflichtend für große sowie kleine und mittelgroße kapitalmarktorientierte Kapitalgesellschaften

¹ Für einen Überblick der Ansätze vgl. Zambon, 2020, S. 65 ff.

² Für einen Überblick über die Veränderungen im politischen Prozess vgl. Liening u. a., 2023.

sowie nach § 264a HGB haftungsbeschränkte Personengesellschaften (Europäische Union, 2022, Art. 1 Nr. 3). Diese Berichtspflicht gilt ebenso für Konzernabschlüsse (Europäische Union, 2013, Art. 29).

Im Zuge der Einführung der CSRD hat die EU entschieden die Größenkriterien zur Klassifikation von Kapitalgesellschaften anzuheben, um der Verhältnismäßigkeit der Berichterstattung und den inflationär gestiegenen Werten der Kriterien Rechnung zu tragen (Europäische Union, 2023). Der Bundestag hat am 22. Februar 2024 entsprechende Änderungen der §§ 267 und 267a HGB beschlossen. Die Änderungen sind für Geschäftsjahre, die nach dem 1. Januar 2024 beginnen anzuwenden, dürfen aber auch schon ab 1. Januar 2023 angewendet werden. Die Größenklassen haben sich wie folgt geändert (Bundestag, 2024, Art. 2):

Tabelle 1: Neue Größenkriterien § 267/267a HGB

Größenklasse	Bilanzsumme	Umsatzerlöse	Mitarbeitende
Sehr klein (neu)	≤ 450.000 €	≤ 900.000 €	≤ 10
Sehr klein (alt)	≤ 350.000 €	≤ 700.000 €	≤ 10
Klein (neu)	≤ 7.500.000 €	≤ 15.000.000 €	≤ 50
Klein (alt)	≤ 6.000.000 €	≤ 12.000.000 €	≤ 50
Mittelgroß (neu)	≤ 25.000.000 €	≤ 50.000.000 €	≤ 250
Mittelgroß (alt)	≤ 20.000.000 €	≤ 40.000.000 €	≤ 250
Groß (neu)	> 25.000.000 €	> 50.000.000 €	> 250
Groß (alt)	> 20.000.000 €	> 40.000.000 €	> 250

Die Größenklassen für die Erstellungspflicht von Konzernabschlüssen wurden entsprechend angepasst (Bundestag, 2024, Art. 2).

Durch die Anhebung der Schwellenwerte werden weniger Unternehmen als groß klassifiziert, d. h. die Anzahl der Unternehmen, die zur Berichterstattung über immaterielle Ressourcen verpflichtet sind, sinkt.

Die Anwendung ist entsprechend der CSRD zeitlich gestaffelt (Europäische Union, 2022, Art. 5):

- **Große, kapitalmarktorientierte Kapitalgesellschaften mit durchschnittlich mehr als 500 Mitarbeitenden** müssen für Geschäftsjahre, die am oder nach dem 1. Januar 2024 beginnen, berichten.
- **Alle übrigen großen Kapitalgesellschaften** müssen für Geschäftsjahre, die am oder nach dem 1. Januar 2025 beginnen, berichten.
- **Kleine und mittelgroße kapitalmarktorientierte Kapitalgesellschaften** müssen für Geschäftsjahre, die am oder nach dem 1. Januar 2026 beginnen, berichten. Die zusätzliche Option, die Berichterstattung um zwei weitere Jahre bis 2028 zu verschieben, gilt nur für die Nachhaltigkeitsberichterstattung, nicht aber für die Berichterstattung über immaterielle Ressourcen (Europäische Union, 2022, Art. 1 Nr. 4).

2.3 Kategorisierung immaterieller Ressourcen

2.3.1 Definition immaterieller Ressourcen

Die CSRD definiert immaterielle Ressourcen als „Ressourcen ohne physische Substanz“ (Europäische Union, 2022, Art. 1 Nr. 2 b)). Diese Formulierung wird auch in § 289 Abs. 3a HGB nF verwendet (Bundesjustizministerium, 2024, Art. 1 Nr. 5 c)). Es handelt sich hierbei um eine negative Abgrenzung. Entsprechend den üblichen Unterscheidungen in der Rechnungslegung dürften zusätzlich finanzielle Ressourcen, die zwar keine physische Substanz besitzen, aber einen direkten monetären Bezug aufweisen, ebenfalls ausgeschlossen werden (Liening et al., 2023, S. 303 f.; AKIWIR, 2001, S. 990). Die Definition ist jedoch erheblich weiter gefasst als der bisher in der Rech-

nungslegung übliche Begriff der immateriellen Vermögensgegenstände, da bewusst über Ressourcen berichtet werden soll, die nicht die Kriterien eines Vermögensgegenstands erfüllen (Europäische Union, 2022, Erwägungsgrund 32; Bundesjustizministerium, 2024, S. 132).

2.3.2 Kategorisierung entsprechend bilanzieller Würdigung

Die immateriellen Ressourcen lassen sich hinsichtlich ihrer Bilanzierung unterscheiden in (Beck'scher Bilanzkommentar, 2024, § 247 Tz. 13 – 17):

- **Erworbene immaterielle Vermögensgegenstände**, die im Rahmen des Einzelerwerbs oder einer Unternehmensübernahme gemäß § 246 HGB und IAS 38 anzusetzen sind (einschließlich erworbenen Goodwills).
- **Selbst geschaffene immaterielle Vermögensgegenstände**, die im Rahmen des Ansatzwahlrechts von § 248 Abs. 2 Satz 1 angesetzt werden können. Nach IAS 38.51-64 müssen selbst geschaffene immaterielle Vermögensgegenstände angesetzt werden, sofern sie die genannten Kriterien erfüllen.
- **Selbst geschaffene Marken**, Drucktitel, Verlagsrechte, Kundenlisten oder vergleichbare immaterielle Vermögensgegenstände, die aufgrund des Ansatzverbots in § 248 Abs. 2 Satz 2 und IAS 38.63 nicht angesetzt werden dürfen.
- **Alle anderen immateriellen Ressourcen** oder Werte, die die Kriterien für den Ansatz als Vermögensgegenstand nach Handelsrecht oder IAS 38 nicht erfüllen (einschließlich selbst geschaffenen Goodwills).

2.3.3 Kategorisierung nach Inhalt

Neben dieser eher formalen Betrachtungsweise können immaterielle Ressourcen auch nach ihrem Inhalt kategorisiert werden. Am verbreitetsten sind die Ansätze des Arbeitskreises „Immaterielle Werte im Rechnungswesen“ (AKIWIR) der Schmalenbach-Gesellschaft, des International Integrated Reporting Council (IIRC) sowie der World Intellectual Capital Initiative (WICI) (von Keitz & Schwedler, 2023, S. 38 ff.; Zambon u. a., 2020, S. 72 ff.; Haller & Fischer, 2023, S. 84 ff.).

Der AKIWIR hat 2001 eine sehr detaillierte und weitgehend überschneidungsfreie Kategorisierung vorgelegt. Es werden dabei insgesamt sieben Kategorien von immateriellen Werten definiert (AKIWIR, 2001, S. 990 f.):

1. **Innovation Capital** umfasst Produkt-, Dienstleistungs- oder Verfahrensinnovationen, z. B. neue Software oder Patente.
2. **Human Capital** umfasst die immateriellen Werte im Personalbereich, insbesondere das Wissen und die Fähigkeiten des Personals sowie des Managements, gute Unternehmenskultur.
3. **Customer Capital** umfasst die immateriellen Werte im Absatzbereich, z. B. gute Kundenbeziehungen und Kundenlisten, Marktanteile und Marken, gute Reputation.
4. **Supplier Capital** umfasst die immateriellen Werte im Beschaffungsbereich, z. B. gute Lieferantenbeziehungen und Lieferantenlisten, Zugang zu knappen Rohstoffen, gute Reputation.
5. **Investor Capital** umfasst die immateriellen Werte im Finanzbereich, z. B. die Möglichkeit zur günstigen Kapitalbeschaffung und gute Investorenbeziehungen, gute Reputation.
6. **Process Capital** umfasst die immateriellen Werte im Organisationsbereich, z. B. gutes Kommunikationsnetzwerk und gute interne Abläufe und Verfahren.
7. **Location Capital** umfasst die immateriellen Werte des Standorts, z. B. günstige Verkehrsanbindung, Steuervorteile und eine wohlwollende Verwaltung.

WICI und IIRC haben 2016 bzw. 2021 ebenfalls Kategorisierungen vorgelegt, die drei Kategorien umfassen und weitgehend deckungsgleich sind; die Kategorien des AKIWIR stellen eine weitere Detaillierung dar und lassen sich bis auf das Location Capital zuordnen (WICI, 2016, S.15; IIRC, 2021, S. 19):³

1. **Organisational Capital (WICI) / Intellectual Capital (IIRC)** umfasst die immateriellen Werte im Organisationsbereich und die Fortentwicklung der Produkte und Dienstleistungen. Dies entspricht dem Innovation Capital und Process Capital.
2. **Human Capital** umfasst die immateriellen Werte im Personalbereich und entspricht direkt der Kategorie des AKIWIR.
3. **Relational Capital (WICI) / Social and Relationship Capital (IIRC)** entspricht den immateriellen Werten im Absatz-, Beschaffungs- und Finanzbereich. Dies entspricht dem Customer Capital, Supplier Capital und Investor Capital.

Bei allen Kategorisierungen wird betont, dass sie flexibel an die Bedürfnisse des berichtenden Unternehmens angepasst werden müssen, da die Zuordnung von immateriellen Werten zu einzelnen Kategorien häufig nicht eindeutig ist und die Kategorien auch nicht überschneidungsfrei sind (AKIWIR, 2001, S. 991; WICI, 2016, S. 14; IIRC, 2021, S.20). Haller & Fischer schlagen eine zweidimensionale Kategorisierung vor. Dabei entspricht die horizontale Unterteilung weitgehend den Kategorien des AKIWIR-Ansatzes, während als vertikale Querschnittskategorien Nachhaltigkeit, Innovationsfähigkeit und Reputation verwendet werden. Der Ansatz reduziert aber nicht die Überschneidungsproblematik und insbesondere die Interdependenzen der vertikalen Kategorien bleiben unklar (Haller & Fischer, 2023, S. 85 f.).

Zusammenfassung

In der folgenden Tabelle werden die beiden Kategorisierungen zusammengefasst und typische Beispiele zugeordnet.

Tabelle 2: Kombination der Kategorisierungen mit Beispielen immaterieller Werte

	Organizational / Intellectual Capital	Human Capital	Relational / Social and Relationship Capital
Ansatzpflicht	Erworbene Software oder Patente	--	Erworbene Marken oder Kundenlisten
Ansatzwahlrecht (HGB) Ansatzpflicht (IFRS)	Selbst geschaffene Software oder Patente	--	--
Ansatzverbot	--	--	Selbst geschaffene Marken oder Kundenlisten und ähnliche Werte
Kein Vermögensgegenstand bzw. keine Ausübung Ansatzwahlrecht HGB	Gute Betriebsabläufe und -verfahren Selbst geschaffene Software oder Patente, die nicht angesetzt werden (HGB) oder die zusätzlichen Kriterien nicht erfüllen (IFRS)	Wissen und Fähigkeiten der Mitarbeitenden, Weiterbildung, Training, gute Unternehmenskultur	Gute Kundenbeziehungen, gute Lieferantenbeziehungen, gute Investorenbeziehungen, gute Reputation

2.4 Berichtspflicht im Lagebericht

Der Referentenentwurf des CSRD-UmsG verpflichtet die Unternehmen in § 289 Abs. 3a HGB nF, „die wichtigsten immateriellen Ressourcen anzugeben. [...] Dabei ist zu erläutern, inwiefern das Geschäftsmodell der Gesellschaft grundlegend von diesen Ressourcen abhängt und inwiefern diese Ressourcen eine Wertschöpfungsquelle für die Gesellschaft darstellen. Wichtigste immaterielle Ressourcen sind Ressourcen ohne physische Substanz, von denen

³ Das <IR> Framework stellt ein umfassendes Konzept zur integrierten Berichterstattung dar, auf das hier aufgrund des Themas nicht weiter eingegangen wird.

das Geschäftsmodell der Gesellschaft grundlegend abhängt und die eine Wertschöpfungsquelle für die Gesellschaft darstellen.“ In § 315 Abs. 3a HGB nF wird eine entsprechende Verpflichtung für Konzernabschlüsse formuliert. Diese Verpflichtung lässt sich wie folgt konkretisieren:

2.4.1 Art und Umfang der Berichterstattung

Die wichtigsten immateriellen Ressourcen sind anzugeben und ihre Bedeutung für das Geschäftsmodell und ihr Wertschöpfungspotenzial sind zu erläutern. Dabei ist es unerheblich, ob die immateriellen Ressourcen bislang schon in der Bilanz erfasst waren oder nicht; es handelt sich hierbei um eine zusätzliche Berichterstattung im Lagebericht. Es besteht eine qualitative Berichtspflicht; eine Bewertung von immateriellen Ressourcen, z. B. selbst geschaffenen Marken, ist nicht erforderlich. Eine Ergänzung der grundsätzlich narrativen Berichterstattung durch finanzielle oder nichtfinanzielle Kennzahlen ist sinnvoll und möglich. Der Berichtsinhalt wird aber nicht weiter konkretisiert; es soll nach aktuellem Stand dazu auch keinen europäischen Standard geben. Umfang und Ausgestaltung liegen daher weitgehend im Ermessen des berichtenden Unternehmens.

Die Berichtspflicht bezieht sich auf den Bestand an immateriellen Ressourcen am Abschlussstichtag. Eine Darstellung der Veränderungen ergibt sich nicht aus dem Gesetzestext. In Kapitel 3 wird darauf weiter eingegangen.

2.4.2 Verortung im Lagebericht

Während für die Nachhaltigkeitsberichterstattung ein eigener Abschnitt im Lagebericht vorgeschrieben ist (§ 289b Abs. 1 HGB nF, § 315b Abs. 1 HGB nF), gibt es keine entsprechende Vorgabe für die Berichterstattung über immaterielle Ressourcen. Die Berichterstattung könnte im Rahmen der Beschreibung des Geschäftsmodells oder bei der Berichterstattung über Forschung und Entwicklung erfolgen. Zu beachten ist zudem, dass es teilweise Überschneidungen mit Nachhaltigkeitsaspekten geben wird, z. B. beim Human Capital (Wulf & Velte, 2024, S. 1018).

2.4.3 Eigenständige Wesentlichkeitsprüfung

Die Regelung wird zwar zusammen mit der Nachhaltigkeitsberichterstattung eingeführt (§§ 289b ff. HGB nF und §§ 315 b und c HGB nF), ist aber kein Teil der Nachhaltigkeitsberichterstattung. Die Regelungen zur doppelten Wesentlichkeit, die für die Nachhaltigkeitsberichterstattung relevant sind, gelten daher nicht für die Berichterstattung über immaterielle Ressourcen (Wulf & Velte, 2024, S. 1018). Es sind lediglich die immateriellen Ressourcen anzugeben und zu erläutern „von denen das Geschäftsmodell der Gesellschaft grundlegend abhängt und die eine Wertschöpfungsquelle für die Gesellschaft darstellen“ (§ 289 Abs. 3a HGB nF). Der Gesetzgeber konkretisiert damit die Vorgehensweise zur Ermittlung der relevanten immateriellen Ressourcen. Allerdings führt er auch Geschäftsmodell und Wertschöpfungsquelle als neue Begrifflichkeiten in das HGB ein, ohne diese zu definieren. Das Geschäftsmodell und die Wertschöpfungsquellen eines Unternehmens sind in hohem Maße unternehmensspezifisch und Ergebnis der Managemententscheidungen. Mit der Entscheidung für ein konkretes Geschäftsmodell entscheidet das Management damit auch über die Grundlage der Berichterstattung. Der aus den IFRS 8 bekannte „management approach“ hält somit auch ins HGB Einzug. Aufgrund ihrer allgemeinen Formulierung gibt die gesetzliche Regelung nur eine Orientierung, ohne eine konkrete Vorgehensweise vorzuschreiben, wie das Geschäftsmodell bzw. die Relevanz der immateriellen Ressourcen für das Geschäftsmodell und ihr Wertschöpfungspotenzial zu ermitteln und zu beschreiben sind.

Das Verständnis für das und die Beschreibung des eigenen Geschäftsmodells im Lagebericht erhalten damit größere Bedeutung in der Berichterstattung als bisher. Für die Beschreibung des Geschäftsmodells empfiehlt es sich einen methodischen Ansatz zu wählen. So kann z. B. das Business Model Canvas von Osterwalder/Pigneur (Osterwalder/Pigneur, 2010) zur Strukturierung verwendet werden: Im Sinne einer „Inventur“ ordnet man die immateriellen Ressourcen den 9 Elementen des Geschäftsmodells zu. In einem weiteren Schritt kann man auf Basis des VRIO-Ansatzes diejenigen immateriellen Ressourcen priorisieren, die aufgrund ihrer Eigenschaften (wertvoll – value, selten – rare, geringe Imitierbarkeit – imitability und Nutzbarkeit in der Organisation – organization) einen strategischen Vorteil und höhere Wertschöpfung versprechen (Barney & Hesterly, 2019, S. 89 ff.).

Unabhängig vom Umfang der Berichterstattung bietet sich hier den Unternehmen aber auch die Chance, sich mit ihren strategischen immateriellen Ressourcen zu beschäftigen, sofern dies nicht schon geschieht. Die Vorbereitung der Berichterstattung durch eine systematische Auseinandersetzung mit den eigenen immateriellen Ressourcen, wie diese zur Wertschöpfung beitragen und in ihrem Wert erhalten werden, kann einen erheblichen Mehrwert für das Management darstellen.

3 Grenzen der neuen Berichtspflicht

3.1 Berichterstattung über den Bestand ist nur eine Seite der Medaille

Wie unter 2.4 dargestellt sind „diejenigen Ressourcen ohne physische Substanz anzugeben, von denen das Geschäftsmodell der Gesellschaft grundlegend abhängt und die eine Wertschöpfungsquelle für die Gesellschaft darstellen, [...]“ (§ 289 Abs. 3a HGB nF) und zu erläutern. Die Verpflichtung bezieht sich damit auf den zum Abschlussstichtag vorhandenen Bestand an immateriellen Ressourcen. Durch diese Berichtspflicht soll die Unvollständigkeit der Bilanz korrigiert werden – Unvollständigkeit ist hier nicht im handelsrechtlichen Sinne zu verstehen, sondern auf die Einbeziehung nicht bilanzierungsfähiger bzw. nicht angesetzter Werte bezogen.

Für den Jahresabschluss wird die Bilanz aber zwingend durch die Gewinn- und Verlustrechnung (GuV) ergänzt, „wobei die GuV als eine die Bilanz ergänzende Detailrechnung zu verstehen ist, mittels derer sich die (Rein-)Vermögensänderung der Bilanz nachvollziehen lässt [...]“ (Coenenberg u.a., 2021, S. 3). Diese Veränderung des Reinvermögens, d. h. Gewinn oder Verlust, gibt als Stromgröße wesentliche Information über den Erfolg eines Unternehmens und seine Quellen. Ziel der Finanzbuchhaltung ist neben dem korrekten (Rein-)Vermögensausweis die periodengerechte Gewinnermittlung, d. h. die korrekte Zurechnung von Aufwendungen und Erträgen zu einzelnen Geschäftsjahren (Coenenberg u.a., 2021, S. 18 ff.). Die immateriellen Ressourcen führen auch in der GuV zu Problemen:

- **Keine Erfassung:** Vermehrungen oder Verminderungen von immateriellen Ressourcen werden in der Finanzbuchhaltung nicht erfasst, d. h. die GuV kann ebenso unvollständig sein wie die Bilanz.

Ein Beispiel hierfür ist die kostenlose Erfassung von Nutzerdaten durch eine Internet-Suchmaschine. Der Nutzer „bezahlt“ die Nutzung der Suchmaschine nicht monetär, sondern durch seine Nutzerdaten, die vom Suchmaschinenbetreiber gespeichert und für weitere Suchen und insbesondere Anzeigenwerbung genutzt werden (Walker, 2021). Da die Datenüberlassung nicht bezahlt oder kompensiert wird, wird sie nicht in der Finanzbuchhaltung erfasst. Immer größere Datenbestände und genauere Nutzerprofile stellen aber wesentliche immaterielle Ressourcen für Suchmaschinenbetreiber dar, da sie dadurch sowohl die Qualität der Suchen als auch die Reichweite ihrer Anzeigenwerbung verbessern können.

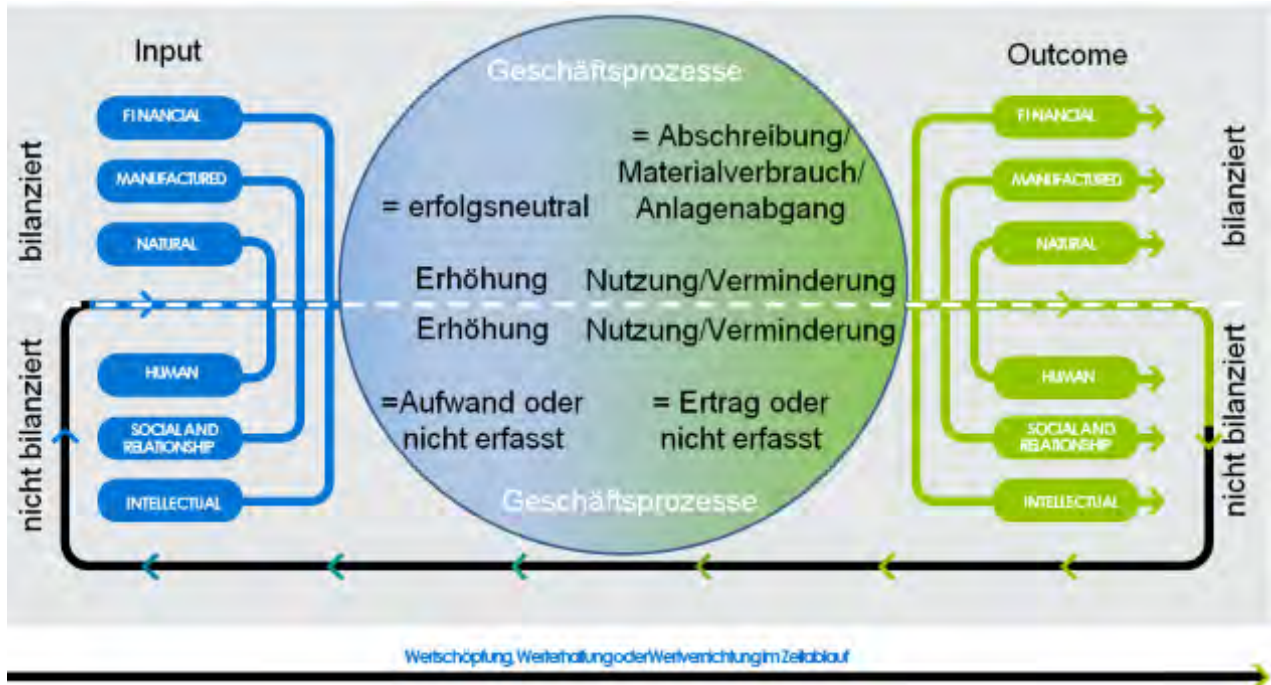
- **Falsche Periodenzuordnung:** Vermehrungen oder Verminderungen von immateriellen Ressourcen werden in der Finanzbuchhaltung erfasst, aber können nicht bilanziell zugeordnet werden, d. h. sie werden als Aufwand oder Ertrag der jeweiligen Periode erfasst.

Ein Beispiel hierfür sind Ausgaben für Mitarbeiterfortbildung. Als monetäre Ausgabe werden sie in der Finanzbuchhaltung als Aufwand der jeweiligen Periode erfasst. In dem Umfang, in dem sie aber die immaterielle Ressource Human Capital vergrößern, stellen sie eine Investition in zukünftige Erträge dar. Der Aufwand der Periode ist somit zu hoch dargestellt, da bei einem vergleichbaren, aber aktivierungsfähigen Sachverhalt der Aufwand über die Nutzungsdauer verrechnet wird.

Die vorgestellten Rahmenkonzepte zur Kategorisierung von immateriellen Ressourcen anerkennen alle, dass die Veränderung der immateriellen Ressourcen die Wertschöpfung des Unternehmens in der Zukunft positiv oder negativ verändern kann. AKIWIR schlägt zur teilweisen Lösung konkrete buchhalterische Ansätze vor (AKIWIR, 2001, S. 991 ff.; IIRC, 2021, S. 21; WICI, 2016, S. 13).

Abbildung 1 verdeutlicht die unterschiedliche Behandlung von bilanzierten und nicht bilanzierten Ressourcen. Die Folge ist, dass nicht nur die Bilanz unvollständig ist, sondern auch die GuV unvollständig und vor allem auch nicht periodengerecht abgegrenzt ist.

Abbildung 1: Fehlerhafte Periodenabgrenzung durch nicht bilanzierte immaterielle Ressourcen



Quelle: In Anlehnung an IIRC, 2021, S. 22.

Um diesem Mangel abzuhelpen, gibt es zwei mögliche Ansatzpunkte:

- Die rechtliche Verpflichtung wird um eine Berichtspflicht über die Veränderungen der immateriellen Ressourcen erweitert. Da das Gesetzgebungsverfahren ohnehin noch nicht abgeschlossen ist und dies wahrscheinlich auch eine Änderung der entsprechenden EU-Richtlinien erfordern würde, ist nicht von einer kurzfristigen Ergänzung auszugehen.
- Im Sinn einer konsistenten und adressatenorientierten Berichterstattung können die Unternehmen selbst ihre Berichterstattung um entsprechende Angaben zu den Veränderungen der immateriellen Ressourcen ergänzen. Dies könnte sich bei entsprechender Verbreitung zur best practice entwickeln und so auch in die einschlägige Kommentierung Eingang finden.

3.2 Berichterstattung über sensible Sachverhalte

Mit der Berichterstattung über immaterielle Ressourcen, die für das Geschäftsmodell grundlegend sind und die zur Wertschöpfung des Unternehmens beitragen, fokussiert die Berichterstattung Kernbereiche der Wettbewerbsfähigkeit des Unternehmens. Immaterielle Ressourcen sind dann grundlegend für das Geschäftsmodell, wenn sie einen Wettbewerbsvorteil gegenüber Konkurrenten bringen und somit strategisch relevant sind. Auch wenn entsprechend dem oben kurz erwähnten VRIO-Ansatz strategisch relevante Ressourcen nicht oder schlecht imitierbar sind, dürfte die Kenntnis über die strategisch relevanten Ressourcen Konkurrenten anregen, über die Relevanz bestimmter immaterieller Ressourcen für ihr eigenes Geschäft nachzudenken. Es zeigt sich hier ganz deutlich der Zielkonflikt zwischen transparenter Berichterstattung für die Berichtsadressaten und dem Erhalt der Wettbewerbsvorteile und dem Schutz der immateriellen Werte durch zusammenfassende, oberflächliche oder ggf. auch nicht erfolgende Berichterstattung.

Der Gesetzgeber hat hier in der CSRD bzw. dem CSRD-UmsG keine Orientierung gegeben. Die bisherigen Schutzvorschriften werden an die Nachhaltigkeitsberichterstattung angepasst, aber nicht erweitert:

- § 286 HGB regelt in welchem Umfang Angaben im Jahresabschluss unterlassen werden können. Für den aktuellen Sachverhalt relevant ist § 286 Abs. 2 HGB; dementsprechend kann die Umsatzaufgliederung im Anhang unterbleiben, „soweit die Aufgliederung nach vernünftiger kaufmännischer Beurteilung geeignet ist, der Kapitalgesellschaft einen erheblichen Nachteil zuzufügen.“ Diese Regelung ist aber explizit auf den Jahresabschluss bezogen. Eine Übertragung im Sinne analoger Anwendung auf den Lagebericht dürfte kaum möglich sein, da es für den Lagebericht eine eigene Regelung gibt.
- § 289e HGB regelt bislang für die nichtfinanzielle Erklärung und künftig für den Nachhaltigkeitsbericht, dass Angaben zu künftigen Entwicklungen oder Sachverhalten, über die Verhandlungen geführt werden, wenn diese Angaben einerseits zu einem erheblichen Nachteil führen können und andererseits das Berichtsziel trotzdem erreicht wird. Diese Regelung ist aber explizit auf die nichtfinanzielle Regelung, künftig den Nachhaltigkeitsbericht, beschränkt und somit nicht auf die Berichterstattung über immaterielle Ressourcen anwendbar (Beck'scher Bilanzkommentar, 2024, § 289e).

Selbst wenn die Regelung auf die Berichterstattung über immaterielle Ressourcen anwendbar wäre, würde sie das Berichtsproblem kaum lösen, da sie nur das Weglassen von Angaben über künftige Ereignisse ermöglicht. Da sich die Berichtspflicht über immaterielle Ressourcen aber auf die zum Abschlussstichtag vorhandenen Ressourcen bezieht, würde die Regelung größtenteils ins Leere laufen.

Die Unternehmen müssen sich daher der Herausforderung stellen, eine angemessene Balance zwischen Transparenz für die Berichtsadressaten und Schutz der eigenen Wettbewerbsvorteile zu erreichen. Es wird interessant sein, die ersten Lageberichte zu analysieren, ob tatsächlich über konkrete strategisch relevante immaterielle Ressourcen berichtet wird oder ob die Berichterstattung eher eine oberflächliche Sammlung von allgemeinen Aussagen sein wird.

4 Zusammenfassung und Ausblick

Mit der CSRD verpflichtet die Europäische Union viele Kapitalgesellschaften zu einer Berichterstattung über ihre strategisch relevanten immateriellen Ressourcen. Dies erweitert die Berichtspflichten neben der neuen Nachhaltigkeitsberichterstattung erheblich. Bei der Umsetzung stellen sich noch viele Fragen, da die Regelung sehr allgemein gehalten und interpretationsbedürftig ist. Dieser Beitrag liefert eine Einordnung und erste Ansätze zum Umgang mit der neuen Verpflichtung. Es werden auch zwei Konfliktfelder aufgezeigt: Zum einen ist eine reine Berichterstattung über den Stand der immateriellen Ressourcen nicht ausreichend, sondern allenfalls ein erster Schritt, um sich dem Thema zu nähern. Zum anderen ist hier eine schwierige Balance zwischen Transparenz und Schutz der eigenen Wettbewerbsvorteile zu finden. Insofern wird es spannend sein zu beobachten, wie die Unternehmen sich dieser neuen Herausforderung stellen und welche best practices sich herausbilden.

Literaturverzeichnis

- AKIWIR (2001): Kategorisierung und bilanzielle Erfassung immaterieller Werte. Der Betrieb 19/2001, S. 989 – 995.
- Barney, J.; Hesterly, W. (2019): Strategic Management and Competitive Advantage. 6. A., Pearson, Harlow/UK.
- Beck'scher Bilanzkommentar (2024), 14. A., C. H. Beck, München.
- BrandFinance (2023): Global Intangible Finance Tracker GIFT 2023. London.
- Bundesjustizministerium (2024): Gesetz zur Umsetzung der Richtlinie (EU) 2022/2464 des Europäischen Parlaments und des Rates vom 14. Dezember 2022 zur Änderung der Verordnung (EU) Nr. 537/2014 und der Richtlinien 2004/109/EG, 2006/43/EG und 2013/34/EU hinsichtlich der Nachhaltigkeitsbericht-erstattung von Unternehmen. Referenten-Entwurf, URL: https://www.bmj.de/SharedDocs/Gesetzgebungsverfahren/DE/2024_CSRD_UmsG.html?nn=110518 (abgerufen am 29. Juli 2024).
- Bundestag, Zweites Gesetz zur Änderung des DWD-Gesetzes sowie zur Änderung handelsrechtlicher Vorschriften, BGBl. 2024 I Nr. 120 vom 16.04.2024.
- Coenenberg, A.; Haller, A.; Schultze, W. (2021): Jahresabschluss und Jahresabschlussanalyse. Schäffer-Poeschel, Stuttgart.
- Europäische Union (2013): Richtlinie 2013/34/EU vom 26. Juni 2013, Abl. L 182/19.
- Europäische Union (2022), Richtlinie (EU) 2022/2464 vom 14. Dezember 2022, Abl. L 322.
- Europäische Union (2023): Delegierte Richtlinie (EU) 2023/2775 vom 17. Oktober 2023, Abl. L, 2023.
- Haller, A.; Fischer, T. (2023): Berichterstattung über Intangibles – Neue Impulse durch die CSRD. KoR IFRS Internationale und kapitalmarktorientierte Rechnungslegung 2/2023, S. 78 – 86.
- IIRC (2021): International <IR> Framework. International Integrated Reporting Council.
- Liening, M.; von Keitz, I.; Wulf, I. (2023): Berichterstattung über immaterielle Ressourcen gemäß Richtlinie (EU) 2022/2464. KoR IFRS Internationale und kapitalmarktorientierte Rechnungslegung 7-8/2023, S. 296 – 308.
- Osterwalder, A.; Pigneur, Y. (2010): Business Model Generation. Wiley, Hoboken NJ.
- Von Keitz, I.; Schwedler, K. (2023): Immaterielle Ressourcen als Werttreiber für (nachhaltiges) Wirtschaften. Bertelsmann Stiftung, Gütersloh.
- Walker, M. (2021): Die Kosten kostenloser Dienste: Personenbezogene Daten als neues Zahlungsmittel. Duncker & Humblot, Berlin.
- WICI (2016): WICI Intangibles Reporting Framework. World Intellectual Capital Initiative.
- Wulf, I.; /Velte, P. (2024): Nachhaltigkeitsberichterstattung nach dem Referentenentwurf eines Umsetzungsgesetzes zur CSRD – eine kritische Würdigung. Der Betrieb 17/2024, S. 1017 – 1024.
- Zambon, S.; Marzo, G.; Girelle, L.; Abela, M.; D'Albore, N. (2020): A literature review on the reporting of intangibles. EFRAG, Brüssel.

The effects of auditor rotation and auditor tenure on audit quality: a meta-analysis

Research Paper

WP Prof. Dr. Markus Widmann

Hochschule für Wirtschaft und Gesellschaft Ludwigshafen, Professur für Allgemeine Betriebswirtschaftslehre, insb. Wirtschaftsprüfung und betriebliche Steuerlehre, Ludwigshafen am Rhein, E-Mail: markus.widmann@hwg-lu.de

Nicolas Flohr

Master-Absolvent der Universität Trier im Studiengang «Financial Management»

Matthias Wolz

Universität Trier, Professur für Allgemeine Betriebswirtschaftslehre, insb. Rechnungswesen und Wirtschaftsprüfung, Trier

1 Introduction

The auditing profession is subject to strict regulation. This is characterized in particular by the fact that in the aftermath of crises, legislators, the public and the media usually criticize existing regulations, discuss further ones and implement new ones. At the level of the European Union (EU), the financial crisis of 2008/2009 mainly led to new regulations for financial statement audits. The intended objective of these regulations should be to increase audit quality. In order to achieve this goal, the legislator from Brussels used in particular such measures that could improve the perceived independence of the auditor. As a result, EU Regulation No. 537/2014 of the European Parliament and of the Council was endorsed. This regulation includes the limitation of the maximum period before mandatory replacement of the responsible audit partner. Long-lasting engagement relationships between the auditor and the client can raise concerns of excessive familiarity between the actors involved. Such an obligation to replace the responsible audit partner after the expiration of a legally prescribed maximum tenure implies that the quality of the audit significantly decreases as the tenure of the auditor grows. Since this measure is used internationally as a regulatory instrument, it is important to check whether it truly has the desired effect. For this purpose, we carry out a meta-analysis to be able to make a statement about the relationship between audit partner rotation and audit quality based on the existing research.

2 Theoretical background

2.1 Preliminary considerations

While mandatory external rotations provide for changes in audit firms (e.g., Widmann et al. (2021a), an audit partner rotation refers to the change in the responsible audit partner within the company during the audit of a particular client. Accordingly, in the case of an audit partner rotation, the auditing firm remains in charge of the audit of the respective company, but the responsible audit partner resigns after a certain period, and another audit partner at the same firm takes his or her position. After the change has been completed, a so-called "cooling-off period" usually begins. During this period, the resigned audit partner may not again participate in the audit of the previous client.

Regarding the independence of the responsible audit partner, the legislator assumes a negative correlation between it and the duration of the mandate. According to the International Ethics Standards Board for Accountants (IESBA), independence is linked to principles of objectivity and integrity. It comprises independence of mind and independence in appearance. The decrease in independence (of mind) over time can be explained by the gradual increase in proximity between the auditor and management. The longer an audit partner is responsible for auditing a company, the more familiar he may become with its management. If personal relationships and dependencies develop in this way, the auditor may lose his critical view of the issues and a sense of scepticism to a certain extent. As a result, the auditor may lack the necessary accuracy and rigor to detect and report truthfully on certain misrepresentations (e.g., Carey/Simnett 2006). Independence in appearance describes how external parties perceive the relationship between auditor and management. Even if no ties have arisen after a long period of service as the responsible audit partner, users of the annual financial statements may still doubt the auditor's opinion, as they no longer believe in his unqualified independence due to conceivable relationships. If, therefore, confidence in the work of the auditor is shattered by a long engagement period, the auditor's opinion will also be of lower value. In addition to the problems arising from the relationship between the auditor and management, a certain audit routines could creep in over time. After many years of service, the auditor may be inclined to always follow the same pattern rather than perform innovative audit procedures. As a result, he may anticipate certain findings and lose the ability to detect new misstatements.

As seen, both the dwindling independence of the auditor and the creeping operational blindness could have a negative effect on the audit and thus its quality. Audit partner rotation is intended to avoid quality-challenging situations such as these. Before the negative effects of a long mandate could arise, the responsible audit partner must change. The newly appointed auditor can continue the audit with a fresh and unbiased attitude and provide a new impetus, which should also send a positive signal to stakeholders.

However, a long tenure can also be linked to a positive effect on audit quality. After all, the longer an auditor audits a client, the more company- and industry-specific knowledge he acquires. Through his experience, he becomes an expert in the audit of the respective company. If an audit partner rotation takes place, his entire knowledge is lost. A new auditor must first learn the ropes and build up knowledge over the years. Given this situation, it is to be expected that qualitatively poorer audits are carried out after an audit partner rotation since a new auditor cannot uncover misrepresentations as well as a long-established auditor can due to the lack of experience. Audit partner rotations can therefore also be viewed critically, as they are always accompanied by a loss of expertise (e.g., Litt et al. 2014).

2.2 Audit quality and its measurement

Regulations such as for audit partner rotations are being introduced to strengthen audit quality. The aim is to maximize audit quality by strengthening independence and preventing operational blindness. However, although “audit quality” is such a central factor, no regulatory or professional body has yet provided an exact definition of the term. Only the explanation by DeAngelo (1981) is widely accepted in the scientific community and has thus established itself as a definition. According to DeAngelo, audit quality should be understood as the combined likelihood that an auditor identifies and truthfully reports on a violation in the client's accounting system. Therefore, audit quality consists of two fundamental aspects: (a) the probability that an auditor discovers a misrepresentation in the client's financial statement and (b) the probability of reporting these breaches. The first aspect can be associated with the auditor's professional competence and technical resources. Given his skills and resources, the auditor is generally in a position to carry out adequate audit procedures to find misstatements with sufficient accuracy and certainty. The second aspect addresses the independence of the auditor, which can be understood as the willingness to report any misstatements found or to work towards their correction. Both professional skills and the willingness to report the findings are crucial aspects of audit quality. Even if auditors perform excellent, the overall quality of the work can only be as good as their independence allows them to report the findings. Since independence can be distinguished into independence of mind and independence in appearance, audit quality also embraces these aspects. Therefore, in line with DeAngelo (1981), audit quality has two facets: audit quality in fact (combines skills and independence of mind) and perceived audit quality given by the auditor's independence in appearance, defined by how the auditor's work is perceived by readers of the financial statements.

Audit quality in fact is extremely difficult and extremely costly to observe; doing so requires calculating the joint probability that an auditor recognizes and reports a misstatement. However, the methods and procedures used in an audit are not known to the financial statement users and, generally, cannot be directly understood. Similarly, non-involved persons have little information about the relationship between the auditor and the client (e.g., DeAngelo 1981). Furthermore, it is practically impossible to observe the auditor's independence of mind. Thus, DeAngelo concludes that a more cost-effective way to measure audit quality is to use so-called surrogates. Surrogates are certain variables that are easier to observe directly and at the same time correlate with audit quality.

A widespread and often used surrogate is the extent of the accounting policy. These policies allow companies a certain degree of latitude in “embellishing” their financial statements to provide a better picture of their statement of financial position and performance than is the case if the financial statements were presented as a true and fair view. Given this premise, a low degree of an accounting policy suggests that the company's financial reporting is of high quality. This in turn can be associated with a high-quality audit. If, after the audit, there is only a low degree of an accounting policy, the auditor can be considered to have successfully prevailed against the client's point of view and to have insisted on a presentation of the actual circumstances (e.g., Widmann/Wolz 2019).

To determine the extent of the accounting policy, it is first necessary to determine the level of earnings without accounting policy interventions. This can be done based on the difference between the resulting comparative situation without an accounting policy and the current situation with an accounting policy. However, the problem here is that the reference situation without the influence of an accounting policy cannot be defined without assumptions and estimates. This is because it is difficult to assess whether a measure and how it is accounted for is motivated by

the accounting policy. To address this problem, various statistical approaches have been established in the international literature, all of which assume that the accounting policy is primarily used to control earnings. This means that accounting policy measures are mainly used by management to influence the annual result. When considering the carrying amounts, it can therefore be assumed that the extent of the accounting policy pursued can be found in the accruals and deferrals in the annual financial statements. Accruals and deferrals are noncash expenses and income. The amount of accruals and deferrals is therefore the difference between the net profit for the period and the cash flow from operating activities. Once the total value of the accruals has been determined, it is necessary to assess what portion has arisen from the motivation to distort the period's actual result. A certain extent of accruals goes hand in hand with the normal business activities of a company and therefore reflects the true earnings situation. To determine the extent of the accounting policy from the total accruals, they must be divided into two elements: discretionary accruals (DA) and nondiscretionary accruals (NDA).

3 Research design and descriptive results

3.1 Literature identification process

We used the EBSCOhost search platform to search for relevant literature in the research field of audit partner rotation and audit quality. After selecting all databases available, a search was conducted for publications containing the keywords "partner rotation" and "partner tenure" in combination with "audit quality". In total, this procedure yielded 209 search results, among which 138 articles from scientific journals could be identified. Four were not written in English and were excluded. After the elimination of duplicates, 113 results remained (see table 1).

Table 1: Literature identification process

Procedure description	Results
Total search results of the literature search on EBSCOhost in all available databases with the keywords "partner rotation" OR "partner tenure" AND "audit quality" (reference date: 4th April 2022)	209
Filtering out all publications that did not appear in scientific journals	-71
Excluding all contributions not written in the English language	-4
Eliminating duplicates	-21
Remaining search results from scientific journals in the English language	113

These 113 research articles were transferred to an Excel list and sorted alphabetically by title following the approach suggested by Hansen et al. (2022). Subsequently, certain characteristic data were extracted from the respective publications and recorded in an overview. In addition to general information about the authors, the scientific journal and the year of publication of the work, these data also included specific information about the type of research carried out. At the beginning, we analysed whether the respective paper actually provided an analysis of the influence of audit partner rotation on audit quality; if not, it was excluded from further processing. Subsequently, the type of contribution was determined. Empirical studies were labelled with a 0, and nonempirical studies were labelled with a 1. In the next step, the extent to which the analysis of audit quality was carried out in the articles was discussed. First, the relevant publications were again marked. This time, articles that dealt with audit quality as an object of investigation received a 1. Then, the surrogate for the investigation of audit quality was taken from the papers. If this was the extent of the accounting policy or the quality of the financial reporting, the model used to determine DA was also noted. Since the literature review covered both "partner rotation" and "partner tenure", it was necessary to assign to the contributions whether the change or the tenure of the responsible audit partner was used as the independent variable. After this assignment, the country, the region beyond, and the period of the examination were recorded. In addition, the characteristic data of the empirical analysis were recorded. These include

the type of association, sample size, β -coefficient, degree of significance, p value, t value, correlation coefficient, and standard deviation of the independent and dependent variables.

3.2 Review and further exclusion

To carry out the meta-analysis and generally improve the comparability of the results of the research contributions, further exclusion was carried out after the literature review described above. A total of 24 of the 113 studies investigated audit quality based on the financial reporting quality or the extent of the accounting policy. Thus, this was the most frequently used surrogate. Therefore, only publications from the literature search that empirically examined the influence of rotation or the tenure of the responsible audit partner on audit quality using the measurement of the extent of the accounting policy were relevant for further processing in this paper. The selection is described in Table 2:

Table 2: Review procedure

Procedure description	Results
Remaining search results from scientific journals in English language	113
Exclusion of all research papers without financial reporting quality or the extent of the accounting policy as a surrogate to investigate audit quality	-89
Exclusion of all research with nonlinear correlation.	-2
Exclusion of all research with unclear results or insufficient data	-4
Remaining research that empirically examined the impact of the rotation or term of the responsible audit partner on audit quality using measures of the degree of accounting policy	18

Among these 18 research papers, three publications embraced two different studies of relevance to our meta-analysis; thus, we obtained a total of 21 relevant empirical studies (see table 3).

Table 3: Literature breakdown categories

Overview	Results
Empirical findings resulting from the 18 research papers	21
<u>thereof:</u> Studies addressing the impact of audit partner rotation on audit quality	8
<u>thereof:</u> Studies using the modified Jones model (Dechow et al. 1995)	3
Studies using the performance-matched-modified Jones model (Kothari et al. 2005)	3
Use of a different model to estimate the extent of accounting policy	2
<u>thereof:</u> Studies addressing the impact of audit partner tenure on audit quality	13
<u>thereof:</u> Studies using the modified Jones model (Dechow et al. 1995)	5
Studies using the performance-matched-modified Jones model (Kothari et al. 2005)	4
Using a different model to estimate the extent of accounting policy	4

3.3 Overview of research findings

In the research area of audit partner rotation, Mohapatra et al. (2021) study the impact of voluntary audit partner rotation on audit quality using several indicators in the Indian audit market. They collect a total of 1,694 observations between 2011 and 2017, taking companies from the "S&P BSE 500" index. Audit quality is measured by the extent of the accounting policy using the modified Jones model of Dechow et al. (1995). Through an examination of the relationship, the authors conclude that changing audit partners has a positive but not significant impact on DA.

In light of the tightening of the audit partner rotation requirement in the U.S. by the Sarbanes-Oxley Act in 2003 and the scarcity of the expert literature in this area, Litt et al. (2014) study the impact of audit partner rotation on audit quality in the United States. Since in the United States, the responsible audit partner is not apparent from the public documentation of an audit, the authors mainly observe external rotations. In these cases, Litt et al. (2014) can be sure that a new audit partner is responsible for the audit of the company's financial statements. They also assume that if a new audit firm is engaged for the minimum term of seven years, audit partner rotation does not take place again until the five years as prescribed in the United States have elapsed. Thus, the authors assume that during a seven-year engagement, one change can be observed directly at the beginning and another change after the fifth year. To generate their dataset, Litt et al. (2014) identify all firms that change their audit firm between 2000 and 2004. They then track the financial reporting of all firms that retain their auditor over the minimum engagement period of seven years. In total, this approach results in a dataset of 2,108 observations. For the determination of audit quality, Litt et al. (2014) use the level of DA, as they assume that profit forecasts are to be achieved or exceeded with the help of DA. To estimate the DA, they use the performance-matched-modified Jones model of Kothari et al. (2005). As a result, Litt et al. (2014) find a highly significant positive relationship between audit partner rotation and DA. In the first two years after a rotation, DA are higher than in the last two years before a rotation. Consequently, audit quality decreases during an audit partner rotation.

Additionally, using the Kothari et al. (2005) model, Kim/Xi (2021) analyse the changes in DA during an audit partner rotation. They conduct the study in China and consider both the rotation of the review auditor and the engagement auditor. Under this dual auditorship practice, both auditors sign the audit report. However, the engagement auditor is usually involved in all phases of the audit process, including fieldwork, whereas the review auditor chiefly conducts a final review of the audit report. While Kim/Xi (2021) find no significant effect on DA for the mandatory rotation of the engagement partner, they find a significant positive correlation for the internal mandatory rotation of the review partner.

Similar to other authors, Horton et al. (2021) use regulatory interventions in the audit market as a starting point for their elaboration. They are interested in the difference between the coexistence of internal and external mandatory rotations. In the U.S., regulators have opted against introducing external rotations and are in favour of introducing internal rotations, whereas in Italy, both have been mandated simultaneously. In their paper, Horton et al. (2021) examine whether this leads to other effects. For the analysis, the authors collect information on companies listed on the Milan Stock Exchange from 1993 to 2012. Since 2006, when EU Directive 2006/43/EC became effective in Italy and resulted in changes to both internal and external mandatory rotation rules, the span from 2006 to 2012 serves as the main study of this dual system. In total, Horton et al. (2021) collect data on 141 firms and make 2,104 observations. As a surrogate for audit quality, they use the quality of financial reporting, which they measure using three different "abnormal accruals". In addition to two other quality measures that serve as a robustness check, they estimate the absolute value of DA, which is the residual amount resulting from a regression analysis that they set up on their own. As a result of their analysis, Horton et al. (2021) find that in the dual system period from 2006 to 2012, internal rotation has a significant negative impact on DA. Accordingly, the considered abnormal accruals decrease due to the change in the responsible audit partner, which ultimately increases audit quality. In the 13-year period prior to 2006, the authors do not find any significant effect in the case of external rotation alone, so Horton et al. (2021) conclude that the positive effect on audit quality during the simultaneous performance of internal and external rotations comes primarily from the rotation of audit partners.

Corbella et al. (2015) investigate the impact of internal audit partner rotation on audit quality. Although the focus of the paper is on the impact of external rotations, the authors also collect data on internal rotations. They conduct their study in Italy. Corbella et al. (2015) manually collect publicly available information on listed companies between 1998 and 2011, resulting in a total of 1,583 observations. The authors use "abnormal working capital accruals" as a surrogate for audit quality. They estimate the amount of these accruals using the difference between the value of the actual working capital in the year under review and the prediction of this value from the ratio of working capital to sales in the previous year. As a result, Corbella et al. (2015) find a highly significant negative relationship between audit partner rotation and abnormal working capital accruals. This means that audit quality increases with a change in the audit partner.

Unlike the previous research contributions, Kalanjati et al. (2019) investigate not only the impact of the rotation of the responsible audit partner but also the impact of his tenure. For this purpose, they collect 688 observations from companies listed on the Indonesia Stock Exchange between 2003 and 2016. The authors measure audit quality using the extent of the accounting policy. For Kalanjati et al. (2019), the absolute DA is the key variable, which they estimate using the modified Jones model. The regression analysis performed shows that a high number of audit partner rotations during an engagement is significantly and negatively related to absolute DA. From this, the authors conclude that the rotation of audit partners has a positive influence on audit quality. A negative but statistically insignificant influence on the absolute DA can also be found for the tenure of the audit partner.

Additionally, from Indonesia, Martani et al. (2021) examine the tenure and rotation of the responsible audit partner. However, the motivation of Martani et al. (2021) is to determine whether there are differences between Big 4 and non-Big 4 companies in this area. For the analysis, Martani et al. (2021) collect data from 215 listed companies between 2013 and 2017, resulting in a total of 1,001 observations after minor exclusions. The authors use DA to determine audit quality using the modified Jones model of Dechow et al. (1995) and its variant according to Kaznik (1999). However, Martani et al. (2021) find no significant correlations with audit quality, regardless of the model used. This applies to both internal rotation and the tenure of the responsible audit partner. When audits performed by Big 4 or non-Big 4 audit firms are considered separately, Martani et al. (2021) only find a significant negative effect for Big 4 firms in terms of audit partner rotation and for non-Big 4 firms in terms of external rotations.

Another research paper that addresses both audit partner rotation and tenure is Chang et al. (2018). This paper refers to the Taiwanese audit market and aims to show the differences in using a linear and a hierarchical-linear regression model. The authors collect information on all companies listed on the Taiwan Economic Journal database and calculate the tenures of the responsible audit partners from 1983 to 2011, resulting in a total of 14,561 observations. Audit quality is determined using the performance-matched-modified Jones model according to Kothari et al. (2005). The linear regression analysis does not indicate a significant relationship between audit partner tenure and audit quality. However, the authors detect a highly significant decrease in DA with respect to audit partner rotation, indicating that it has a positive effect on audit quality.

Buntara/Adhariani (2019) publish another study from Indonesia that focuses on the relationship between audit partner tenure and audit quality. For their study, they collect data from 2010 to 2014 on companies listed in the Kompas 100 index in 2016. Buntara/Adhariani (2019) determine audit quality by the extent of the accounting policy and DA, respectively, using the modified Jones model of Dechow et al. (1995). The authors find that long auditor-client relationships are significantly and positively related to DA. Thus, they conclude that long tenures of the responsible audit partner have a decreasing effect on audit quality.

Garcia-Blandon et al. (2020) and Fargher et al. (2008) find results similar to those in Buntara/Adhariani (2019). Garcia-Blandon et al. (2020) investigate the influence of audit partner tenure on the level of DA using the modified Jones model of Dechow et al. (1995) in Spain and find a significant positive correlation. Fargher et al. (2008), on the other hand, conduct their study in Australia and use a cross-sectional modified Jones model to determine DA. They also find a highly significant positive impact of audit partner tenure on DA, which equates to a decrease in audit quality.

Ball et al. (2015) also discover a significant and positive correlation. In Australia, they investigate the effects of long mandate tenure on audit quality to assess the justification of external and internal mandatory rotations. For this purpose, they collect information from 266 companies in the S&P/ASX Top 500 index for 2006. Ball et al. (2015) use financial reporting quality as a surrogate for audit quality. However, they do not measure this via DA but using the difference in the reported annual result for the reporting year 2005 in the financial statements of 2006 and 2005. The authors justify their approach with the fact that the publication of annual financial statements in accordance with the International Financial Reporting Standards (IFRS) became mandatory as of January 1, 2005, for the companies examined. For this purpose, the companies had to change their accounting methodology and calculate new valuations. According to the regulation, these should be made as accurate as possible. However, it is conceivable that at the time of the new introduction and the preparation of the first financial statements in accordance with IFRS, some inaccuracies may have reduced the quality of the published report. However, these possible inaccuracies may have been discovered and corrected by the time of the next annual financial statements in 2006, with the result that the previous year's figures reported for the 2005 annual result in the 2006 report differ from those of the annual result presented in the 2005 report. The results of Ball et al. (2015) show that long auditor-client relationships reduce audit quality.

In contrast, other research reach completely different results. Singh et al. (2019) find for their Australian dataset that a long tenure of the responsible auditor has a highly significant and negative effect on DA. This result is provided by both the modified Jones model of Dechow et al. (1995) and the performance-matched-modified Jones model of Kothari et al. (2005). This finding is consistent with that in Manry et al. (2008) and Chen et al. (2008). Manry et al. (2008) conduct a corresponding study in the U.S. to investigate the auditor rotation regulation introduced by the Sarbanes-Oxley Act using a cross-sectional modified Jones model to estimate DA. They find that audit quality significantly increases with the length of the audit partner's tenure. Chen et al. (2008) reach to the same conclusion with their study in Taiwan. However, they use a slightly modified version of the performance-matched-modified Jones model of Kothari et al. (2005) to determine the DA.

In China, Liu et al. (2021) find a similar effect. They investigate whether long auditor-client relationships have an impact on audit quality. For this purpose, they collect a total of 12,805 observations from 1998 to 2009. As a surrogate for audit quality, they take DA, which they determine with the performance-matched-modified Jones model of Kothari et al. (2005). They find a significant negative influence of internal audit partner tenure on the level of DA for normal auditor-client relationships.

In addition to the research papers that show a significant relationship between the tenure of the responsible audit partner and audit quality, there are also studies that do not find a significant relationship. One of these studies is by Carey/Simnett (2006). They use three different evaluation criteria for their analysis. First, they observe whether the longer the auditor works for a client, the less the responsible audit partner tends to doubt the company's ability to continue as a going concern ("going concern audit opinion"). On the other hand, they investigate whether certain performance targets are more likely to be achieved or missed with increasing audit partner tenure. In addition, they also focus on the change in abnormal working capital accruals. The study takes place in Australia and is based on a total of 1,021 observations. The results show that responsible audit partners who have been working for a client for more than seven years are less likely to question the continuation of financially distressed companies and to issue a going concern audit opinion. The authors find no significant correlation between hitting or missing certain earnings targets and long audit partner tenures. They reach a similar conclusion for abnormal working capital accruals; therefore, according to this measure of accounting policy, audit quality is not influenced by long-term auditor-client relationships. This finding is supported by Baatwah (2016) and Gipper et al. (2021). Baatwah (2016) uses 573 observations in Oman to examine whether long auditor tenures have an impact on the level of DA or the audit opinion. While as noted above, no significant relationship is found for DA, Baatwah (2016) finds that responsible audit partners are more likely to issue a modified audit opinion with increasing audit tenure, indicating higher audit quality. Gipper et al. (2021), in their study of the U.S. audit market, also do not find a significant influence of long engagement periods on audit quality, which they determine on the basis of absolute accruals. The authors merely

see that audit fees fall and working hours rise during a mandatory rotation. Both subsequently regulate themselves again over the tenure.

3.4 Summarizing results of the literature review process

An initial examination of the data collected and processed in the resulting 18 research papers reveals that all of the papers were published in the 2000s, starting in 2008. Of course, one reason for this is the fact that, based on the selection of literature, all of these papers are studies of financial reporting quality primarily based on the determination of DA using the Jones model or its variants. The original Jones model dates back to 1991 (Jones 1991). However, this model was not used in its original form in any of the studies presented here. Rather, modifications according to Dechow et al. (1995) and Kothari et al. (2005) were used. Based solely on the publication years of these research articles, it stands to reason that the publications collected in this paper were published after the turn of the millennium. On the other hand, the timing of the publications was also related to the historical background and the respective events. A particularly formative event for the regulation of the auditing profession and thus of audit partner rotation in particular was the accounting scandal of the U.S. energy company Enron. Following the insolvency of Enron due to immense accounting fraud and the collapse of the audit firm Arthur Andersen, the U.S. legislature decided in 2002 to regulate auditing more strictly via the Sarbanes-Oxley Act. Across the borders of the United States, many regulators in other countries subsequently adopted the Sarbanes-Oxley Act as a model or reason to also reform their auditing regulations. Furthermore, in recent years, there have been additional plans for change, as was observed in the EU during the financial crisis of 2008/09 and in Germany after the Wirecard scandal. In addition to these possible explanations for the fact that the years of publication are not far behind, it can generally be stated based on the data that the study of the influence of internal rotations on audit quality by measuring the extent of the accounting policy is still a young field of research in which activity is gradually increasing. As shown in Figure 1, ten of the 18 research papers have been published in the last three years. A slight trend is appearing that indicates that more attention is now being paid to this topic, and the analysis of the relationship is intensifying.

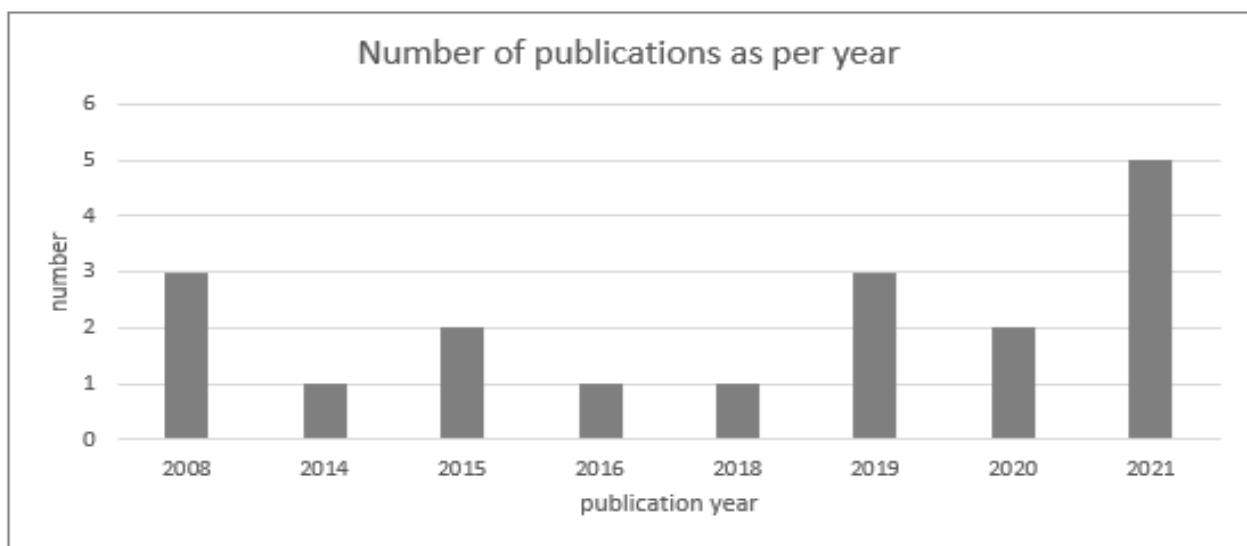


Abb. 1: Number of publications per year

The recent increase in research interest also explains the focus on the years under investigation. Since the vast majority of research contributions have taken as their motivation an examination of the justification or effectiveness of the newly introduced regulations in their countries, the analyses focus on the years after the respective regulations came into force. Figure 2 shows that the most extensively studied period is between 2003 and 2014. In total, the 18 studies collected cover the period from 1990 to 2017.

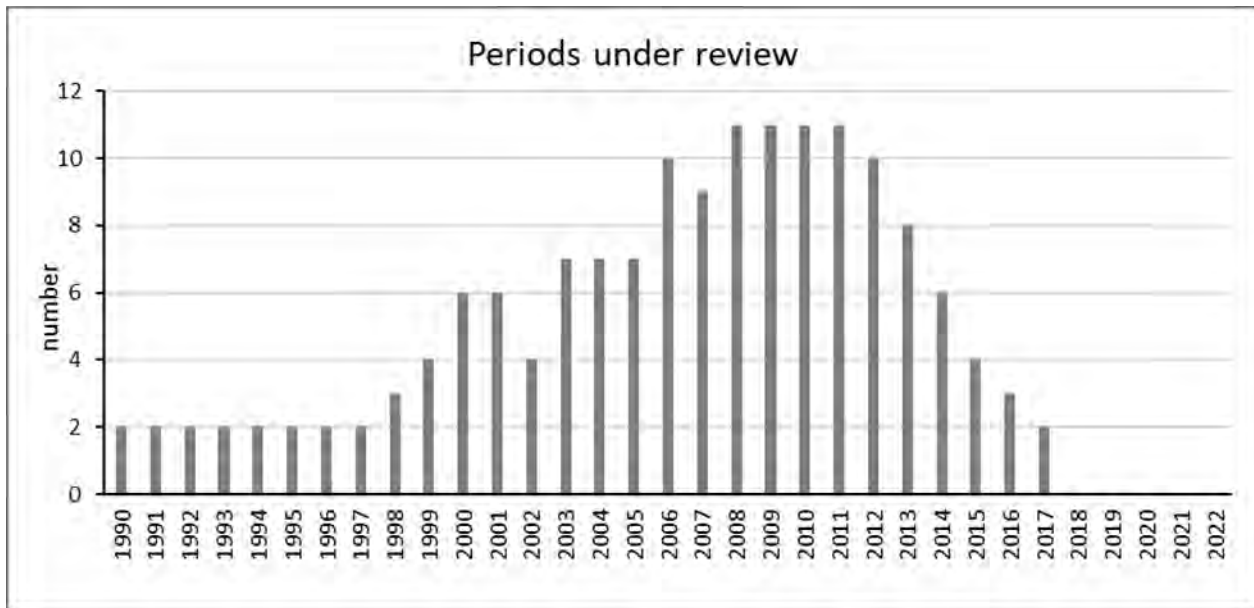


Abb. 2: Periods under review in audit partner rotation studies (number of studies that use data from the respective year)

These data show that the influence of internal rotations on audit quality and accounting policy is under review almost globally. Among the 18 research contributions, only the continents of Africa and South America are not represented. However, there are regional differences in the intensity of the research. While nine publications on the topic come from Asia, only three papers each come from Australia, Europe and North America, as shown in Figure 3.

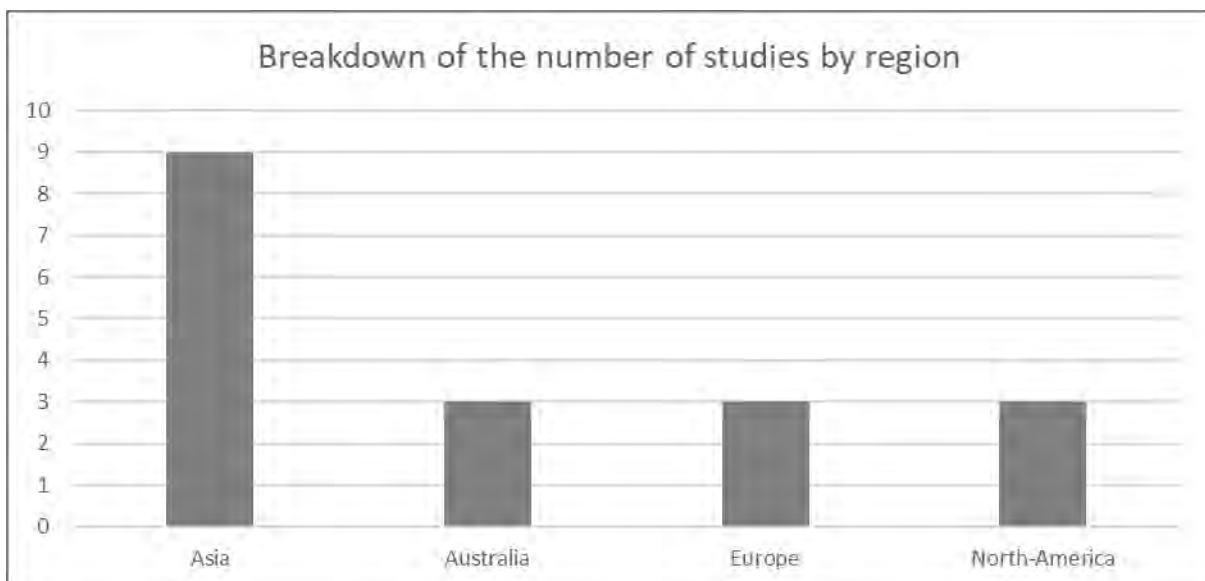


Abb. 3: Breakdown of the number of studies by region

It can be assumed that the range of disclosed information – which varies significantly in different regions – may be responsible for the different publication figures. For example, in countries such as Indonesia, China and Australia, the names of the signatory audit partners are included in the financial statements so that they can be identified and their tenure determined. As described above, according to Litt et al. (2014), this is not the case in the United States, making it more difficult for researchers there to find useful information for an investigation.

Figures 4 and 5 present the results of the studies using the approach by Widmann et al. (2021b). The x-axis shows the direction and significance of the effect of audit partner rotation or audit partner tenure on the audit quality found in the examined studies. Negative numbers indicate a negative correlation, and positive numbers indicate a positive

correlation. The respective scale in both areas from 1 to 4 shows the significance of the effects. While 1 represents an insignificant result, 2 through 4 indicate significant results at the 10, 5 and 1 percent levels ($\alpha=0.10$, $\alpha=0.05$ and $\alpha=0.01$). When examining the results of the studies on the influence of auditor rotation on the extent of audit quality, it can be seen that four of the eight analyses found a negative, significant correlation. Among these, three are highly significant. Of the remaining four results, three show a positive but nonsignificant correlation, and only one shows a highly significant positive correlation. These initial observations indicate that the change in audit partner generally leads to a lower degree of accounting policy and thus increases audit quality.

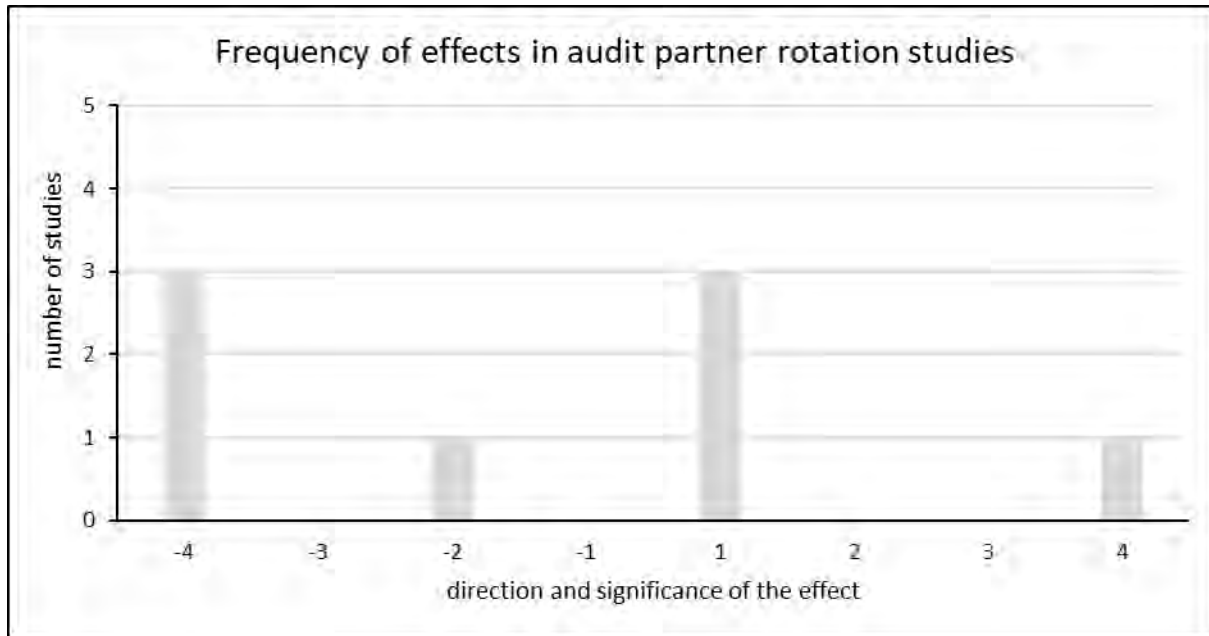


Abb. 4: Frequency of effects in audit partner rotation studies

The results of the studies on the tenure of the responsible audit partner are somewhat broader than those on the rotation of the audit partner. Four of the 13 analyses show a significant negative correlation between tenure and the extent of accounting policies. Another four studies, however, conclude that there is a significant positive correlation. The remaining five studies do not find any significant results. Four of them only detect a negative and one a positive influence of audit partner tenure on the extent of the accounting policy. Based on these initial observations, it is not possible to interpret whether and how audit partner tenure influences audit quality across studies. Overall, more negative than positive correlations were identified.

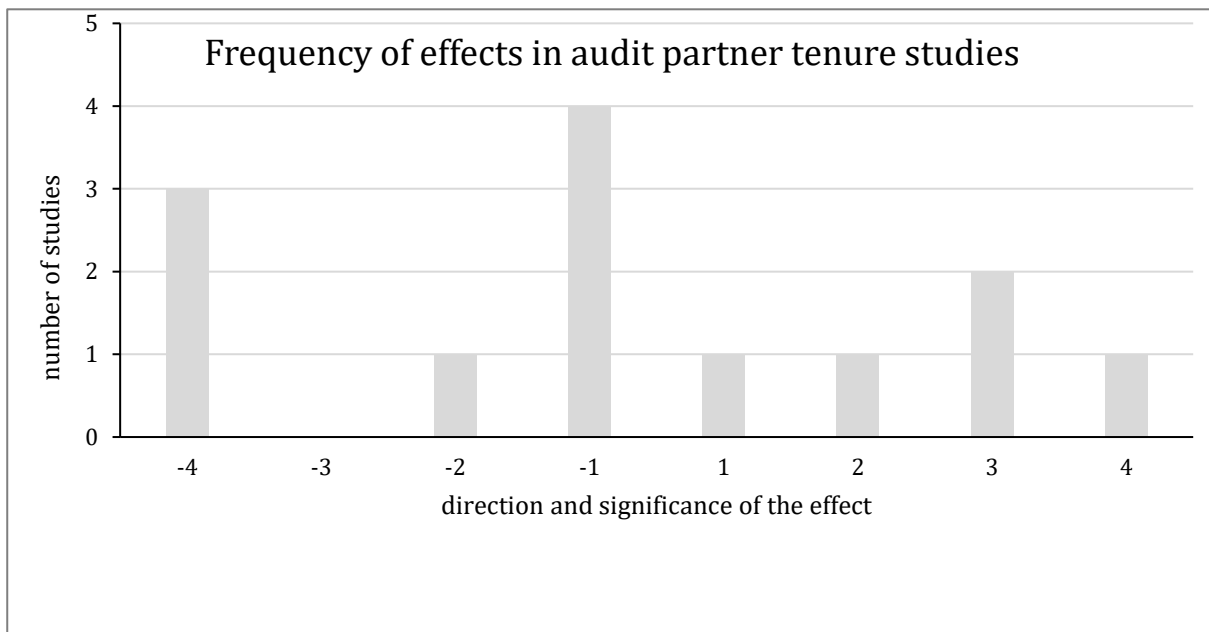


Abb. 5: Frequency of effects in audit partner tenure studies

4 Meta-analysis and discussion

4.1 Describing the approach

The meta-analysis conducted in this paper is based on the "bare-bones meta-analysis" by Schmidt/Hunter (2015). In this method, sampling errors of the individual studies under review are corrected during the analysis. Sampling errors are, in relation to the actual conditions in the total population, incorrect study results because only a certain proportion of all relevant study subjects can be observed in follow-up studies. Thus, due to the use of samples, sample-specific biases in the results eventually appear. However, since these do not affect the correlations systematically and additively, they cannot be corrected for a single study but can be for several studies. If one considers that with the inclusion of many research papers in a meta-analysis, the resulting sample under review grows and that both positive and negative sampling errors are made in the individual research papers, the sampling error tends to zero via better coverage of the entire population. Although this means that the sampling error increasingly evens out as the number of studies increases, the error still needs to be corrected because, due to squaring, only the total amount of negative and positive deviations is significant in the variance. The error must also be corrected because it is a flaw in the study that distorts the result of the meta-analysis. Schmidt/Hunter (2015) call these study deficiencies "artefacts". Since the goal of a meta-analysis is to describe the distribution of the true correlations between a given independent variable and a given dependent variable, it is necessary to remove artefacts from the studies. The sampling error is the most significant artefact. For simplicity and to avoid going beyond the scope of this paper, the "bare-bones meta-analysis" was performed, in which only the sampling error is corrected, and it is assumed that there are no other artefacts requiring correction.

The grouping of the studies for the meta-analysis is based on the classification in Table 3. Accordingly, for all of the following calculations, on the one hand, those studies are summarized that investigate either the influence of rotation or the influence of tenure; on the other hand, subgroups are formed in each of these two groups using the modified Jones model of Dechow et al. (1995) or the performance-matched-modified Jones model of Kothari et al. (2005).

Since meta-analyses, as already explained, are concerned with the description of a correlation, the first step is to determine the weighted arithmetic mean of the respective correlation coefficients between audit partner rotation or

tenure and audit quality across all studies. The weighted arithmetic mean is used because Schmidt/Hunter (2015) state that this is the best measure for estimating the correlation, assuming a constant population correlation. Even with cross-study variances, it still provides the most accurate results.

4.2 Empirical results

The results of the meta-analysis are shown in Table 4 below.

Table 4: Results of the meta-analysis

Variable	Effect of audit partner rotation			Effect of audit partner tenure		
	Overall	Dechow et al. (1995) model	Kothari et al. (2005) model	Overall	Dechow et al. (1995) model	Kothari et al. (2005) model
$\bar{r}$	-0,0025	-0,0105	0,0015	-0,0940	0,0051	-0,0239
σ_r^2	0,0005	0,0006	0,0004	0,0125	0,0044	0,0002
σ_e^2	0,0002	0,0009	0,0001	0,0002	0,0015	0,0001
σ_p^2	0,0003	-0,0003	0,0003	0,0123	0,0029	0,0001
CI lower limit	-0,0029***	-0,0121***	0,0007**	-0,1044***	0,0015**	-0,0242**
CI upper limit	-0,0021***	-0,0089***	0,0023**	-0,0836***	0,0086	-0,0235***
X^2	19	2	13	687	14	7
K	8	3	3	13	5	4
$\bar{N}$	4.574	1.128	10.175	4.153	661	7.482

The results of the entire groups make it is noticeable that there is a negative correlation for both the influence of auditor rotation and the influence of auditor tenure. This means that, on weighted average, the included studies show that the rotation and tenure of the audit partner reduce the extent of the accounting policy and thus improve audit quality. The confidence intervals (CI) of the two groups also exclude the effect size zero with a significance level of 0.01 (99 percent certainty and one percent error probability), such that the negative correlation found is highly significant. However, the magnitude of $\bar{r}$ and the ranges of CI indicate that the overall strength of the correlation is low ($\bar{r} = -0.0025$ and -0.0940).

In the subgroups of the two analyses, which are each divided according to the Jones model used in the studies, it can be observed that the results between the subgroups do not match in either case. While the studies with the modified Jones model of Dechow et al. (1995) indicate an overall negative correlation ($\bar{r} = -0.0105$) for audit partner rotation, the studies with the performance-matched-modified Jones model of Kothari et al. (2005) indicate a positive correlation ($\bar{r} = 0.0015$). For audit partner tenure, the situation is exactly the opposite. Here, the studies with the modified Jones model of Dechow et al. (1995) indicate a positive correlation ($\bar{r} = 0.0051$), whereas the studies with the performance-matched modified Jones model of Kothari et al. (2005) provide a negative correlation ($\bar{r} = -0.0239$). If we now look at the confidence intervals, we find that the negative correlations determined are highly significant in each case (significance level = 0.01) and the positive correlations determined are significant (significance level = 0.05). Thus, for both auditor rotation and auditor tenure, the subgroups provide, on the one hand, the result that audit quality increases with a change in the audit partner and its tenure and, on the other hand, the completely opposite result that audit quality decreases with the aforementioned processes. To understand why the results of the subgroups paradoxically contradict each other and thus also partly do not correspond to the results of the overall groups, a closer look must be taken at the compositions of the analyses.

Table 5: Results of meta-analysis on the influence of audit partner rotation

Results of meta-analysis on the influence of audit partner rotation					
#	Study	N	r	99% CI	
				lower limit	upper limit
1	Mohapatra et al. (2021)	1.694	0,0139	-0,0487	0,0764
2	Kalanjati et al. (2019)	688	-0,0290	-0,1267	0,0693
3	Martani et al. (2021)	1.001	-0,0390	-0,1200	0,0425
	meta-analysis "Dechow"	1.128	-0,0105	-0,0121	-0,0089
4	Litt et al. (2014)	2.108	0,0770	0,0210	0,1325
5	Kim und Xi (2021)	13.856	-0,0020	-0,0239	0,0199
6	Chang et al. (2018)	14.561	-0,0061	-0,0274	0,0152
	meta-analysis "Kothari"	10.175	0,0015	0,0007	0,0023
7	Corbella et al. (2015)	1.583	-0,0327	-0,0972	0,0321
8	Horton et al. (2021)	1.100	-0,0450	-0,1222	0,0327
	meta-analysis overall	4.574	-0,0025	-0,0029	-0,0021

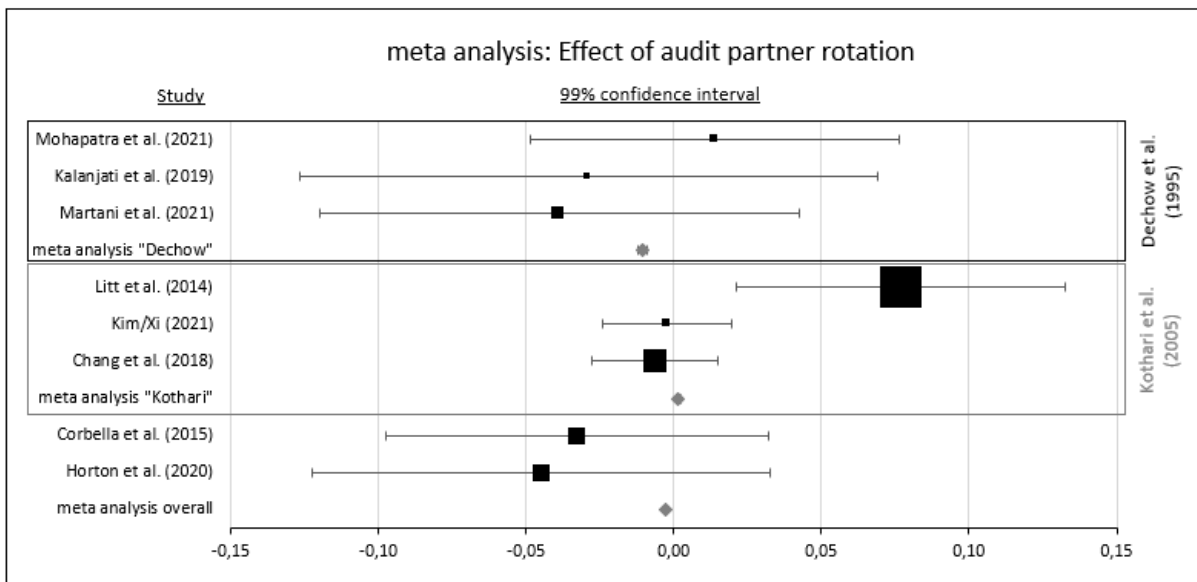


Abb. 6: Forest plot: Results of meta-analysis on the influence of audit partner rotation

Table 5 shows the studies that make up the meta-analyses conducted on the influence of auditor rotation on audit quality. Mohapatra et al. (2021), Kalanjati et al. (2019), and Martani et al. (2021) use the modified Jones model of Dechow et al. (1995) and form the subgroup for the "Dechow" meta-analysis. On the other hand, Litt et al. (2014), Kim/Xi (2021), and Chang et al. (2018) use the performance-matched modified Jones model based on Kothari et al. (2005) in their work; therefore, their results are included in the "Kothari" meta-analysis. The remaining two studies by Corbella et al. (2015) and Horton et al. (2021) use other models to determine the extent of the accounting policy. However, all studies listed in Table 5 are considered in the overall meta-analysis on the impact of internal rotations of the responsible audit partner on audit quality. The individual sample sizes, correlations and 99 percent confidence intervals for the respective studies and the meta-analyses conducted are also provided in Table 5. In Figure 7, the correlations and the confidence intervals are shown graphically as a forest plot. When looking at the table and the forest plot, it is particularly noticeable that the study by Litt et al. (2014) has a particularly strong influence on the meta-analysis. The reason for this is that it has the largest effect size of the entire group, with a calculated correlation of 0.0770, and with a sample of 2,108 observations, it is also one of the largest studies in terms of numbers. The extremely strong influence of Litt et al. (2014) causes the detected positive correlation in the meta-analysis of the Kothari subgroup in the field of audit partner rotation, which differs from the results of the whole group and the Dechow subgroup. As seen in the forest plot, the result of Litt et al. (2014) deviates strikingly

far from the results of the other studies. Due to the comparatively large divergence, Litt et al. (2014) can be seen as an outlier that has a distorting effect on the results of the meta-analysis in the area of audit partner rotation. Similar findings can be seen in the meta-analysis on the influence of audit partner tenure.

Table 6: Results of meta-analysis on the influence of audit partner tenure

Results of meta-analysis on the influence of audit partner tenure					
#	Study	N	r	99% CI	
				lower limit	upper limit
1	Buntara und Adhariani (2019)	295	0,1286	-0,0214	0,2730
2	Baatwah (2016)	508	-0,0100	-0,1240	0,1043
3	Kalanjati et al. (2019)	688	-0,0220	-0,1198	0,0763
4	Martani et al. (2021)	1.001	0,0640	-0,0175	0,1446
5	Garcia-Blandon et al. (2020)	813	-0,0800	-0,2693	0,1152
	meta-analysis "Dechow"	661	0,0051	0,0015	0,0086
6	Chen et al. (2008)	5.213	-0,0007	-0,0364	0,0350
7	Liu et al. (2021)	8.903	-0,0195	-0,0468	0,0078
8	Singh et al. (2019)	1.250	-0,0800	-0,1519	-0,0072
9	Chang et al. (2018)	14.561	-0,0300	-0,0925	0,0328
	meta-analysis "Kothari"	7.482	-0,0239	-0,0242	-0,0235
10	Manry et al. (2008)	90	-0,2420	-0,4801	0,0293
11	Ball et al. (2015)	266	0,0630	-0,1561	0,2762
12	Gipper et al. (2021)	17.903	-0,2484	-0,2802	-0,2160
13	Fargher et al. (2008)	2.495	0,0299	-0,1015	0,1602
	meta-analysis overall	4.153	-0,0940	-0,1044	-0,0836

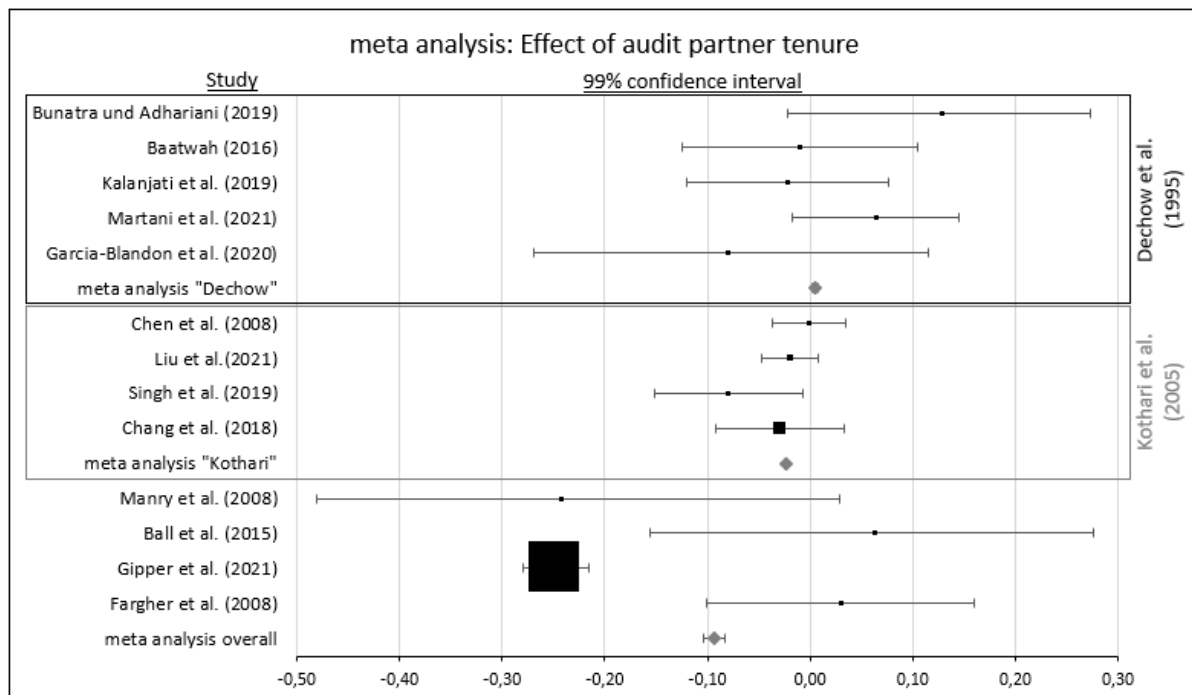


Abb. 7: Forest plot: Results of meta-analysis on the influence of audit partner tenure

Table 6 shows the composition of the meta-analyses on the influence of audit partner tenure on audit quality and indicates the values for the sample sizes, the correlations, and the 99 percent confidence intervals of the individual studies. The groupings made can be explained as in Table 5. Figure 8 presents this again in the form of a forest plot.

The positive correlation observed in the meta-analysis of the Dechow subgroup can be attributed primarily to the relatively strong positive correlation of Buntara/Adhariani (2019) ($r = 0.1286$). However, the positive correlation discovered in Martani et al. (2021) ($r = 0.0640$) and the comparatively large sample ($N = 1,001$) also lead to the deviating result of the Dechow subgroup. Here, the result of Buntara/Adhariani (2019) ($r = 0.1286$) could be understood as an outlier. However, the weight of the study within the meta-analysis is only small due to the relatively small sample size ($N = 295$). The influence of Gipper et al. (2021) on the meta-analysis is clearly more striking when looking at the forest plot in Figure 8. Gipper et al. (2021) has both the highest effect size ($r = -0.2484$) and the largest sample ($N = 17,903$), giving it a particularly high weighting in the meta-analysis. Due to this effect, Gipper et al. (2021) can be identified as an outlier in the meta-analysis on the influence of audit partner tenure, as is Litt et al. (2014) in the meta-analysis on the influence of audit partner rotation. Judging by the effect size of the correlation, Manry et al. (2008) ($r = -0.2420$) could also be considered an outlier. However, as in Buntara/Adhariani (2019), the extremely small sample size ($N = 90$) makes the study's impact on the meta-analysis relatively small.

If the results of Litt et al. (2014) and Gipper et al. (2021) are assumed to have a confounding effect on the results of the meta-analysis, it is reasonable to remove them from the meta-analysis to assess the studies in isolation from the identified outliers.

Table 7: Results of meta-analysis (adjusted for outliers)

Results of meta-analysis (adjusted for outliers)			
Variable	Audit partner rotation excluding Litt et al. (2014)		Audit partner tenure excluding Gipper et al. (2021)
	Overall	Kothari	Overall
$\bar{r}$	-0,0073	-0,0041	-0,0174
σ_r^2	0,0002	0,0000	0,0069
σ_e^2	0,0002	0,0001	0,0003
σ_p^2	0,0000	0,0000	0,0066
CI lower limit	-0,0074***	-0,0057***	-0,0233***
CI upper limit	-0,0073***	-0,0025***	-0,0115***
X^2	6	1	249
K	7	2	12
N	4.926	14.209	3.007

Table 7 shows the results of the meta-analysis excluding Litt et al. (2014) and Gipper et al. (2021). Because they are included only in the analyses on the influence of audit partner rotation (total and Kothari) and the overall analysis on the influence of audit tenure, their exclusion does not change the remaining results. After excluding Litt et al. (2014), the Kothari subgroup provides a highly significant negative correlation in the analysis of audit partner rotation. This also further strengthens the observed negative correlation for the overall group. Thus, a uniformly beneficial effect on audit quality can now be determined for audit partner rotation. Regarding audit tenure, the result of a highly significant negative correlation with audit quality at the overall group level remains despite the exclusion of Gipper et al. (2021). However, the effect size of the correlation decreases (from $\bar{r} = -0.0940$ to $\bar{r} = -0.0174$), and the 99 percent confidence interval shifts closer to zero.

Another variable that changes significantly as a result of the adjustment for outliers and that has not been addressed thus far is homogeneity. If we look at the chi-square values in Table 4 and compare them with the chi-square distribution at a usual significance level of 0.05, we find that, with the exception of the Dechow subgroup for audit partner rotation and the Kothari subgroup for audit tenure, all other groups show heterogeneity. This can also be seen from the forest plots in Figures 7 and 8. While the confidence intervals and the correlation coefficients for the homogeneous groups are close to each other and overlap, respectively, they show a larger dispersion for the heterogeneous groups. Due to the exclusion of Litt et al. (2014) in the meta-analysis on the influence of audit partner rotation, the chi-square value (total: from $\chi^2 = 19$ to $\chi^2 = 6$; Kothari: from $\chi^2 = 13$ to $\chi^2 = 1$) is reduced to such an

extent that the previously heterogeneous groups become homogeneous. Due to the existing or statistically not excluded homogeneity of the studies, their results may be compared with each other, indicating that the results for audit partner rotation are valid. Additionally, in the meta-analysis on the influence of audit tenure, the chi-square value is reduced considerably (from $\chi^2 = 687$ to $\chi^2 = 249$) by the exclusion of Gipper et al. (2021). However, this value still exceeds the critical value of the chi-square distribution, indicating that only the results of the Kothari subgroup are meaningful in the meta-analysis for audit partner tenure.

4.3 Discussion

As our results show, the meta-analysis indicates that the internal rotation of the responsible audit partner has a positive influence on audit quality. After a rotation has taken place, the level of accounting policy finally decreases, giving financial statement readers a better picture of the actual prevailing asset and income situation. In line with the previous thoughts on the potential impact of audit partner rotation, this finding suggests that the potential improvement in the area of independence outweighs the potential deterioration in terms of lost client-specific expertise. The "fresh look" at things brought by the newly appointed audit partner could explain the improvement in audit quality. Similarly, the new auditor might perform a more accurate audit due to inexperience, or the client might be more inclined to display high-quality accounting due to the new person in charge. Overall, based on the results of the meta-analysis, it can be said that audit partner rotation achieves the regulators' intended goal of increasing audit quality, even if the impact is only small considering its strength.

Since audit partner rotation is inextricably linked to the client-related tenure of the responsible audit partner's work or explicitly aims to limit it, it is also important to examine its influence on audit quality. Although the majority of studies in this area are not comparable due to the heterogeneity observed, the valid result of the Kothari subgroup and a slight trend across studies suggest that audit quality increases with increasing audit tenure. If this assumption were true, this influence in combination with the observed quality-increasing effect of audit partner rotation would be quite surprising. Finally, if the increase in audit quality due to rotation is attributed to an improvement in independence, independence would have to deteriorate beforehand during the previous auditor's tenure. The qualification or client-specific expertise of the audit partner is likely to play only a subordinate role; otherwise, the increase in audit quality during a change in auditor cannot be rationally explained. If both rotation and tenure had a positive influence on audit quality, it would have to increase immeasurably over time. One possible explanation for the observations made could be that the auditor actually does not change in the background during an internal rotation but continues to lead the audit of the client with his accumulated expertise and only has different audit partner signs on his behalf. However, it is questionable and unclear whether such a practice is applied regularly and across countries. Another possible explanation could be a general and fundamental improvement in audit quality over time, for example, through more extensive technical possibilities, more innovative audit methods or better training of auditors. However, due to the heterogeneity of the studies, as already mentioned, no generalized statement can be made, indicating that the tenure of the audit partner could also not have a generally valid influence on audit quality.

Regarding the heterogeneity observed, it makes sense to ask why this heterogeneity exists. Conceivably, this heterogeneity results from so-called moderator variables that, in addition to the independent variable used (the number of years the audit partner has been working for a client), also have an effect on the dependent variable and vary from study to study. Possible moderators could be, for example, region, country, or legal framework, such as litigation. In a globally conducted meta-analysis such as this one, the studies come from different countries, which have a different effect on the results of the studies depending on the conditions there. If the respective conditions are the same or at least similar, the studies are comparable and thus homogeneous. The identification of existing moderators, which could be causal for the heterogeneity, is extremely difficult in view of the small number of studies and the still young state of the research since no further subdivisions can be made, especially at the level of the subgroups, due to the lack of breadth.

If we compare the results of the meta-analysis with the correlations found in the individual studies, which are shown graphically in Figures 5 and 6, we see that the observations made are entirely consistent. Both Figure 5 and the

meta-analysis show that audit partner rotation has a negative impact on the extent of the accounting policy, even if this impact is quite small. In the case of audit partner tenure, on the other hand, a wider spread of results can be observed in both Figure 6 and the meta-analysis, indicating that the influence on audit quality is more difficult to assess. Only a slight influence on the extent of the accounting policy, which tends to be negative, emerges.

Regarding the usefulness of this work, the meta-analysis carried out provides information on the current state of research in the field of audit partner rotation in connection with audit quality. The meta-analysis provides an overview of the studies conducted thus far, their characteristics, the observations made, and the interstudy effect relationships. Thus, this work can serve researchers by collecting investigations as current status reports of the technical literature existing today. The generalization of the results also provides the first impression of the overall effect of internal rotation on audit quality. The findings of the meta-analysis also provide legislators and regulators with important empirically based information on the effects of audit partner rotation. The results can be used to inform policy decisions with relevant scientific observations. They can also be used to evaluate the changes that have been made and to inform decision making when market interventions in this area are still to come.

In addition to its utility, this work has certain limitations. One limitation is the number of studies included in the meta-analysis. The results should always be viewed with the caveat that they are based on only 21 studies (from 18 papers). Accordingly, the findings obtained here should neither be over nor underestimated. Due to the still young research field and the relatively small body of literature, the results of the meta-analysis are to be understood as the current state of the art. A further limitation arises from the nature and conduct of the meta-analysis. All relevant data were extracted manually from the studies and transferred to an Excel list; thus, there might be transformation deficiencies. The quality of financial reporting was used as a surrogate for audit quality, which was determined based on the extent of the accounting policy. The extent of the accounting policy was in turn predominantly determined by the level of DA. The statements made in this paper about audit quality are thus based only on studies of this surrogate. Consideration of different surrogates may yield different results. In addition, a "bare-bones meta-analysis" according to Schmidt/Hunter (2015) was performed to correct for sampling errors in the studies. This work does not claim to examine the studies for possible further artefacts and to correct them. This work also conducted univariate analysis to indicate the general direction of the relationship through the weighted arithmetic mean of the correlation coefficients. Further regression or structural analyses are not provided in this paper. Finally, a possible "publication bias" underlying the studies relevant to the meta-analysis is another limitation. Publication bias is the phenomenon in which research articles with certain results, such as nonsignificant associations, are published less frequently. The fact that these results are not reported means that they cannot be included in the meta-analysis. Finally, the lack of possible further research may bias the results of the meta-analysis.

5 Concluding remarks and outlook

We carried out a meta-analysis of whether audit partner rotation has an influence on audit quality and is therefore suitable as a regulatory instrument. For this purpose, the focus was on the audit partner rotation itself as well as on audit partner tenure, which was ultimately limited by the mandatory rotation.

After discussing the concept of audit quality, we used the currently most commonly used surrogate – discretionary accruals – to limit the study selection when identifying the relevant literature for conducting the meta-analysis.

We distinguished between the modified Jones model of Dechow et al. (1995) and the performance-matched modified Jones model of Kothari et al. (2005). The results presented show that in the case of audit partner rotation, a cross-study negative effect on the extent of discretionary accruals can be identified. Accordingly, the rotation of the responsible audit partner increases audit quality. It can thus be said that the objective pursued by the legislator or the market regulator with audit partner rotation seems effective. Regarding the influence of audit partner tenure on audit quality, no statement can be made for the entire group due to the heterogeneity of the studies. However, the result of the Kothari subgroup indicates a negative correlation with the extent of discretionary accruals. Thus, auditor tenure also increases audit quality, which rather argues against audit partner rotation.

Based on the described limitations, an outlook for future research on audit partner rotation can be given. Above all, this work should be a fundamental appeal to the research community to intensify its investigations into this field. Many audit markets around the world still have regulations on audit partner rotations; however, their impact on audit quality has not yet been scientifically analysed. New studies would bring new insights that could be used in a future meta-analysis to better determine the overall effect. Furthermore, heterogeneity could be investigated by identifying possible moderator variables. Given the current state of the research, in more advanced work, the meta-analysis could be modified and conducted differently. Instead of the exclusive correction of the sampling error, an additional correction of further artefacts is conceivable. Similarly, a meta-regression analysis or meta-analytic structural equation modelling can be performed. Instead of financial reporting quality or the extent of the accounting policy, it would be possible for future meta-analyses to investigate the influence of audit partner rotation on audit quality using other surrogates. Such studies could also provide information on whether the use of different surrogates leads to diverging effects.

Literaturverzeichnis

- Baatwah S R (2016) Audit Tenure and Financial Reporting in Oman: Does Rotation affect the Quality? Risk governance & control: financial markets & institutions 6:16-27. <https://doi.org/10.22495/rcgv6i3c1art2>.
- Ball F, Tyler J, Wells P (2015) Is audit quality impacted by auditor relationships? Journal of Contemporary Accounting & Economics 11:166-181. <https://doi.org/10.1016/j.jcae.2015.05.002>.
- Buntara A, Adhariani D (2019) Audit Tenure and Audit Quality: the Renewal Sense of Comfort?. Australasian Accounting, Business and Finance Journal 13:46-62. <https://doi.org/10.14453/aabfj.v13i4.4>.
- Carey P, Simnett R (2006) Audit Partner Tenure and Audit Quality. The Accounting Review 81:653-676. <https://doi.org/10.2308/accr.2006.81.3.653>
- Chang Y-S, Lin Y-J, Liu L-L (Sunny), Shiue M-J, Wheatley C M. (2018) The Use of Hierarchical Linear Modeling to Address Lack-of-Independence in Empirical Auditing Research. European Accounting Review, 27: 185-196. <https://doi.org/10.1080/09638180.2016.1248459>.
- Chen C-Y, Lin C-J, Lin Y-C (2008) Audit Partner Tenure, Audit Firm Tenure, and Discretionary Accruals: Does Long Auditor Tenure Impair Earnings Quality? Contemporary Accounting Research 25: 415-445. <https://doi.org/10.1506/car.25.2.5>.
- Corbella S, Florio C, Gotti G, Mastrolia S A (2015) Audit Firm Rotation, Audit Fees and Audit Quality: The Experience of Italian Public Companies. Journal of International Accounting, Auditing and Taxation 25:46-66. <https://doi.org/10.1016/j.intaccudtax.2015.10.003>.
- DeAngelo, L A (1981) Auditor Size and Audit Quality. Journal of Accounting and Economics 3:183-199. [https://doi.org/10.1016/0165-4101\(81\)90002-1](https://doi.org/10.1016/0165-4101(81)90002-1).
- Dechow P M, Sloan R G, Sweeney A P (1995) Detecting Earnings Management. The Accounting Review 70:193-225. <https://doi.org/10.1111/j.1475-679X.2012.00449.x>.
- Fargher N, Lee H-Y, Mande V (2008) The effect of audit partner tenure on client managers' accounting discretion. Managerial Auditing Journal 23:161-186. <https://doi.org/10.1108/02686900810839857>.
- Garcia-Blandon J, Argiles J M, Ravenda D (2020) On the Relationship between Audit Tenure and Fees Paid to the Audit Firm and Audit Quality. Accounting in Europe 17:78-103. <https://doi.org/10.1080/17449480.2019.1669808>.
- Gipper B, Hail L, Leuz C (2021) On the Economics of Mandatory Audit Partner Rotation and Tenure: Evidence from PCAOB Data. The Accounting Review 96:303-331. <https://doi.org/10.2308/tar-2018-0347>.
- Hansen C, Steinmetz H, Block J (2022) How to conduct a mega-analysis in eight steps: a practical guide. Manag Rev Q 72:1-19. <https://doi.org/10.1007/s11301-021-00247-4>.
- Healy P M, Palepu, K G (2003) The Fall of Enron. Journal of Economic Perspectives 17:3-26. <https://doi.org/10.1257/089533003765888403>.
- Horton J, Livne G, Pettiniccchio A (2021) Empirical Evidence on Audit Quality under a Dual Mandatory Auditor Rotation Rule. European Accounting Review 30:1-29. <https://doi.org/10.1080/09638180.2020.1747513>.
- Jones J J (1991) Earnings Management During Import Relief Investigations. Journal of Accounting Research 29:193-228. <https://doi.org/10.2307/2491047>.
- Kalanjati D S, Nasution D, Jonnergård K, Sutedjo S (2019) Auditor rotations and audit quality: A perspective from cumulative number of audit partner and audit firm rotations. Asian Review of Accounting 27:639-660. <https://doi.org/10.1108/ARA-10-2018-0182>.
- Kaznik R (1999) On the association between voluntary disclosure and earnings management. Journal of Accounting Research 37:57-81. <https://doi.org/10.2307/2491396>.
- Kim S H, Xi J (2021) Relative Influences of Review and Engagement Auditor Rotation on Audit Quality: Evidence from China. Asian Journal of Business and Accounting 14:171-205. <https://doi.org/10.22452/ajba.vol14no1.7>.
- Kothari S.P, Leone A J, Wasley C E (2005) Performance matched discretionary accrual measures. Journal of Accounting and Economics 39:163-197. <https://doi.org/10.1016/j.jacceco.2004.11.002>.

- Litt B, Sharma D, Simpson T, Tanyi P N (2014) Audit Partner Rotation and Financial Reporting Quality. *Auditing: A Journal of Practice & Theory* 33:59-86. <https://doi.org/10.2308/ajpt-50753>.
- Liu Q, Zhao L, Tian L, Xie J (2021) Close auditor-client relationships: adverse effects and the potential mitigating role of partner rotation. *Managerial Auditing Journal* 36:889-919. <https://doi.org/10.1108/MAJ-07-2020-2770>.
- Manry D L, Mock T J, Turner J L (2008) Does Increased Audit Partner Tenure Reduce Audit Quality? *Journal of Accounting, Auditing & Finance* 23:553-572. <https://doi.org/10.1177/0148558X0802300406>.
- Martani D, Rahmah N A, Fitriany F, Anggraita V (2021): Impact of audit tenure and audit rotation on the audit quality: Big 4 vs non big 4. *Cogent Economics & Finance* 9. <https://doi.org/10.1080/23322039.2021.1901395>.
- Mohapatra P, Dayanandan A, Kuntluru S, A A (2021) Audit partner rotation, and its impact on audit quality: Evidence from India. *Cogent Economics & Finance* 9. <https://doi.org/10.1080/23322039.2021.1938379>.
- Schmidt F L, Hunter J E (2015) *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Sage Publications. <https://dx.doi.org/10.4135/9781483398105>.
- Singh A, Singh H, Sultana N, Evans J (2019) Independent and joint effects of audit partner tenure and non-audit fees on audit quality. *Journal of Contemporary Accounting and Economics* 15:186-205. <https://doi.org/10.1016/j.jcae.2019.04.005>.
- Widmann M, Wolz M (2019) Prüfungsqualität im Spannungsfeld von Joint Audits und externer Rotation. *Zeitschrift für Corporate Governance* 5:222-228. <https://doi.org/y 10.37307/j.1868-7792.2019.05.08>.
- Widmann M, Follert F, Wolz, M (2021a) On the political decision of audit market regulation: Empirical evidence of audit firm tenure and maximum durations within the European Union. *Economies* 9:79. <https://doi.org/10.3390/economies9020079>.
- Widmann M, Follert F, Wolz, M (2021b) What is it going to cost? Empirical evidence from a systematic literature review of audit fee determinants. *Manag Rev Q* 71:455-489. <https://doi.org/10.1007/s11301-020-00190-w>.

3

Risk Management



Corporate Social Responsibility: Determinanten, Motivation und Implementierung

Extended Abstract

Nicola Alves-Illi

Hochschule Luzern, Institut für Finanzdienstleistungen Zug IFZ, Rotkreuz, E-Mail: nicola.alves-illi@hslu.ch

Abstract

Corporate Social Responsibility (CSR) hat in den letzten Jahren in der Wissenschaft, Wirtschaft und Politik an Aufmerksamkeit gewonnen (Behringer et al., 2024; Galletta et al., 2022; Martiny et al., 2024; Schweizerische Eidgenossenschaft, 2022; Stüttgen & Mattmann, 2023). Trotz wachsendem Interesse für das Thema besteht über viele Aspekte des komplexen und multidimensionalen Konzepts der CSR noch kein Konsens. Diese Aspekte reichen beispielsweise von der Definition über die Determinanten von und die Motivation für CSR bis hin zu den Implementierungsansätzen. Die vorgeschlagene empirische Forschung soll einen Beitrag zur bestehenden Literatur über die Determinanten, die Motivation und die Umsetzung von CSR leisten. *Forschungsprojekt 1* zielt darauf ab, die Determinanten von CSR im Bankensektor zu ermitteln. Die Motivation für CSR auf der Grundlage der Legitimitäts- und Signaling-Theorie und die Nutzung von CSR als Signal an die Stakeholder im Schweizer Markt stehen im Fokus von *Forschungsprojekt 2*. In *Forschungsprojekt 3* sollen die CSR-Implementierungsansätze von Schweizer KMU untersucht werden.

1 Einleitung

Corporate Social Responsibility (CSR) hat in den letzten Jahren in der Wissenschaft, Wirtschaft und Politik an Aufmerksamkeit gewonnen (Behringer et al., 2024; Galletta et al., 2022; Martiny et al., 2024; Schweizerische Eidgenossenschaft, 2022; Stüttgen & Mattmann, 2023). Trotz wachsendem Interesse für das Thema besteht über viele Aspekte des komplexen und multidimensionalen Konzepts der CSR noch kein Konsens. Diese Aspekte reichen beispielsweise von der Definition über die Determinanten von und die Motivation für CSR bis hin zu den Implementierungsansätzen. Die vorgeschlagene empirische Forschung soll einen Beitrag zur bestehenden Literatur über die Determinanten, die Motivation und die Umsetzung von CSR leisten.

2 Forschungsprojekt 1 - Determinanten von Corporate Social Responsibility (CSR)

In *Forschungsprojekt 1* sollen die Determinanten von CSR untersucht werden, wobei der Fokus auf dem Bankensektor liegt. Unternehmensspezifische Determinanten von Corporate Social Performance (CSP) wie Unternehmensgrösse (Al Mamun et al., 2013; Cai et al., 2016; Cassely et al., 2021; Chih et al., 2010; Crespi & Migliavacca, 2020; Drempetic et al., 2020; Khaled et al., 2021; Scholtens, 2009; Semenescu & Curmei, 2015), finanzielle Performance (Cai et al., 2016; Crespi & Migliavacca, 2020; Khaled et al., 2021), Leverage (Cai et al., 2016; Khaled et al., 2021; Reverte, 2009), Forschung und Entwicklung (Cassely et al., 2021; Fu, 2023) oder Eigentumsstruktur (Labelle et al., 2018; Rees & Rodionova, 2015; W. Wang et al., 2023; Y. Wang et al., 2023) wurden in vergangenen Forschungsarbeiten identifiziert. Auch der Einfluss länderspezifischer Determinanten wie die wirtschaftliche Entwicklung (Cai et al., 2016; Ho et al., 2012), Institutionen (Cai et al., 2016; Venugopal et al., 2023), Kultur (Ho et al., 2012; Ioannou & Serafeim, 2012), Governance-Landschaft (Mooneeapen et al., 2022) und politische Massnahmen (Chih et al., 2010; Xue et al., 2023; D. Zhang et al., 2023) wurden aufgezeigt. Bei den genannten Determinanten besteht jedoch wenig Konsens über die Richtung oder gar das Vorhandensein einer signifikanten Wirkung. Darüber hinaus gibt es nach wie vor nur wenige Forschungsarbeiten, die einen umfassenden Blick auf die Determinanten von CSR werfen. Trotz der Bedeutung des Bankensektors für die Wirtschaft aufgrund seiner Rolle als Finanzintermediär und Vermittler von Investitionen und Transaktionen ist die Literaturlandschaft, welche sich mit den Determinanten von CSR im Bankensektor befasst, karg. Die wenigen Studien, welche sich auf die Determinanten von CSR in der Finanzindustrie konzentrieren (siehe Bătae et al., 2020; Birindelli et al., 2018; Chih et al., 2010; Crespi & Migliavacca, 2020; El Khoury et al., 2023), deuten auf Forschungslücken hin. Um diese Forschungslücken zu schliessen, zielt die vorgeschlagene Untersuchung darauf ab, die unternehmens-, branchen-, länderspezifischen und makroökonomischen Determinanten sowie die zeitlichen Auswirkungen von CSR im Bankensektor anhand einer quantitativen Analyse von Sekundärdaten zu untersuchen.

3 Forschungsprojekt 2 - Motivation von CSR

Forschungsprojekt 2 untersucht die Motivation für CSR-Bemühungen von Unternehmen und insbesondere den Einsatz von CSR-Kommunikation als Signal an Stakeholder bei börsennotierten Schweizer Unternehmen. Die Legitimitätstheorie und die Signaling-Theorie sind zwei Theorien, welche auf das Konzept der CSR angewendet werden können und die unterschiedlichen Ausprägungen von CSR zwischen Unternehmen, Branchen und Ländern sowie die Motivation für CSR erklären. Laut der Legitimitätstheorie wird CSR als Mittel eingesetzt, um die Vereinbarkeit eines Unternehmens mit den Erwartungen der Stakeholder zu kommunizieren (Branco & Rodrigues, 2008b) und somit als Legitimationsinstrument genutzt. Die Ergebnisse vergangener Studien sprechen für die Legitimitätstheorie in der Anwendung auf das Konzept der CSR (Branco & Rodrigues, 2008b; Drempetic et al., 2020; Illi, 2019; Mooneeapen et al., 2022; Reverte, 2009; Sánchez-Torné et al., 2020; Schaltegger & Hörisch, 2017). In ähnlicher Weise legt die Signaling-Theorie nahe, dass die Kommunikation von CSR durch Unternehmen ein Signal darstellt, welches ihr Engagement für CSR zeigt und dadurch Reputationsrisiken verringert (Conte et al., 2023). CSR kann Stakeholdern die nicht-beobachtbaren Eigenschaften eines Unternehmens signalisieren (Bhattacharya et al., 2021; Marano et al., 2017; Semenescu & Curmei, 2015; D. Wang et al., 2023; Zerbini, 2017; T. Zhang et al., 2022). Die beiden Theorien schliessen sich nicht gegenseitig aus, sondern ergänzen sich vielmehr in der Anwendung auf das

Konzept der CSR. Aufgrund der derzeit noch relativ lockeren Regulierung der CSR-Berichterstattungspflichten für Unternehmen, bietet der Schweizer Markt die Möglichkeit, CSR-Signale durch die freiwillige Offenlegung zu erfassen. Mit einem ähnlichen Ansatz wie Branco & Rodrigues (2008a) und Baldini et al. (2018) soll das *Forschungsprojekt 2* die Auswirkung der Sichtbarkeit eines Unternehmens auf die Offenlegung von CSR untersuchen, allerdings mit einem Fokus auf börsennotierte Schweizer Unternehmen, für die eine Forschungslücke identifiziert wurde. Dabei geht die Untersuchung der Frage nach, ob Unternehmen, die sichtbarer sind, ein grösseres Interesse daran haben, ihre CSR-Bemühungen zu signalisieren. Gleichzeitig ergibt sich auch das Ziel, die Legitimitätstheorie und Signaling-Theorie in Bezug auf CSR zu bestätigen oder zu verwerfen. Die quantitative Analyse soll anhand von Sekundärdaten die Auswirkung von der Sichtbarkeit von Unternehmen auf deren CSR-Kommunikation untersuchen.

4 Forschungsprojekt 3 – CSR-Implementierungsansätze in KMU

Forschungsprojekt 3 zielt darauf ab, die verschiedenen Arten von CSR-Ansätzen zu untersuchen und Unterschiede und potenzielle Muster bei der Implementierung verschiedener CSR-Praktiken in Schweizer KMU zu ermitteln. Der Grossteil der Forschung im Bereich CSR konzentriert sich auf Grossunternehmen. Verschiedene Autoren argumentieren jedoch, dass KMU aufgrund ihrer einzigartigen Merkmale im Zusammenhang mit CSR nicht wie Grossunternehmen behandelt werden können (Baumann-Pauly et al., 2013; Coppa & Sriramesh, 2013; Garrido-Ruso et al., 2024; Gholami et al., 2022; Metzker & Zvarikova, 2021; Rózsa et al., 2022; Vo, 2011). Dies rechtfertigt die weitere, speziell auf KMU ausgerichtete Untersuchung. Insbesondere Oduro et al. (2024) fordern in einer aktuellen Literaturübersicht über CSR in KMU die Erforschung von CSR-Praktiken in europäischen KMU. Die Ergebnisse vergangener Studien haben ergeben, dass KMU in der Tendenz und im Vergleich zu Grossunternehmen proaktive (Looser & Wehrmeyer, 2015), informelle (Coppa & Sriramesh, 2013) und intern-ausgerichtete (Baumann-Pauly et al., 2013; Coppa & Sriramesh, 2013) CSR-Ansätze anwenden. Trotz der Bedeutung der KMU für die Schweizer Wirtschaft gibt es nach wie vor wenig Forschung zur CSR von Schweizer KMU (mit Ausnahme von Baumann-Pauly et al., 2013). Darüber hinaus kann argumentiert werden, dass KMU in Bezug auf ihr Verständnis von und ihre Herangehensweise an CSR als heterogene Gruppe betrachtet werden sollten (Aragón et al., 2016; Déniz & Suárez, 2005; Johnson & Schaltegger, 2016; Vo, 2011). Konkret wird eine Forschungslücke in der differenzierten Untersuchung von CSR-Ansätzen und der Umsetzung entsprechender CSR-Praktiken bei Schweizer KMU identifiziert, wobei die Heterogenität von KMU zu berücksichtigen ist. Das Ziel des Projekts ist es, die Unterschiede in den CSR-Ansätzen und Umsetzung von CSR-Praktiken in Schweizer KMU zu untersuchen. Die qualitative Analyse soll auf der Basis semi-strukturierter Interviews mit Schweizer KMU durchgeführt werden.

Literaturverzeichnis

- Al Mamun, Md., Sohag, K., & Akhter, A. (2013). A dynamic panel analysis of the financial determinants of CSR in Bangladeshi banking industry. *Asian Economic and Financial Review*, 3(5), 560–578.
- Aragón, C., Narvaiza, L., & Altuna, M. (2016). Why and how does social responsibility differ among SMEs? A social capital systemic approach. *Journal of Business Ethics*, 138(2), 365–384. <https://doi.org/10.1007/s10551-015-2632-2>
- Baldini, M., Dal Maso, L., Liberatore, G., Mazzi, F., & Terzani, S. (2018). Role of country- and firm-level determinants in environmental, social, and governance disclosure. *Journal of Business Ethics*, 150(1), 79–98. <https://doi.org/10.1007/s10551-016-3139-1>
- Bătae, O. M., Dragomir, V. D., & Feleagă, L. (2020). Environmental, social, governance (ESG), and financial performance of European banks. *Journal of Accounting and Management Information Systems*, 19(3), 480–501. <https://doi.org/10.24818/jamis.2020.03003>
- Baumann-Pauly, D., Wickert, C., Spence, L. J., & Scherer, A. G. (2013). Organizing Corporate Social Responsibility in small and large firms: Size matters. *Journal of Business Ethics*, 115(4), 693–705. <https://doi.org/10.1007/s10551-013-1827-7>
- Behringer, S., Laun, U., & Miesner, J. (2024). *ESG Reporting Survey 2024*. Institut für Finanzdienstleistungen Zug IFZ.
- Bhattacharya, A., Good, V., Sardashti, H., & Peloza, J. (2021). Beyond warm glow: The risk-mitigating effect of Corporate Social Responsibility (CSR). *Journal of Business Ethics*, 171(2), 317–336. <https://doi.org/10.1007/s10551-020-04445-0>
- Birindelli, G., Dell’Atti, S., Iannuzzi, A. P., & Savioli, M. (2018). Composition and activity of the board of directors: Impact on ESG performance in the banking system. *Sustainability*, 10(12), 4699. <https://doi.org/10.3390/su10124699>
- Branco, M. C., & Rodrigues, L. L. (2008a). Factors influencing social responsibility disclosure by Portuguese companies. *Journal of Business Ethics*, 83(4), 685–701. <https://doi.org/10.1007/s10551-007-9658-z>
- Branco, M. C., & Rodrigues, L. L. (2008b). Social responsibility disclosure: A study of proxies for the public visibility of Portuguese banks. *The British Accounting Review*, 40(2), 161–181. <https://doi.org/10.1016/j.bar.2008.02.004>
- Cai, Y., Pan, C. H., & Statman, M. (2016). Why do countries matter so much in corporate social performance? *Journal of Corporate Finance*, 41, 591–609. <https://doi.org/10.1016/j.jcorpfin.2016.09.004>
- Cassely, L., Ben Larbi, S., Revelli, C., & Lacroux, A. (2021). Corporate social performance (CSP) in time of economic crisis. *Sustainability Accounting, Management and Policy Journal*, 12(5), 913–942. <https://doi.org/10.1108/SAMPJ-07-2020-0262>
- Chih, H.-L., Chih, H.-H., & Chen, T.-Y. (2010). On the determinants of Corporate Social Responsibility. International evidence on the financial industry. *Journal of Business Ethics*, 93(1), 115–135. <https://doi.org/10.1007/s10551-009-0186-x>
- Conte, F., Sardanelli, D., Voller, A., & Siano, A. (2023). CSR signaling in controversial and noncontroversial industries: CSR policies, governance structures, and transparency tools. *European Management Journal*, 41(2), 274–281. <https://doi.org/10.1016/j.emj.2021.12.003>
- Coppa, M., & Sriramesh, K. (2013). Corporate social responsibility among SMEs in Italy. *Public Relations Review*, 39(1), 30–39. <https://doi.org/10.1016/j.pubrev.2012.09.009>
- Crespi, F., & Migliavacca, M. (2020). The determinants of ESG rating in the financial industry: The same old story or a different tale? *Sustainability*, 12(16), 6398–6418. <https://doi.org/10.3390/su12166398>
- Déniz, M. D. L. C. D., & Suárez, M. K. C. (2005). Corporate Social Responsibility and family business in Spain. *Journal of Business Ethics*, 56(1), 27–41. <https://doi.org/10.1007/s10551-004-3237-3>

- Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, 167(2), 333–360. <https://doi.org/10.1007/s10551-019-04164-1>
- El Khoury, R., Nasrallah, N., & Alareeni, B. (2023). The determinants of ESG in the banking sector of MENA region: A trend or necessity? *Competitiveness Review: An International Business Journal*, 33(1), 7–29. <https://doi.org/10.1108/CR-09-2021-0118>
- Fu, L. (2023). Why bad news can be good news: The signaling feedback effect of negative media coverage of corporate irresponsibility. *Organization & Environment*, 36(1), 98–125. <https://doi.org/10.1177/10860266221108704>
- Galletta, S., Mazzù, S., & Naciti, V. (2022). A bibliometric analysis of ESG performance in the banking industry: From the current status to future directions. *Research in International Business and Finance*, 62, 101684. <https://doi.org/10.1016/j.ribaf.2022.101684>
- Garrido-Ruso, M., Otero-González, L., López-Penabad, M., & Santomil, P. D. (2024). Does ESG implementation influence performance and risk in SMEs? *Corporate Social Responsibility and Environmental Management*, 1–21. <https://doi.org/10.1002/csr.2783>
- Gholami, A., Murray, P. A., & Sands, J. (2022). Environmental, social, governance & financial performance disclosure for large firms: Is this different for SME firms? *Sustainability*, 14(10), 6019. <https://doi.org/10.3390/su14106019>
- Ho, F. N., Wang, H.-M. D., & Vitell, S. J. (2012). A global analysis of corporate social performance: The effects of cultural and geographic environments. *Journal of Business Ethics*, 107(4), 423–433. <https://doi.org/10.1007/s10551-011-1047-y>
- Illi, N. (2019). Determinants of CSR and ESG engagement. *Master Thesis*.
- Ioannou, I., & Serafeim, G. (2012). What drives corporate social performance? The role of nation-level institutions. *Journal of International Business Studies*, 43(9), 834–864. <https://doi.org/10.1057/jibs.2012.26>
- Johnson, M. P., & Schaltegger, S. (2016). Two decades of sustainability management tools for SMEs: How far have we come? *Journal of Small Business Management*, 54(2), 481–505. <https://doi.org/10.1111/jsbm.12154>
- Khaled, R., Ali, H., & Mohamed, E. K. A. (2021). The Sustainable Development Goals and corporate sustainability performance: Mapping, extent and determinants. *Journal of Cleaner Production*, 311, 127599. <https://doi.org/10.1016/j.jclepro.2021.127599>
- Labelle, R., Hafsi, T., Francoeur, C., & Ben Amar, W. (2018). Family firms' corporate social performance: A calculated quest for socioemotional wealth. *Journal of Business Ethics*, 148(3), 511–525. <https://doi.org/10.1007/s10551-015-2982-9>
- Looser, S., & Wehrmeyer, W. (2015). An emerging template of CSR in Switzerland. *Corporate Ownership and Control*, 12(3), 541–560. <https://doi.org/10.22495/cocv12i3c5p6>
- Marano, V., Tashman, P., & Kostova, T. (2017). Escaping the iron cage: Liabilities of origin and CSR reporting of emerging market multinational enterprises. *Journal of International Business Studies*, 48(3), 386–408. <https://doi.org/10.1057/jibs.2016.17>
- Martiny, A., Tagliatalata, J., Testa, F., & Iraldo, F. (2024). Determinants of environmental social and governance (ESG) performance: A systematic literature review. *Journal of Cleaner Production*, 456, 142213. <https://doi.org/10.1016/j.jclepro.2024.142213>
- Metzker, Z., & Zvarikova, K. (2021). The perception of company employees by SMEs with CSR concept implementation. *International Journal of Entrepreneurial Knowledge*, 9(1), 81–96. <https://doi.org/10.37335/ijek.v9i1.128>
- Mooneepen, O., Abhayawansa, S., & Mamode Khan, N. (2022). The influence of the country governance environment on corporate environmental, social and governance (ESG) performance. *Sustainability Accounting, Management and Policy Journal*, 13(4), 953–985. <https://doi.org/10.1108/SAMPJ-07-2021-0298>

- Oduro, S., Bruno, L., & Maccario, G. (2024). Corporate social responsibility (CSR) in SMEs: What we know, what we don't know, and what we should know. *Journal of Small Business & Entrepreneurship*, 36(2), 207–238. <https://doi.org/10.1080/08276331.2021.1951064>
- Rees, W., & Rodionova, T. (2015). The influence of family ownership on Corporate Social Responsibility: An international analysis of publicly listed companies. *Corporate Governance: An International Review*, 23(3), 184–202. <https://doi.org/10.1111/corg.12086>
- Reverte, C. (2009). Determinants of Corporate Social Responsibility disclosure ratings by Spanish listed firms. *Journal of Business Ethics*, 88(2), 351–366. <https://doi.org/10.1007/s10551-008-9968-9>
- Rózsa, Z., Holúbek, J., Veselá, Z., & Soboleva, O. (2022). Antecedents and barriers which drive SMEs in relation to corporate social responsibility? Literature review. *International Journal of Entrepreneurial Knowledge*, 10(2), 107–122. <https://doi.org/10.37335/ijek.v10i2.174>
- Sánchez-Torné, I., Morán-Álvarez, J. C., & Pérez-López, J. A. (2020). The importance of corporate social responsibility in achieving high corporate reputation. *Corporate Social Responsibility and Environmental Management*, 27(6), 2692–2700. <https://doi.org/10.1002/csr.1993>
- Schaltegger, S., & Hörisch, J. (2017). In search of the dominant rationale in sustainability management: Legitimacy- or profit-seeking? *Journal of Business Ethics*, 145(2), 259–276. <https://doi.org/10.1007/s10551-015-2854-3>
- Scholtens, B. (2009). Corporate Social Responsibility in the international banking industry. *Journal of Business Ethics*, 86(2), 159–175. <https://doi.org/10.1007/s10551-008-9841-x>
- Schweizerische Eidgenossenschaft. (2022). *Prüfungsbericht des EJPD vom 23. Februar 2022. Analyse Vorschläge EU-Richtlinien über Sorgfaltspflichten und Berichterstattung zur Nachhaltigkeit und möglicher Anpassungsbedarf im Schweizer Recht. Bericht Entwürfe Nachhaltigkeitspflichten EU und geltendes Recht Schweiz*. Eidgenössisches Justiz- und Polizeidepartement EJPD. <https://www.news.admin.ch/news/message/attachments/81650.pdf>
- Semenescu, A., & Curmei, C. V. (2015). Using CSR to mitigate information asymmetry in the banking sector. *Management & Marketing*, 10(4), 316–329. <https://doi.org/10.1515/mmcks-2015-0021>
- Stüttgen, M., & Mattmann, B. (2023). *IFZ Sustainable Investments Studie 2023*. Institut für Finanzdienstleistungen Zug IFZ.
- Venugopal, A., Al-Shammari, M., Thotapalli, P. M. V., & Fuad, M. (2023). Liabilities of origin and its influence on firm's corporate social performance? A study of emerging market multinational corporations. *Thunderbird International Business Review*, 65(5), 519–531. <https://doi.org/10.1002/tie.22354>
- Vo, L. C. (2011). Corporate social responsibility and SMEs: A literature review and agenda for future research. *Problems and Perspectives in Management*, 9(4), 89–97.
- Wang, D., Wei, J., Noorderhaven, N., & Liu, Y. (2023). Signaling effects of CSR performance on cross-border alliance formation. *Journal of Business Ethics*, 186(4), 831–850. <https://doi.org/10.1007/s10551-023-05432-x>
- Wang, W., Sun, Z., Wang, W., Hua, Q., & Wu, F. (2023). The impact of environmental uncertainty on ESG performance: Emotional vs. rational. *Journal of Cleaner Production*, 397, 136528. <https://doi.org/10.1016/j.jclepro.2023.136528>
- Wang, Y., Lin, Y., Fu, X., & Chen, S. (2023). Institutional ownership heterogeneity and ESG performance: Evidence from China. *Finance Research Letters*, 51, 103448. <https://doi.org/10.1016/j.frl.2022.103448>
- Xue, Q., Wang, H., & Bai, C. (2023). Local green finance policies and corporate ESG performance. *International Review of Finance*, 23(4), 721–749. <https://doi.org/10.1111/irfi.12417>
- Zerbini, F. (2017). CSR initiatives as market signals: A review and research agenda. *Journal of Business Ethics*, 146(1), 1–23. <https://doi.org/10.1007/s10551-015-2922-8>
- Zhang, D., Meng, L., & Zhang, J. (2023). Environmental subsidy disruption, skill premiums and ESG performance. *International Review of Financial Analysis*, 90, 102862. <https://doi.org/10.1016/j.irfa.2023.102862>

Zhang, T., Zhang, Z., & Yang, J. (2022). When does Corporate Social Responsibility backfire in acquisitions? Signal incongruence and acquirer returns. *Journal of Business Ethics*, 175(1), 45–58. <https://doi.org/10.1007/s10551-020-04583-5>

Ganzheitliches Risikomanagement und Nachhaltigkeit in universitären Forschungsprojekten: Ein integraler Ansatz zur Maximierung des gesellschaftlichen Mehrwerts.

Extended Abstract

Christina Angela Groß, M.Eng.

Universität der Bundeswehr München, Institut für Energie- und Antriebstechnik, Neubiberg, E-Mail: christina.gross@unibw.de

Christian Georg Zimmermann, M.Eng.

Universität der Bundeswehr München, Institut für Projektmanagement und Bauwirtschaft, Neubiberg, E-Mail: christian.zimmermann@unibw.de

Prof. Dr.-Ing. Christian Trapp

Universität der Bundeswehr München, Institut für Energie- und Antriebstechnik, Neubiberg, E-Mail: christian.trapp@unibw.de

Univ.-Prof. Dipl.-Ing. Dr. techn. Philip Sander

Universität der Bundeswehr München, Institut für Projektmanagement und Bauwirtschaft, Neubiberg, E-Mail: philip.sander@unibw.de

Abstract

Im Kontext der sich dynamisch verändernden gesetzlichen und marktrelevanten Anforderungen an die Ergebnisse in Projekten, steht das Thema Nachhaltigkeit im Fokus. Besonders in universitären Forschungsprojekten, welche primär Innovationen erschaffen wollen, liegt ein Fokus auf der Generierung eines Mehrwerts für die Gesellschaft. Eine Herausforderung ist es sicherzustellen, dass das Management dazu beiträgt, agil zu handeln, um die Anforderungen an die Ergebnisse unter Einhaltung von Zeit und Budget zu berücksichtigen. Um diesem Problem zu begegnen, wurden gesellschaftliche Risiken zusätzlich zu den technischen Risiken in einem Modell vereint, welches auch die Illusion von vollständiger Berechenbarkeit im Projektmanagement hinterfragt und methodisch damit umgeht. Diese ganzheitliche Methode ermöglicht eine effektive Steuerung der Forschungsaktivitäten und maximiert den Forschungsoutput.

1 Einleitung

Auch universitäre Forschungsprojekte wollen Innovationen erschaffen, um den Herausforderungen der heutigen Gesellschaft zu begegnen. Daher wird es immer wichtiger, auch in Projekten eine angepasste Vorgehensweise zu integrieren, um die Erfolgchancen von Projekten zu erhöhen. Das Projektmanagement hat zum Ziel, Zeit, Kosten und Leistung in der Balance zu halten, da die Änderung eines Wertes Auswirkungen auf die anderen hat (Wagner, 2010). Bewiesen ist, dass das Projektmanagement dazu beitragen kann, bis zu 20 % Zeit und Kosten einzusparen (Braehmer, 2005, S. 9). Zudem kann ein effektives Projektmanagement als einer der Erfolgsfaktoren für universitäre Forschungsprojekte angesehen werden. Zudem kann ein effektives Projektmanagement als einer der Erfolgsfaktoren für universitäre Forschungsprojekte angesehen werden, um gutes Projektmonitoring, -planung und eine effektive Kommunikation sicherzustellen (Barnes et al., 2002).

1.1 Problemstellung

Da die Forschung vor allem durch technische Komplexität und fachliche Spezifikation geprägt ist, spielen Unsicherheiten durch schnelle Anforderungsänderungen eine wichtige Rolle. Diese Anforderungsänderungen entstehen vor allem durch unerwartete und paradoxe Ergebnisse der Forschung. Betrachtet man nun das Projektmanagement dieser Projekte, ist der richtige Umgang mit Risiken und Unerwartetem besonders relevant. In universitären Forschungsprojekten besteht eine Dualität zwischen der Förderung von Innovationen und dem Erreichen der angestrebten Ergebnisse. Einerseits müssen Risiken eingegangen werden, um Innovationen zu ermöglichen; andererseits müssen diese Risiken reduziert werden, um die gewünschten Ergebnisse innerhalb des vorgegebenen Zeit- und Budgetrahmens zu erreichen (Ernø-Kjølhede, 2000).

1.2 Relevanz des gesellschaftlichen Nutzens in der Forschung

Ein universitäres Forschungsprojekt zeichnet sich, bezogen auf das Projektmanagement in der Frage des gesellschaftlichen Nutzens, durch Einzigartigkeit aus. Es besteht oftmals keine unmittelbare Kundenbeziehung, aus welcher detaillierte Ziele und Anforderungen hervorgehen. Ziele universitärer Forschungsprojekte können aus Werten oder Normen, resultieren. Ein Beispiel wäre das UN-Ziel 7 – bezahlbare und saubere Energie (Vereinte Nationen, 2024). Dieses Ziel soll den Zugang zu bezahlbarer und sauberer Energie sicherstellen. Des Weiteren, ist in den vergangenen Jahrzehnten eine engere Kopplung zwischen Wissenschaft und Gesellschaft beobachtet worden (Metschl, 2016, S. 52). Auf Basis dieses Kontextes stellt sich folgende Forschungsfrage: Wie lässt sich der gesellschaftliche Mehrwert in universitären Forschungsprojekten optimal integrieren und maximieren, während zugleich die Herausforderung der vollständigen Berechenbarkeit besteht?

2 Theoretische Fundierung

2.1 Methodik

Anhand des exemplarischen Projekts „Munich Mobility Research Campus“ an der Universität der Bundeswehr in München wurde eine Einzelfallstudie nach Yin durchgeführt (Thahabi, 2010, S. 111–115). Basierend auf einer vorangegangenen Literaturrecherche, Interviews und vorhandenen Dokumenten wurde ein House of Quality aufgestellt, um technische Zielwerte und kundenspezifische Bedürfnisse abzubilden. Gemeinsam mit den Stakeholdern wurden die Kennzahlen und Gewichtungen bestimmt. Das beispielhafte Programm RIAAT dient zur Visualisierung der Ergebnisse (RiskConsult GmbH, 2024). In dieser Abhandlung wird der Fokus auf den Umgang mit den gesellschaftlichen Risiken gelegt.

2.2 Gesellschaftliche Risiken identifizieren und bewerten

Um frühzeitig gesellschaftliche Risiken zu erkennen, ist es wichtig, diese in die Zielsetzung zu integrieren. Eine frühzeitige Involvierung möglicher Stakeholder und Zielgruppen ist erwünscht (Wissenschaftsrat, 2021). Der Wissens- und Technologietransfer kann das Risiko für den gesellschaftlichen Nutzen identifizieren, indem neue Perspektiven in die Annahmen inkludiert werden. Dies ist beispielsweise durch Kooperationen mit Industriepartnern

und Gründungsunterstützungen möglich (Etzkowitz & Leydesdorff, 2000). Zudem kann die Kommunikation und Einbindung der Studierenden in die Lehre frühzeitig praxisrelevante Fähigkeiten vermitteln und die Risiken zur Vorbeientwicklung senken. Bewertet werden kann der gesellschaftliche Nutzen durch qualitative oder auch quantitative Methoden. Zur Bewertung eignen sich beispielsweise Kennzahlen, Interviews mit Fokusgruppen, Erfahrungsberichte oder auch die partizipative Einbindung aller Stakeholder (van Besselaar et al., 2018).

2.3 Methodenauswahl

Romeikes Einschätzung zufolge, welche Methoden nach ihrem Einsatz für verschiedene Risikoarten klassifiziert, liefern sowohl die stochastische als auch die deterministische Risikoanalyse die besten Ergebnisse (Romeike, 2018, S. 61–62). Es wurde eine fünfstufige Skala zur Bewertung verwendet. Die Methoden-Matrix bezieht sich auf strategische, finanzwirtschaftliche, politische, rechtliche, Corporate Government, sowie Organisations-, Leistungs- und Reputationsrisiken. Eine deterministische Szenarioanalyse besteht darin, dass sie nur einen Wert zur Bestimmung für Planungen verwendet. Durch die einwertige Planung kann es zu Scheingenauigkeiten kommen (Gleißner & Kalwait, 2017). Die stochastische Szenarioanalyse zeichnet sich dadurch aus, dass mithilfe unterschiedlicher Werte, Bandbreiten für die Planung berücksichtigt werden. Es können Asymmetrien mit verschiedenen Verteilungsfunktionen dargestellt werden und auch zusätzliche Gewichte. Durch Verteilungsfunktionen kann die Realität besser beschrieben werden als durch einen einzigen deterministischen Wert (Sander, 2023). Dies wiederum ermöglicht auch eine Vorhersage über die vorzeitige oder verspätete Zielerreichung. Je früher Verzögerungen sowohl in zeitlicher als auch in finanzieller Hinsicht erkannt werden, desto schneller können Gegenmaßnahmen ergriffen werden, um den gesellschaftlichen Nutzen zu maximieren.

Durch die hohen Risiken, welche universitäre Forschungsprojekte für die Zielerreichung mit sich bringen, wurde sich zur genaueren Vorhersage für eine stochastische Szenarioanalyse entschieden. Daher wird eine 3-Punkt-Schätzung mithilfe einer minimalen, erwarteten und maximalen Wertung verwendet. Zur Aggregation der Risiken und zur Erkennung von Abhängigkeiten werden die Verteilungsfunktionen mit einer Monte-Carlo-Simulation simuliert (Gleißner & Wolfrum, 2019). Durch die Kombination von Risiken ist es möglich bestandsgefährdende Entwicklungen frühzeitig zu erkennen (Gleißner, 2016).

3 Ergebnisse

Mithilfe der Verteilungsdichten, welche durch die Monte-Carlo-Simulation erzeugt werden, können nun die aggregierten Auswirkungen der Risiken und Unsicherheiten auf Zeit und Kosten dargestellt werden. Durch das Beispiel des Meilensteins „Verbrennungsmotor“, ist es möglich, die Risiken und Unsicherheiten auszuwerten. Anschließend wird die Unterschreitungswahrscheinlichkeit des geplanten Zieldatums betrachtet. In diesem Beispiel werden Budget und Kosten nicht grafisch dargestellt; stattdessen liegt der Fokus auf der zeitlichen Planung.

Nach Integration aller relevanten Risiken zur Erstellung eines Prototyps eines innovativen Verbrennungsmotors ist ersichtlich, dass das Zieldatum zu 55 % eingehalten werden kann. Das technische Risiko hat indirekten Einfluss auf die für die Zielgruppe der Gesellschaft notwendigen Zielsetzungen. Dies bedeutet, dass in diesem Fall sowohl die Testung alternativer Kraftstoffe als auch das Verbrauchsverhalten beeinflusst werden können. Daher besteht ein erhöhtes Risiko, dass diese Ziele nicht erreicht werden, da eine fehlerhafte Applizierung des Steuergeräts möglicherweise zu einem erhöhten Kraftstoffverbrauch durch falsche Ansteuerung führen kann. Ein solcher Fehler würde einen negativen Einfluss auf die angestrebte energieeffiziente Nutzung haben. Zudem wären auch Motorschäden durch nicht konkrete Regelung möglich. Dies könnte Testungen von alternativen Kraftstoffen verzögern, was sich wiederum negativ auf das Projektende auswirken könnte oder den gesellschaftlichen Mehrwert reduzieren würde.

Daher wurde eine Alternativstrategie entwickelt, welche das vorhandene Risiko in der Eintrittswahrscheinlichkeit um 30 % senkt. Die Lösung ist entweder eine alternative Software zu verwenden oder neue Komponenten zu beschaffen. Durch Anpassung des technischen Risikos, welches auch einen Einfluss auf das gesellschaftliche Risiko hat, kann nun eine prognostizierte Unterschreitungswahrscheinlichkeit von 63 % erreicht werden. Der Vergleich zwischen Ziel- und Plandatum vor und nach Berücksichtigung der gesellschaftlichen Risiken und der damit verbundenen Alternativstrategien kann in Abbildung 1 gesehen werden.

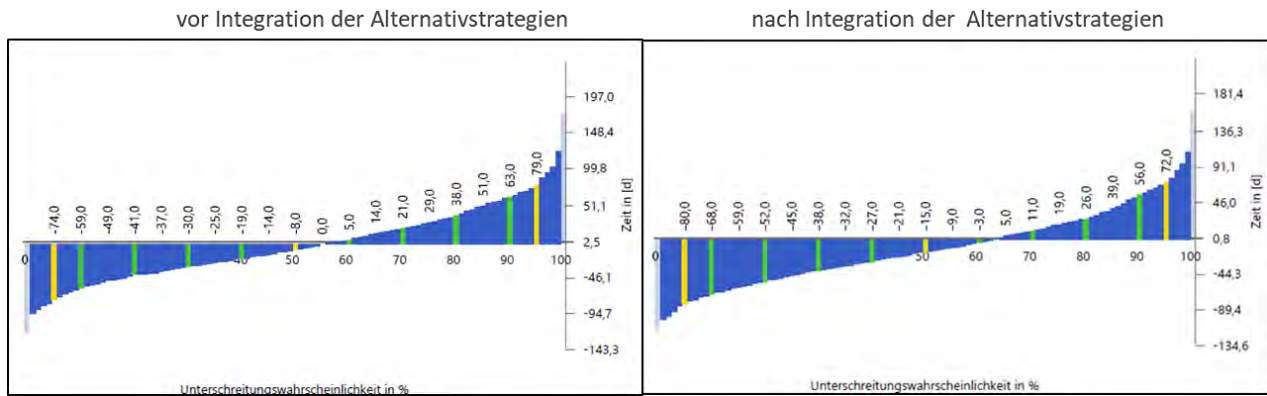


Abbildung 1: Auswirkung auf den Terminplan

4 Zusammenfassung und Ausblick

Durch die Berücksichtigung gesellschaftlicher Risiken wird zum einen der Fokus auf nachhaltige Entwicklung gefördert und zum anderen sichergestellt, dass der gesellschaftliche Nutzen auch bei Anforderungsänderungen während des laufenden Projekts nicht außer Acht gelassen wird. Dieser Impuls bietet die Möglichkeit wie im Beispiel dargelegt, universitäre Forschungsprojekte in ihrer Nachhaltigkeit zu fördern und ermöglicht eine Maximierung des Forschungsausgangs.

Dieses Forschungsthema zeigt zudem auf, wie der Illusion der Berechenbarkeit im Projektmanagement begegnet und Ressourcen unter sich ändernden Rahmenbedingungen optimal genutzt werden können. Als nächsten Schritt werden die gesellschaftlichen Risiken in ein integrales Projektmanagementkonzept integriert und getestet.

Danksagung

Diese Forschung ist Teil des Projekts MORE - Munich Mobility Research Campus. Das Projekt wird gefördert durch dtec.bw, finanziert durch die Europäische Union - NextGenerationEU.

Literaturverzeichnis

- Barnes, T., Pashby, I. & Gibbons, A. (2002). Effective University – Industry Interaction. *European Management Journal*, 20(3), 272–285. [https://doi.org/10.1016/S0263-2373\(02\)00044-0](https://doi.org/10.1016/S0263-2373(02)00044-0)
- Braehmer, U. (2005). *Projektmanagement für kleine und mittlere Unternehmen: Schnelle Resultate mit knappen Ressourcen*. Hanser Verlag, München, Wien.
- Ernø-Kjølhed, E. (2000). *Project management theory and the management of research projects*. Department of Management, Politics and Philosophy. <https://research-api.cbs.dk/ws/portalfiles/portal/58880359/6308.pdf>. Abgerufen am 21.05.2024.
- Etzkowitz, H. & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations. *Research Policy*, 29(2), 109–123. [https://doi.org/10.1016/S0048-7333\(99\)00055-4](https://doi.org/10.1016/S0048-7333(99)00055-4)
- Gleißner, W. (2016). *Bandbreitenplanung, Planungssicherheit und Monte-Carlo-Simulation mehrerer Planjahre: Risikoaggregation – auch über die Zeit*. Controllern Magazin
- Gleißner, W. & Kalwait, R. (2017). Einführung: Controlling und Risikomanagement. In: Gleißner, W., Klein, A. (Hrsg.) *Risikomanagement und Controlling: Chancen und Risiken erfassen, bewerten und in die Entscheidungsfindung integrieren* (2. Auflage). Haufe Gruppe, Freiburg, München, Stuttgart. <https://www.haufe.de/>
- Gleißner, W. & Wolfrum, M. (2019). *Risikoaggregation und Monte-Carlo-Simulation*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-24274-9>
- Metschl, U. (2016). *Vom Wert der Wissenschaft und vom Nutzen der Forschung: Zur gesellschaftlichen Rolle akademischer Wissenschaft*. Springer VS, Wiesbaden. <http://www.springer.com/>
- RiskConsult GmbH. (2024). *RIAAT know your risk*. <https://www.riskcon.at/software/riaat>
- Romeike, F. (2018). *Risikomanagement*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-13952-0>
- Sander, P. (2023). *Risikomanagement in Bauprojekten*. <https://drive.google.com/file/d/1L7qKTJbT9fP5m2eG0Kgce0-8qelShMLW/view>. Abgerufen am 22.05.2024.
- Thahabi, E. M. (2010). *Die Bildung von strategischen Geschäften in mittleren, international tätigen Unternehmen*. Gabler Verlag, Wiesbaden.
- van Besselaar, P. den, Flecha, R. & Radauer, A. (2018). *Monitoring the impact of EU Framework Programmes: Expert report*. Publications Office of the European Union, Luxembourg.
- Vereinte Nationen. (2024). *Bezahlbare und saubere Energie*. <https://unric.org/de/17ziele/sdg-7/>. Abgerufen am 21.05.2024
- Wagner, S. (2010). *Kennzahlen für das Messeprojektmanagement: Durch Leistungstransparenz erfolgreich am Markt der Messedienstleister*. Diplomica Verlag, Hamburg.
- Wissenschaftsrat. (2021). *Wissenschaftskommunikation*. https://www.wissenschaftsrat.de/download/2021/9367-21.pdf?__blob=publicationFile&v=10. Abgerufen am 12.07.2024.

Integriertes Risikomanagement für Alters- und Pflegeheime

Research Paper

Prof. Dr. Wilfried Lux

OST Otschweizer Fachhochschule, IFL Institut für Finance und Law, St.Gallen,
E-Mail: wilfried.lux@ost.ch

Sebastian Müller

OST Otschweizer Fachhochschule, IPM Institut für Informations- und Prozessmanagement, St.Gallen,
E-Mail: sebastian.mueller@ost.ch

Kevin Bieber

OST Otschweizer Fachhochschule, IPM Institut für Informations- und Prozessmanagement, St.Gallen,
E-Mail: kevin.bieber@ost.ch

Abstract

Das vorliegende Research Paper basiert auf einem abgeschlossenen Forschungsprojekt der OST Otschweizer Fachhochschule, welches von der Funk-Stiftung gefördert wurde. Die Ergebnisse bestehen zum einen in einem Leitfaden zum integrierten Risikomanagement für Alters- und Pflegeheime, das strategische, finanzielle, personelle, technische, prozessbezogene, infrastrukturelle und Reputationsrisiken beinhaltet. Ergänzt wurde der Leitfaden durch eine mobile Applikation, die die wichtigsten Inhalte des Leitfadens enthält. Die Resultate wurden neben umfangreicher Desk Research mittels qualitativer Forschung in Form von Experteninterviews, Fokusgruppen und einem Symposium erarbeitet und validiert. Das Risikomanagement in Alters- und Pflegeheimen hat viele Gemeinsamkeiten mit anderen Organisationen. Es weist aber auch eine Reihe von Besonderheiten auf. Insbesondere sind in diesem Zusammenhang zu nennen: Hauswirtschaft und Küche (z.B. Hygienevorschriften), Arzneimittelereignisse (z.B. falsche Medikamente) und Personal (v.a. physische, physikalische, chemische und psychische Belastungen der Mitarbeitenden sowie der Fachkräftemangel).

1 Einleitung

Die zunehmende Komplexität und Dynamik in der Umgebung sowie innerhalb von Alters- und Pflegeheimen führen zu einem Anstieg und einer Intensivierung von Risiken. Demografische Entwicklungen, wie die gestiegene Lebenserwartung und die damit einhergehende Multimorbidität, erhöhen den Druck auf Organisationen im Gesundheitswesen erheblich (Blüher & Kuhlmei, 2016). Hinzu kommt die wachsende Komplexität des Leistungsangebots, der medizinische Fortschritt und der sich verschärfende Mangel an qualifiziertem Personal (Fachkräftemangel (Biernoth, 2016)). Die Bevölkerung erwartet zunehmend qualitativ hochwertigere und kostengünstigere Leistungen. Auch die Sicherheitsanforderungen und die Relevanz präventiver Massnahmen nehmen zu. Beispielsweise gewann die Infektionsprävention im Rahmen der COVID-Pandemie an Brisanz, aber auch klassische Massnahmen, wie die Sturzprophylaxe, stellen die Pflegeeinrichtungen im Zuge der Risikoermittlung immer wieder vor Herausforderungen.

In der Praxis ist der Professionalisierungsgrad des Risikomanagements jedoch noch nicht ausgereift. Gerade in kleinen Alters- und Pflegeheimen fehlen oft die Ressourcen für eine ganzheitliche Auseinandersetzung mit dem Thema Risikomanagement. Es findet dann oft eine Beschränkung auf pflegerische Risiken statt (vgl. z.B. (Frodl, 2012)). Zudem gibt es keine allgemeingültigen Standards bzw. Good oder Best Practices für Alters- und Pflegeheime zum Thema ganzheitliches Risikomanagement. Regionale und nationale Unterschiede bei den Finanzierungsstrukturen erschweren solche Standards zusätzlich, da dadurch insbesondere bei betriebswirtschaftlichen Risiken Unterschiede bestehen können. So werden in der DACH-Region (Deutschland, Österreich, Schweiz) die Erträge von Pflegeheimen durch sogenannte „Pflegestufen“ bzw. „Pflegegrade“ reguliert. Bewohnende werden einer bestimmten „Pflegestufe“ (Schweiz: 1-12; Österreich: 1-7) bzw. einem „Pflegegrad“ (Deutschland: 1-5) zugeordnet, die jeweils einer spezifischen Tagespauschale entsprechen.

Eine betriebswirtschaftliche Führung ist daher unerlässlich. Diese beinhaltet ein Chancen- und Risikokalkül, das stets in Relation zu den strategischen Zielen und den unternehmerischen Entscheidungen steht. Infolgedessen hat auch das Risikomanagement in Alters- und Pflegeheimen in jüngerer Zeit an Bedeutung gewonnen. Dieses sollte sich jedoch nicht im Detail mit einer Vielzahl von Risiken befassen, sondern sich auf die zehn bis zwanzig bedeutendsten Risiken – die sogenannten „Schlüsselrisiken“ – konzentrieren. Die übrigen Risiken werden in einer Liste geführt, die regelmässig (z.B. einmal jährlich) auf ihre Aktualität hin überprüft wird. Bei vielen Risiken besteht das Potential, dass sich die Eintrittswahrscheinlichkeit, das Schadensausmass und die zu treffenden Massnahmen verändern. Ein integriertes Risikomanagement für Alters- und Pflegeheime umfasst daher die Bereiche Pflege, Medizin, Wirtschaft sowie Prozesse und Infrastruktur. Alternativ können auch die Kategorien Bewohnende, Angehörige, Personal und Organisation zur Betrachtung der Risiken genutzt werden.

Diese Aspekte sind essenziell für ein ganzheitliches Verständnis und eine effektive Steuerung der Risiken, die in Alters- und Pflegeheimen auftreten können.

2 Stand der Forschung

Im Bereich Risikomanagement gibt es weit verbreitete und generell akzeptierte Grundlagen, welche branchenunabhängig angewendet werden. Beispielsweise gibt die Norm ISO 31000 branchenunabhängige Leitlinien für das Risikomanagement in Organisationen vor. Sie beschreiben das Vorgehen im Risikomanagement generisch, um eine Anpassung auf unterschiedliche Unternehmenskontexte zu ermöglichen. Zusätzlich wird beschrieben, wie konkrete Aufgaben im Risikomanagement durchgeführt werden können. So beinhaltet die Norm unter anderem Informationen, wie Unternehmen eine Risikobewertung durchführen können. Neben ISO 31000 als allgemeinen Standard, gibt es auch noch das COSO ERM-Rahmenwerk, welches seinen Ursprung in der Finanzberichterstattung hat und ebenfalls verbreitet ist. Es beschreibt wie der ISO-Standard, wie Unternehmen Risiken managen können. Neben den Grundlagen, welche sich nur auf das Risikomanagement beziehen, gibt es auch andere Fachbereiche, in welchen das Risikomanagement einen hohen Stellenwert hat. Im Projektmanagement, welches sich mit der Planung und Durchführung von Projekten beschäftigt, ist das Risikomanagement integraler Bestandteil. Deshalb finden sich in

der Fachliteratur zum Projektmanagement oftmals Informationen zum Umgang mit Risiken. In der Schweiz ist unter anderem Hermes 5.1 verbreitet und im deutschsprachigen Raum sind unter anderem die Standards IPMA und PMI verbreitet. Die Grundlagen des Risikomanagements beziehen sich auf die organisatorische Einordnung und Abgrenzung zu anderen Bereichen (wie zum Beispiel das Qualitätsmanagement und Krisenmanagement), den Risikomanagement-Prozess (Identifikation, Analyse, Bewertung, Bewältigung und Überwachung) sowie auf Methoden der Risikobeurteilung (vgl. dazu Brühwiler (2016), Gleißner (2022), Diederichs (2023), Vanini (2012)).

Der Fokus beim zugrundeliegenden Forschungsprojekt lag speziell auf der Anwendung von Risikomanagement in Alters- und Pflegeheimen, welche aber oft von den oben ausgeführten Grundlagen beeinflusst werden. Es gibt relativ wenig Literatur zum Risikomanagement in dieser Branche. Zu nennen sind in diesem Zusammenhang die Werke von Kahla-Witzsch und Platzer (2018) sowie Paula (2006).

Aktuelle wissenschaftliche Forschung im Bereich Risikomanagement in Alters- und Pflegeheimen fokussiert oft auf eine einzelne eingeführte Massnahme und untersucht deren Wirkung mit einem quantitativen Ansatz. Dabei wird häufig eine Qualitätssteigerung im medizinischen oder pflegerischen Bereich angestrebt. Exemplarisch aufgeführt werden die Studien von Konnyu et al. (2023), welche sich auf Diabetes konzentrieren, jene von Scopetti et al. (2021) mit Fokus auf Präventions- und Kontrollmassnahmen in Verbindung mit COVID-19 und die Übersichtsarbeit von Toles et al. (2021), welche sich im Allgemeinen mit qualitätssteigernden Massnahmen in Alters- und Pflegeheimen beschäftigt, allerdings mit dem Ziel der Verbesserung der Resultate der pflegerischen Behandlungen.

Für verwandte Bereiche, insbesondere für Spitäler, lässt sich Literatur finden, welche das Thema ganzheitliches Risikomanagement gut für Praktiker aufgearbeitet. Kahla-Witzsch (2023) zeigt in ihrem Praxisleitfaden auf, wie organisationales Risikomanagement in Gesundheitseinrichtungen allgemein eingeführt werden kann. In der Monographie von Hunziker et al. (2022) wird ein in der Praxis validierter Leitfaden zur Umsetzung eines ganzheitlichen Risikomanagements in Spitälern und Krankenhäusern zur Verfügung gestellt. Im Sammelband von Euteneier (2015) werden die wesentlichen Elemente des klinischen Risikomanagement (medizinisch, managementbezogen, ökonomisch und juristisch) umfassend und aus verschiedenen Blickwinkeln dargestellt. Auch werden praktische Lösungen und notwendige Werkzeuge (z.B. Checklisten, Critical Incident Reporting System (CIRS; Berichtserstattungssysteme zur Meldung von kritischen Ereignissen oder Beinahe-Schäden), Ursachenanalysen, Methoden der Risikofassung und -bewertung) vorgestellt. Weiterhin existiert auch für Spitäler viel Literatur zum Thema Risikomanagement, welche auf den klinischen Teil fokussiert und die Patientensicherheit in den Vordergrund stellt. Als Beispiel dient der Sammelband von Zapp (2020), in welchem das Critical Incident Reporting System und das Global Trigger Tool (GTT; retrospektive Screening-Methode, bei der Patientenakten auf Trigger beziehungsweise aufgetretene Patientenschäden hin untersucht werden) analysiert und verglichen werden. Auch hier liegt der Schwerpunkt auf der Anwendungsorientierung, und es wird mit Praxisbeispielen gearbeitet. Für Spitäler lässt sich viel Literatur finden, in welcher Risikomanagement Praktiken für einzelne Bereiche (z.B. Anästhesie (Speer et al., 2023)) oder noch kleinteiligere Einheiten (z.B. Gasabschaltung im High-Care-Bereich (Heller et al., 2023)) detailliert beschrieben werden.

Praxisleitfäden für Risikomanagement in Alters- und Pflegeheimen existieren für Teilbereiche. So behandelt die Monographie von Saßen (2019) die Einführung von Fehler-Ereignis-Meldesystemen (Critical Incident Reporting System) in Alters- und Pflegeheimen. Die einzelnen Autoren im Sammelband von Blonski (2014) – Welters, Kämmer, Borutta, Offermann, Hanke, Reiner, Baars, Elser und Althammer - beschreiben unter anderem die Bereiche gesetzliche Anforderungen, pflegerisches Risikomanagement, die Arzneimittelversorgung, Hauswirtschaft und Küche, Arbeitssicherheit und Gesundheitsschutz, Brandschutz sowie Datenschutz und Datensicherheit.

Die vorliegende Studie füllt hier insofern eine Forschungslücke und ist innovativ, als sie das Thema ganzheitlich angeht. Der innovative Charakter liegt nicht in den Grundlagen des Risikomanagement, sondern in der konkreten Anwendung in der Praxis und der Sammlung von Anforderungen und Praxiswissen von Personen aus Pflegeeinrichtungen. Es werden betriebswirtschaftliche, technische, prozessuale sowie pflegerische Aspekte berücksichtigt

und in einem Leitfaden integriert. Auf diese Weise werden unnötige Schnittstellen vermieden und ein interdisziplinärer Ansatz verfolgt.

3 Methodik

Im Rahmen des zugrundeliegenden Forschungsprojektes¹ und der daraus entstandenen Studie wurde ein qualitativer Forschungsansatz verfolgt, um tiefgehende Einblicke in die Ausgestaltung des Risikomanagements in Alters- und Pflegeheimen zu gewinnen. Während es bei quantitativer Forschung möglich ist, eine grosse Stichprobengrösse zu befragen und allgemeingültige Aussagen zu treffen, erlaubt die qualitative Forschung eine tiefere Untersuchung spezifischer Themen, zum Beispiel durch Interviews mit ausgewählten Experten.

Die nachfolgende Grafik (siehe Abbildung 1) zeigt einen Überblick der Methodik des Forschungsprojekts auf und ermöglicht es, einen Überblick zu erhalten. In der Abbildung ist ersichtlich, dass sich das Forschungsteam der Thematik zu Beginn durch eine umfangreiche Sekundärrecherche genähert hat. Basierend auf den Resultaten wurden Interviews geführt und ein Leitfaden entwickelt, welcher in den nachfolgenden Schritten überprüft und verbessert wurde. Im Rahmen der Primärforschung wurden mehrere Methoden eingesetzt, namentlich Experteninterviews, einem Symposium und abschliessenden Fokusgruppen.

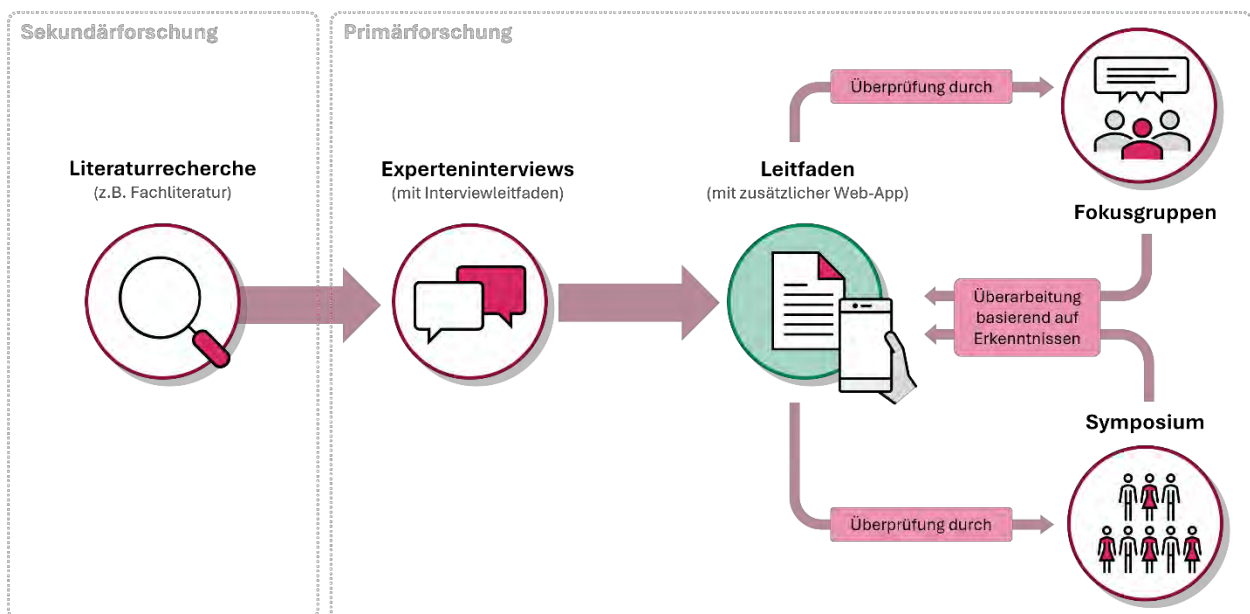


Abbildung 1: Übersicht Methodik des Forschungsprojekts (eigene Darstellung)

Die nachfolgenden Unterkapitel dienen dazu, Einblicke in die Vorgehensweise der einzelnen Schritte zu erhalten.

3.1 Sekundärrecherche

In einem ersten Schritt wurden umfangreiche Literatur- und Internetrecherchen durchgeführt, um den aktuellen Stand der Forschung zu identifizieren und zu beschreiben. Dabei legte das Forschungsteam den Fokus auf die Themen Risikomanagement in dieser Domäne, Gesundheitswesen, Pflegeheime sowie angrenzende Themen wie strategisches und Qualitätsmanagement.

¹ Das Forschungsprojekt «Integriertes Risikomanagement für Alters- und Pflegeheime» wurde von der Funk-Stiftung in Hamburg finanziell unterstützt.

3.2 Experteninterviews

Um ein vertieftes Verständnis der Ausgestaltung des Risikomanagements in den Alters- und Pflegeheimen zu erhalten, wurden im Anschluss an die Sekundärrecherche Experteninterviews durchgeführt. Ziel der Interviews war es, Einblicke in die bestehenden Strategien und Herausforderungen im Risikomanagement dieser Institutionen zu gewinnen und bestehende Erkenntnisse aus der Sekundärrecherche zu validieren.

Der Interviewleitfaden wurde auf Basis einer Literaturrecherche unter Einbezug von Expertenfeedback entwickelt. Er umfasste Fragen zu den Hauptthemen Risk Governance, Risikomanagement-Prozess und zu speziellen Themen des Risikomanagements (Notfall- und Krisenmanagement, Informationssicherheit und Datenschutz sowie Zutrittschutz und Bewachung).

Die Experten wurden basierend auf ihrer Position in der Heimleitung oder der Verbände ausgewählt. Gesucht wurden Personen, welche aktuell für die Themen Risikomanagement oder Qualität verantwortlich waren. Die Auswahl erfolgte durch Kontakte des Instituts für Pflegewissenschaften der OST - Ostschweizer Fachhochschule. Insgesamt wurden sieben Experteninterviews durchgeführt (fünf einzelne Alters- und Pflegeheime, zwei Verbände). Die Interviews waren halbstrukturiert, was es ermöglichte, auf spezifische Antworten einzugehen und tiefergehende Informationen zu sammeln (Bryman, 2016). Die Interviews dauerten jeweils zwischen 58 und 110 Minuten und wurden persönlich in den jeweiligen Institutionen durchgeführt. Alle Interviews wurden mit Zustimmung der Teilnehmenden aufgezeichnet und die wichtigsten Aussagen stichwortartig festgehalten. Die Ergebnisse wurden thematisch zusammengefasst und analysiert, um Gemeinsamkeiten und Differenzen zu identifizieren.

Die teilnehmenden Alters- und Pflegeheime hatten zum Zeitpunkt der durchgeführten Interviews zwischen 45 und 230 Vollzeitstellen und betreuten zwischen 49 und 230 Bewohnende. Drei der Institutionen waren eher kleiner (< 100 Bewohnende), die anderen beiden eher grösser (> 200 Bewohnende). Die drei kleineren Institutionen sind als Verein organisiert, die beiden grösseren Heime werden von einer oder mehreren Gemeinden getragen. Vier Institutionen wurden zwischen 1970 und 1980 gegründet, eine nach 2010 (Fusion aus vier Vorgängerinstitutionen). Die den beiden Verbänden angeschlossenen Heime beschäftigen insgesamt rund 10'000 Mitarbeitende.

Basierend auf den Resultaten der Sekundärrecherche und den Experteninterviews wurde anschliessend eine erste Version eines Leitfadens erarbeitet, welcher beim Risikomanagement unterstützen soll. Dieser Leitfaden wurde sowohl als Dokument konzipiert, sowie zusätzlich als Web-App umgesetzt. Durch die Web-App sollte die Zugänglichkeit erhöht werden, indem ein niederschwelliger Zugang zu einem komplexen Thema geschaffen wird und digitale Templates (z.B. zur Bewertung der Risiken) einfacher verfügbar gemacht werden.

3.3 Symposium

Um die Ergebnisse aus der Sekundärrecherche und den Experteninterviews breit abzustützen und zu validieren, wurde ein halbtägiges Symposium in St.Gallen organisiert. Dazu wurden Heimleitungen aus der Region DACH (Deutschland, Österreich, Schweiz) eingeladen. Die Einladung erfolgte über Kontakte der beteiligten Institute und der Funk-Gruppe. Die Teilnahme war kostenlos. Neben einem an das World Café anschliessenden Input über aktuelle Herausforderungen für Pflegeeinrichtungen und einer abschliessenden Podiumsdiskussion wurden im Rahmen eines World Cafés an sechs Tischen (je 2 Tische pro Thema), moderiert durch Mitglieder des Projektteams, während 1.5 Stunden die Inputs der etwa 40 Teilnehmenden zu den folgenden Themen eingeholt:

- **Feedback zum Leitfaden:** Was ist gut? Was kann noch verbessert werden?
- **Feedback zur mobilen Applikation:** Werden die Inhalte praxistauglich dargestellt? Was kann noch verbessert werden?
- **Herausforderungen Einführung Risikomanagement:** Welches sind die Herausforderungen bei der Einführung von Risikomanagement in Pflegeheimen?

Die Antworten wurden systematisch erfasst und im Rahmen der Nachbereitung dazu genutzt, den bereits bestehenden Leitfaden weiterzuentwickeln.

3.4 Fokusgruppen

Als letzte Validierungsstufe wurden Fokusgruppen (Robson & McCartan, 2016) mit Heimleitungen aus der Region DACH gebildet. Um die Aussagekraft zu erhöhen, wurden dabei nicht die gleichen Personen wie bei den Experteninterviews ausgewählt. Die Sitzungen dauerten fünf Stunden und fanden persönlich an der OST - Ostschweizer Fachhochschule statt. Trotz krankheitsbedingter Ausfälle nahmen 3 Heimleitende an der ersten und eine Person an der zweiten Fokusgruppe teil, wobei die fehlenden Heimleitenden schriftliches Feedback einreichten. Die Diskussionen zielten darauf ab, die Praxistauglichkeit der Leitfadeneinhalte kritisch zu hinterfragen. Die Leitfadeneinhalte wurden den Teilnehmenden im Vorfeld zugestellt.

Gleich wie beim Symposium, wurden die Antworten systematisch erfasst und dazu genutzt, den bereits bestehenden Leitfaden und die dazugehörige Web-App weiterzuentwickeln.

4 Ergebnisse aus den Untersuchungen

Die wichtigsten empirischen Ergebnisse stammen aus den durchgeführten Experteninterviews und aus dem veranstalteten Symposium. Die Resultate werden in diesem Kapitel zusammengefasst und analysiert.

4.1 Experteninterviews

4.1.1 Risk Governance

In Bezug auf Risk Governance (unterteilt in Risikopolitik, normative Grundlagen, Risikokultur, Organisation des Risikomanagements und Reifegrad des Risikomanagements) fällt generell auf, dass dieser Bereich sowohl in den kleineren als auch in den grösseren Heimen wenig Beachtung erhält. In sämtlichen kleineren Heimen liegt die Organisation und Verantwortung des Risikomanagements bei der Heimleitung. In einem der grösseren Heime ist dies ebenfalls der Fall, das andere grössere Heim weist die Bearbeitung der operativen Risiken dem Qualitätsbeauftragten zu. Ein systematisches Vorgehen ist bei keinem der Heime erkennbar. Im Gespräch werden oft einzelne, meist medizinische oder pflegerische Risiken genannt, welche meist über interne Standards und Richtlinien bekannt sind (z.B. Umgang mit resistenten Bakterien). Im Rahmen des monatlichen oder jährlichen Reportings an die Gemeinden oder Vorstände wird meist über Auslastung und Personalbestand sowie damit verbundene Risiken informiert, ein Heim erwähnt dabei den Einsatz einer Risikomatrix für operative Risiken. Vereinzelt wird zu dem Thema auf externe Beratungsunternehmen zurückgegriffen.

Die beiden Verbände stellen den angeschlossenen Heimen Richtlinien und Handbücher zur Verfügung zum Thema Risikomanagement. Auch werden ERFA-Gruppen (Erfahrungsaustausch-Gruppen) organisiert. Einer der beiden Verbände bietet auch spezifische Weiterbildungsangebote an und führt alle zwei Jahre ein kriterienbasiertes Audit durch die Regionalleitung durch, in welchem der Umgang mit den Risiken angeschaut wird.

4.1.2 Risikomanagement-Prozess

Zwei der fünf Heime verfügen über keinen Risikomanagement-Prozess. Die weiteren drei setzen auf einen regelmässigen und geplanten Austausch im Vorstand oder mit externen Beratungsunternehmen. Für die Identifikation von Risiken setzen alle Heime auf Gespräche, teils informell und teils strukturiert, unter Einbezug von Verwaltungsrat, Vorstand, den internen Führungspersonen oder externen Beratungsunternehmen.

Gefragt nach den relevantesten Risiken betonen sämtliche Heime Personalmangel, eine unzureichende Belegungsquote, Engpässe bei Lieferanten sowie eine angespannte Finanzierungssituation. Kritisiert wird generell die feh-

lende Unterstützung der Politik. Auch vermehrt festgestellt werden erhöhte Ansprüche und Erwartungen von Angehörigen und der allgemeine Trend zum betreuten oder allgemein eigenbestimmteren Wohnen. Auch pflegerische Risiken (z.B. Sturz) werden wiederholt genannt.

Bei der Analyse und Bewertung von Risiken wird erwähnt, dass oft reaktiv gehandelt wird und keine strukturierte Vorgehensweise besteht. Strategische Risiken werden teilweise im Verwaltungsrat oder bei der periodischen Strategieüberarbeitung besprochen. Zur Risikobeurteilung setzen die Heime hauptsächlich auf Kreativitätstechniken (Brainstorming, World-Café) oder andere Methoden (Delphi-Technik). Auch werden die gemeldeten Fälle im Critical Incident Reporting System ausgewertet und Massnahmen abgeleitet. Die Risiken werden anschliessend den Mitarbeitenden und dem Vorstand oder dem Verwaltungsrat schriftlich oder mündlich kommuniziert. Eine systematische Risikoüberwachung, welche eine ganzheitliche Aufnahme von Risiken und eine regelmässige Neubeurteilung dieser Risiken umfasst, fehlt in allen fünf Heimen. Die Risikobewältigung erfolgt bei eher unbekannten Risiken meistens reaktiv und fallbezogen oder bei bereits bekannten Risiken proaktiv. Als Beispiele für Themen, welche aktuell bearbeitet werden, wurden die Brandbekämpfung, die Stellenplanung, Medikamentenfehler und Übergriffe genannt. Während der COVID-19 Pandemie wurden Krisenstäbe eingesetzt. Ein einheitliches Vorgehen und die Thematisierung von betriebswirtschaftlichen Risiken fehlen in sämtlichen untersuchten Heimen.

Als Stolpersteine bzw. Hinderungsgründe bei der Umsetzung eines Risikomanagements nannten die befragten Heimleitenden, dass ein generisches Risikomanagement nicht als zielführend erachtet wird, da ein fallbezogenes sinnvoller erscheint. Auch wird vom Vorstand teilweise keine Risikostrategie eingefordert. Zudem wird argumentiert, dass auch mit einem existierenden Risikomanagement nicht alle Risiken (rechtzeitig) erkannt werden. Als Beispiel dient auch hier das Pandemierisiko.

4.1.3 Spezielle Themen des Risikomanagements

Angesprochen auf das Notfall- und Krisenmanagement erwähnten sämtliche Heimleitenden, dass im Zuge der COVID-19 Pandemie Notfallpläne entwickelt wurden. Davor existierten diese überwiegend für Brandfälle und (kurzfristige) Personalausfälle. Für pflegerische Risiken existieren Richtlinien (z.B. für vermisste Personen) und Standards (z.B. für Stürze und Verletzungen von Mitarbeitenden oder Bewohnenden). In Bezug auf die Themen Informationssicherheit und Datenschutz ist ein Teil der Heime eingebunden in die entsprechenden Dienste der Gemeinden oder setzt auf die Zusammenarbeit mit weiteren Gesundheitsinstitutionen. In jenen Heimen, bei welchen dies aktuell (noch) nicht der Fall ist, wurde das Risiko erkannt, und Massnahmen werden erarbeitet. So wurde beispielsweise als Sofortmassnahme in einem der Heime ein zweites WLAN für Lernende aufgeschaltet, in einem anderen Heim wurden striktere Zugriffsberechtigungen zum Beispiel in der Buchhaltung eingeführt. Beim Zutrittsschutz und der Bewachung setzen die Heime auf Kameraüberwachung, externe Sicherheitsdienste und elektronische Schliesssysteme. Insbesondere seitens der Angehörigen hat das Sicherheitsbedürfnis zugenommen (Weglaufschutz, Ortungssysteme, Sturzerkennung etc.), während dies von den Bewohnenden, vor allem durch die Auswirkungen auf die individuelle Freiheit, nicht so hoch gewichtet wird.

Zusammenfassend fällt auf, dass die Entwicklung des Risikomanagements bei den meisten Gesprächspartnern noch nicht sehr weit fortgeschritten war, sondern vielfach intuitiv und neben der täglichen Arbeit erledigt wurde oder in das bestehende Qualitätsmanagement integriert war. Auch die Identifikation der wichtigen Risiken erfolgte meist nicht systematisch. Daher wurde die Entwicklung eines Risikokatalogs sehr begrüsst. Eingesetzte Methoden beschränkten sich vor allem auf das Brainstorming und Schadenfallanalysen (Critical Incident Reporting System). Sämtliche befragten Experten betonten die Wichtigkeit eines ganzheitlichen Risikomanagements, verfügen jedoch aktuell über wenig oder keine Ressourcen für die Umsetzung, weshalb sie sich auf die (subjektiven) Einzelrisiken beschränkten.

4.2 Symposium und World Café

4.2.1 Verbesserungen beim Leitfaden

An den beiden entsprechenden Tischen des World Café haben Heimleitende verschiedene Verbesserungsvorschläge für den im Rahmen des Projekts erarbeiteten Leitfaden formuliert. Diese Vorschläge zielen darauf ab, den Leitfaden praxisnäher und effektiver zu gestalten.

Ein zentraler Punkt war die Notwendigkeit, in den Heimen eine Kultur zu fördern, die den offenen Umgang mit Risiken und Fehlern unterstützt. Die Teilnehmenden betonten, dass eine solche Kultur es erleichtert, aus Fehlern zu lernen und proaktiv Risiken zu identifizieren und zu managen. Zudem soll eine prozessorientierte Herangehensweise die Anwendbarkeit des Leitfadens im Alltag erleichtern und die Umsetzung der Risikomanagement-Massnahmen effizienter gestalten. Die Heimleitenden sprachen sich auch für eine noch stärkere Berücksichtigung branchenspezifischer Besonderheiten aus (z.B. spezifische Liste der Risiken). Ein noch stärkerer Branchenfokus würde es ermöglichen, gezieltere und relevantere Massnahmen zu entwickeln, die den spezifischen Herausforderungen gerecht werden. Gleichzeitig muss beachtet werden, dass Hilfsmittel, welche unterschiedliche Unternehmen unterstützen sollen, nicht zu spezifisch sein können. Ansonsten wäre die Allgemeingültigkeit beeinträchtigt. Ein weiterer wichtiger Aspekt war die Forderung nach einem noch stärkeren Praxisbezug des Leitfadens. Dies könnte durch die Integration von Fallbeispielen, Good oder Best Practices und praxisnahen Anleitungen erreicht werden, die den Nutzenden konkrete Hilfestellungen bieten und den Transfer in den Arbeitsalltag erleichtern. Abschliessend wurde die Notwendigkeit betont, das Risikomanagement organisatorisch besser zu verankern. Dies beinhaltet die Implementierung klarer Strukturen und Verantwortlichkeiten innerhalb der Organisation, um sicherzustellen, dass das Risikomanagement nicht als isolierte Aufgabe, sondern als integraler Bestandteil der Organisationskultur betrachtet wird.

4.2.2 Verbesserungen bei der Web-App

An den letzten beiden Tischen des World Café wurden Verbesserungen bei der begleitenden Web-App diskutiert.

Ein zentraler Verbesserungsvorschlag war die Notwendigkeit einer stärkeren Interaktion und kürzeren, prägnanteren Texten innerhalb der App. Durch interaktive Elemente könnte die Nutzerfreundlichkeit erhöht und die Anwendung für die Nutzer ansprechender gestaltet werden. Bei den Texten wurde wiederum auf fehlende zeitliche Ressourcen hingewiesen, und darauf, dass niemand längere Texte an seinem Smartphone lesen möchte (sondern am Computer). Des Weiteren wurde eine graphische Überarbeitung der zur Verfügung gestellten Templates angeregt. Dabei ging es hauptsächlich um die Übersichtlichkeit und das Hervorheben von zentralen Elementen. Ein letzter wichtiger Punkt war die Integration einer Suchfunktion. Diese Funktion würde es den Nutzern ermöglichen, spezifische Informationen und Templates schneller und effizienter zu finden, was die Bedienbarkeit der Web-App erheblich erleichtern würde.

4.2.3 Herausforderungen beim Risikomanagement

An zwei weiteren Tischen des World Café wurden die bestehenden Herausforderungen im Risikomanagement diskutiert.

Ein wesentliches Problem stellt das Fehlen klar definierter Prozesse dar. Ohne strukturierte Abläufe ist es schwierig, Risiken systematisch zu identifizieren und zu managen, was die Effektivität des gesamten Risikomanagements erheblich beeinträchtigt. Ein weiteres Hindernis ist die schlechte Datenqualität. Ungenaue oder unvollständige Daten erschweren es, fundierte Entscheidungen zu treffen und Risiken präzise zu bewerten. Zusätzlich wurde die mangelnde Verfügbarkeit von Zeit als bedeutende Herausforderung identifiziert. Viele Mitarbeitende sind bereits stark ausgelastet und finden kaum Gelegenheit, sich intensiv mit dem Risikomanagement auseinanderzusetzen. Diese Zeitknappheit verhindert eine gründliche Auseinandersetzung mit potenziellen Risiken und die Implementierung entsprechender Massnahmen. Hier setzt das Projekt an, indem es Hilfestellung gibt, beispielsweise in Form von

Erklärungen oder Templates, welche den initialen Zeiteinsatz und den Zeiteinsatz beim fortlaufenden Risikomanagement reduziert. Zudem wurde das fehlende Risikobewusstsein, insbesondere in Bezug auf betriebswirtschaftliche Risiken, innerhalb der Organisationen angesprochen. Wenn das Bewusstsein für Risiken und deren Management nicht ausreichend entwickelt ist, wird es schwierig, eine Kultur des proaktiven Risikomanagements zu etablieren. Zu den Hauptherausforderungen gehören darüber hinaus spezifische Themen im Zusammenhang mit Bewohnenden, Mitarbeitenden und betriebswirtschaftlichen Aspekten wie Finanzen. Die Betreuung und Sicherheit der Bewohnenden erforderten ein hohes Mass an Aufmerksamkeit und Ressourcen, während die Mitarbeitenden oft unter hohen Belastungen stehen und deren Gesundheit und Zufriedenheit ebenfalls ein kritischer Faktor sind. Betriebswirtschaftliche Themen, insbesondere finanzielle Herausforderungen, können zudem die Implementierung effektiver Risikomanagementstrategien behindern, weil ungenügende Mittel für das Risikomanagement und die Umsetzung der Massnahmen vorhanden sind.

5 Leitfaden «Integriertes Risikomanagement für Alters- und Pflegeheime»

Das Projektergebnis, ein Leitfaden inklusiv begleitender Web-App und diversen Templates, soll Hilfe beim Aufbau eines effizienten und effektiven Risikomanagementsystems bieten und unterstützt als praxisnahes sowie praxiserprobtes Konzept die Zielgruppe der Risikoverantwortlichen bei ihrer Arbeit. Der Leitfaden beinhaltet sowohl methodische als auch inhaltliche Empfehlungen und Templates zum Aufbau und zur Führung eines ganzheitlichen Risikomanagementsystems in Pflegeeinrichtungen. Die mobile App enthält die wichtigsten Inhalte und Templates in komprimierter Form und unterstützt bei der Durchführung des Risikomanagements. In der Web-App können die Templates heruntergeladen und geteilt werden.

5.1 Leitfaden

5.1.1 Grundlagen

Im ersten Teil des Leitfadens werden die wesentlichen Grundlagen des Risikomanagements behandelt, welche für die Praxis in Pflegeeinrichtungen von Bedeutung sind. Dies wurde explizit von den Interviewpartnern und den Teilnehmenden des Symposiums auch so gewünscht, weil insbesondere in kleineren Einrichtungen die Grundlagen oft nicht bekannt sind.

Neben Begriffsdefinitionen und Hinweisen auf die Schnittstellen zu verwandten Themengebieten (Strategisches Management, Qualitätsmanagement, Compliance Management, Informationssicherheit, Versicherungs- und Schadenmanagement, Krisenstab und Leadership) wird insbesondere auf die in den Experteninterviews und dem Symposium angesprochenen Schwerpunkte eingegangen. So wurde auch das Thema Risiko- und Fehlerkultur aufgegriffen. Eine offene und transparente Fehlerkultur ist essenziell für ein effektives Risikomanagement. Es soll ein Umfeld geschaffen werden, in dem Fehler und Risiken offen angesprochen und analysiert werden können. Dies fördert eine lernende Organisation, in der Pflegeeinrichtungen aus Fehlern lernen und ihre Prozesse kontinuierlich verbessern, um zukünftige Risiken zu minimieren. Eine gelebte und positive Fehlerkultur stärkt das Vertrauen in die Führung, ermöglicht frühzeitige Risikoerkennung und fördert Innovation. Führungskräfte spielen hierbei eine wichtige Rolle, indem sie ehrlich sind, klare Erwartungen setzen und Vertrauen in die Mitarbeitenden zeigen. (Bestvater, 2022; Ebner, 2008)

In der Schweiz, Deutschland und Österreich wird durch gesetzliche Grundlagen vorgeschrieben, dass Pflegeeinrichtungen ein effektives Risikomanagementsystem implementieren und aufrechterhalten, dies mit dem Ziel der Qualitätssicherung in der Pflege und der Minimierung von wirtschaftlichen und personellen Risiken. Für den Leitfaden wurden die jeweils wichtigsten Punkte für die drei Länder durch je einen Juristen aufgearbeitet und für Praktikerinnen und Praktiker zusammengefasst. Dabei fallen Unterschiede zwischen den drei Ländern auf, insbesondere in der Art und Weise der Regelungen, der Verantwortlichkeiten und der Prüfungsmechanismen. Die Schweiz weist kantonale Unterschiede und eine Vielfalt an Organisationsformen auf, Deutschland legt einen starken Fokus auf regelmässige Qualitätsprüfungen und finanzielle Sanktionen bis hin zu Ordnungswidrigkeitenverfahren mit hohen

Bussgeldern, während Österreich einen freiwilligen Corporate Governance Kodex sowie spezifische Anforderungen an interne Kontrollsysteme und Prüfungsausschüsse hat.

5.1.2 Integriertes Risikomanagement: Aufbau und Prozess

Die Implementierung des Risikomanagements in Pflegeeinrichtungen gemäss des Leitfadens sollte mit der organisatorischen Verankerung beginnen. Dies beinhaltet die klare Zuordnung von Verantwortlichkeiten und die Schaffung einer geeigneten Organisationsstruktur. Besonders wichtig ist es, dass die Geschäftsführung und das obere Management aktiv eingebunden sind, da deren Unterstützung entscheidend für den Erfolg des Risikomanagements ist. In kleineren Pflegeeinrichtungen kann dies durch die Geschäftsführung selbst erfolgen, während in grösseren Einrichtungen die Schaffung einer eigenen Abteilung oder die Benennung eines Risikomanagers sinnvoll ist. Es ist wichtig, dass das Risikomanagement eng mit anderen Managementbereichen, wie dem Qualitätsmanagement und dem Compliance Management, verzahnt ist, um Synergien zu nutzen und eine umfassende Risikoüberwachung zu gewährleisten.

Der Risikomanagementprozess besteht gemäss ISO 31000 aus mehreren aufeinanderfolgenden Schritten: Risikoidentifikation, Risikoanalyse und -bewertung, Risikobewältigung sowie Risikoüberwachung und -kommunikation. Die nachfolgende Abbildung zeigt die Schritte des Risikomanagementprozesses visuell auf und wurde auch während des Forschungsprojekts für die Kommunikation mit den Anspruchsgruppen genutzt. Sie ermöglicht es aufzuzeigen, dass die Risikoüberwachung und -kommunikation nicht der letzte Schritt im Prozess ist, sondern eine Aktivität, welche den gesamten Prozess begleitet.



Abbildung 2: Generelle Vorgehensweise im Risikomanagement (eigene Darstellung, basierend auf ISO 31000)

Der erste Schritt umfasst die systematische Erkennung potenzieller Risiken, die die Ziele der Pflegeeinrichtung gefährden könnten. Als Methoden zur Risikoidentifikation werden Risikokataloge, Brainstorming, Dokumentenanalyse, World Café, Wertkettenanalyse, Prozesskettenanalyse und die Fehlerbaumanalyse genannt. Nach der Identifikation sollen die Risiken hinsichtlich ihrer Eintrittswahrscheinlichkeit und ihres Schadenspotenzials bewertet werden. Instrumente dazu sind beispielsweise Risikomatrix, Szenariotechnik und die Fehlerbaumanalyse. Die Bewertung hilft, die Risiken zu priorisieren und diejenigen zu identifizieren, die besonderer Aufmerksamkeit bedürfen. Zur Bewältigung der identifizierten und bewerteten Risiken stehen verschiedene Strategien zur Verfügung: Risikovermeidung, Risikoverminderung, Risikotransfer und Risikoakzeptanz. Die Wahl der geeigneten Strategie hängt von der Art des Risikos und den spezifischen Umständen der Pflegeeinrichtung ab. Wichtige Massnahmen können organisatorische Änderungen, technische Schutzmassnahmen oder der Abschluss von Versicherungen sein. Abschliessend ist auch die Überwachung von Restrisiken entscheidend, da sie sich auf das Unternehmensergebnis auswirken können. Eine effektive Risikoüberwachung erfordert die Verwendung geeigneter Kennzahlen und Frühwarnindikatoren. Die Risikokommunikation sollte alle Ebenen in der Organisation kontinuierlich über die Risikosituation informieren und das Risikobewusstsein fördern. Die Risikoberichterstattung umfasst Massnahmen zur Erfassung, Verarbeitung und Speicherung von risikorelevanten Informationen und richtet sich an interne und externe

Empfänger. Externe Berichte sind oft Teil des Jahresberichts und vom Umfang her prägnanter und kürzer als interne Berichte. Diese sind detaillierter und variieren je nach Zielgruppe. Eine transparente Risikoberichterstattung ist entscheidend für das Risikomanagement, damit sich alle Mitarbeitenden mit den Risiken auseinandersetzen und die Massnahmen mittragen.

5.2 Templates

Zur Unterstützung des Risikomanagementprozesses werden verschiedene Templates bereitgestellt. Die Nutzung standardisierter Tools erleichtert die Umsetzung des Risikomanagements und sorgt für Konsistenz und Vergleichbarkeit der Ergebnisse. Gemäss den Hinweisen der Heimleitenden in den Interviews und dem Symposium wurden die Templates mit praktischen Beispielen vorab ausgefüllt, an welchen man sich orientieren kann.

5.2.1 Risikomatrix / Risikokatalog

Risikokataloge (vgl. Gleißner (2022, S. 282–287) sind gebräuchliche Tools zur Bewertung und Darstellung von Risiken. Sie verwenden eine zweidimensionale Matrix, die Eintrittswahrscheinlichkeit und Schadenspotential auf den Achsen abbildet. Optional kann eine dritte Dimension hinzugefügt werden, dargestellt durch die Grösse der Kreise, welche dann beispielsweise eine Aussage über die Vorhersehbarkeit eines Risikos macht. Abbildung 3 zeigt auf, wie eine solche Matrix aussehen kann anhand von drei Beispielrisiken mit unterschiedlichen Schadenpotentialen und Eintrittswahrscheinlichkeit. Das Aussehen einer Risikomatrix kann stark variieren. Farben wie Rot, Gelb und Grün können beispielsweise genutzt werden, um aufzuzeigen, wie kritisch ein Risiko durch die Kombination von Schadenpotential und Eintrittswahrscheinlichkeit ist. Es gibt aber auch Unternehmen, welche keine Farben nutzen und Darstellungen mit weiteren Abstufungen bevorzugen.

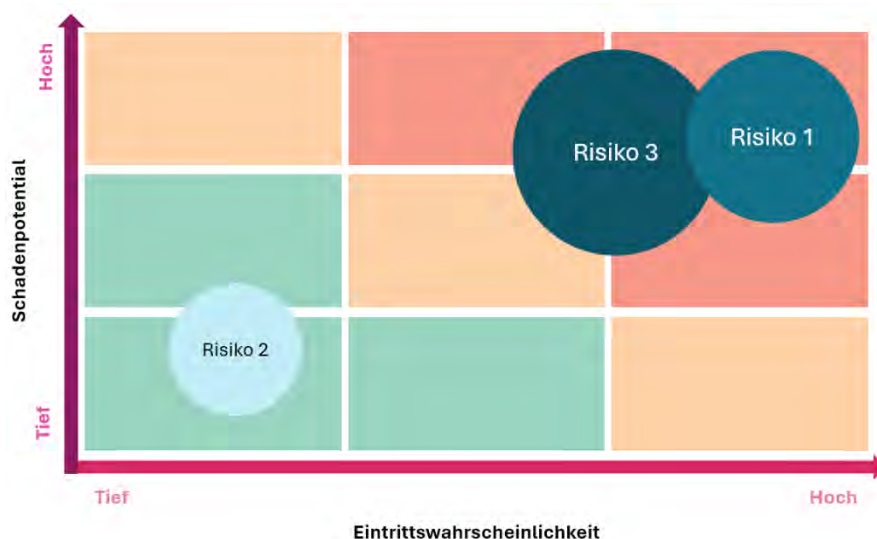


Abbildung 3: Exemplarische Risikomatrix (eigene Darstellung basierend auf (Gleißner, 2022))

Während die Risikomatrix dazu dient einen schnellen Überblick über vorhanden Risiken zu ermöglichen, finden Risikokataloge häufig in der Risikoberichterstattung Anwendung. Bei der Anwendung herausfordernd ist nicht das Eintragen der Risiken in die Matrix, sondern die Identifikation und Bewertung der Einzelrisiken. Im Projekt wurden 87 mögliche Einzelrisiken als Orientierungshilfe für Praktikerinnen und Praktiker identifiziert. Einige davon sind beispielhaft in folgender Tabelle aufgeführt:

Tabelle 1: Exemplarische Einzelrisiken

Risikokategorie Risikofeld	Einzelrisiko (Beispiele)
Strategisch	Fehlende Lern- und Entwicklungsperspektive
	Fehlende Fehlerkultur
Externe Faktoren (politisch, rechtlich-regulatorisch, technisch, ökologisch, volkswirtschaftlich)	Fehlende Attraktivität der Langzeitpflege insgesamt für Mitarbeitende
	Verschärfter Wettbewerb in der Stadt; neue private Anbieter, Ausbau ambulanter Leistungen («Länger zuhause leben»: Spitex, 24-Stunden-Betreuung, Entlastungsangebote)
	Neue Wohnformen (betreutes Wohnen, Alterswohnungen, Alters-WG, behindertengerechte Neubauwohnungen)
Finanziell	Schlechte Belegung führt zu Umsatzeinbussen bei gleich hohen Kosten
	Zunehmender Kostendruck
IT, Daten, Digitalisierung	Hackerangriff auf IT-Systeme
Personal	Fachkräftemangel in der Pflege, ungenügende Fachkräftequote
	Zeitdruck auf Pflegepersonal (Qualitätseinbussen, psychische Probleme)
	Übergriffe von Bewohnenden
	Mängel bei der Ergonomie: Heben und Tragen von Lasten, Bewohnendentransfers, Bildschirmarbeitsplätze, Arbeitsumgebung
	Machtmissbrauch
Bewohnende und Angehörige	Steigender Einfluss und höhere Erwartungen der Angehörigen
	Fehlende Massnahmen zur Vermeidung von Pflegefehlern (z. B. State of the Art durch Expertenstandards, Leitlinien, evidenzbasierte Pflege)
	Übergriffe vom Personal
	Späterer Heimeintritt, kürzerer Aufenthalt, höherer Betreuungs- und Pflegeaufwand, komplexe Pflegesituation
Prozesse	Fehlende Transparenz bei dokumentierten Willensäusserungen (z. B. Patientenverfügungen)
	Fehlendes Meldesystem bzgl. kritischer Ereignisse, Beinahe- oder Fast-Fehlern
Infrastruktur/ Gebäude	Ausgleiten, Kollisionen, Einklemmen und sonstige Verletzungen
	Fehlende Rücklagen bzgl. Instandsetzungen, Erneuerungen oder Renovationen
Kommunikation/ Reputation	Wenig bis keine Integration ins Quartier
	Geringe Vernetzung mit Fachpersonen, Partnern, Zuweisern, ...
	Negative Rezension auf Google etc. und fehlende Erfahrung im Umgang mit diesen
	Fehlende Verantwortungsperson von externer Kommunikation und Marketing

Diese stehen in einer Exceltabelle zur Verfügung. Eine Risikomatrix wird automatisch generiert, sobald die Felder für Schadenspotential und Eintrittswahrscheinlichkeit ausgefüllt wurden. Als Hilfestellung für die Bewertung von Schadenspotential und Eintrittswahrscheinlichkeit auf einer Skala von 1-5 werden verbale und exemplarische Hilfestellungen geboten.

5.2.2 Wertkettenanalyse

Die Wertkettenanalyse (vgl. Diederichs (2023, S. 200)) strukturiert Unternehmensaktivitäten, um Wettbewerbsvorteile und Risiken zu erkennen. Primäre Aktivitäten tragen zur Wertschöpfung bei, während sekundäre diese unterstützen. Das Business Modell Canvas, ein mögliches Analysemodell, berücksichtigt Kundenbeziehungen, Ressourcen, Partner, Finanzen und externe Einflüsse. Die nachfolgende Abbildung zeigt, wie eine solche Wertkettenanalyse in der Praxis aussehen kann. Inhalte werden entweder in Stichworten oder kurzen Sätzen in die entsprechenden Felder eingefügt.

■ Risk Management

Gesellschaft	Politik/Recht	Ökologie	Technologie	Wirtschaft
- Höhere Ansprüche der Angehörigen - Wertewandel in der Gesellschaft (weniger Bereitschaft ins Sozialwesen einzusteigen)	Keine starke Lobby in der Politik, wenig Unterstützung	...	- Digitalisierung - Roboter in der Altenpflege	Regulierte Finanzierung
Schlüsselpartner (Lieferengpässe)	Schlüsselaktivitäten	Wertangebote	Kundenbeziehungen	Kundensegmente
- Lebensmittellieferanten - IT-Lieferant - Spitäler	- Pflege - Hauswirtschaft	- Persönliche Betreuung - Individuell (Zentrale) Lage - Bezug zum Quartier - Hoher Wohnstandard, grosse Zimmer - Möglichkeit für Kurzaufenthalte - Umfassendes Angebot (z.B. Alterswohnungen, betreutes Wohnen, geriatrische Reha, Spitex)	- Direkt - Über Angehörige	- Senioren (Übergriffe) - Demenzranke
Wettbewerb	Schlüsselressourcen		Kanäle	
- Andere Pflegeheime in der Region - Spitex	- Fachkräfte (Fachkräftemangel) - Finanzen (begrenzt) - IT (IT-Sicherheit)		- Spitäler - Reha - Spitex - Zuweiser	
Kostenstruktur		Einnahmequellen		
- Fixkosten (Gehälter, Mieten) - Variable Kosten - Grössenvorteile		- Pflegepauschalen (versicherung, Staat, Bewohner) - Einnahmen aus Infrastruktur (Finanzierung schwierig)		

Abbildung 4: Wertkettenanalyse mit exemplarischen Einträgen (eigene Darstellung in Anlehnung an (Diederichs, 2023))

Im Vergleich zum Business Model Canvas Originalmodell von Osterwalder und Pigneur (2010) wurden die externen Sphären des St. Galler Management Modells «Gesellschaft», «Politik/Recht», «Ökologie», «Technologie» und «Wirtschaft» ergänzt. Im Beispiel (siehe Abbildung 4) wurden diese im oberen Bereich über den eigentlichen Inhalt des Business Model Canvas platziert. Die Vorlage wird über die begleitende Web-App als Excel-Datei zur Verfügung gestellt.

5.2.3 Prozesskettenanalyse

Mithilfe der Prozesskettenanalyse (vgl. Diederichs (2023, S. 201–206) und Kahla-Witzsch und Platzer (2018, S. 104–106)) lassen sich Risiken durch die Simulation und Analyse von bestimmten geschäftlichen Prozessen identifizieren. Dies geschieht unter Einbezug der prozessverantwortlichen Mitarbeitenden. Im Projekt wurde als Hilfestellung für Praktikerinnen und Praktiker am Beispiel des Prozesses «Versorgung einer Bewohnerin / eines Bewohners» aufgezeigt (siehe Abbildung 5), wie ein Resultat konkret aussehen kann. Dadurch, dass wie im Beispiel ersichtlich, die verschiedenen Prozessschritte und eine Prozessübersicht aufgeführt sind, wird erwartet, dass eine Identifikation der einzelnen Risiken einfacher möglich ist. Die Vorlage wird über die begleitende Web-App im PowerPoint-Format zur Verfügung gestellt.

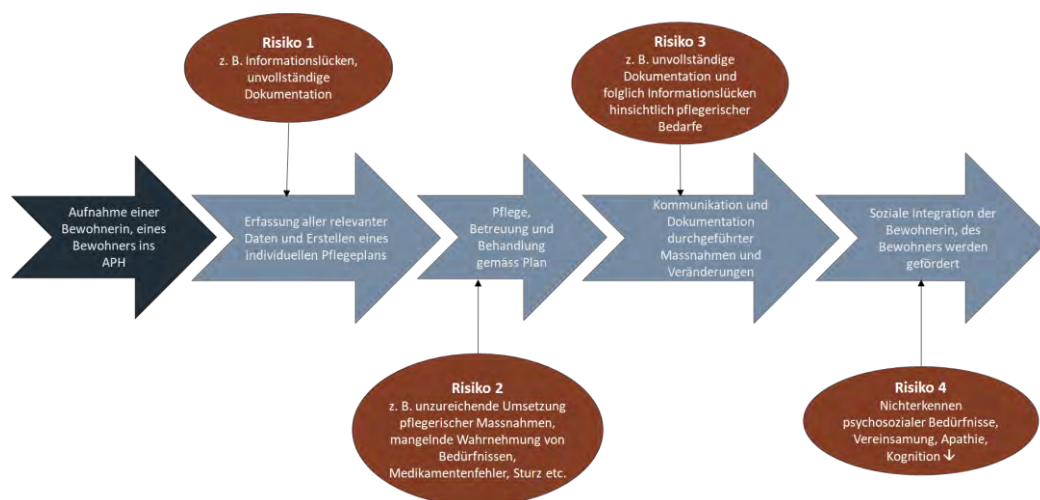


Abbildung 5: Exemplarische Prozesskettenanalyse (eigene Darstellung basierend auf (Diederichs, 2023))

5.2.4 Fehlerbaum

Die Fehlerbaumanalyse (vgl. Gleißner (2022, S. 154) oder Diederichs (2023, S. 234–239)) beginnt mit einem unerwünschten Ereignis auf oberster Ebene und untersucht, wie Einzelfunktionen auf unteren Ebenen dazu beitragen könnten. So entsteht eine Fehlermöglichkeitenbaum-Struktur, die durch logische Verknüpfungen (UND, ODER) Kombinationen für das unerwünschte Ereignis identifiziert. Sie dient der Risikoidentifikation und -bewertung. Ein Template wird über die begleitende Web-App in PowerPoint zur Verfügung gestellt, exemplarisch wurde ein Fehlerbaum für das unerwünschte Top-Ereignis «Sturz» ausgearbeitet und in der nachfolgenden Abbildung visualisiert. In der Grafik (Siehe Abbildung 6) ist ersichtlich, dass das Top-Ereignis „Sturz“ durch mehrere zugrunde liegende Ursachen (Primäre Störungen und Sekundäre Störungen) eintreten kann. Durch die Hierarchie wird ersichtlich, welche Wirkzusammenhänge bestehen und welche Störungen einen Einfluss auf Ereignisse haben.

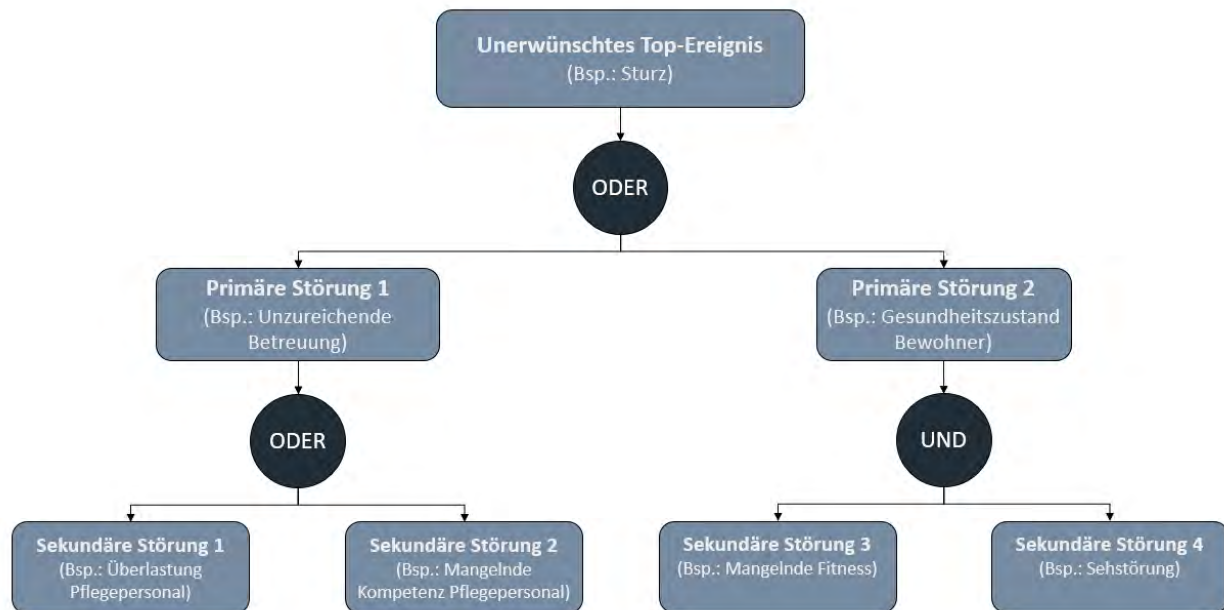


Abbildung 6: Template für die Fehlerbaumanalyse (eigene Darstellung basierend auf (Gleißner, 2022) und (Diederichs, 2023))

5.2.5 Klinisch-haftungsrechtliche Risiken

Für die Adressierung von klinisch-haftungsrechtlichen Risiken wird ein selbst erstelltes Template über die begleitende Web-App im Excel-Format zur Verfügung gestellt. Dieses enthält bereits 25 exemplarische pflegerische Risiken der Kategorien Behandlung, Dokumentation, Hygiene, Medikation, Schadenmanagement, Technik, aber auch organisatorische haftungsrechtliche Risiken aus den Kategorien Personalmanagement und Qualitätsmanagement. Neben einer Beschreibung des Risikos bietet das Template für die 25 ausgewählten Risiken auch Empfehlungen und mögliche Massnahmen. Die Vorlage soll auch von den Verantwortlichen in den einzelnen Heimen um spezifische Risiken und Massnahmen ergänzt werden. Diese können zum Beispiel aus dem ausgefüllten Risikokatalog abgeleitet werden. Exemplarisch wird in Abbildung 7 das Thema Fehlerkultur aufgeführt, dieses wurde in den durchgeführten Interviews und dem Symposium mehrfach erwähnt.

Hauptkategorie	Thema	Keyword	Frage	Beschreibung
Qualitätsmanagement	Fehlerkultur	Meldesystem	Existiert ein anonymes (Beinahe-) Fehlermeldesystem?	Die an der stationären Versorgung teilnehmenden Leistungserbringenden sind zur Sicherung und Weiterentwicklung der Qualität nach den sozialrechtlichen Vorschriften verpflichtet, ein einrichtungsinternes Risikomanagement einzuführen und weiterzuentwickeln (in Dtl bspw.: §§ 135a Abs. 2 Nr. 2, 136 SGB V). Qualitätsmanagement ist eine Führungsaufgabe und muss für die Einrichtung, insbesondere ihre Leitung und die Mitarbeitenden sowie die Bewohner*innen, effektiv und effizient sein. Um eine Sicherheitskultur zu fördern, erfordert ein gelebtes Risikomanagement die Einrichtung eines Critical Incident Reporting Systems (CIRS). Unter CIRS versteht man Fehlerberichts- und Lernsysteme, in denen vertraulich über kritische Ereignisse und Beinahe-Schäden berichtet wird. Ein gelebtes CIRS erhöht die Patientensicherheit, indem aus Fehlern gelernt wird. Das Qualitätsmanagement sorgt dann in Kooperation mit den betroffenen Fachabteilungen/Verantwortlichen für die Entwicklung und Umsetzung von Präventionsmaßnahmen, damit Beinahe-Schäden und tatsächliche Schädeneintritte an Patient*innen zukünftig vermieden werden können.
Ist-Situation		Empfehlung		Empfohlene Maßnahme
Bisher existiert kein Meldesystem, mit welchem man anonym Beinahe-Schäden im Heimalltag melden und auswerten kann.		Um die Bewohner*innensicherheit zu erhöhen, sollte das CIRS besser bespielt werden. Dazu empfiehlt sich die Durchführung regelmäßiger Bewertungen und Besprechungen von CIRS-Einträgen, um den Mitarbeitenden zu vermitteln, dass die Eintragungen im Alltag auch zu tatsächlichen und positiven Konsequenzen im Arbeitsalltag führen.		Sensibilisierung der Mitarbeitenden für das System (ggf. durch Schulungsangebot) Niederlegung eines einheitlichen Verhaltensleitfadens bei Beinaheschäden

Abbildung 7: Ausschnitt aus dem Template für klinisch-haftungsrechtliche Risiken, Beispiel zum Thema Fehlerkultur (eigene Darstellung)

5.2.6 Formular zur Meldung unerwünschter Ereignisse

Unfälle, Vorfälle, Risiken für Vorfälle oder Fehlfunktionen im Betrieb, welche schädigende Konsequenzen für die Bewohnenden, Besuchenden, Begleitpersonen oder für das Eigentum des Betriebs bzw. des Unternehmens haben könnten, sind als unerwünschte Ereignisse zu bewerten und entsprechend zu bearbeiten. Auch wenn in den Interviews und beim Symposium viele Heimleitende betonten, dass ein Critical Incident Reporting System bereits im Heim eingeführt wurde, wird aufgrund der Wichtigkeit der Thematik über die Web-App eine Word-Vorlage zur Meldung solcher unerwünschten Ereignisse zur Verfügung gestellt. Im Formular gilt es, die betroffene Person oder den betroffenen Bereich zu beschreiben, das unerwünschte Ereignis einer Kategorie (z.B. Fehlmedikation, Unfall, Einbruch, sexuelle Belästigung) einzuordnen und das Ereignis zu dokumentieren. Auch Sofortmassnahmen und die Bearbeitung des Ereignisses sollen festgehalten werden. Wichtig ist der Abschnitt im Formular, in welchem Massnahmen zur Vermeidung des Ereignisses in Zukunft verbindlich (Verantwortlichkeit und Zeitpunkt) festgehalten werden. Diese Vorlage unterstützt vor allem bei dem Umgang mit eingetretenen Risiken im Bereich der Pflege oder des allgemeinen Betriebs und deckt nicht alle betriebswirtschaftlichen Risiken ab.

5.2.7 Risikobericht

Für eine transparente Kommunikation sind ein Risikobericht und eine Präsentation der Risiken von entscheidender Bedeutung. Eine periodische Berichterstattung stellt sicher, dass die Risikoanalyse aktuell bleibt und Anpassungen an sich verändernde Bedingungen vorgenommen werden können. Als Templates werden eine Word-Vorlage für einen Risikobericht und eine PowerPoint-Vorlage für die Präsentation des Risikoberichts zur Verfügung gestellt. Darin finden sich auch Platzhalter für die Ergänzung durch weitere bereits verwendete Templates, zum Beispiel für die Risikomatrix.

5.3 Web-App

Die Inhalte des rund 80-seitigen Leitfadens werden in einer gekürzten Form in einer Web-App kostenlos zur Verfügung gestellt. In der Abbildung 8 sind exemplarisch drei Bereiche der Web-App dargestellt und werden folglich von links nach rechts erläutert. Die Startseite ermöglicht einen einfachen Einstieg und erlaubt den Zugriff auf die zentralen Inhalte. Auf Unterseiten mit Text- und Bildinhalten ist grundlegendes Wissen rund um das Thema Risikomanagement abrufbar. Auf der Unterseite mit den Templates sind alle Vorlagen mit einer kurzen Erklärung und einem Hinweis zum Format als Download verfügbar. Wie in den Screenshots ersichtlich, wurde der Fokus auf eine einfache, auf das wesentliche reduzierte, Oberfläche gelegt.

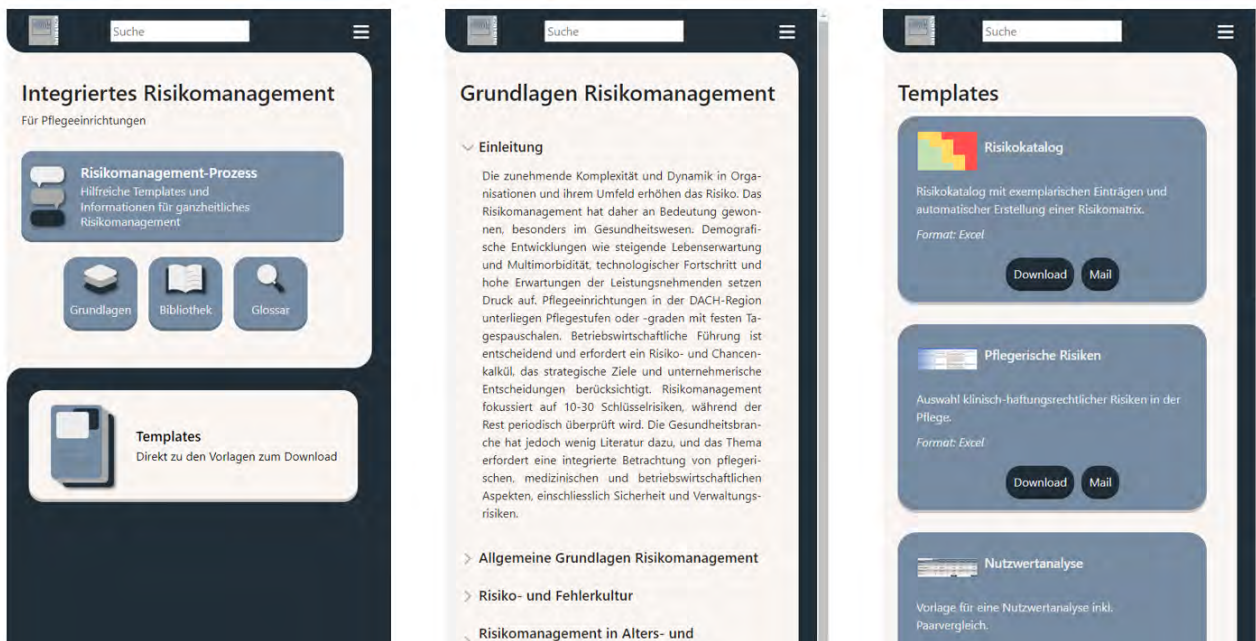


Abbildung 8: Screenshots mit Inhalten der Web-App (eigene Darstellung)

Die Web-App ist unter www.ost-risikomanagement.ch erreichbar. Die Web-App dient auch zur Distribution der Templates, welche entweder direkt heruntergeladen werden oder als Link per Mail versendet werden können. Die Struktur der Inhalte orientiert sich dabei am Risikomanagement-Prozess nach ISO-Norm 31000 und somit auch am Aufbau des Leitfadens. Als Web-App können die Inhalte und Templates sowohl auf mobilen Geräten als auch am PC abgerufen werden.

6 Fazit und Ausblick

Das Ergebnis des abgeschlossenen Forschungsprojektes – gefördert durch die Funk-Stiftung – ist der Leitfaden zur Einführung eines integrierten Risikomanagements in stationären Alters- und Pflegeheimen sowie die Web-App mit einer gekürzten Version des Leitfadens. Zudem wurden Templates entwickelt, welche Praktikerinnen und Praktikern helfen sollen, das Risikomanagement einzuführen. Der innovative Charakter des Projekts ist nicht der generelle Umgang mit dem Thema Risikomanagement. Dafür existiert bereits Standardliteratur (vgl. Kap. 2), und das Vorgehen ist etabliert. Innovativ ist jedoch die Erforschung der gelebten Praxis inklusive der Bedürfnisse und Herausforderungen von Personen, welche in Pflegeheimen für das Risikomanagement verantwortlich sind. Die Resultate des Projekts dienen folglich als Hilfestellung für alle Bereiche des Risikomanagements und geben einen Einblick in diese spezifische Branche.

Während des Projektes wurde festgestellt, dass die Einstellung von Praktikerinnen und Praktikern zu einem professionellen Risikomanagement und dementsprechend dem Leitfaden zur Einführung unterschiedlich ist. Während die einen es als Hilfestellung begrüßen, lehnen es andere wegen des zusätzlichen administrativen Aufwands ab. Es bleibt zu hoffen, dass die Praxis den Leitfaden einsetzt und ihre Erfahrungen einbringt. Auf diese Weise können dann tatsächlich Best oder zumindest Good Practice abgeleitet und entwickelt werden.

Zukünftig wird es wahrscheinlich in der Forschung und Praxis weitere Entwicklungen geben. Aspekte der Risikokultur, die in diesem Projekt bereits kurz behandelt wurden, werden vermutlich an Bedeutung gewinnen. Die bereits in dieser Studie angestossene Integration verschiedener Bereiche wird voraussichtlich ausgeweitet auf Themen wie Nachhaltigkeit, Digitalisierung und künstliche Intelligenz. Insbesondere das letztgenannte Thema wird in Zukunft eine besondere Bedeutung einnehmen. Zum einen wird es dabei unterstützen, Risiken und Fehler frühzeitig zu erkennen und zu vermeiden. Zum anderen kann künstliche Intelligenz teilweise Arbeiten von Fachkräften übernehmen (z.B. Dokumentation) und dadurch die grosse Herausforderung des Fachkräftemangels in der Pflege etwas abmildern.

Literaturverzeichnis

- Bestvater, K. (2022). *Fehlerkultur in Organisationen* (Bd. 33). Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-36653-7>
- Biernoth, M. (2016). Fachkräftemangel. In M. Biernoth (Hrsg.), *Employer- und Behavioral Branding im Gesundheitswesen* (S. 4–23). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-12866-1_2
- Blonski, H. (Hrsg.). (2014). *Pflege. Risikomanagement in der stationären Altenhilfe: Anforderungen, Methoden, Erfahrungen*. Schlütersche.
- Blüher, S. & Kuhlmei, A. (2016). Demographischer Wandel, Altern und Gesundheit. In M. Richter & K. Hurrelmann (Hrsg.), *Soziologie von Gesundheit und Krankheit* (S. 313–324). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-11010-9_21
- Brühwiler, B. (2016). *Risikomanagement als Führungsaufgabe: Umsetzung bei strategischen Entscheidungen und operationellen Prozessen* (4., aktualisierte Auflage). Risikomanagement. Haupt Verlag.
- Bryman, A. (2016). *Social research methods* (Fifth edition). Oxford University Press. <https://www.loc.gov/catdir/enhancements/fy1617/2015940141-b.html>
- Diederichs, M. (2023). *Risikomanagement und Risikocontrolling* [Dissertation, Verlag Franz Vahlen]. K10plus.
- Ebner, G. (2008). *Fehler - Lernen - Unternehmen: Wie Sie die Fehlerkultur und Lernreife Ihrer Organisation wahrnehmen und gestalten*. Lang.
- Euteneier, A. (Hrsg.). (2015). *Erfolgskonzepte Praxis- & Krankenhaus-Management. Handbuch klinisches Risikomanagement: Grundlagen, Konzepte, Lösungen - medizinisch, ökonomisch, juristisch*. Springer. <http://www.springer.com/>
- Frodl, A. (2012). Risikomanagement im Gesundheitsbetrieb. In A. Frodl (Hrsg.), *Controlling im Gesundheitsbetrieb* (S. 145–154). Gabler Verlag. https://doi.org/10.1007/978-3-8349-7086-2_7
- Gleißner, W. (2022). *Grundlagen des Risikomanagements: Handbuch für ein Management unter Unsicherheit* (4., vollständig überarbeitete und erweiterte Auflage). Management competence. Verlag Franz Vahlen.
- Heller, A. R., Eberlein-Gonska, M., Held, H. C. & Koch, T. (2023). Systematisches Risikomanagement für eine geplante Gasabschaltung im High-Care-Bereich eines Universitätsklinikums [Systematic risk management for a planned gas shutdown in the high-care facility of a university hospital]. *Die Anaesthesiologie*, 72(4), 282–292. <https://doi.org/10.1007/s00101-023-01254-8>
- Hunziker, S., Henrizi, P., Hilsbos, A. & Schneeberger, T. (2022). *Quick Guide Ganzheitliches Risk Management im Krankenhaus*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-36849-4>
- Kahla-Witzsch, H. A. (2023). *Organisations-Risikomanagement für Gesundheitseinrichtungen: Praxisleitfaden zur Implementierung und Weiterentwicklung* (1. Auflage). QuickInfo. Austrian Standards plus GmbH.
- Kahla-Witzsch, H. A. & Platzer, O. (2018). *Risikomanagement für die Pflege: Ein praktischer Leitfaden* (2., überarbeitete Auflage). Pflegemanagement. Verlag W. Kohlhammer. http://www.kohlhammer.de/wms/instances/KOB/appDE/nav_product.php?product=978-3-17-031983-7
- Konnyu, K. J., Yogasingam, S., Lépine, J., Sullivan, K., Alabousi, M., Edwards, A., Hillmer, M., Karunananthan, S., Lavis, J. N., Linklater, S., Manns, B. J., Moher, D., Mortazhejri, S., Nazarali, S., Paprica, P. A., Ramsay, T., Ryan, P. M., Sargious, P., Shojania, K. G., . . . Ivers, N. (2023). Quality improvement strategies for diabetes care: Effects on outcomes for adults living with diabetes. *The Cochrane database of systematic reviews*, 5(5), CD014513. <https://doi.org/10.1002/14651858.CD014513>
- Osterwalder, A. & Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers*. Wiley.
- Paula, H. (2006). *Patientensicherheit und Risikomanagement: Im Pflege- und Krankenhausalltag*. Springer Berlin Heidelberg. <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1504019>
- Robson, C. & McCartan, K. (2016). *Real world research: A resource for users of social research methods in applied settings* (Fourth edition). Wiley. <https://permalink.obvsg.at/AC14480602>

- Saßen, S. (2019). *Risikomanagement: Fehler vermeiden, melden, analysieren und bewältigen*. Altenheim. Vincentz Network.
- Scopetti, M., Santurro, A., Tartaglia, R., Frati, P. & Fineschi, V. (2021). Expanding frontiers of risk management: care safety in nursing home during COVID-19 pandemic. *International journal for quality in health care : journal of the International Society for Quality in Health Care*, 33(1). <https://doi.org/10.1093/intqhc/mzaa085>
- Speer, T., Mühlbradt, T., Fastner, C. & Schröder, S. (2023). Safety-II: ein systemischer Ansatz für ein effektives klinisches Risikomanagement [Safety-II: a systemic approach for an effective clinical risk management]. *Die Anaesthesiologie*, 72(1), 48–56. <https://doi.org/10.1007/s00101-022-01215-7>
- Toles, M., Colón-Emeric, C., Moreton, E., Frey, L. & Leeman, J. (2021). Quality improvement studies in nursing homes: a scoping review. *BMC health services research*, 21(1), 803. <https://doi.org/10.1186/s12913-021-06803-8>
- Vanini, U. (2012). *Risikomanagement: Grundlagen, Instrumente, Unternehmenspraxis*. Schäffer-Poeschel. <http://www.socialnet.de/rezensionen/isbn.php?isbn=978-3-7910-3126-2>
- Zapp, W. (2020). *Qualitäts- und Risikomanagement im Krankenhaus*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-31491-0>

The risk trilemma of how to hedge gas-fired power plants in times of war

Research Paper

Cedric Schneider (Master of Science, WINGS)

Hochschule Ruhr West, Duisburger Strasse 100, 45479 Mülheim an der Ruhr, Deutschland,
E-Mail: cedric.schneider00@gmx.de

Prof. Dr. Michael Römmich

Hochschule Ruhr West, Duisburger Strasse 100, 45479 Mülheim an der Ruhr, Deutschland,
E-Mail: michael.roemmich@hs-ruhrwest.de

Dr. Andreas Heinen

E.ON Energy Markets GmbH, Brüsseler Platz, 45131 Essen, Deutschland, E-Mail: Andreas.Heinen@eon.com

The views, opinions, positions or strategies expressed in this article are those of the author and do not necessarily represent the views, opinions, positions or strategies of, and should not be attributed to E.ON.

Abstract

European power generators optimizing their asset value typically face a risk trilemma consisting of market, credit and liquidity risk. During the energy crisis as induced by the Russian war in the Ukraine, the relevance of optimization decisions between those risk categories was amplified through extreme price volatility and unprecedented price spikes. This paper analyses hedge decisions for the example of a German based gas-fired power plant owner who reduces market risks with futures at exchanges. Hedge volumes are calculated at positive generation margins for power, gas and emission allowances. Liquidity risk over a three-year hedging horizon is quantified using historical simulations. The relevance of specific liquidity risk components is highlighted, notably the effect of cross margining between commodity exposures and initial margin requirements at exchanges. The paper concludes with practical recommendations for liquidity risk management including threshold considerations in alternative bilateral agreements between counterparties.

1 Introduction

1.1 Price spikes and volatility in the energy market crisis

In the second half of 2021, market volatility of traded wholesale energy products already increased notably with the build-up of political tensions preceding the Russian invasion of Ukraine later in February 2022. A tight supply and demand balance specifically for gas led to a wholesale price rally for the whole European energy system as merit order induced correlation impacts into other commodities created enormous price spikes notably in the second half of 2022. Low availability of nuclear power plants in France and gas prices were driving most European power prices as gas-fired power plants were often the marginal technology type clearing electricity demand. In forward traded markets dependencies were such that statistical measures indicated nearly maximum correlation between power and gas prices. Fig. 1 shows German wholesale forward prices resp. futures for electricity plotted against gas prices on the same trading day.

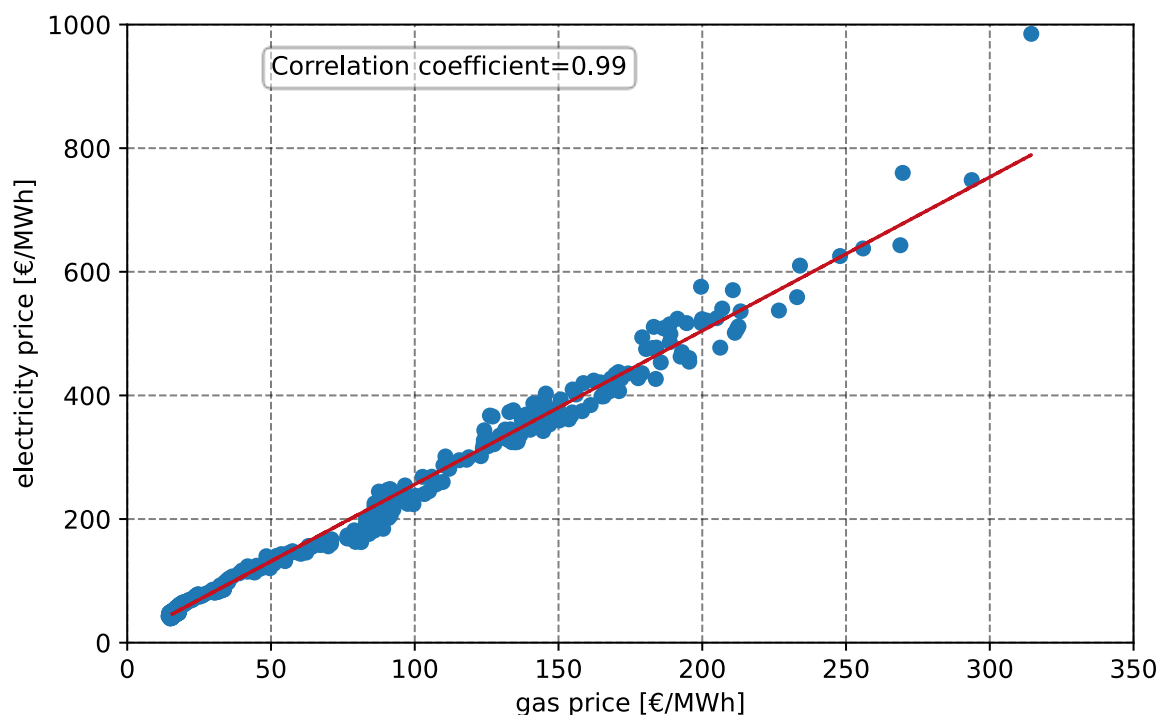


Figure 1: Electricity and gas prices plotted against each other (blue dots) based on German year-ahead futures (Cal 2023 Baseload/Cal 2023 THE NAT Gas) for both commodities on corresponding trading days (source EEX). The red line represents a linear parametric fit indicating a near perfect correlation.

The subsequent decrease of gas prices in 2023 is commonly attributed to additional liquified natural gas imports to Europe, build-up of gas storage reserves, energy saving measures and favourable winter weather. The combination of efforts and favourable conditions allowed to replace the declining pipeline-based gas imports from Russia. Strong up and downtrends, spikes in prices as well as a higher price volatility can pose significant financial risk to market participants.

Fig. 2 illustrates the dynamics between 2021-23 on German front-year baseload power prices and their volatility using a standard volatility measure with exponentially weighted moving averaging (ewma) of price returns. The rapid increase of wholesale prices and volatility caused stress for individual market players and the whole energy system as such. Electricity baseload prices are used throughout for consideration of a sufficiently robust price signal given traded volumes on the most liquid contract type. Generally, through the crisis liquidity of traded products decreased compared to prior years (Monitoringbericht 2023, p.157 for electricity and p.255 for gas). With hindsight

the spike in price volatility in early October 2021 can be considered as a foreshock to later developments unfolding with the start of the Russian invasion. While spikes in prices itself and in volatility appear at similar times, magnitudes can differ and play a significant role in risk management.

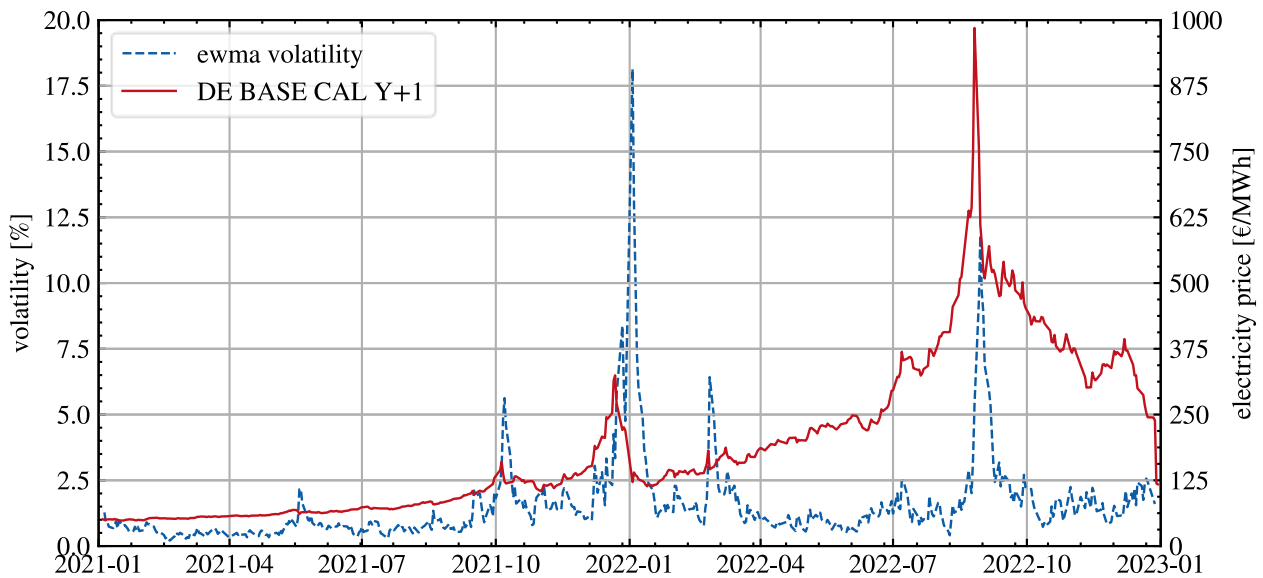


Figure 2: Historic evolution of German Baseload Electricity prices for the year-ahead (Y+1, red solid line) shown on the right axis in €/MWh along with its volatility on the left axis (blue dotted line). Volatility is calculated using a standard exponentially weighted measure of price returns.

1.2 Risk Management trilemma: market, credit and liquidity risk

This brief overview highlights the risk trilemma: The trade-off between hedging market risk on the one hand and exposure to credit and/or liquidity risk on the other. So, what is the optimal level of hedging against market price risk for a power producer's planned (future) production? A theoretical basis for the interaction between hedging activities and planned production of power producers can be found by using the microeconomic theory of the firm. A one-period microeconomic model (Holthausen 1979) provides two main results:

Separation Property: With perfect hedging instruments, i.e., a complete set of exchange-traded futures with no transaction costs, the electricity producer can separate decisions about production and hedging decisions.

Full Hedge Property: Assuming unbiased futures markets, i.e., the expected spot market price is equal to today's futures price, the power producer will hedge its planned production level fully by selling the equivalent amount futures contracts. In this theoretical situation, the generator can lock in all expected profits today, resulting in a net zero position without market price risk.

Römmich (2018) showed that, taking into account the transaction costs of trading, that both separation and full hedge properties of the Holthausen model do no longer apply: Introduction of transaction costs by market bid-ask spreads of hedging instruments reduces the hedge ratio, i.e., the planned production volumes to be hedged. The power producer still hedges significant volumes and therefore enters into large short positions on the German power futures market.

Transaction costs and liquidity in wholesale markets are the main drivers for power generators to their hedging decisions over a longer time horizon, e.g., up to three trading years ahead (see e.g., company reports RWE 2020,

Uniper 2022, Vattenfall 2023). The typical outcome would be one hedging transaction per commodity leg per exchange trading day. Under this assumption the following decision tree and subsequently the case study was developed with related work by Lang and Madlener (2010) and Lehrbass (2021).

Hedges of electricity, gas and CO₂ emission allowance volumes follow a purpose of securing generation spreads, thereby reducing earnings volatility. From a profit and loss perspective, the relevant price developments can thus be assessed in a different way. For a power plant operator, the difference between the electricity price p_{el} and fuel price p_{gas} (and emission costs p_{CO_2}) is the relevant parameter as it indicates the short term profitability of the plant. The electrical efficiency η and the emission factor f_{CO_2} of the power plant must be considered for the calculation. For a gas-fired power plant this is called clean spark spread, which can be calculated as shown in formula 1.

$$css = p_{el} - \frac{1}{\eta} \cdot (p_{gas} + p_{CO_2} \cdot f_{CO_2}) \quad (1)$$

The clean spark spread in the German region is typically calculated in €/MWh as this is the standard unit for energy trading. Fig. 3 shows the evolution of the clean spark spread of a gas-fired power plant at a certain electric efficiency matching the parameters of the case study described in Sec. 2. Generally, the clean spark spread reflects the difference of revenues from power generation with electricity as a positive revenue stream and gas plus CO₂ emission allowances as asset-specific fuel cost, typically driven by a single electrical efficiency parameter. A positive clean spark spread indicates profitable operations based on short-run marginal costs.

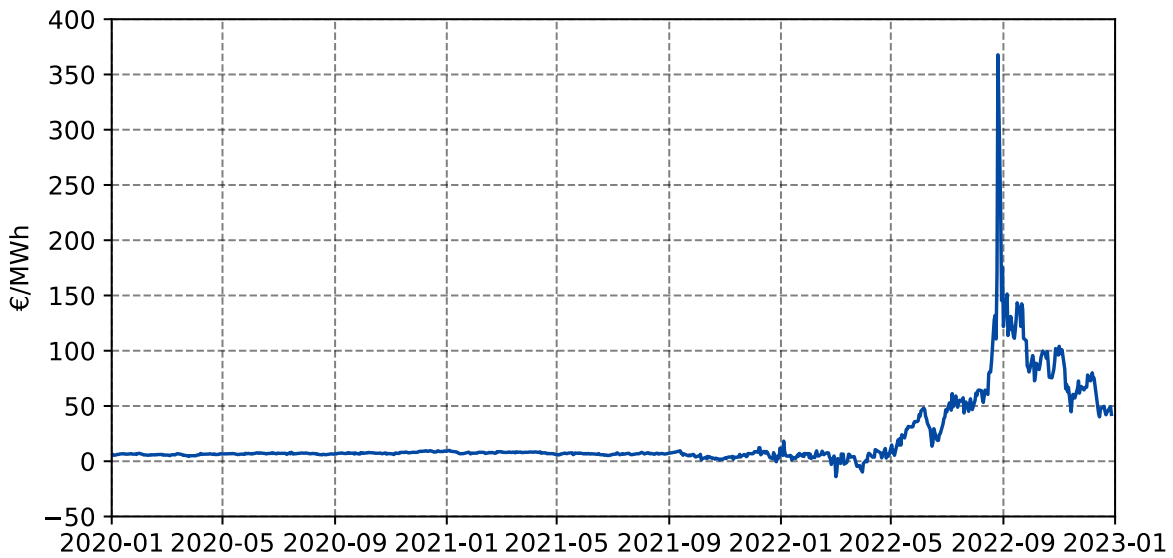


Figure 3: Clean spark spread of the gas-fired power plant fleet.

An owner of a gas-fired power plant is thus typically faced with a decision on how to secure profitability of generation volumes and margins in volatile markets. Selling electricity generation volumes simultaneously with buying the required gas and emission allowances reduces price risk and locks in generation margins accordingly. If hedges are carried out with forward transactions e.g., “over-the-counter” (OTC) in broker-based bilateral counterparty matches, price risks are effectively transferred to credit risk and/or liquidity risk depending on collateralisation agreements with a counterparty. In contrast, exchange traded futures almost always require full collateralisation, i.e., cash margining with an initial margin and a variation margin covering exposure changes from price changes. While detailed regulatory implications e.g., based on the European Market Infrastructure Regulation (EMIR) are not part of this research paper, the authors comment on transparency improvements on margin calls to be achieved with amendments to EMIR decided by the European Union in Sec. 2.1.

Operators would make large profits by hedging with futures during price peaks. In the historic environment the clean spark spread also increased strongly in absolute terms driven by fuel supply and generation capacity shortages and an overall high correlation of commodity prices across European markets. However, at the same time, initial margin requirements would also peak and likely create liquidity stress if only using exchange trading. Initial margin and variation margin requirements are calculated in detail in Sec. 2.

From a credit risk perspective, price spikes would typically trigger to additional risk as the root cause of price spikes may link to tight supply and demand fundamentals, balance sheet stress for relevant counterparties or generally nervous market environment due to political events or other uncertainties. While generators may have an interest to hedge in OTC markets with counterparties under credit risk arrangements, there is a dilemma in identifying adequate counterparties. At highest market prices, it is less likely to find natural buyers for the spread exposure and the asset owner likely needs to distribute transactions over different counterparties. Natural buyers of electricity who need electricity volumes in downstream customer portfolios would likely pause their own forward hedging at prices exceeding anticipated regulated price cap mechanisms or prices affordable to customers. As such it can be assumed that hedging at highest price spikes can be difficult in practice or requires the use of additional instruments.

Fig. 4 illustrates the resulting holistic decision tree for the asset owner over time. Initially, the asset owner may decide to hedge or not to hedge price risk. If hedges are OTC outside exchanges, credit risk may be unsecured, partly secured or fully covered if cash collateralisation is agreed between counterparties such that a transition to liquidity risk is possible. If hedges are via exchanges, then margining is inevitable thereby creating liquidity risk. The risk trilemma in this refers to the inability to eliminate all risk entirely, i.e., accepting one risk while minimising others (Sparla et al. 2023).

In a certain way, the risk trilemma is maximised under stress of extreme market events. However, hedge and channel decisions do not need to be static and can be partly or fully reversed. If hedges were carried out at exchanges, liquidity risk may be transferred to credit risk with so-called Exchange-for-Physical (EFP) transactions between two counterparties having (opposite) exchange positions. Such trilateral arrangements allow two counterparties to transact positions at zero market risk with the objective of influencing only the financial exposure in margining arrangements. EFPs as trilateral transactions are thus beneficial compared to unilateral counter trades of an individual counterparty posing timing risk of open positions. Both counterparties are then able to eliminate initial margin requirements on the share of exchange traded position reductions, albeit at the expense of “freezing” variation margin accumulated after initial hedges were carried out (Financial Stability Board 2023, p. 22-23).

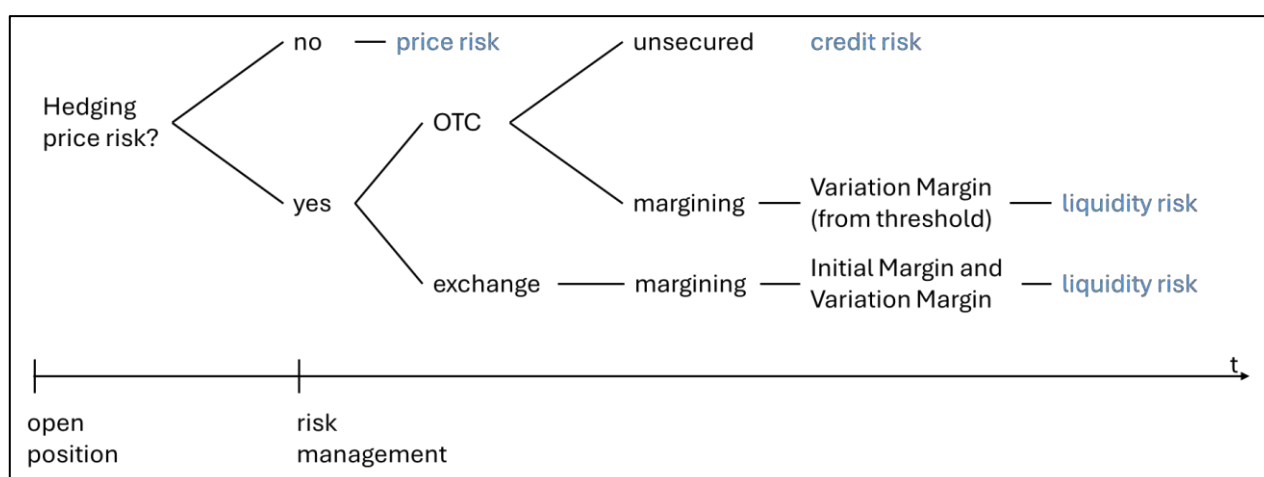


Figure 4: Decision tree for a generation asset owner over a time horizon t reflecting a risk trilemma in managed and accepted risks.

Full reversal of credit risk positions is typically not possible and depends on the willingness of a counterparty to terminate previous transactions. As such, credit risks may only be reduced through counter transactions at a later

point in time and at different market prices. Counter transactions would effectively create a residual fixed counterparty exposure which remains static until delivery or at a credit event. The use of Credit Default Swaps is not part of this analysis.

1.3 Hedge path impact on resulting credit or liquidity risk

In practice, market liquidity, i.e., traded volumes, of required wholesale products is limited such that asset owners choose a hedge path distributing hedge volumes over a longer period of time e.g., up to three trading years ahead with three hedging transaction per exchange trading day (RWE 2020, Uniper 2022, Vattenfall 2023).

Focussing on liquidity risk effects in the case study, it is sufficient to consider the impact of a linear hedge path. The larger the portfolio of a generation asset owner is relative to available market liquidity, the less flexibility in significant deviations from an extended path can be assumed based on the least liquid commodity. I.e., for a proportionally large portfolio, accelerated hedging specifically well ahead of delivery periods is usually strongly constrained. In 2022 traded volumes in exchange and OTC markets together for German electricity were annually ca. 1,800 TWh for the front year, ca. 500 TWh for the second year ahead and less than 150 TWh for the third year ahead (Monitoringbericht 2023, p. 157-159). While hedging of commodities takes place simultaneously in the modelled case, this paper rather discusses in detail which impact the choice of trading venues may have on resulting net cash requirements.

Finally, non-availability of sufficiently granular hedge products requires the asset owner to focus on block products. One prime example regards the hourly structure of electricity prices based on spot auction clearing mechanisms. In forward traded wholesale markets or with standard definitions of exchange traded futures, hourly structured products are not available. Choosing yearly baseload products instead reflects a common choice of proxy hedging for asset owners as the bulk of outright commodity exposure is effectively reduced, especially if assumed production volume forecasts are sufficiently stable. In this analysis, yearly products are not further cascaded into more granular hedge products closer to delivery as intra-year seasonality effects are less relevant for the case study selected.

2 Case Study – to hedge and not to hedge - price risk versus liquidity risk of a power producing company

2.1 Portfolio specifics and modelling approach

For the case study, the effects of a hedging path over three years from 2020 to 2022 are calculated for a stylised owner of a gas and steam power plant portfolio, whereby only exchange trading with the resulting liquidity risk is assumed. Similar to Lang and Madlener (2010), hedging over a three-year period, with one reference year was considered. In reality, hedging is carried out in parallel for several years, so that, for example, hedging transactions for the reference years 2025, 2026 and 2027 take place in parallel in 2024. In order to clarify the relationships between different variables, the case study has been reduced to one reference year allowing for direct identification of effects in input and output time series without additional effects of multi-year averaging. Fig. 5 provides an overview of the progression and weighting of the three scenarios. For the historic price scenario used, there were no significant differences in relevant output parameters of liquidity risk identified. As such, for the sake of simplicity, only the linear hedge is discussed below.

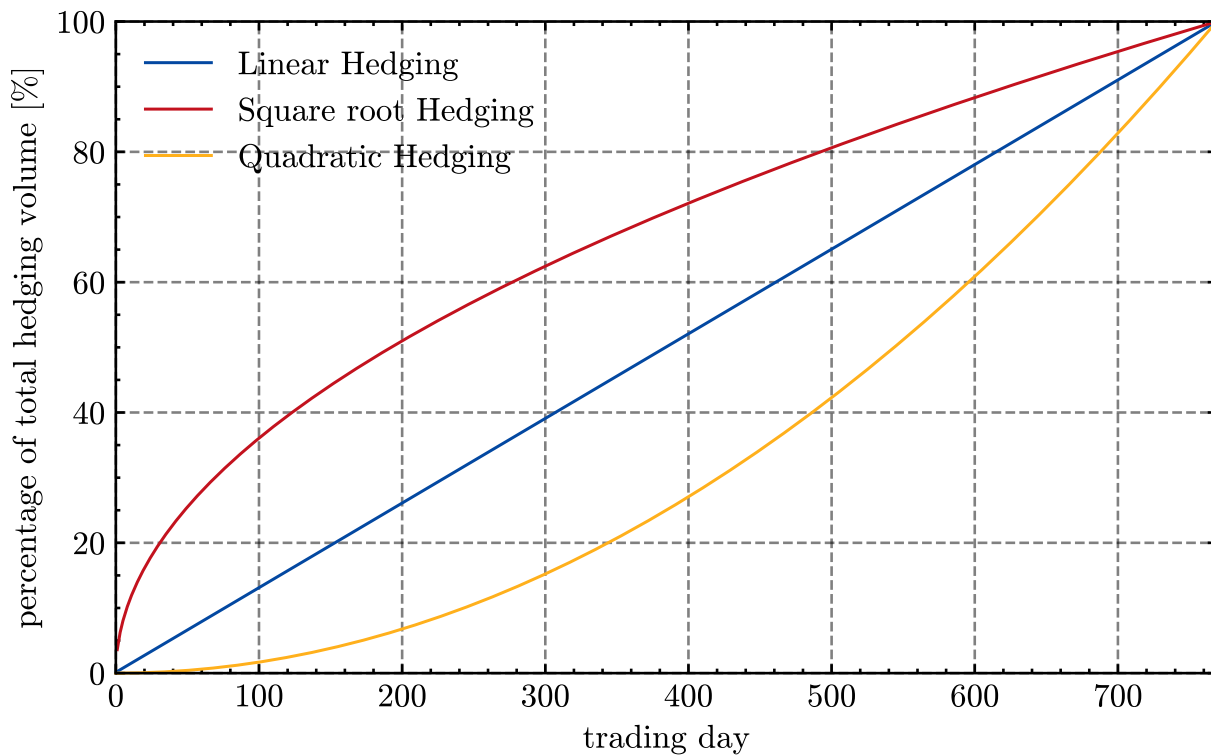


Figure 5: Overview of the assumed hedging scenarios

The case study further assumes that the power plant fleet can be represented by combined-cycle gas turbine power plants of the same electrical efficiency, set at a realistic value of 54%. As such the portfolio effectively resembles a single asset allowing for scaling of input and output factors with capacity or volume assumptions. The emission factor for the production of CO₂ equivalent through gas combustion is assumed to be 0.056 t CO₂/GJ or 0.2016 t CO₂/MWh using the standard values for natural gas fuel emissions (EBeV 2022). It is assumed that the hedging simultaneously includes the sale of electricity base load quantities and the purchase of gas and emission certificates on each trading day. While deviations from such a strategy may occur in practice, simultaneous hedging allows the asset owner to lock-in desired clean spark spreads from electricity generation commercially. The fixed amount of electricity to be hedged for the reference year 2023 is 50 TWh (50 billion kWh). The volume for gas and CO₂ allowances results from the consideration of efficiency and the emission factor. The production and hedging volumes in our case study are based on data published by German energy suppliers for the 2022 financial year (RWE: 157 TWh [RWE Geschäftsbericht 2022], Uniper: 61 TWh [Uniper Geschäftsbericht, 2022], EnBW: 42 TWh [EnBW Geschäftsbericht, 2022]). In 2022, the five companies with the highest volumes accounted for around 63.5% of generation in the German market area including Luxembourg (Monitoringbericht 2023, S. 17). A generation volume of 50 TWh also corresponds to approx. 10 % of the electricity generated or consumed in Germany in 2022 (Monitoringbericht 2023), meaning that the portfolio in our case study represents the size of a big player in electricity generation.

With the linear hedging strategy over a period of the three years or 769 trading days, the same amount of electricity, gas and CO₂ is transacted every day, i.e., without additional consideration of corresponding price levels in hedge decisions. The only commercial condition is that hedging only takes place if there is a positive clean spark spread. As such the asset owner will operate profitably on variable marginal costs.

We quantitatively model the impact of variation margin calls as well as initial margin calls and their volatility. If all transactions are carried out at the same exchange, then cross margining between individual commodity exposures

is assumed. A simulation of not benefitting from cross margining in resulting cash requirements links to choosing different trading venues for the individual commodity hedges. For example, electricity could be traded on the European Energy Exchange (EEX), whereas gas is currently primarily traded on the Intercontinental Currency Exchange (ICE) due to the higher volume with a common feature of “liquidity attracting liquidity”. Underlying price data was extracted from Bloomberg databases.

The calculations for the variation margin are also based on Lang and Madlener (2010) and apply the following logic: First, the clean spark spread (CSS) is calculated. For a positive CSS for a trading day, the hedging volume is calculated according to the hedging strategy. For the linear hedge, 1/769 of the 50 TWh of electricity is sold on each trading day and the required gas and CO₂ volume is purchased accordingly. The variation margin can then be calculated for each trading day s by multiplying the already hedged volume v by the difference between the settlement price of the last trading day p_{s-1} and the current day p_s (see formulas 2.1 - 2.3).

$$VM_{electricity,s} = -\sum_1^{s-1} v_{electricity} \cdot (p_{electricity,s} - p_{electricity,s-1}) \quad (2.1)$$

$$VM_{gas,s} = \sum_1^{s-1} v_{gas} \cdot (p_{gas,s} - p_{gas,s-1}) \quad (2.2)$$

$$VM_{CO_2,s} = \sum_1^{s-1} v_{CO_2} \cdot (p_{CO_2,s} - p_{CO_2,s-1}) \quad (2.3)$$

Positive values of the variation margin VM mean that it will be received, negative values signal a payment. Because electricity is sold and therefore a margin must be paid if prices rise, the calculation is negated accordingly. To show the status of the margin account, all daily margins are cumulated. In the case of cross-margining, the margins are offset against each other, and a net margin is formed.

Exchanges currently do not publish methods for calculating the initial margin in all relevant details and there is no static or standard industry method set. In principle, however, initial margins must be deposited when a position is opened on the stock exchange and depends on the volatility of prices as the key measure impacting close out risk. We have examined calculation methods from two different exchanges. The first is a formulaic simplified approximation of the EEX margin model (see formula 3). This takes into account the volatility, in this definition using relative returns $\sigma_{rel,i}$, the price level p_i and a risk factor RF . In addition, a variety of weighted past price data is included in the calculation (European Commodity Clearing 2022).

$$IM_i = p_i \cdot \sigma_{rel,i} \cdot RF \quad (3)$$

The formula can be considered as generic in the sense of requiring a definition of a reasonable risk factor RF and a volatility metric reflecting the assumed underlying close out risk. The risk factor may be approximated using a confidence level of a price risk distribution over a short holding period. If a standard normal distribution is assumed and a holding period of around two trading days with prices following Geometric Brownian Motion (GBM), then a risk factor RF could be defined completely based on a chosen confidence level. At a 99% confidence level, using GBM for a holding period impact (Hull 2022), RF could thus be approximated as $RF_{99\%,2d-GBM}$ as per formula (4), i.e., leading to a risk factor of around 3.29.

$$RF_{99\%,2d-GBM} = \sqrt{2} \cdot z_{99\%} \approx \sqrt{2} \cdot 2.33 \approx 3.29 \quad (4)$$

Formula (3) in combination with formula (4) and a calculation of the volatility $\sigma_{rel,i}$ then provides specific values of the initial margin for a commodity [€/MWh]. If multiplied by the respective trading volume, an absolute value in [€] is obtained.

Some exchanges may also provide tools to simulate initial margin requirements. In case of the Intercontinental Exchange (ICE), a web-based clearing analytics simulation allows to estimate initial margins and a value at risk metric (ICE Clearing Analytics). While methods are generally subject to change, even for identical trade data, results at different exchanges may differ in individual choices of details in volatility calculations or risk factors if linked to a similar effective equation as in formula (3). Depending on future market scenarios such differences between exchanges or trade venues can be relevant for market participants decisions in trading achieving lower or higher liquidity risk.

During the energy crisis in December 2022, the European Union has recognised a need to increase transparency through clearing of central counterparties (CCP) in a proposal for regulation amending the European Market Infrastructure Regulation (“EMIR 3.0”, 2022). Political agreement was reached in February 2024. Amendments to Article 38 and 41 shall ensure higher transparency for clients with better visibility and predictability of margin calls as well as continuous revisions of margin call levels reflecting e.g., market conditions. Model robustness shall be improved with Article 49 amendments requesting CCPs to apply procedures ensuring shorter model update cycles with less complexity and more certainty in outcomes. A recent study prepared for Energy Traders Europe by Frontier Economics and Luther (2024) supports these developments amongst several other initiatives. Following implementation and practical experience, approximations of initial margin requirements should then allow better forecast predictability for users in day-to-day cash management and risk mitigation measures of liquidity risk.

2.2 Dependencies in variation margin and initial margin results

As described above, the power plant operator hedges electricity, gas and CO₂ emission allowances on each trading day. The resulting exchange transactions trigger variation and initial margin with a different amount on each trading day and for each commodity. Fig. 6 shows the development of cumulative margins for the case where trading takes place on one exchange such that netting of all individual exposures for electricity, gas and CO₂ allowances is possible (cross margining is used in red curves).

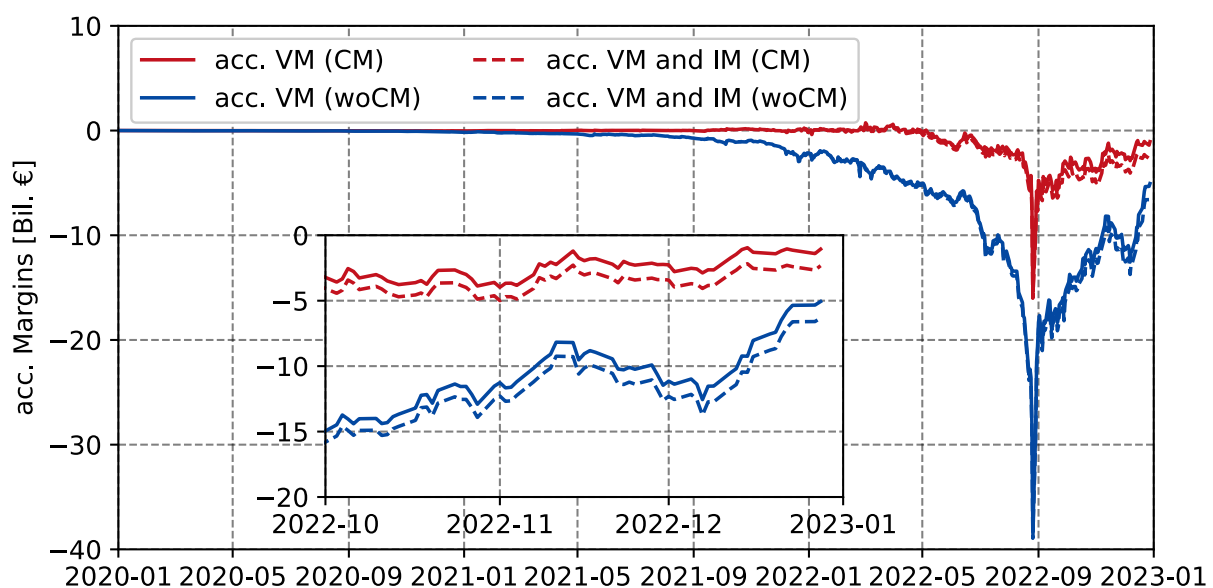


Figure 6: Calculated accumulated Variation Margin (solid lines) and the total of Variation and Initial Margin (dashed lines). If cross commodity netting of exposures is possible (red), then margining requirements are substantially smaller than without

For the case where different channels are used such that cross margining is not enabled (blue curves), individual exposures add up to a significantly larger total. A distinction can also be made between pure variation margin (solid lines) and the additional consideration of initial margin (dashed lines). It should be noted again that a negative value means a payment of margin or a deposit of cash in a margin account, while positive values mean a receipt of margin.

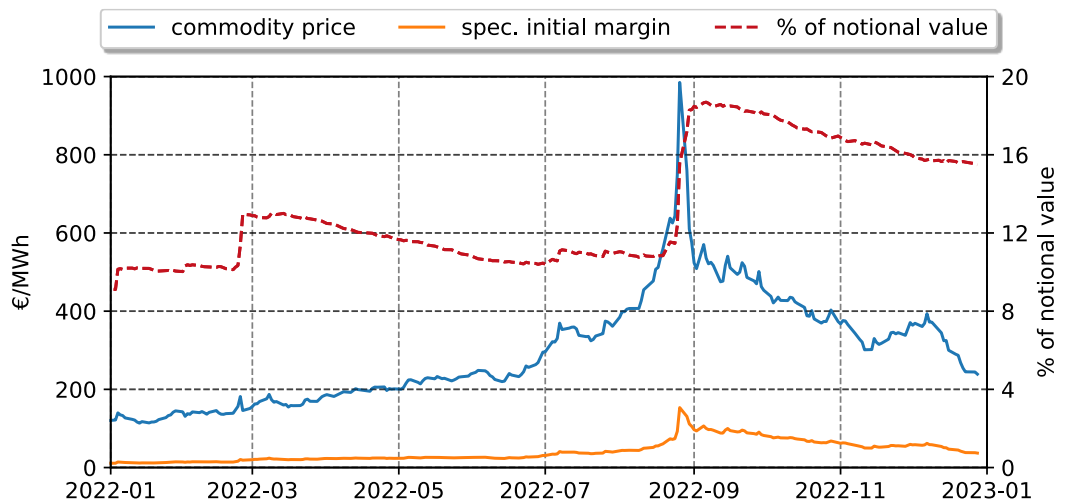
While margin calls remain relatively low initially, they start to rise sharply from the Q4 2021 onwards in line with price developments and accumulated hedge volumes affecting variation margins. As the net exposure is proportionally dominated by electricity, the operator will have to make high margin payments from then on. The absolute maximum of cash to be deposited (minimum of the curve) occurs on 26/08/2022, the day of the price peaks for electricity (€985/MWh) and gas (€314/MWh) as markets were shaken up with the events around the destruction of North Stream 1/2 pipelines.

With cross margining, around €16bn needs to be deposited on the margin account on this day, with a variation margin of €8.5bn to be paid on this day. Without cross margining, these figures would rise to €38bn and €10bn. Afterwards, prices fall steadily until the end of the year, easing the operator's liquidity situation, which however remains tight for the rest of the year.

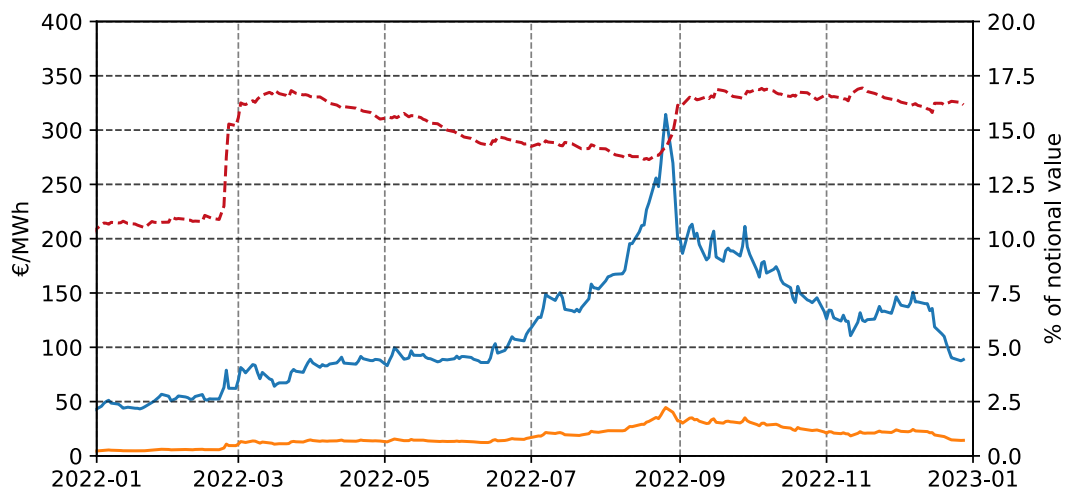
In absolute terms as shown in Fig. 6, initial margins appear small compared to spiking variation margins driven by spiking commodity prices. As such it is useful to illustrate initial margins also relative to underlying commodity prices. In fact, consideration of initial margin requirements in hedging decisions is particularly relevant at high wholesale prices. Fig. 7a-c illustrate the development of calculated initial margins alongside underlying commodity prices for electricity, gas and CO₂ emission allowances. Expressed as a percentage of notional values of underlying commodity prices, i.e., initial margins divided by the commodity price, we can demonstrate that initial margins have varied between 10-20% for electricity with higher volatility leading to higher initial margins (Fig 7a). During price spikes, initial margins for electricity were thus around two times higher already in relative terms compared to calmer periods. In addition, higher absolute price levels amplify the impact. At a five times higher price level and at a 20% relative initial margin, cash needs would be around ten times higher at the time of concluding a (new) transaction. For gas, initial margins were also driven to a higher level at the outbreak of the war end of February 2022. Thereafter initial margins remain at an elevated level (Fig 7b). In contrast, initial margin requirements on CO₂ emission allowances remained more controlled throughout the crisis because of sideways price movements and a more controlled volatility (Fig 7c). Note that relative initial margins expressed as % of notional values may feature jumps between levels at the time of price spikes followed by a relaxation period. This is due to dividing initial margins which are a function of a smoothing volatility parameter of relative price returns in our analysis by commodity prices itself. Other volatility models may thus present different dynamics in detail. Overall conclusions on the relevance of initial margins particularly in tight markets with elevated price and volatility levels remain unchanged.

Comparing the impact of variation margins and initial margins, a structural difference can be highlighted. Initial margins reflect a direct cash out at the time of concluding a transaction. The absolute exposure a generator may face is then influenced by new transactions and highest at perceived price spikes. In contrast, variation margin reflects cash risk with highly uncertain direction and magnitude. For electricity as the dominant driver of overall variation margins in the case study, gradual trends towards high prices lead to enormous cash outs. Rapidly declining prices after a price spike then reduce variation margins not only due to a lower delta to initial hedges but also due to negative variation margin from more recent transactions compensating positive variation margin from older transactions if hedges can be carried out constantly at the (same) exchanges. These effects are the reason for the rapid decline of variation margin in Fig. 6 while initial margins remain higher proportionally.

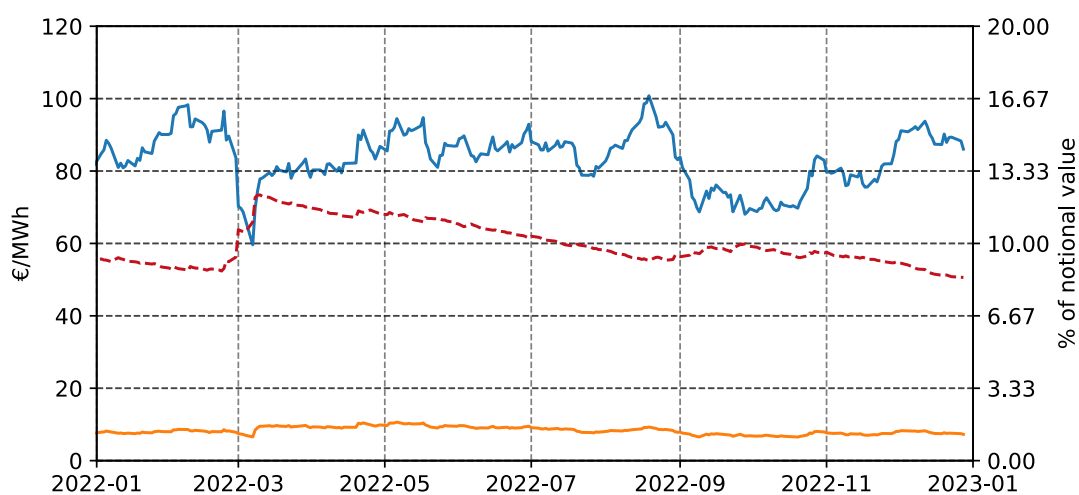
Relative to the portfolio size, cash pay-outs in the historic simulation at peaks close to €20bn are in any case threatening the modelled asset owner. While the price evolution in an unprecedented amount may be regarded as extreme, it is unclear if a repeat of such a scenario would reflect a once in hundred years or more than once in twenty years event. This uncertainty has in any case implications for any generator who may solely rely on collateralised trading activities.



a. Electricity



b. Gas



c. CO₂

Figure 7: Initial margin requirements (orange solid line) over time for commodities a. electricity, b. gas and c. CO₂ emission allowances along with underlying commodity prices (solid blue line) on the left axis in €/MWh. On the right-hand axis, initial margin requirements are shown in % of the notional commodity value (red dashed line).

3 Practical advice on how to cope with the Risk Trilemma

3.1 Scenarios impacting the risk trilemma and optimisation of trading channels

Overall, we show that hedging activities via futures/forwards lead to exchanging electricity and fuel price risks for (il-)liquidity risks due to clearing obligations or exceeding a collateral threshold agreed with a counterparty to secure credit risks. This risk trilemma of price/liquidity/credit risk became highly relevant notably at the end of 2021 and throughout the majority of 2022. The deterioration of Europe's pipeline based natural gas supply has dominated both market price movements and the decision-making of business leaders and policymakers. Based on the case study, the following advice can be given on how to deal with the risk trilemma.

Power plant operators should have a baseline approach to energy trading within the company's risk management system but be aware that in extreme situations it is necessary to deviate from this to optimise risk/reward trade-offs. One crucial foundation is a mix of uncollateralised OTC transactions and collateralised OTC or exchange transactions that is both cost effective in calm periods and reasonably robust against disruptions. Planning of mitigations actions should be done already in calm periods, i.e., at lower market volatility. This may include the definition of desired (negotiation) steps with counterparties at an early stage to avoid either trade stops on all uncollateralised channels or excessive margin requirements in collateralised channels.

In the event of unusual market developments, power plant operators and other market participants should be familiar with the tools for transformation within the risk trilemma, e.g., EFPs. Cross-margining should be used whenever possible to minimise the liquidity impact of margin calls. However, in some cases this may be difficult due to market liquidity. One example is the difference in gas trading at different exchanges. Volumes on the ICE were significantly higher than on EEX. At least in case of different exchange channels, clearing bank arrangements become more involved.

Thresholds on unsecured credit risk should be set highest for counterparties which are not at risk of defaulting in extreme prices either due to balance sheet robustness relative to exposures or due to critical relevance in the physical system. Monitoring and pattern recognition can help to identify market developments at an early stage and are therefore useful as early warning tools. At the same time, global (political) events must be closely observed, as they have an increasing influence on the development of wholesale markets and can fundamentally change the situation within a brief period.

3.2 Relevance of risk metrics and their dynamic adjustment

Quantitative assessments of risk factors and potentially changing risk drivers will typically require careful consideration of risk metrics. As evident in initial margin assumptions as per formula (2), different exchanges may require different initial margins. Next to intuitive risk factors, there is a relevance in statistical methods e.g., in volatility parameters. While volatility calibration might be an art in itself with the utilisation of absolute, relative or logarithmic returns, the same applies to sampling windows of historic data or reverse engineering of implied volatilities from traded options prices, where available. In our view, it is relevant for generators to be aware of how metric choices may impact results of assumed margins. The practical importance lies in the observed differences between exchange venues on initial margins to be posted at deal closure or their evolution over time with updated market parameters leading to one over the other exchange requiring higher initial margins.

Compared to initial margins, variation margins are more clearly defined with the mark-to-market difference of contract versus market prices. However, any risk assessment of how variation margins may evolve over time require equal consideration of risk drivers and value-at-risk metrics linked to the risk appetite of the generator. The same applies to credit risks where bilateral trading is preferred or commercially beneficial. Ultimately, generators and other market participants will aim for a commercial risk/reward optimisation linked to profitability objectives. For all market risk or hedge consideration to be sufficiently robust against realistic and also more extreme scenarios, decisions on attractive price levels or residual risks will thus require at best a very integrated consideration of risk trilemma drivers and risk metrics.

References

- Bouveret, A., Di Nello, D., Gutierrez Curos, J., Haferkorn, M. & Viollet, F. (2022). TRV Risk Analysis EU natural gas derivatives markets: risks and trends. ESMA. https://www.esma.europa.eu/sites/default/files/2023-05/ESMA50-165-2483_TRV-EU_natural_gas_derivatives_markets.pdf
- Bundesministerium der Justiz (2020), Verordnung über die Emissionsberichterstattung nach dem Brennstoffemissionshandelsgesetz für die Jahre 2021 und 2022 (Emissionsberichterstattungsverordnung 2022 - EBeV 2022): EBeV 2022, 2020.
- Danielsson, J. (2011). Financial Risk Forecasting. Wiley
- EnBW Integrierter Geschäftsbericht. (2022). <https://www.enbw.com/integrierter-geschaeftsbericht-2022/>
- E. ON SE, "Group Equity Story: Group Equity Story. March 2012", 2012. Zugriff am: 9. Januar 2024. <https://www.eon.com/en/investor-relations/presentations.html>
- European Commodity Clearing AG (2022), "ECC Derivative Market Margining", European Commodity Clearing AG, Apr. 2022.
- European Market Infrastructure Regulation "EMIR 3.0" (2022), Brussels, 7 Dec 2022, "REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - amending Regulations (EU) No 648/2012, (EU) No 575/2013 and (EU) 2017/1131 as regards measures to mitigate excessive exposures to third-country central counterparties and improve the efficiency of Union clearing markets"; <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0697>
- Financial Stability Board (2023). The Financial Stability Aspects of Commodities Markets. <https://www.fsb.org/wp-content/uploads/P200223-2.pdf>
- Frontier Economics and Luther (2024), report summary prepared for Energy Traders Europe, April 2024, "Principles of energy market regulation – securing efficient and resilient energy trading"; detailed report of this title for download from the Energy Traders Europe, European Union, 19 April 2024, <https://www.energytraderseurope.org/documents/principles-of/>
- Holthausen, D. (1979). Hedging and the Competitive Firm Under Price Uncertainty. American Economic Review. Nr. 5, P. 989–995. <https://econpapers.repec.org/RePEc:aea:aecrev:v:69:y:1979:i:5:p:989-95>
- Hull, J. (2022). *Optionen, Futures und andere Derivate*, 11. Aufl. München: Pearson.
- Lang, J. & Madlener, R. (2010). Relevance of Risk Capital and Margining for the Valuation of Power Plants: Cash Requirements for Credit Risk Mitigation. ENERDAY 2010. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1620514
- Lehrbass, F. (2021). Corporate liquidity risk management: Coping with Corona and the clearing obligations. International Journal of Management Research and Economics. Nr. 1. P. 1-26. <https://doi.org/10.51483/IJMRE.1.1.2021.1-26>
- Monitoringbericht (2023), Bundesnetzagentur und Bundeskartellamt, Stand 29. November 2023 <https://data.bundesnetzagentur.de/Bundesnetzagentur/SharedDocs/Mediathek/Monitoringberichte/MonitoringberichtEnergie2023.pdf>
- RWE AG (2022), "RWE Geschäftsbericht 2022", Zugriff am: 20. Dezember 2023. [Online]. Verfügbar unter: <https://www.rwe.com/-/media/RWE/documents/05-investor-relations/finanzkalender-und-veroeffentlichungen/2022-GJ/2023-03-21-rwe-geschaeftsbericht-2022.pdf>
- RWE AG (2020), "Q1-Q3 2020 Results", Essen. Zugriff am: 24. Juni 2024. https://www.rwe.com/-/media/RWE/documents/05-investor-relations/finanzkalender-und-veroeffentlichungen/2020-Q3/RWE-Presentation-Q1-Q3-2020.pdf?sc_lang=en
- Römmich, M. (2018). Optimale Produktions- und Hedging-Entscheidungen auf dem Großhandelsmarkt für Strom. PhD-Thesis. University Essen. <https://nbn-resolving.org/urn:nbn:de:hbz:464-20180309-134949-9>
- Schwabe, Ley und Greiner (2023), "TreasuryLog", Wien.

- Sparla, T., Lehrbass, F., Römmich, M. (2023), "Risk Management in Times of War: Case Study - to hedge or not to hedge", in *Rethinking Finance*, Bd. 05/2023, S. 26–32.
- Uniper Geschäftsbericht. (2022). https://www.uniper.energy/system/files/2023-03/2023-03-17_GJ_2022_Uniper_Gesch%C3%A4ftsbericht_.pdf
- Uniper “Capital Markets Story July 2022”, https://www.uniper.energy/sites/default/files/2022-07/2022-07_Uniper_Capital_Markets_Story_1.pdf Vattenfall “Interim Report January – June 2023”, Stockholm, 2023, https://group.vattenfall.com/siteassets/corporate/investors/interim_reports/2023/q2_report_2023.pdf
- Ziggel, D., Behrens, T., Weiß, G. & Wied, D. (2014). A new set of improved Value-at-Risk backtests. *Journal of Banking & Finance*. Nr. 48. P. 29-41. <https://doi.org/10.1016/j.jbankfin.2014.07.005>

Einfluss weiblicher Nachfolge auf Risiken in mittelständischen Familienunternehmen

Extended Abstract

Prof. Dr. habil. Patrick Ulrich

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: Patrick.ulrich@hs-aalen.de

Natalie Sorg, M.A.

Hochschule Aalen, Aalener Institut für Unternehmensführung, E-Mail: Natalie.sorg@hs-aalen.de

Abstract

Familienunternehmen sind weltweit verbreitet und spielen eine zentrale Rolle in der globalen Wirtschaft. Die Nachfolge stellt einen kritischen Punkt dar, der erhebliche Risiken birgt. Dieses extended Abstract untersucht die geschlechtsspezifischen Unterschiede im Umgang mit Risiken bei der Unternehmensnachfolge, da Nachfolgerinnen besonderen Herausforderungen gegenüberstehen. Die Ergebnisse zeigen, dass Frauen in Führungspositionen, abhängig vom Unternehmenskontext, unterschiedlich risikobereit sind im Vergleich zu Männern. Frauen bringen vielfältige Perspektiven und Herangehensweisen ein, wodurch Risiken effektiver identifiziert und gemanagt werden können. Gleichzeitig können durch die Einbindung von Frauen in die Nachfolge neue Risiken entstehen, beispielsweise in Bezug auf die Akzeptanz durch Stakeholder. Weitere empirische Untersuchungen sind erforderlich, um die spezifischen Risiken, Kontextfaktoren und situativen Anforderungen zu identifizieren, die diese geschlechtsspezifischen Unterschiede beeinflussen, um so die Langzeitstabilität von Familienunternehmen zu sichern.

1 Einführung

Familienunternehmen sind weltweit die häufigste Form von Unternehmensorganisationen und spielen eine zentrale Rolle in der Weltwirtschaft (Massis & Rondi, 2020). Sie weisen spezifische Merkmale auf, die ihre Risiken beeinflussen, insbesondere in den Bereichen Nachfolge, Governance, Eigentum, Entscheidungsfindung und Unternehmensleistung (Ward, 2004).

Die Nachfolge ist ein besonders risikoreiches Thema, da sie einen kritischen Punkt im Lebenszyklus eines Familienunternehmens darstellt. Im Gegensatz zu Publikumsgesellschaften, in denen Leitungsfunktionen regelmäßig wechseln, fehlt es Familienunternehmen oft an Erfahrungswerten aufgrund der konstanten Unternehmensleitung (Duller et al., 2023). Eine unklare Rollenverteilung sowie mangelnde Kommunikation können bei der Nachfolge in Familienunternehmen zu Streitigkeiten oder Rivalitäten führen, die eine schwerwiegende Beschädigung des Unternehmens zur Folge haben können (Visser & van Scheers, 2018). Diese Faktoren zeigen, dass der Erfolg des Nachfolgeprozesses wesentlich von einer intensiven und vielseitigen Vorbereitung abhängt, da er risikoreich und häufig zum Scheitern verurteilt ist.

Gemäß dem Institut für Mittelstandsforschung (IfM) Bonn stehen in den Jahren 2022 bis 2026 rund 190.000 Unternehmen zur Übergabe an (Fels et al., 2021). Diese Zahl veranschaulicht die steigende Anzahl von Unternehmensübergaben und betont die Relevanz einer strategischen Nachfolgeplanung. In Anbetracht dieser Entwicklungen gewinnt die Auseinandersetzung mit der Unternehmensnachfolge an Bedeutung, da zum einen die Zahl der verfügbaren Nachfolger begrenzt ist und zum anderen unvorhersehbare Übergaben aufgrund des demografischen Wandels zunehmen (Schwartz, 2023). Die Alterung der Gesellschaft sowie der Rückgang jüngerer Nachfolger unterstreichen die Dringlichkeit einer sorgfältigen Nachfolgeplanung, um potenzielle Risiken zu minimieren.

Sofern das Unternehmen weiterhin innerhalb der Familie fortgeführt werden soll, müssen traditionelle Nachfolgemuster, die insbesondere Männer und die Primogenitur bevorzugen, überwunden werden (Lyons et al., 2024). Diese Veränderungen führen zu einem Wandel in der Nachfolge, bei dem zunehmend auch weibliche Nachkommen in das Familienunternehmen integriert werden. Die Präferenz von Familienunternehmen, das Unternehmen innerhalb der Familie weiterzuführen, erfordert eine intensive Auseinandersetzung mit der Thematik der weiblichen Nachfolge und den damit verbundenen Risiken.

Die verstärkte Einbindung von Töchtern in die Nachfolge kann für Familienunternehmen neue Chancen, aber auch spezifische Risiken mit sich bringen, die es frühzeitig zu identifizieren gilt. Eine umfassende Risikoidentifikation ist entscheidend, um geeignete Strategien zur Risikominimierung zu entwickeln und eine erfolgreiche familieninterne Nachfolge sicherzustellen.

Die zentrale Forschungsfrage lautet daher: *Wie beeinflusst die Einbindung von Frauen in die Nachfolge die Risiken in mittelständischen Familienunternehmen?*

2 Methodik

Die methodische Grundlage dieser Untersuchung bildet eine systematische Literaturrecherche, die bis Dezember 2024 durchgeführt werden soll. Ziel dieser Recherche ist es, alle relevanten Arbeiten zur weiblichen Nachfolge in Familienunternehmen und deren Auswirkungen auf die unternehmerischen Risiken zu analysieren.

Für die systematische Literaturrecherche werden die Datenbanken Econbiz, JSTOR und Scopus genutzt. Diese Datenbanken wurden aufgrund ihrer umfassenden Abdeckung wirtschaftswissenschaftlicher, managementbezogener und interdisziplinärer Forschungsarbeiten ausgewählt. Die Suche konzentriert sich auf Veröffentlichungen ab dem Jahr 2017, um die neuesten Forschungsergebnisse zu berücksichtigen. In Anlehnung an Campopiano et al., 2017 erfolgt die Suche nach den folgenden Schlüsselwörtern im Titel oder Abstract: „family business“ oder „family firm“

und „gender“ oder „woman“ oder „female“ oder „wife“ oder „daughter“ oder „sister“ oder „spouse“. Publikationen, die sich nicht mit wirtschaftswissenschaftlichen oder managementbezogenen Themen befassen, die nicht spezifisch auf Nachfolge oder Risiken eingehen, sowie solche, die sich stark von Deutschland unterscheidende Regionen behandeln (z. B. Nachfolgerinnen im arabischen Raum), werden von der weiteren Betrachtung ausgeschlossen. Weiterhin sollen nur Artikel in deutscher und englischer Sprache berücksichtigt werden. Darüber hinaus soll ein Backward Reference Searching durchgeführt werden, um weitere relevante Publikationen zu identifizieren, die für die Forschung von Bedeutung sein könnten.

Zusätzlich zu dieser Literaturrecherche sind Interviews mit Nachfolgerinnen und Unternehmern geplant, die sich insbesondere auf die Wahrnehmung der Herausforderungen und Risiken der Nachfolge in Familienunternehmen durch Frauen konzentrieren. Diese qualitativen Interviews sollen tiefere Einblicke in die persönlichen Erfahrungen und Perspektiven der Beteiligten geben und die aus der Literatur gewonnenen Erkenntnisse ergänzen. Die Auswahl der Unternehmen erfolgt anhand theoretisch hergeleiteter quantitativer und qualitativer Mittelstandscharakteristika. Dieses Vorgehen bietet die notwendige Flexibilität und Tiefe, um detaillierte Einblicke in die subjektiven Sichtweisen der Beteiligten zu erhalten und sicherzustellen, dass die Interviewergebnisse die Risiken und Dynamiken widerspiegeln, mit denen mittelständische Familienunternehmen im Rahmen der weiblichen Nachfolge konfrontiert sind. Darüber hinaus ermöglicht die Befragung eine spezifische Betrachtung verschiedener Typen von Familienunternehmen, um zu ermitteln, welche Risiken in unterschiedlichen Unternehmenskontexten auftreten.

3 (Vorläufige) Ergebnisse

Aktuelle Forschungsergebnisse beleuchten die Entwicklung von Risiken in Familienunternehmen, insbesondere im Kontext der Beteiligung von Frauen an der Unternehmensnachfolge. Die Erwartungen an die nächste Generation in Familienunternehmen sind häufig von traditionellen männlichen Rollenbildern geprägt, die Macht und Dominanz betonen (Kammerlander & Wilmes, 2024). Dies kann dazu führen, dass Töchter, die diesen Stereotypen nicht entsprechen, von der Unternehmensnachfolge ausgeschlossen werden oder das Risiko besteht, dass ihre Eignung für die Nachfolge übersehen wird.

Frauen werden oft als weniger selbstbewusst in Bezug auf ihre unternehmerischen Fähigkeiten wahrgenommen und als risikoaverser eingeschätzt, was sich auf ihren Einstieg oder Wiedereinstieg in die Geschäftswelt auswirken kann (Maseda et al., 2022). In Fällen, in denen die Tochter die Nachfolge antritt, tendiert der Gründervater dazu, die Führungsrolle stärker zu kontrollieren und bevorzugt in risikoarme Innovationen zu investieren (Zhu & Kang, 2022). Dies verstärkt den Eindruck, dass Töchter als Führungskräfte weniger geeignet erscheinen.

Die Führungsfähigkeiten von Töchtern in Familienunternehmen werden von Stakeholdern häufig nicht automatisch anerkannt, was zu einer unzureichenden Einbindung und Sozialisierung führen kann (Hytti et al., 2017). Bei externen Partnern wie Banken oder Lieferanten besteht oftmals eine gewisse Skepsis hinsichtlich der Führungsqualitäten von Nachfolgerinnen, sodass diese ihre Kompetenz wiederholt unter Beweis stellen müssen, um als gleichwertige Verhandlungspartnerinnen anerkannt zu werden (Kubíček & Machek, 2019). Diskriminierungen und Sticheleien durch männliche Kollegen können bei Nachfolgerinnen zu einer Beeinträchtigung des Selbstbewusstseins und zu Unsicherheit führen, was das Risiko von Krisen in Familienunternehmen erhöht (Kammerlander & Wilmes, 2024).

Die Skepsis von Stakeholdern und externen Partnern gegenüber den Führungskompetenzen von Nachfolgerinnen in Familienunternehmen birgt das Risiko, dass diese in ihrer Rolle nicht vollständig akzeptiert und integriert werden. Dies kann zu einer ineffektiven Führungsstruktur führen, in der Nachfolgerinnen ihre Kompetenz ständig unter Beweis stellen müssen, um anerkannt zu werden (Kubíček & Machek, 2019). Dies beeinträchtigt nicht nur die Effizienz der Unternehmensführung, sondern kann auch die Entscheidungsfindung verlangsamen und die Umsetzung wichtiger strategischer Maßnahmen behindern. Darüber hinaus kann die mangelnde Akzeptanz und das Infragestellen ihrer Führungsqualitäten das Selbstvertrauen der Nachfolgerinnen untergraben, was das Risiko negativer Auswirkungen auf die Gesamtleistung und die Zukunft des Unternehmens erhöhen kann (Jiang et al., 2024).

Die Forschungsliteratur legt nahe, dass die Präsenz von Geschäftsführerinnen einen positiven Einfluss auf das Risikoverhalten von Familienunternehmen hat. Dies steht im Widerspruch zur traditionellen Annahme, dass Frauen eine geringere Risikobereitschaft aufweisen (Tao-Schuchardt & Kammerlander, 2024). Es konnte gezeigt werden, dass der Anteil unabhängiger weiblicher Mitglieder im Vorstand das Wagnisrisiko erhöht, ohne das Risiko der Leistungsgefährdung zu beeinträchtigen. Im Gegensatz dazu erhöht der Anteil nicht-unabhängiger weiblicher Vorstandsmitglieder das Risiko der Leistungsgefährdung signifikant, insbesondere in Familienunternehmen (Poletti-Hughes & Briano-Turrent, 2019).

Die Präsenz von Frauen in leitenden Positionen innerhalb von Familienunternehmen kann zu einer Veränderung des Risikomanagementverhaltens führen, indem eine größere Vielfalt an Sichtweisen und Ansätzen bei der Risikobewertung Berücksichtigung findet (Maseda et al., 2022). Die Erhöhung der Geschlechterdiversität führt in der Konsequenz zu einer breiteren Risikoeinschätzung sowie einer gesteigerten Innovationsbereitschaft, wodurch proaktivere und kreativere Risikomanagementstrategien gefördert werden (Tao-Schuchardt & Kammerlander, 2024). Die Beteiligung von Frauen an Entscheidungsprozessen trägt zu einem umfassenderen Verständnis und Management von Risiken bei, was wiederum die Fähigkeit von Unternehmen verbessert, flexibel und effektiv auf Herausforderungen zu reagieren.

4 Diskussion und Zusammenfassung

Die vorliegenden Ausführungen verdeutlichen, dass Frauen als Nachfolgerinnen spezifischen Risiken ausgesetzt sind, wobei sich diese sowohl aus der Perspektive der Frauen als auch aus der Perspektive der Unternehmen betrachten lassen.

Der Umstand, dass Töchter oft eine Karriere außerhalb des Familienunternehmens bevorzugen, könnte auf eine unterschiedliche Risikobereitschaft hinweisen, deren Auswirkungen noch unklar sind (Halkias et al., 2010). Aktuelle Studien von Zona et al. (2023) zeigen, dass Frauen unter bestimmten Bedingungen genauso risikobereit oder sogar risikofreudiger sein können als Männer. Frauen in Führungspositionen weisen abhängig von ihrer sozialen Identität und den Erwartungen an ihre Rolle unterschiedliche Risikopräferenzen auf, wobei geschlechtsspezifische Unterschiede in der Risikobereitschaft stark vom sozialen und kulturellen Kontext abhängen.

Die Rolle von Töchtern in der Unternehmensnachfolge kann aufgrund ihrer tendenziellen Risikoaversion als vorteilhaft erachtet werden, da sie zur Sicherung des Unternehmensüberlebens beitragen könnte. Andererseits kann der erschwerte Zugang zu Finanzmitteln in Kombination mit einer übermäßigen Vorsicht, die Frauen oft davon abhält, finanzielle Mittel zu beantragen, die Expansion des Unternehmens erschweren und somit dessen Fortbestand gefährden (Kubíček & Machek, 2019).

Die dargestellten Forschungsergebnisse verdeutlichen, dass die Einbindung von männlichen und weiblichen Nachfolgern unterschiedliche Risiken birgt, wodurch die Notwendigkeit weiterer Forschungsarbeiten in diesem Bereich aufgezeigt wird. Obgleich die Risiken von Frauen in der Unternehmensführung hinsichtlich ihrer Auswirkungen auf Innovation, finanzielle Leistung und Entscheidungsfindung bereits untersucht wurden, bleiben wichtige Fragen offen. Dazu gehören empirische Untersuchungen, die sich beispielhaft mit dem Akzeptanzrisiko von Stakeholdern, den Risiken der Doppelbelastung von Frauen und spezifischen Herausforderungen sowie den bisher wenig beachteten Risiken bei Mutter-Tochter-Übergaben befassen.

Es ist erforderlich, geschlechtsspezifische Unterschiede im Umgang mit Risiken und deren Wechselwirkungen eingehend zu analysieren, um effektive Risikomanagementstrategien für eine erfolgreiche und sichere familieninterne Unternehmensnachfolge zu entwickeln und somit die Langzeitstabilität von Familienunternehmen zu gewährleisten.

Literatur

- Campopiano, G., Massis, A. de, Rinaldi, F. R. & Sciascia, S. (2017). *Women's involvement in family firms: Progress and challenges for future research*. *Journal of Family Business Strategy*, 8(4), 200–212. <https://doi.org/10.1016/j.jfbs.2017.09.001>
- Duller, C., Hiebl, M., Kuttner, M., Mayr, S. & Mitter, C. (2023). *Herausforderungen im Management von Familienunternehmen: Werte - Wachstum - Controlling - Krisenbewältigung* (1st ed. 2023). Springer Fachmedien Wiesbaden; Springer Gabler. <https://doi.org/10.1007/978-3-658-41978-3>
- Fels, M., Suprinovič, O., Schlömer-Laufen, N. & Kay, R. (2021). *Unternehmensnachfolgen in Deutschland 2022 bis 2026* (Daten und Fakten Nr. 27). Bonn: Institut für Mittelstandsforschung (IfM) Bonn. <https://www.econstor.eu/handle/10419/248848>
- Halkias, D., Thurman, P. W., Harkiolakis, N., Katsioloudes, M., Stavrou, E. T., Swiercz, P. M. & Fragoudakis, M. (2010). Father-daughter succession issues in family business among regional economies of Asia. *International Journal of Entrepreneurial Venturing*, 2(3/4), Artikel 37116, 320. <https://doi.org/10.1504/IJEV.2010.037116>
- Hytti, U., Alsos, G. A., Heinonen, J. & Ljunggren, E. (2017). Navigating the family business: A gendered analysis of identity construction of daughters. *International Small Business Journal: Researching Entrepreneurship*, 35(6), 665–686. <https://doi.org/10.1177/0266242616675924>
- Jiang, Y., Cheng, W. & Xie, X. (2024). The Dismissal of New Female CEOs: A Role Congruity Perspective. *Journal of Business Ethics*. Vorab-Onlinepublikation. <https://doi.org/10.1007/s10551-024-05630-1>
- Kammerlander, N. & Wilmes, R. (2024). *Nachfolgerinnen in Familienunternehmen: Untersuchung zu den Chancen und Herausforderungen für Nachfolgerinnen in Familienunternehmen*.
- Kubiček, A. & Machek, O. (2019). Gender-related factors in family business succession: a systematic literature review. *Review of Managerial Science*, 13(5), 963–1002. <https://doi.org/10.1007/s11846-018-0278-z>
- Lyons, R., Ahmed, F. U., Clinton, E., O’Gorman, C. & Gillanders, R. (2024). The impact of parental emotional support on the succession intentions of next-generation family business members. *Entrepreneurship & Regional Development*, 36(3-4), 516–534. <https://doi.org/10.1080/08985626.2023.2233460>
- Maseda, A., Iturralde, T., Cooper, S. & Aparicio, G. (2022). Mapping women's involvement in family firms: A review based on bibliographic coupling analysis. *International Journal of Management Reviews*, 24(2), 279–305. <https://doi.org/10.1111/ijmr.12278>
- Massis, A. de & Rondi, E. (2020). Covid-19 and the Future of Family Business Research. *Journal of Management Studies*, 57(8), 1727–1731. <https://doi.org/10.1111/joms.12632>
- Poletti-Hughes, J. & Briano-Turrent, G. C. (2019). Gender diversity on the board of directors and corporate risk: A behavioural agency theory perspective. *International Review of Financial Analysis*, 62, 80–90. <https://doi.org/10.1016/j.irfa.2019.02.004>
- Schwartz, M. (2023). Nachfolge-Monitoring Mittelstand 2022: Knappheit an Nachfolgekandidaten nimmt zu, Misserfolge dürften häufiger werden. *KfW Research Fokus Volkswirtschaft*(Nr. 424), Artikel 28.
- Tao-Schuchardt, M. & Kammerlander, N. (2024). Board diversity in family firms across cultures: A contingency analysis on the effects of gender and tenure diversity on firm performance. *Journal of Family Business Strategy*, 15(2), 100554. <https://doi.org/10.1016/j.jfbs.2023.100554>
- Visser, T. & van Scheers, L. (2018). Can family business managers manage family business risks? *Management Journal of Contemporary Management Issues*, 23(1), 123–137. <https://doi.org/10.30924/mjcmi/2018.23.1.123>
- Ward, T. B. (2004). Cognition, creativity, and entrepreneurship. *Journal of Business Venturing*, 19(2), 173–188. [https://doi.org/10.1016/S0883-9026\(03\)00005-3](https://doi.org/10.1016/S0883-9026(03)00005-3)
- Zhu, Z. & Kang, Y. (2022). A Far-Reaching Parental Love? Co-Governance of Intergenerational Succession and Innovation Activities in Chinese Family Firms. *Management and Organization Review*, 18(2), 358–394. <https://doi.org/10.1017/mor.2021.45>

Zona, F., Pesci, C. & Zamarian, M. (2023). CEO risk preferences in family firms: Combining socioemotional wealth and gender identity perspectives. *Journal of Family Business Strategy*, 100553. <https://doi.org/10.1016/j.jfbs.2023.100553>

4

Finanzen



Sustainability Reporting - A Global Analysis of Sustainability Reporting and its Impact on Cost of Capital - A Bibliographic and Content Analysis

Research Paper

Prof. Dr. Guido Sopp CFE

Westfälische Hochschule Zwickau, Fakultät Wirtschaftswissenschaften, E-Mail: guido.sopp@fh-zwickau.de

Björn Bunzel MBA

Westfälische Hochschule Zwickau, Fakultät Wirtschaftswissenschaften, E-Mail: bjoern.bunzel.llw@fh-zwickau.de

Abstract

In recent years, public pressure on companies regarding sustainable corporate performance (CSR performance) and appropriate CSR disclosure has steadily increased. The aim of this study is to add to the existing research on the relationship between CSR performance, CSR disclosure and cost of capital, to analyse possible trends and to identify the impact of policy initiatives. In using bibliographic and content analysis, we found that while the number of research papers on CSR disclosure, performance and cost of capital is increasing, research on the individual dimensions of ESG is limited. We also found that studies focusing on specific industries are rather rare compared to studies covering all types of industries. Considering the very different impact of the diverse environment-, social- and governance-related regulations on the various economic sectors, we believe it would be useful to evaluate a more sector-oriented analysis of the individual dimensions of ESG in future research.

1 Introduction

There is a longstanding academic interest in the relation of disclosure and cost of capital. Studies like those of Diamond and Verrecchia (1991), Botosan (1997), Leuz and Verrecchia (2000) and Botosan and Plumlee (2002) focus on the impact of disclosure on cost of capital in general or, more specifically, on the impact of financial disclosure on cost of capital. In the last two decades, there is an increasing academic interest in CSR performance or CSR disclosure and cost of capital. While CSR performance relates to a firm's internal actions and should provide a measure for CSR activities and socially responsible behaviour (Hinze and Sump, 2019), CSR disclosure refers to the transparency around environmental, social and governance matters. Other frequently used synonyms for CSR disclosure are ESG reporting and sustainability reporting. A recent bibliometric review of Benameur et al. (2022) evidenced more than 1,000 studies dealing with sustainability reporting, written by more than 2,000 scholars and covering almost 70 countries in the years 2000 to 2022. They observe "exponential growth" in the body of sustainability reporting related studies. Similarly, however with a more limited scope, Bosi et al. (2022) identified 358 relevant publications in the years 1998 to 2022. They highlight a "clear and positive trend in the publication of ESG and sustainability reports" over time, with a gradual increase from 2006 to 2018 and a peak in 2021.

The earlier review study of Friede et al. (2015) is broader in scope and provides aggregated evidence from more than 2,000 empirical studies in the area of ESG and financial performance. The scope is not limited to sustainability reporting, but rather focuses on E-, S- and G-performance, Corporate Social Responsibility at large and Corporate Governance.

Compared to other bibliometric reviews, Friede et al. (2015) cover the longest historical period, starting from as early as 1970. Similar to Benameur et al. (2022) and Bosi et al. (2022) they confirm a strong increase of the number of studies over time. Even though the high intensity of academic research in the area of CSR disclosure, CSR performance and cost of capital might indicate little need for additional contributions, the body of literature still seems to lack more focused, granular and content-based analyses at the E, S and G sub-levels in corporate CSR disclosures. We found that, while there are a number of recent review studies, they define distinct research objectives. The following Table 1 illustrates the latest review studies, their respective research methods and focus, the database being used and the applicable search terms.

We observe that the search terms used in those studies are rather broad, with a specific focus on CSR reporting or a combination of CSR reporting and performance. However, none of the most recent studies explicitly introduces search terms related to both CSR reporting *and* cost of capital. This study combines bibliometric and content analyses to further fill this gap.

More specifically, we aim to answer the following research questions:

Q1: What is the current state of research on the association of CSR disclosure and cost of capital in terms of research method, industry focus, geographical focus, and reference periods being used?

Q2: What proportion of studies explores the relationship between CSR disclosure and (a) cost of equity, (b) cost of debt and (c) combined cost of capital (e.g., Weighted Average Cost of Capital, WACC)?

Q3: What proportion of studies further breaks down CSR disclosures into environmental (E), social (S) and governance (G) disclosures?

Q4: What findings can be observed from the studies regarding the directional relationship between CSR disclosure and cost of capital, also distinguishing the different dimensions of (a) CSR disclosure and (b) cost of capital?

Table 1: Overview of most recent bibliographic studies and literature reviews, Created by authors

Author(s)	Research method and focus	Database	Search terms
Benameur et al. (2022)	bibliometric and content analyses of literature on sustainability reporting	Scopus	“Sustainability Reporting,” “Sustainability Disclosure,” “Corporate Social Responsibility Reporting,” “Corporate Social Responsibility Disclosure,” “CSR Reporting,” “CSR Disclosure”
Bosi et al. (2022)	bibliometric analysis of literature on environmental, social, and governance (ESG) and sustainability reporting	Scopus	“environment” AND “social” AND “governance” AND “sustainability” AND “reporting”
Osobajo et al. (2022)	bibliometric literature review of peer-reviewed studies on sustainability performance reporting	Web of Science and Scopus	“corporate sustainability reporting”, “socially responsible investment”, “global reporting initiative”, “sustainability index”, and “global reporting framework”
Khan (2022)	scientometric analysis of the sustainable corporate finance literature, analysis of the conceptual structure of studies in this field and classification into dominant research streams, meta-analysis and summary of the empirical findings	Scopus	"Environmental social and governance performance", "Environmental social and governance index", "Environmental social and governance ratings", "Environmental social and governance score", "Corporate social responsibility disclosure", "ESG performance", "ESG index", "ESG ratings", "ESG score", "CSR disclosure"
Nobanee et al. (2021)	bibliometric analysis of literature on sustainability practices and risk	Scopus	“sustainable” and “risk”

2 Research design

We analyse the development, the conceptual structure, the thematic evolution and the findings of academic research on CSR reporting *and* cost of capital based on Scopus peer-reviewed articles between 2010 and 2023. Following Block and Fisch (2020), bibliographic studies can be defined as systematic literature reviews that analyse bibliographic data with bibliometric methods. The focus is on the structure of a particular research field, while narrative literature reviews rather aim at summarising the content of studies. Our analysis goes beyond mere descriptive reporting on prior research, as we further interpret and discuss our findings and single out specific studies, if deemed necessary to support our conclusions. Unlike other bibliographic studies, we do not perform network or citation analysis, but focus on content and conclusions. The choice of combining bibliometric and content analysis offers a complementary approach to traditional literature reviews on the one and pure bibliographic studies on the other hand.

Our analysis follows a six-phase process: The first phase concerns the database selection and the creation of keywords for the database search (chapter 2.1). The second phase covers data collection, i.e. retrieving the data from Scopus (chapter 2.2), while the third phase deals with data extraction, that is extracting relevant information from Scopus and the individual research papers (chapter 2.3). In the fourth phase, we enhanced the extracted data by numerical codes (chapter 2.4) to enable a structured analysis in phase five (chapter 2.5). The sixth and final phase deals with the visualisation, interpretation, and conclusion (chapter 3).

2.1 Phase I: Selection of the database and creation of keyword chains

Selection of the database

The data for the analysis was taken from the Scopus March 2024 online database. This database is used in the study because, as explained by Echchakoui (2020), and others, it provides comprehensive coverage of articles from scientific publications on the topic. Scopus contains a wide range of scientifically analysable information about the publications such as access type, year, author name, subject, text type, source title, institution, nation, font and language. Another advantage of Scopus is that in addition to articles, the database contains other documents such as conference proceedings as well as books and that Scopus covers publications from developing countries more comprehensively (Echchakoui 2020).

Creating the keyword chain

Keywords were identified for the database analysis to enable a comprehensive literature search and to analyse the subject area broadly, as well as to ensure the largest possible population on the topic of sustainability reporting in connection with the cost of capital. On the other hand, searching with an adequate set of keywords narrows down the number of literature and allows for a more in-depth manual content analysis where necessary.

The search terms shown in Table 2 below were defined for the basic search function for sustainability reporting. This was stored as a search function for articles, abstracts, and keywords.

Table 2: Keyword structure sustainability reporting, Created by authors

Keyword	Related Keywords
sustainability reporting	sustainability disclosure; csr reporting; csr disclosure; non financial reporting; non financial disclosure

To expand the context of the search function, the references to the cost of capital and its equivalent keywords was added (Table 3).

Table 3: Keyword structure cost of capital, Created by authors

Keyword	Related Keywords
cost of capital	cost of debt; cost of finance; cost of equity

2.2 Phase II: Data collection

This study was conducted through four phases of data collection, consisting of (1) identification, (2) screening, (3) eligibility and (4) inclusion. For illustration, we used the PRISMA flow diagram, which is considered the standard protocol accepted by researchers in academia (Bosi et al. 2022).

To identify relevant publications, we used the keyword string as explained in chapter 2.1. The technical specification in the Scopus search mask was done as follows (Table 4).

Table 4: Keyword structure to the search function for “sustainability reporting”, Created by authors

Keyword	Related Keywords
"*sustainability reporting*"	"*sustainability disclosure*" OR "*csr reporting*" OR "*csr disclosure*" OR "*non financial reporting*" OR "*non financial disclosure*"

No filters were applied to the subject area and document type. English was selected as the language of publication. For data collection, the results in the publication backlog were limited to final articles in order to exclude possible unavailability.

This search entry yielded a total of 4,841 results in the database. After limiting to final articles, 4,672 results were displayed.

The technical specifications of the second chain of keywords are as follows (Table 5).

Table 5: Keyword structure to the search function for “cost of capital”, Created by authors

Keyword	Related Keywords
cost of capital	"*cost of debt*" OR "*cost of finance*" OR "*cost of equity*"

Execution of the search function in the database across the complete string of terms for sustainability reporting and the cost of capital provided 64 results. Technically, we used the following string: TITLE-ABS-KEY ("*sustainability reporting*" OR "*sustainability disclosure*" OR "*csr reporting*" OR "*csr disclosure*" OR "*non financial reporting*" OR "*non financial disclosure*") AND TITLE-ABS-KEY ("*cost of capital*" OR "*cost of debt*" OR "*cost of finance*" OR "*cost of equity*" OR "*cost of investing*") AND (LIMIT-TO (PUBSTAGE , "final")).

In the screening phase we first checked the availability of publications in a digital format. We found that all except for three studies were digitally available. As we could not gain access to the three remaining publications, we excluded them from the sample.

In the next step, we systematically screened the remaining 61 publications as regards eventual misclassifications. Therefore, we read the abstracts and, if necessary, selectively also the entire paper. We found that a total of seven

studies did not correspond to our search criteria. Three of those studies were only available in non-English languages, while the other four studies did not fit to the thematic focus of our study. Therefore, the final sample size comprises therefore 54 publications.

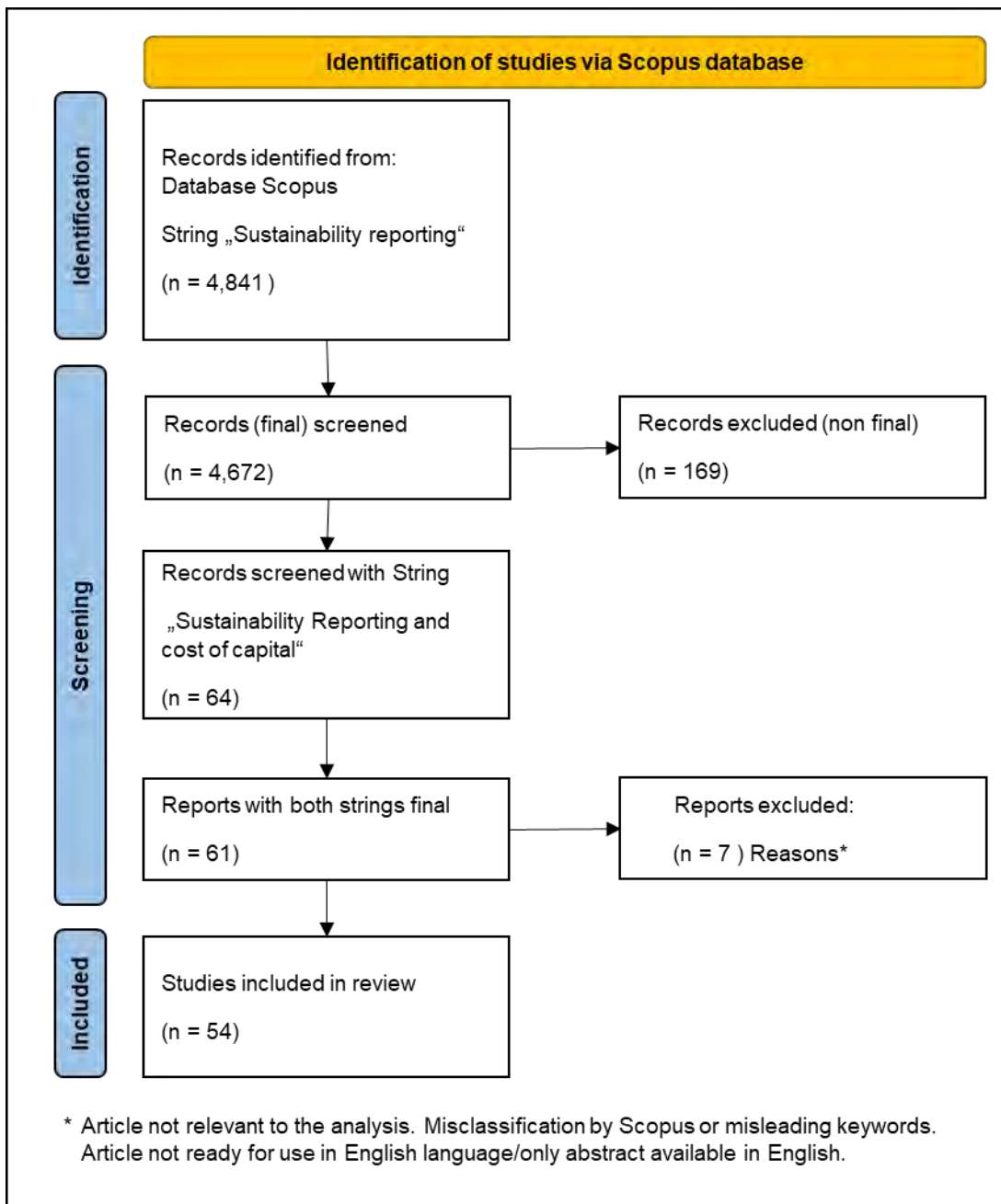


Figure 1: PRISMA Flow Diagram, Created by authors

We found several studies that refer to CSR performance instead of disclosure or reporting. Both dimensions are related, but different: In the CSR disclosure dimension, the degree and/or depth (“richness”) of reporting are explored. In turn, CSR performance relates to a firm’s internal actions and should provide a measure for CSR activities and socially responsible behaviour (Hinze and Sump, 2019). CSR disclosure should ideally provide a faithful picture of CSR performance. Against the background of voluntary (instead of mandatory) application of CSR reporting frameworks, such as the GRI, in most jurisdictions and the different possible levels of application within those frameworks, however, CSR disclosure does not necessarily reflect the true CSR performance of companies (also refer to García-Sánchez et al., 2021).

In this vein, Gutsche et al. (2017) argue that CSR disclosure mediates CSR performance through to a positivity bias – a relatively high amount of CSR disclosure is misinterpreted as good CSR performance – and conclude “words currently speak louder than actions”. García-Sánchez et al. (2021) specifically examine the gap between CSR disclosure and CSR performance (also referred to as “CSR decoupling”) and find, *inter alia*, that wider gaps result in greater cost of capital and reduced access to finance.

We considered excluding studies with a pure focus on CSR performance, as these studies do not perfectly match our research objectives. Unfortunately, we found that there is no common understanding of the terms CSR disclosure and CSR performance across the studies. In our view, it is challenging, if not arbitrary, to separate CSR disclosure from CSR performance studies. For example, ESG scores provided by Bloomberg were sometimes used as a measure of CSR disclosure, while other studies used them as a measure of CSR performance. We therefore decided to include the CSR performance studies in the sample but using separate numerical codes to allow for a more distinct analysis (see chapter 2.4).

Similarly, some studies more generally investigate the relation of CSR disclosure or CSR performance to some financial performance measures, not necessarily cost of capital. We decided to keep those studies in the sample with separate classification codes. The reason is that in some studies, cost of capital-analysis is combined with some other financial performance measure, and we didn’t want to lose information about those studies.

Few studies have a different, but still related research focus. Examples include studies that focus on CSR assurance, on the impact of integrated reporting (rather than standalone CSR reporting) on cost of capital and studies that explore the association of reporting the company’s annual Communication on Progress under the UN Global Compact (UNGC) and the cost of capital. We decided to retain those studies in the sample and do not expect any negative impacts on our analysis. Our coding system (chapter 2.4) allows for a sufficient degree of granularity, and we can therefore individualise our analysis using the Excel filter function.

2.3 Phase III: Data extraction

A tabular analysis was created in Excel for data preparation and for analysing the final sample. The following categorisation criteria were selected for this purpose:

Author(s); Publication date (year); Publication name; Sample companies; Geographical focus, Industry focus; Details sample; Reference period, Reference period start (year); Reference period end (year); Research question(s); Methodology/Design; Details methodology; Classification(s) research objective(s); Data source; Conclusion, Classification(s) conclusion(s); Future research – recommendations, Limitations; Journal; DOI.

Basic bibliometric data, such as author(s) name, publication date, publication name, Journal and DOI were extracted from Scopus. This information is used for bibliographic analysis of the evolution of relevant publications over time, publication by media type and associated field of study based on the categorisation of scientific journals.

The remaining information was manually collected from the publications and entered into the Excel file. We added further information in the “Details” columns for the sample companies and methodologies to allow for a further “deep dive”, if necessary, in the analytical phase.

Apart from the methodology/design, all information was either retrieved from Scopus or directly extracted from the individual studies. In this vein, we managed to limit the need for interpretations and judgment. As regards the research methodology, we decided on a rather high-level distinction of empirical studies and non-empirical studies (bibliometric studies, conceptual studies, and social sciences studies). This distinction is important, as for the content-based analysis of research objectives and research conclusions we limit the sample to empirical studies (for details see chapter 2.5).

2.4 Phase IV: Coding

In order to support a structural analysis of the publications in terms of research questions and conclusions we decided to use a coding system to categorise the different studies. We used double-digit numerical codes for categorisation of research questions and three-digit numerical codes for the categorisation of the conclusions reached in the primary studies.

For the categorisation of research questions, we decided to use double-digit numerical codes in order to make visible the relation of CSR on the one hand and cost of capital on the other hand. The first digit provides a categorisation in the cost of capital dimension, the second digit reflects the categorisation of the CSR leg of studies. As an example, the first digit number “1” stands for studies that focus on a combined cost of capital measure, like WACC, the first digit number “2” signals that the study examines the cost of debt relation, the first digit number “3” refers to cost of equity-studies. Based on this coding, a structured analysis as regards research question Q2 becomes possible.

We decided to use the second digit to differentiate studies that use a general measure for CSR disclosure from studies that more granularly examine the environmental, social or governance dimension. This allows us to address the research question Q3.

We allowed for multiple assignments, as some studies serve more than one research objective. For example, studies that explore the relationship of CSR disclosure, cost of equity, cost of debt and a combined cost of capital measure were assigned the numerical codes “11”, “21” and “31”.

We further differentiated studies that focus on the association of CSR disclosure and cost of capital (coding “11” to “34”) from studies that refer to CSR performance (coding “41” to “74”). Indeed, the differentiation of disclosure and performance is not straightforward, as performance can only be assessed based on the available set of information, including assessments by (sustainability) rating agencies and other intermediaries (such as financial analysts). As an example, Bloomberg highlights that its scores “are based on publicly available, company-disclosed data, and do not rely on estimates or an analyst’s opinion. As a result, the scores can be updated in a timely manner as companies release new data throughout the year. Scores also consider disclosure of quantitative data as a dimension of performance.” (see Bloomberg, 2023).

Further to this, CSR performance can be interpreted in several ways. First, and foremost, performance can refer to the impact of environmental, social, or governance factors on the company’s financial performance, such as revenue streams, operating costs, cost of capital or asset value. This is the common interpretation used by major rating agencies, such as Sustainalytics and Refinitiv, and in the ESG score methodology of Bloomberg (see Bloomberg, 2023). This dimension is sometimes referred to as the “outside-in” perspective of sustainability. Another possible approach is to define CSR performance from an “inside-out” perspective. In that case, the focus is on the impact of the business activities on environmental and/or social matters.

As Bloomberg ESG scores are based on publicly available data and do not rely on estimates or analysts’ opinions, most studies that use Bloomberg scores interpret these scores as a measure of CSR disclosure. In turn, the former Kinder, Lydenberg, and Domini (hereinafter, KLD) STATS (today MSCI ESG Research) are widely considered as a measure of CSR performance. We recognise that the methodologies of leading ESG rating agencies and data providers change from time to time. As of today, the Bloomberg ESG score is considered a “measure a company’s management of financially material ESG issues” (see Bloomberg, 2023). Similarly, MSCI ESG Ratings “aim to measure a company’s management of financially relevant ESG risks and opportunities” (see <https://www.msci.com/our-solutions/esg-investing/esg-ratings>). Therefore, we would rather consider both ratings measures of CSR performance. However, we acknowledge that it is hard to reconcile the methodologies being used at specific points in time with the reference periods being used in the different studies. We therefore concluded to base our categorisation and coding of studies on the “self-declaration” provided in the underlying primary studies.

We further distinguish those studies that investigate the relation to cost of capital and, therefore, are in the centre of our analysis, from studies that use a different financial performance measure. For the latter, we used the numerical codes “71” to “81”.

During the screening phase of the data collection, we found that most studies indeed explored the association of CSR disclosure/CSR performance and cost of capital. However, some studies investigate the opposite directional relationship, that is the impact of actual cost of capital on the extent/quality of CSR disclosures. The underlying hypothesis could be that higher cost of capital might provide incentives for management to improve the degree and quality of CSR disclosures, thereby reducing information asymmetries and signalling sustainability orientation towards the market participants. As the approach is different, the conclusions are not comparable to those of studies that analyse the association between CSR disclosure and cost of capital. We therefore assigned a specific code (“91”) to those kinds of studies.

We use a separate category and code for empirical studies that do not specifically investigate CSR disclosure/CSR performance and cost of capital or some other financial performance measure (“other empirical study”, numerical code “92”). Finally, for other than empirical studies we define a separate category and code. The reason is that those studies cannot be analysed in a structured manner using the coding system for conclusions (Table 7).

The differentiation of categories and codes at the most granular allows us to perform well-suited analyses and adapt, depending on the underlying research objective. Table 6 provides a complete list of the categorisation of research objectives and related numerical codes.

Table 6: Categories and numerical codes for research objectives, Created by authors

Research objective	Numerical code
CSR disclosures and Cost of Capital	11
Environmental disclosures (E) and Cost of Capital	12
Social disclosures (S) and Cost of Capital	13
Governance disclosures (G) and Cost of Capital	14
CSR disclosures and Cost of Debt	21
Environmental disclosures (E) and Cost of Debt	22
Social disclosures (S) and Cost of Debt	23
Governance disclosures (G) and Cost of Debt	24
CSR disclosures and Cost of Equity	31
Environmental disclosures (E) and Cost of Equity	32
Social disclosures (S) and Cost of Equity	33
Governance disclosures (G) and Cost of Equity	34
CSR performance and Cost of Capital	41
Environmental performance and Cost of Capital	42
Social performance and Cost of Capital	43
Governance performance and Cost of Capital	44
CSR performance and Cost of Debt	51
Environmental performance and Cost of Debt	52
Social performance and Cost of Debt	53
Governance performance and Cost of Debt	54
CSR performance and Cost of Equity	61
Environmental performance and Cost of Equity	62
Social performance and Cost of Equity	63
Governance performance and Cost of Equity	64
CSR performance and Financial Performance	71
Environmental performance and Financial Performance	72
Social performance and Financial Performance	73
Governance performance and Financial Performance	74
CSR disclosure and Financial Performance	81
Cost of Equity and CSR Disclosure	91
Other empirical study	92
No empirical study	100
No relation to CSR reporting and/or Cost of Capital	200

For the analysis of the research conclusions of primary studies we developed a similar coding system. In order to allow for a targeted approach in line with research question Q4, we limited our coding system to studies that explore the relation of CSR disclosure/CSR performance and cost of capital.

As to ensure alignment and reconciliation with the coding system used for the categorisation according to the research objective, we implement a three-digit coding system, where the first two digits refer to the numerical code assigned based on research objectives and the last digit is defined based on the findings of the studies.

For the coding of research conclusions, we gave up the differentiation of CSR disclosure and CSR performance and merged the respective categories. The reason is, as further described in chapters 2.2 and 2.4, that the distinction of

CSR disclosure and CSR performance made in the various empirical studies is not consistent. Retaining those distinctions might impair our analyses of the findings.

Following other review studies, like Friede et al. (2015), we distinguish positive, negative and neutral impacts to the cost of capital measure. A positive impact¹ is assumed when an increase (decrease) of the CSR disclosure or performance measure induces an increase (decrease) of the cost of capital measure. In turn, negative impact means that an increase (decrease) of the CSR disclosure or performance measure results in a decrease (increase) of the cost of capital measure. The association is deemed “neutral”, if there could not be found a statistically significant impact from an increase (decrease) of the CSR disclosure or performance measure on cost of capital.

Based on the observations in previous review studies we expect a higher share of negative relationships than positive ones. We therefore assign the last digit number “0” in case of negative relationships. “1” is used to signal observed positive relationships and “2” in case there could not be evidenced a statistically significant relation. Table 7 contains a complete list of the classifications of conclusions and the assigned classification codes.

Table 7: Categories and numerical codes for research conclusions, Created by authors

Classification Conclusion	Classification code
CSR disclosure/CSR performance and Cost of Capital (WACC) are negatively associated	110
CSR disclosure/CSR performance and Cost of Capital (WACC) are positively associated	111
There is no statistically relevant relation of CSR disclosure/CSR performance and Cost of Capital (WACC)	112
Environmental disclosure and Cost of Capital are negatively associated	120
Environmental disclosure and Cost of Capital are positively associated	121
There is no statistically relevant relation of environmental disclosure and Cost of Capital	122
Social disclosure and Cost of Capital are negatively associated	130
Social disclosure and Cost of Capital are positively associated	131
There is no statistically relevant relation of social disclosure and Cost of Capital	132
Governance disclosure and Cost of Capital are negatively associated	140
Governance disclosure and Cost of Capital are positively associated	141
There is no statistically relevant relation of Governance disclosure and Cost of Capital	142
CSR disclosure/CSR performance and Cost of Debt are negatively associated	210
CSR disclosure/CSR performance and Cost of Debt are positively associated	211
There is no statistically relevant relation of CSR disclosure/CSR performance and Cost of Debt	212
Environmental disclosure and Cost of Debt are negatively associated	220
Environmental disclosure and Cost of Debt are positively associated	221
There is no statistically relevant relation of environmental disclosure and Cost of Debt	222
Social disclosure and Cost of Debt are negatively associated	230
Social disclosure and Cost of Debt are positively associated	231
There is no statistically relevant relation of social disclosure and Cost of Debt	232
Governance disclosure and Cost of Debt are negatively associated	240
Governance disclosure and Cost of Debt are positively associated	241
There is no statistically relevant relation of Governance disclosure and Cost of Debt	242
CSR disclosure/CSR performance and Cost of Equity are negatively associated	310
CSR disclosure/CSR performance and Cost of Equity are positively associated	311
There is no statistically relevant relation of CSR disclosure/CSR performance and Cost of Equity	312
Environmental disclosure and Cost of Equity are negatively associated	320
Environmental disclosure and Cost of Equity are positively associated	321
There is no statistically relevant relation of environmental disclosure and Cost of Equity	322
Social disclosure and Cost of Equity are negatively associated	330
Social disclosure and Cost of Equity are positively associated	331
There is no statistically relevant relation of social disclosure and Cost of Equity	332
Governance disclosure and Cost of Equity are negatively associated	340
Governance disclosure and Cost of Equity are positively associated	341
There is no statistically relevant relation of Governance disclosure and Cost of Equity	342

Given that some studies follow multiple research objectives, we allowed for assignment of multiple codes for research conclusions. As an example, if the research is on the association of CSR disclosure and cost of equity and cost of debt and the study evidences a negative association for cost of equity, while the impact on cost of debt is not statistically significant, the resulting numerical codes would be “310” and “212”.

2.5 Phase V: Analysis

¹ The terms impact, relationship and association are used synonymously in this study.

The raw qualitative and quantitative data collected on the publications were tabulated and analysed in Excel, with tables being created and data quantified for the relevant fields of investigation.

We performed the following kinds of analyses:

- a. Bibliometric analysis of the type of publication by publication medium.
- b. Bibliometric analysis of the journals in which the studies have been published by field of study.
- c. Bibliometric analysis of the year of publication.
- d. Bibliometric analysis of the type of publication by research method using own classification categories.
- e. Bibliometric analysis of the geographical distribution of the sample companies used in the different studies.
- f. Bibliometric analysis of the industry dispersion by industry sectors.
- g. Bibliometric analysis of the reference periods used in the studies, by geographical areas.
- h. Content analysis of the research objectives.
- i. Content analysis of the research conclusions.

The first three types of bibliometric analyses under a. to c. above were produced using the analytical functionalities of Scopus. By using these functionalities, we implicitly adhere to the sample resulting from the search functions in Scopus, including the ten studies we manually excluded from the scope at a later stage (see chapter 2.2). We accept this limitation as the total number of studies concerned is relatively low and we don't expect bias in our analyses from including those studies.

As regards the content analysis of the research objectives under h. above, we investigated the structure of research objectives using the coding system described in chapter 2.4 (analysis h1.). We further explored the most common measures used for CSR disclosure/performance and the data providers behind (h2.), provide a breakdown of studies that more specifically focus on the individual dimensions of ESG (h3.) and assess how many studies serve multiple research objectives (h4.).

Concerning the findings of the empirical studies under i. above, we perform three analyses: First, we analyse for cost of equity-studies the proportion of studies that show a negative, positive or neutral relationship of CSR disclosure/performance and cost of equity (analysis i1.). Second, we perform the same kind of analysis for the cost of debt-studies (i2.). And third, we investigate for the E-, S- and G-dimensions the relation of disclosure/performance and the cost of capital measures (i3.).

3 Results

3.1 Bibliometric analysis

3.1.1 Type of publication by media type

Scopus analytical functions reveal that the vast majority of the sample studies have been published as articles (more than 80%), while conference papers, reviews and book chapters are less relevant. This is not surprising as we restrict our search to peer-reviewed publications and typically, books are not subject to peer-review. Furthermore, we narrow down our search to final publications, which might explain the relatively low proportion of conference papers. Figure 2 displays the detailed proportions of publications by media type.

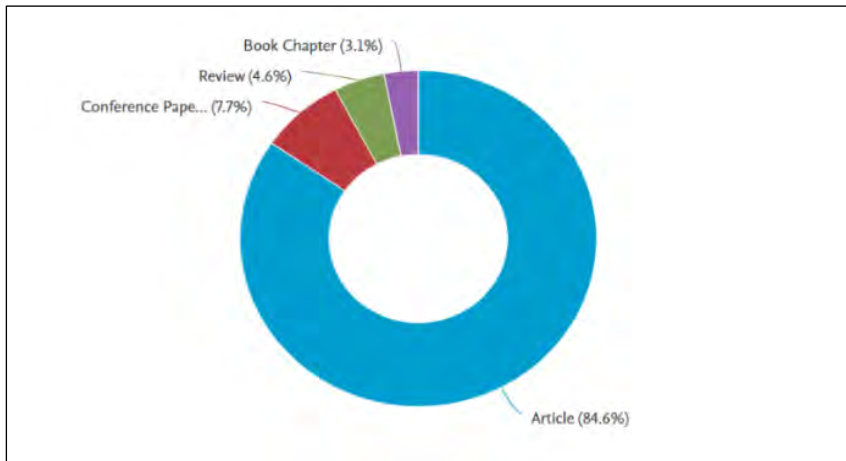


Figure 2: Publications by media type, Created by authors

3.1.2 Type of journal by field of study

In order to gain insights into the typical subject areas of the journals where the articles have been published, we used Scopus journal-level metrics for our analysis. Scopus offers a broad range of different metrics in that regard. We limit our analysis to the fields of study associated with respective journals. Academics might use this information to form their decisions in favour of or against a submission for publication in a specific field of study.

We found that most studies have been published in business and management journals (33%), followed by economic journals and social sciences journals (both around 15%) and environmental journals (12%) (see Figure 3 below).

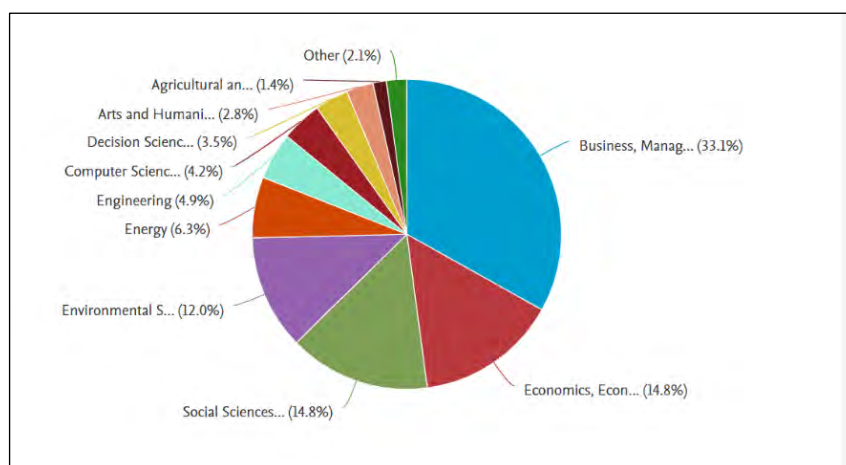


Figure 3: Journals by subject area, Created by authors

3.1.3 Publications by year

The final analysis we conducted based on Scopus metrics is the dispersion of publications over time. Based on this analysis, we are able to identify publication trends. In line with other review studies, such as Friede et al. (2015), Benameur et al. (2022) and Bosi et al. (2022), overall, we observed an increasing number of publications over time (see Figure 4 below). However, rather than a steady increase, we found considerable variations, with peaks in the years 2017 and 2022 and valleys in 2016, 2018 and 2023.

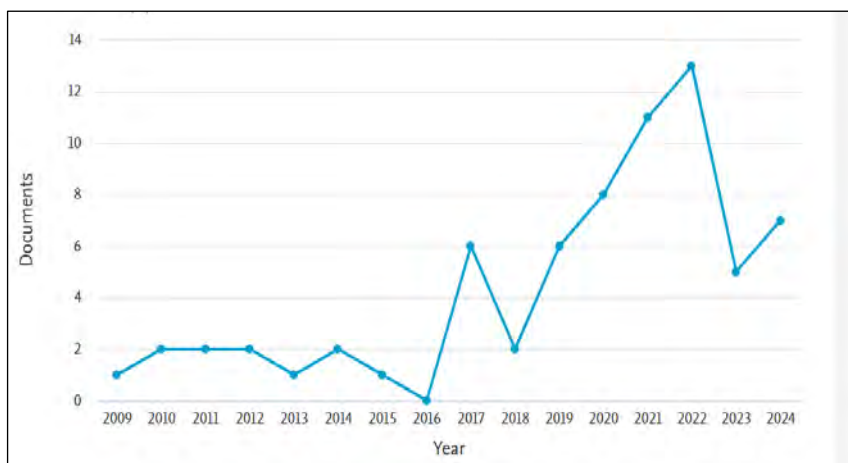


Figure 4: Publications by year, Created by authors

3.1.4 Type of publication by research method

We broadly categorise studies into empirical and non-empirical studies. By far the majority of studies (50 out of 54, more than 92%) are empirical studies. The remaining four studies consist of two bibliometric studies and one conceptual study and social sciences study, respectively. We define social sciences studies as studies that discuss certain aspects of sustainability reporting and cost of capital from an either ethical, behavioural or cultural perspective. Conceptual studies explore concepts for further (usually empirical) studies to examine the relation between sustainability reporting and cost of capital. As a common characteristic, social sciences studies and conceptual studies do not use empirical approaches. Figure 5 shows the distribution of publications by research method.

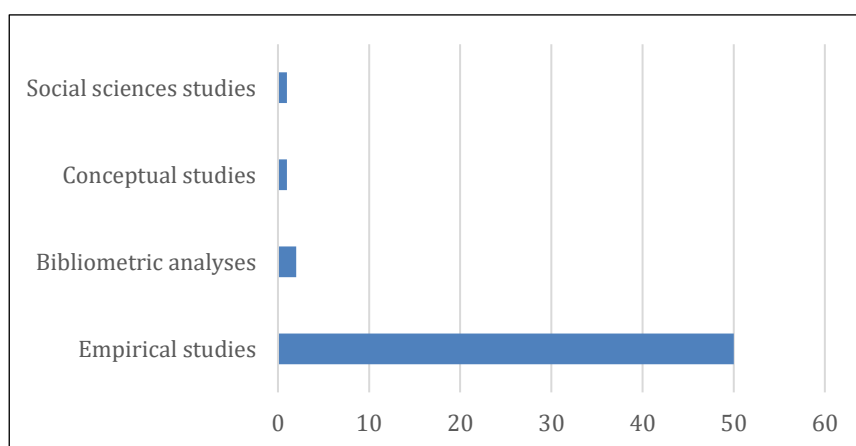


Figure 5: Type of publication by research method, Created by authors

The two bibliometric studies are those of Zaid (2023) and Bosi et al. (2022). Zaid (2023) investigates research articles on climate-related disclosures in the period 2020 to 2023 in light of the newly issued IFRS Foundation Sustainability Standards. The study of Bosi et al. (2022) examines a large body of literature on environmental, social, and governance (ESG) and sustainability reporting over the years 1998 to 2022.

The conceptual study of Shad et al. (2019) conceptualises a framework that examines the moderating effect of sustainability reporting practices on the relationship between enterprise risk management (ERM) implementation and business performance. According to the authors, this framework could be tested and validated in future empirical studies.

Finally, Aras and Crowther (2008) approach the topic of sustainability reporting from a rather theoretical, even somehow philosophical, perspective. They argue that the terms “sustainability” or “sustainable development” have the effect of obfuscating the real situation regarding the effect of corporate activity upon the external environment and the consequent implications for the future. We classify this study as a social sciences study.

3.1.5 Geographical distribution of sample companies

In the following step of the analysis, the geographical composition of the sample companies used in the studies was considered. In addition to the 50 empirical studies, which each clearly articulate the geographical scope of the sample companies, one additional study restricts the research to a certain geographical scope: The conceptual study of Shad et al. (2019) is developed with regard to Malaysian companies in the oil and gas sector.

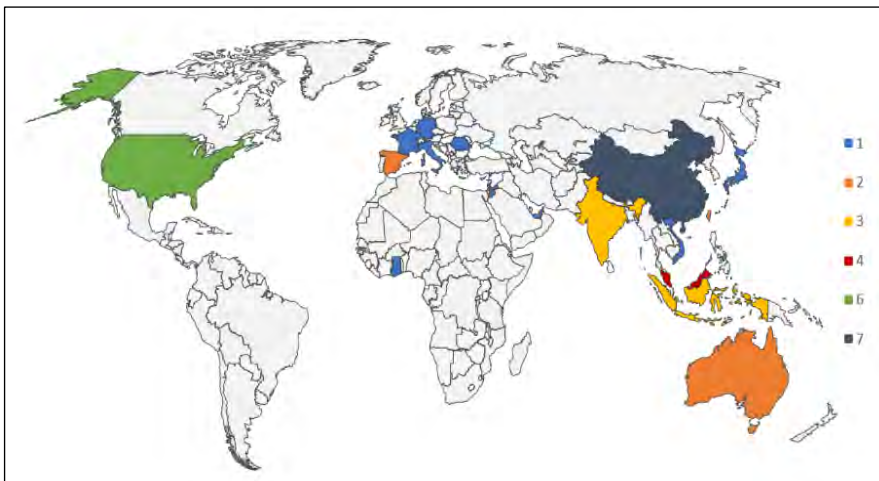


Figure 6 Geographical scope and frequency of the sample companies, Created by authors

The geographical scope of the studies is widely dispersed. The majority of studies (13.73%) refer to China, followed by the United States (11.76%), Malaysia (7.84%), India and Indonesia (5.88%, each). Some studies use global settings or multinational settings (each 13.73%, in absolute amounts seven studies, each). 15.68% of studies concern European companies, covering Spain (3.92%), France, Germany, Italy, and Romania (1.96%, each) and Europe as a whole (3.92%).

3.1.6 Industrial focus of sample companies

Both, anecdotal evidence and academic research, show that the extent of the impact of environmental, social and governance matters and associated risks differs by industry sector (see for example Brooks and Oikonomou, 2018). Usually, manufacturing firms and extractive industries are considered more affected by environmental matters, while for example the textile industry is often linked to both, social and environmental matters. It therefore comes not as surprise, that the European standard setter for sustainability reporting, the European Financial Reporting Advisory Group (EFRAG), prioritises certain sectors in developing sector specific standards for sustainability reporting, namely: Mining, Quarrying and Coal mining; Oil and Gas; Road Transportation; and Agriculture, Farming and Fisheries (see EFRAG, <https://www.efrag.org/lab5>).

Against this background, it is surprising that for the vast majority of studies (total of 39 out of 52, equals 75%) there are no sectoral restrictions, apart from excluding financial firms such as banks and insurance companies. We would have expected that a larger proportion of studies focus on the most affected industries. For the remaining studies, there is no clear focus on certain industries. Manufacturing and oil and gas were specifically explored by two studies (manufacturing companies in India and Indonesia and both studies focussing on oil and gas refer to Malaysian companies). Further, two studies focus on Agriculture, both with a global setting. Basic materials, financial firms, food and beverage, broader industrial firms, and textile were defined as focus sectors in one study, each. Two further studies have a focus on multiple selected sectors or so-called „high pollution industries“. Figure 7 provides an overview of the industrial focus of the primary studies. Interestingly, all but three of those industrial focus studies are also limited to certain countries. Only the studies that focus on the agriculture sector apply a global setting.

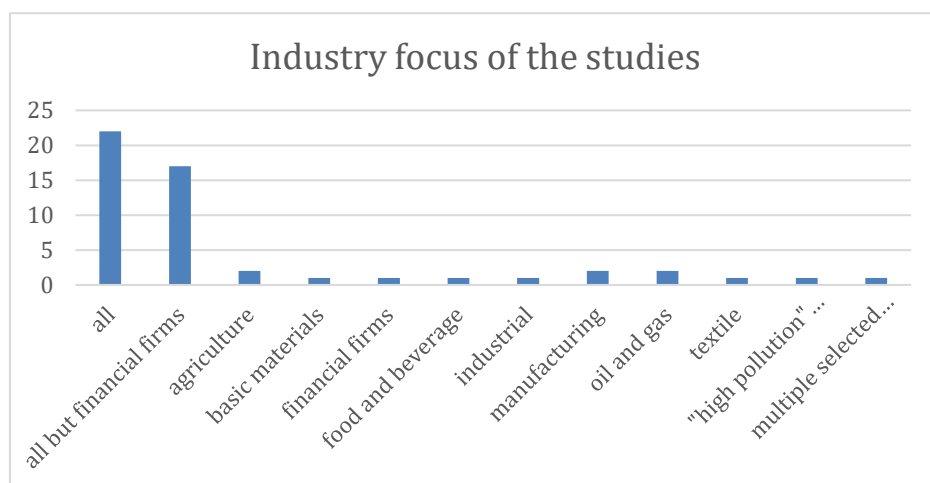


Figure 7: Industrial focus of primary studies, Created by authors

3.1.7 Reference period

We analysed the reference period used in the studies. This helps to understand the degree of coverage of the different time periods in empirical research on sustainability reporting and cost of capital. In combination with our findings as regards geographical and industrial focus, with this analysis we are in a better position to identify potential areas of future research.

Except for two studies that use a reference period starting in the early 1990s, all reference periods start after the year 2000. 39 out of 52 studies (75%) use a reference period with a start date 2005 or later. This reflects the increasing importance of sustainability reporting and the continuous development of sustainability reporting frameworks, starting with the release of the first version of GRI Guidelines in the year 2000. In 2016, the GRI moved from providing guidance to issuing standards – a major step in the evolution of global sustainability frameworks. In the year after, 2017, the Task Force on Climate-related Financial Disclosures (TCFD) issued the widely recognised TCFD recommendations. 2017 was also the year when the Non-financial reporting Directive (Directive 2014/95/EU) was first to be applied in Europe. The impact of these new or amended frameworks is visible in sustainability reports starting from the year 2018.

We therefore expected that a significant share of studies would implicitly or explicitly cover the impact of those newly introduced or amended frameworks by defining a reference period up until at least the year 2018. Indeed, we find that 19 out of 52 studies (equals 36.5%) use reference periods that include at least the year 2018. However, while it is common to measure the quality of sustainability disclosure against the GRI guidelines or standards, respectively, none of the studies explicitly explores the transitional effect of the introduction of the new frameworks on the cost of capital.

The average length of the reference periods used is around six and a half years, the maximum length is 18 years. The median length of reference periods is six years, almost equal to the average length. Very few studies (two out

of 52) use a reference period of only two consecutive years, which would allow for singling out transitional impacts. Figure 8 illustrates the dispersion of reference periods for the different geographical research areas.

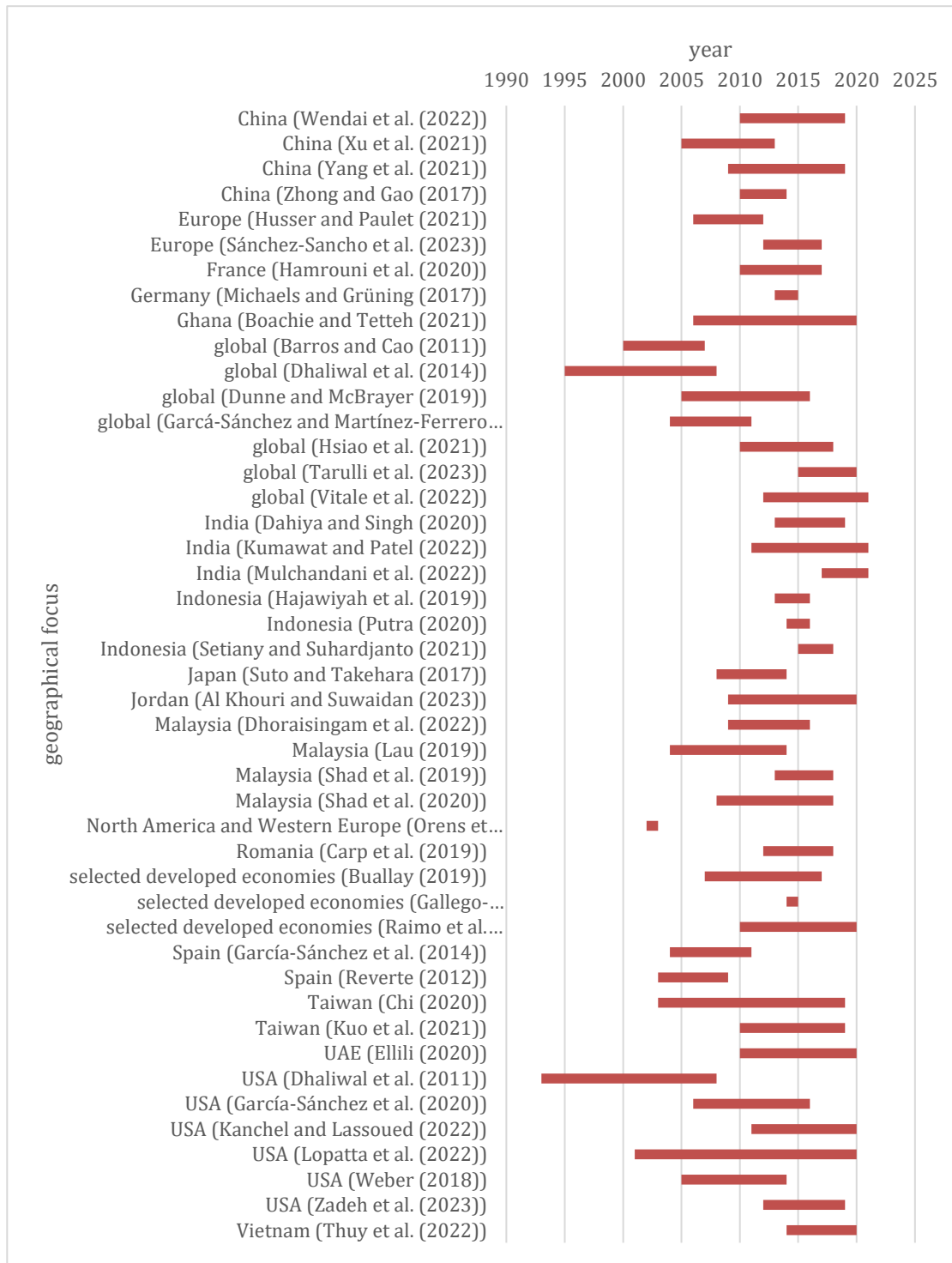


Figure 8: Overview of reference periods by geographical focus, Created by authors

We expected that studies using reference periods with relatively early start dates would more frequently concern mature economies, where sustainability reporting practices have evolved for a longer period of time. Except for the study of Dhaliwal et al. (2011), covering listed US firms in the period from 1993 to 2007, this presumption, however, does not hold true. We found no significant patterns in terms of geographical focus and the reference periods being selected.

3.2 Content analysis

3.2.1 Research focus

We further analysed the focus areas of the empirical studies. Based on the defined search criteria and the combination of keywords, we expected that all empirical studies at least partially examine the relation of CSR disclosures and cost of capital. While this is true for 80% of the studies (in total 40), another 20% of studies define different research objectives:

- Six studies explore the relation of CSR disclosure/performance and some other financial performance measures. We found the following financial performance measures being used in our sample: Return on Assets, Return on Equity, Return on Sales, market performance (Tobin's Q) and investment efficiency.
- Seven studies have a different, but still related research focus:
 - Two studies explore the impact of CSR assurance on cost of capital.
 - One study, each, (1) examines the relation of CSR disclosure and company-level cash holdings;
 - (2) focuses on the impact of applying the International Integrated Reporting Framework (IIRF) on cost of capital;
 - (3) analyses how aspects of national institutions affect CSR reporting on an international level;
 - (4) explores the impact of media pressure on CSR practices;
 - (5) investigates the extent by which being a “signatory company” under the UN Global Compact (UNGC) reduces information risk, and thus leads to better market returns, lower cost of debt, and lower cost of equity; and
 - (6) analyses the role of independent directors regarding CSR disclosure.

We found that more than three quarters of the studies explore the relation of CSR disclosure and cost of capital. 42% of the studies specifically focus on how CSR disclosures affect cost of equity, another 28% deal with CSR disclosures and cost of debt and 6% use Weighted-Average Cost of Capital (WACC) as cost of capital measure (see Figure 9). The Bloomberg ESG score is by far the most important measure of CSR disclosure (in total 11 studies). Some studies use simpler metrics, such as binary codes to reflect the existence (or non-existence) of CSR reports or self-constructed indices.

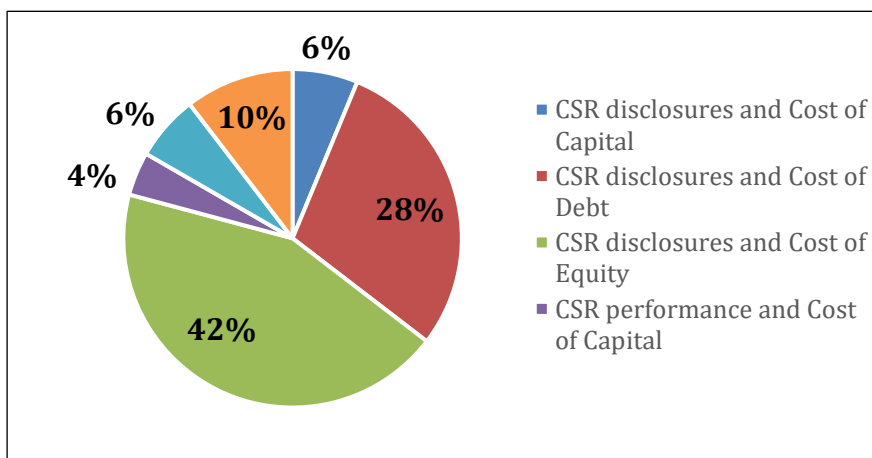


Figure 9: Research focus of empirical studies²

² % of all empirical studies, multiple assignments possible, does not add up to 100%.

20% of the studies focus on the relation of CSR performance and some cost of capital measure, with 10% deal with cost of equity, 6% cover cost of debt, and another 4% use WACC. The most important sources for CSR performance measures are KLD STATS (today MSCI ESG Research) and Thomson Reuters/Refinitiv ESG scores (3 studies, each). The following Table 8 illustrates the distribution of data providers and respective measures used for CSR disclosure and performance.

Table 8: CSR disclosure and performance measures, Created by authors

CSR disclosure/performance measure	number	number E-scores	number S-scores	number G-scores
Bloomberg ESG score	11	7	6	6
Thomson Reuters (Refinitiv)	3	1	1	1
KLD Stats	3			

6% of studies (not captured in the Figure 9 above, 3 in total) explore the relation of CSR disclosure or CSR performance and some other financial performance measures like Return on Assets, Return on Equity, Return on Sales, market performance (Tobin's Q) and investment efficiency. 4% of studies focus on the relation of cost of capital and CSR disclosure. Those two studies examine whether the previous year cost of equity has some relation to CSR disclosure level in subsequent years.

We further analysed to what extent studies specifically focused on the three dimensions of ESG – the environmental dimension (“E”), the social dimension (“S”) and the governance dimension (“G”). We hereby limited our analysis to the 40 studies that examine the relation of CSR disclosure or CSR performance and cost of capital.

14 out of those 40 studies (35%) specifically investigate the impact of environmental disclosure or environmental performance on cost of capital. Interestingly, while there are few studies that limit their analysis to the environmental dimension, there are no studies with an exclusive focus on either of the two, the social or the governance dimension. Ten out of the 13 studies add the social dimension to the analysis, seven have a separate focus on governance. (see Figure 10)

Seven out of eleven studies that use Bloomberg ESG scores as a measure of CSR disclosure and CSR performance, respectively, use the breakdown to E-, S-, and G-subscores provided by Bloomberg, although to a varying degree. Six studies separately analyse E-, S-, and G-scores next to the total ESG score, one study focuses on the environmental dimension, only.

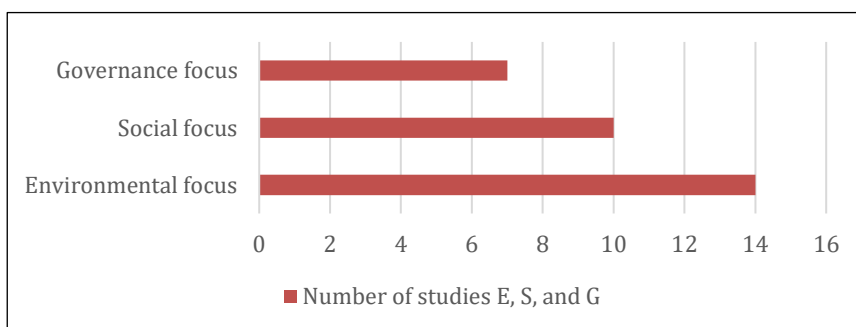


Figure 10: Number of studies with focus on the environmental (“E”), the social (“S”) and the governance (“G”) dimension

As shown by the above analysis, some studies combine different kinds of investigations. We found that 45% (in total 18 studies) define multiple research questions. Most common is the combination of different sustainability measures, breaking down the total ESG score into E-, S-, and G-subscores. 32% of studies (in total 14) narrow down the analysis by type of sustainability matter. Furthermore 15% of studies (in total 6) explore the relation of CSR disclosure/CSR performance, cost of equity *and* cost of debt.

The following **Table 9** gives an overview of the number and share of the different combinations of research objectives used in the studies.

Table 9: Combination of research objectives, Created by authors

Combination	number	percentage
Combination of research questions (minimum 2 codes)	18	45,00%
CSR disclosure, cost of equity and cost of debt	5	12,50%
CSR disclosure, cost of equity, cost of debt and cost of capital	1	2,50%
CSR performance, cost of equity and cost of debt	1	2,50%
CSR performance, cost of equity, cost of debt and cost of capital	1	2,50%

A further breakdown within the environmental, social or governance categories cannot be observed directly from the research objectives. However, Vitale et al. (2023) include in their analytical models as independent variables sustainability scores that assess particular corporate environmental and social issues, such as: sustainable internal business production and procurement processes (i.e. ResourceUseScore); Greenhouse Gases (GHGs) emissions mitigation (i.e. EmissionScore); environmental corporate innovation (i.e. Env Innov Score); the integration of customer's health and safety, integrity, and data privacy in quality goods and services productions (i.e. ProductRespScore); human resources sustainable management (i.e. Workforce Score, Human Right Score); and corporate citizenship (i.e. Community Score).

3.2.2 Conclusions from empirical studies

We systematically reviewed the conclusions reached in empirical studies. We therefore further distinguished between cost of equity-studies and cost of debt-studies. Studies, that investigated both cost of capital measures, have been included in both categories. We excluded the study of García-Sánchez et al. (2020), because of the specific research focus on the gap of CSR performance and CSR disclosure. The final sample for analysis consists of 39 studies. 25 studies investigate the impact of CSR disclosure or CSR performance on cost of equity, 17 studies explore the relation to cost of debt.

Cost of equity: Based on previous studies, there exists mixed observations as regards the connection of CSR disclosures and cost of equity. While some studies confirm a negative association of CSR disclosure and cost of equity – i.e., an increase in CSR disclosure volume or quality results in cost of equity reductions – other studies provide evidence for the opposite. We found that the vast majority of studies confirm a negative relationship of CSR disclosure/CSR performance and cost of equity. While 84% of studies (in total 21) hold that view, only 12% (in total 3) observe a positive relationship. 4% (in total 1 study) cannot confirm any significant statistical relationship (see Figure 11).

We further analysed any potential trends over time, geographical and industry specifics. We couldn't find any specific pattern. The four studies that could not confirm a negative relationship have been published in the years 2018, 2020, 2022 and 2023, respectively. Weber (2018) analyses the CSR reporting of US firms in the period 2005-2013, Dahiya and Singh (2020) disclosure of Indian manufacturing firms in the period 2013-2018. Lopatta et al. (2022) have a rather distinct research focus on the role of individual managers in the relationship between sustainability performance, sustainability reporting and cost of equity. They focus on US companies in the period 2001-2019. Finally, Hutagaol-Martowidjojo et al. (2023) investigate firms' ESG disclosures for basic materials companies in Asia between 2014 and 2019.

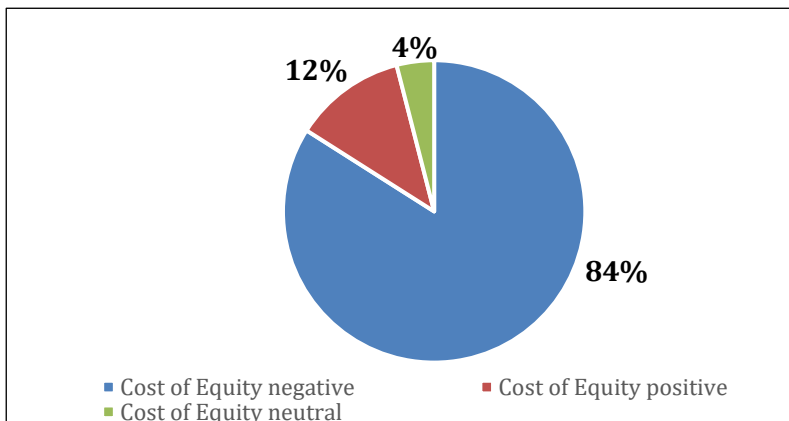


Figure 11: Impact of CSR disclosure or CSR performance on cost of equity, Created by authors

Cost of debt: Based on the findings in empirical studies, the relation of CSR disclosure/CSR performance and cost of debt seems to be weaker compared to cost of equity, but still predominantly negative. 71% of studies (in total 12) confirm this finding. In turn, 12% of studies that investigate cost of debt (in total two studies) find a positive relationship and 18% (three studies) cannot confirm any significant statistical relation or found mixed results. The results are illustrated in Figure 12.

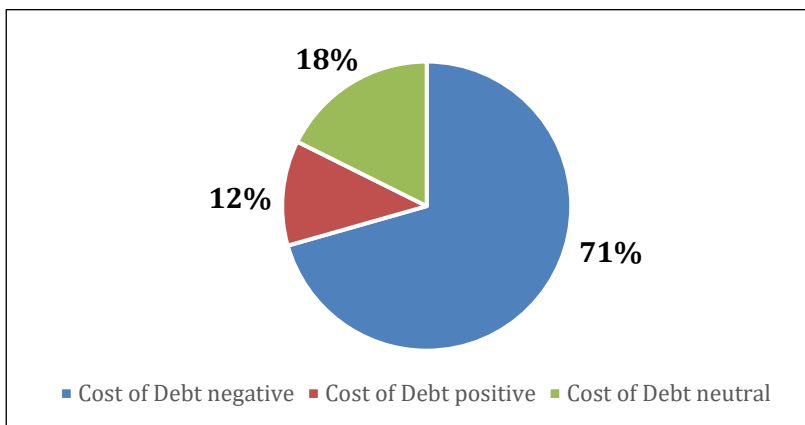


Figure 12: The impact of CSR disclosure or CSR performance on cost of debt, Created by authors

Again, there is no specific pattern observable:

- Orens et al. (2010) analyse the relation of web-based non-financial disclosure and cost of debt for a sample of European and North American firms in the year 2002. While for North American firms there could not be confirmed a significant negative relationship, for Continental European firms this association has been confirmed.
- Suto and Takehara (2017) investigate the relation of CSR performance and cost of debt for Japanese firms in the period 2008-2013 and rejected the hypothesis of a negative association for that period. However, as they further divided the period into two sub-periods (2008-2010, including the financial crisis, and 2011-2013, post financial crisis), the results were mixed. While for the former period they found positive relations, for the latter period there was no significance observed, suggesting that debtors were more aware of CSR performance when estimating risk in more recent periods (Suto and Takehara (2017), page 810).
- Chi et al. (2020) examine for Taiwanese listed companies in the years 2003-2018 the association of CSR disclosure and credit ratings (cost of debt) for private and public firms. They found that CSR disclosure is associated with more favourable credit ratings and a lower cost of debt for public firms. However, for private firms the results are different as they even appear to be penalised for reporting CSR.

- Boachie and Tetteh (2021) focus on listed food and beverage, manufacturing and oil and gas firms on the Ghana stock exchange in the period 2006-2019. They found an inverse relationship of high CSR scores and cost of debt.
- Vitale et al. (2023) explore the relation of mandatory non-financial disclosure and financial performance measures, including cost of debt, for global agricultural companies in the years 2012-2020. The findings are mixed. Regulations on environmental disclosure positively affect the cost of debt in the short-term. Likewise, sustainable-oriented practices (expressed by the variables ResourceUseScore, EmissionScore, EnvInnovScore and ProductRespScore) tend to increase the cost of debt. In turn, regulations on social disclosure induced a direct negative effect on the cost of debt.

Cost of capital (WACC): For the few studies that use a combined cost of capital measure, more specifically in all cases the WACC, there cannot be derived a clear picture. One study, each, confirms a negative and positive association to CSR disclosure, while three studies did not find a statistically significant relation at all. Kanchel and Lassoued (2022) investigate for US firms in the period 2011-2019, whether there is a ratcheting effect over time. They found that the three dimensions of E, S, and G do not have the same effect. While governance disclosure decreases the cost of capital during the first years, in later years, the effect becomes positive. Similarly, over time, social disclosure increases the cost of capital. However, environmental disclosure shows a negative effect on the cost of capital during the first years but no significant effect later in time.

Environmental, Social and Governance dimension: More specifically, we also analysed the findings of those studies that focused on one or more of the three dimensions of E, S, and G. Apart from the study of Kanchel and Lassoued (2022) mentioned above, 13 studies explicitly focus on the environmental dimension, while nine studies deal with the social dimension and seven studies with the governance dimension. The conclusions reached as regards the association to cost of equity, cost of debt and cost of capital are mixed. Table 10 below shows the directional relationship of each of the ESG dimensions and the cost of capital measures. Because of the limited number of studies that specifically explore those associations, we disclose absolute numbers instead of percentages.

For environmental disclosures, most studies confirm a negative relation to cost of equity, cost of debt and cost of capital, respectively, while some studies find no significant relation. For social disclosures and governance disclosures, the findings are mixed and do not provide clear evidence for either a negative or positive association.

Table 10: Directional relationships for E, S and G disclosures, Created by authors

	Cost of Equity			Cost of Debt			Cost of Capital (WACC)		
	negative	positive	neutral	negative	positive	neutral	negative	positive	neutral
Environmental disclosure	3	0	2	5	0	2	1	0	0
Social disclosure	1	0	2	2	1	2	0	0	1
Governance disclosure	2	0	0	1	1	2	1	0	0

4 Conclusion

The aim of this study was to determine the areas of application as well as the scope and extent of the data available on sustainability reporting in relation to cost of capital. We used a combination of bibliometric and content-based analysis and formulated the following research questions:

Q1: What is the current state of research on the association of CSR disclosure and cost of capital in terms of research method, industry focus, geographical focus, and reference periods being used?

Q2: What proportion of studies explores the relationship between CSR disclosure and (a) cost of equity, (b) cost of debt and (c) combined cost of capital (e.g. Weighted Average Cost of Capital, WACC)?

Q3: What proportion of studies further breaks down CSR disclosures into environmental (E), social (S) and governance disclosures (G)?

Q4: What findings can be observed from the studies regarding the directional relationship between CSR disclosure and cost of capital, also distinguishing the different dimensions of (a) CSR disclosure and (b) cost of capital?

Concerning Q1, we found that empirical studies are clearly dominating the academic research. In line with other review studies, we can confirm an increasing interest of scholars in investigating the relation of CSR disclosure/performance and cost of capital. The geographical focus of research is broad. Surprisingly, only few studies define a specific focus industry. Future research could therefore be more targeted as regards specific industry sectors. We observed that the reference periods being used are rather long and cover periods up to 18 years. One-off impacts stemming from changes in regulation, market behaviour or the environment might not be sufficiently reflected.

Regarding Q2, we found that most studies refer to cost of equity, followed by cost of debt. Studies using a combined cost of capital measure seem to be less relevant. Analysis of the association of CSR disclosure or performance and cost of capital require a solid measure for CSR disclosure quality or performance. Based on our sample, Bloomberg ESG score is by far the most frequently used measure, followed by Thomson Reuters/Refinitiv and KLD STATS (today MSCI ESG Research).

In relation to Q3, 35% of empirical studies put emphasis on environmental disclosures or performance. The share of studies that specifically investigate the social and/or governance dimension is considerably lower. We identified only one study that made use of a further breakdown within the environmental, social or governance categories, disentangling factors like resource use and greenhouse gas emissions disclosures. Specifically in light of the rising public and political awareness of corporate sustainability and the more and more stringent environmental, social and governance-related requirements we would expect that more targeted research around specific environmental factors and their impact on cost of capital will evolve.

Finally, under Q4 we analysed the findings of the studies. In line with previous research, we can confirm based on our sample that a high percentage of studies (84% for cost of equity-studies and 71% for cost of debt-studies) evidence a significant negative relationship of CSR disclosure/performance and cost of capital. This supports the value creation theory. The relation seems to be weaker at the E-, S- and G-disclosure sublevel, indicating that investors consider the overall CSR disclosure/performance rather than the individual contributions.

Bibliography (References)

- Bansal, M., Samad, T.A. and Bashir, H.A. (2021), The sustainability reporting-firm performance nexus: evidence from a threshold model, *Journal of Global Responsibility*, Vol. 12 No. 4, pp. 491-512.
- Benameur, K.B., Mostafa, M.M., Hassanein, A., Shariff, M.Z. and Al-Shattarat, W. (2022), Sustainability reporting scholarly research: a bibliometric review and a future research agenda, *Management Review Quarterly*, p. 74.
- Block, J.H. and Fisch, C. (2020), Eight tips and questions for your bibliographic study in business and management research. *Manage Rev Q* 70(3), pp. 307–312.
- Bloomberg (2023), ESG scores, Overview & FAQ, Updated November 2023.
- Bosi, M.K., Lajuni, N., Wellfren, A.C. and Lim, T.S. (2022), Sustainability Reporting through Environmental, Social, and Governance: A Bibliometric Review, *Sustainability* 2022, Vol. 14 No. 19, 12071.
- Boachie, C. and Tetteh, J.E. (2021), Are the saving decision patterns of bank customers also clouded by psychological biases? Evidence from Ghana, *Review of Behavioral Finance*, Vol. 15 No 1, pp. 1-18.
- Botosan, C. A. (1997), Disclosure level and the cost of equity capital, *The Accounting Review*, Vol. 72, pp. 323–349.
- Botosan, C. A. and Plumlee, M. A. (2002), A re-examination of disclosure level and the expected cost of equity capital, *Journal of Accounting Research*, Vol. 40, pp. 21–40.
- Chi, W., Wu, S-J. and Zheng, Z. (2020), Determinants and consequences of voluntary corporate social responsibility disclosure: Evidence from private firms, *The British Accounting Review*, Vol. 52 No. 6.
- Dahiya, R., Saha, R., Shashi, Cerchione, R. and Singh, R. (2020), Effect of ethical leadership and corporate social responsibility on firm performance: A systematic review. *Corporate Social Responsibility and Environmental Management*, Vol. 27 No. 2, pp. 409–429.
- Diamond, D. and Verrecchia, R. (1991), Disclosure, liquidity and the cost of equity capital, *The Journal of Finance*, Vol. 46, pp. 1325–1360.
- Echchakoui, S. (2020), Why and how to merge Scopus and Web of Science during bibliometric analysis: the case of sales force literature from 1912 to 2019, *J Market Anal*, Vol. 8, pp. 165–184.
- Friede, G., Busch, T. and Bassen, A. (2015), ESG and financial performance: aggregated evidence from more than 2000 empirical studies, *J Sustain Finance Invest*, Vol. 5 No. 4, pp. 210–233.
- García-Sánchez, I-M., Hussain, N., Khan, S. A. and Martinez-Ferrero, J. (2021), Do Markets Punish or Reward Corporate Social Responsibility Decoupling? *Business & Society*, Vol. 60 No. 6, pp. 1431–1467.
- Gutsche, R., Schulz, J.F. and Gratwohl, M. (2017), Firm-value effects of CSR disclosure and CSR performance, paper presented at EFMA Annual Meeting, Athens, 30 June, available at: www.alexandria.unisg.ch/publications/251910.
- Hinze, A. K. and Sump, F. (2019), Corporate social responsibility and financial analysts: A review of the literature, *Sustainability Accounting, Management and Policy Journal*, Vol. 10, pp. 183–207.
- Leuz, C. and Verrecchia, R. (2000), The economic consequences of increased disclosure, *Journal of Accounting Research*, Vol. 38, pp. 91–124.
- Hutagaol-Martowidjojo Y., Tohang, V. and Payung, E.P.T. (2023), The Moderating Role of Policy Intervention on the Relationship of Environment, Social, and Governance (ESG) and Cost of Equity Capital: A Study in Basic Materials Companies in Asia, *Australasian Accounting, Business and Finance Journal*, Vol. 17 No. 5, pp. 4-14.
- Khanchel, I. and Lassoued, N. (2022), ESG Disclosure and the Cost of Capital: Is There a Ratcheting Effect over Time? *Sustainability*, Vol. 14 No. 15, 9237.
- Lopatta, K., Buchholz, F. and Kornblum, W. (2022), Controlling und Corporate Sustainability: eine exemplarische Darstellung anhand der Kreditwirtschaft, In: Becker, W. and Ulrich, P. (eds) *Handbuch Controlling*, Springer Gabler, Wiesbaden.

- Mc Williams, A. (2000), Corporate social responsibility and financial performance: Correlation or misspecification? *Strategic Management Journal*, Vol. 21, pp. 603-609.
- Orens, R., Aerts, W. and Cormier, D. (2010), Web-Based Non-Financial Disclosure and Cost of Finance, *Journal of Business Finance & Accounting*, Vol. 37 No. 9&10, pp. 1057–1093.
- Souissi, M. and Khalif, H. (2012), Meta-analytic review of disclosure level and cost of equity capital, *International Journal of Accounting and Information Management*, Vol. 20 No. 1, p. 52.
- Suto, M. and Takehara, H. (2017), CSR and cost of capital: evidence from Japan, *Social Responsibility Journal*, Vol. 13 No. 4, pp. 798-816.
- Tarulli, A., Morrone, D., Conte, D., Bussoli, C. and Russo, A. (2023), The relevance of non-financial disclosure in influencing the cost of capital: Empirical evidence from the agri-food sector. *Business Strategy and the Environment*, Vol. 32 No. 4, pp. 1739–1751.
- UN (2021), Shifting to a Circular Economy Essential to Achieving Paris Agreement Goals, <https://unfccc.int/news/shifting-to-a-circular-economy-essential-to-achieving-paris-agreement-goals>
- Vitale, G., Cupertino, S. and Riccaboni, A. (2023), The effects of mandatory non-financial reporting on financial performance. A multidimensional investigation on global agri-food companies, *British Food Journal*, Vol. 125 No. 13, pp. 99-124.
- Weber, J.L. (2018), Corporate social responsibility disclosure level, external assurance and cost of equity capital, *Journal of Financial Reporting and Accounting*, Vol. 16 No. 4, pp. 694-724.

Optimierungsmöglichkeiten im Finanzbereich: Einsatz von ChatGPT & Co. und Prompt Engineering – Nutzungsverhalten, Anwendungsbereiche & Herausforderungen

Research Paper

Sonia Frenkenberger, BA

Fachhochschule Oberösterreich, Studiengang Controlling, Rechnungswesen und Finanzmanagement, Steyr,
E-Mail: sonia.frenkenberger@fh-steyr.at

FH-Prof. Dr. MMag. Susanne Leitner-Hanetseder

Fachhochschule Oberösterreich, Studiengang Controlling, Rechnungswesen und Finanzmanagement, Steyr,
E-Mail: susanne.leitner-hanetseder@fh-steyr.at

Abstract

Die Veröffentlichung von ChatGPT im November 2022 markierte einen Meilenstein für Large Language Models (LLMs). Diese Modelle verarbeiten große Datenmengen effizient und generieren kontextbezogene Antworten. Durch benutzer:innenfreundliche Oberflächen fördern sie die Akzeptanz und Anwendung im beruflichen Alltag. Der Literatur zu Folge gibt es im Finanzbereich vielseitige Anwendungsmöglichkeiten (zB automatisierte Berichterstellung). Es stellt sich die Frage, inwieweit ChatGPT & Co. von Mitarbeiter:innen im Finanzbereich bereits genutzt werden, welche Herausforderungen in der Anwendung gesehen werden und, ob Techniken wie Prompt Engineering, um den Output von ChatGPT & Co. zu optimieren, zum Einsatz kommen. Um die Nutzungshäufigkeit und Anwendungsbereiche von ChatGPT & Co. und in dem Zusammenhang Prompt Engineering im beruflichen Kontext zu erheben sowie die größten wahrgenommenen Herausforderungen zu identifizieren, wurde eine empirische Studie durchgeführt. Die Ergebnisse zeigen, dass diese Technologien bereits für den Wissenserwerb, E-Mail-Formulierung, Erstellung von Zusammenfassungen und Entscheidungsunterstützung genutzt werden und dass Prompt Engineering intuitiv angewandt wird. Zudem ist in Bezug auf den Umfang der beruflichen Nutzung und dem Alter oder Geschlecht kein signifikanter Zusammenhang nachzuweisen. Es scheint jedoch unerlässlich, Prompt Engineering in die Aus- und Weiterbildung im Finanzbereich zu integrieren, um das volle Optimierungspotenzial auszuschöpfen.

1 Einführung

Die Veröffentlichung von ChatGPT (Chat Generative Pre-trained Transformer), einem Large Language Model (LLM) kombiniert mit einem Chatbot, im November 2022, markierte einen signifikanten Meilenstein in der Anerkennung der Fähigkeiten von Large Language Models seitens der allgemeinen Öffentlichkeit (Teubner et al., 2023). Language Models (=Sprachmodelle) im Allgemeinen sind ein integraler Bestandteil des Natural Language Processing (NLP) und nutzen statistische Methoden, um die Abfolge von Wörtern in Texten vorherzusagen (Kulkarni et al., 2023). Die Anwendung dieser Modelle erfordert kein tiefgehendes technisches Wissen und ermöglichen durch ihre intuitive Schnittstelle eine natürliche Sprachinteraktion. Nutzer:innen können ohne spezielle Befehle oder Programmiersprachen in natürlicher Sprache mit dem LLM kommunizieren, was die Zugänglichkeit erheblich erhöht (McTear & Ashurkina, 2024). Die Fähigkeit der LLMs, den Kontext von Anweisungen (Prompts) präzise zu erfassen und kontextbezogene Antworten zu generieren, fördert die breite Akzeptanz und Anwendung der Technologie in verschiedenen Bereichen wie dem Finanzbereich.

So wird ein bedeutender Anwendungsbereich von LLMs im Finanzbereich in der **Automatisierung der Text- und Finanzberichterstellung** gesehen. LLMs sind in der Lage Texte zu synthetisieren und Berichte zu erstellen. Basierend auf bestehenden Informationen eignen sich LLMs für die Erstellung von Präsentationen, Anhängen, Lageberichten oder nichtfinanziellen Berichten, da diese einen starken Textbezug aufweisen (Leitner-Hanetseder et al., 2023; Wegenstein & Waniczek, 2023; Hacker, 2023). In der Literatur wird auch darauf hingewiesen, dass generative KI vielfältige Anwendungsmöglichkeiten in der Kommunikation mit Stakeholder:innen bietet. So können E-Mails anhand von Stichwörtern ausformuliert werden. In Bezug auf **Routineanfragen** ist der Automatisierungsgrad noch viel höher, wenn diese vom LLM kategorisiert und **automatisiert innerhalb kürzester Zeit beantwortet werden**, was zu einer Effizienzsteigerung und verbesserten Kund:innensupport führt (Hacker, 2023; Leitner-Hanetseder et al., 2023; Wegenstein & Waniczek, 2023). Ein weiterer in der Literatur oft hervorgehobener Nutzen von LLMs im Finanzbereich ist der **Wissenserwerb**. Generative KI kann umfassende Informationen sammeln und zusammenfassen (zB zu Neuerungen von Rechnungslegungsstandards bzw. Berichterstattungsanforderungen) (Schulze et al., 2023; Singh & Singh, 2023; Zadorozhnyi et al., 2023). Fachkräfte können durch LLMs schnell Antworten zu Fragestellungen erhalten, die ansonsten zeitaufwendig zu recherchieren wären bzw. könnten auch personalisierte Schulungsmaterialien erhalten (Hüsch et al., 2023). LLMs sind weiters in der Lage bereits vorhandene Texte in andere Sprachen zu übersetzen (Pröhl et al., 2023; Schulze et al., 2023; Singh & Singh, 2023). Dies stellt eines der primären Ziele der Entwicklung von Künstlicher Intelligenz dar, welches bereits während des Kalten Krieges verfolgt wurde (Wennker, 2020). Durch das entsprechende Training des LLMs ist es heutzutage zudem möglich, Texte **im spezifischen Jargon des Unternehmens** (zB für die Lageberichterstattung eines spezifischen Unternehmens) oder **der Nutzer:innen** (zB für Infoschreiben an Klient:innen einer Steuerberatung, die ein Grundverständnis von regulatorischen Regeln in einfacher Sprache vermitteln (Zhao & Wang, 2023)) zu formulieren (Schulze et al., 2023; Wegenstein & Waniczek, 2023;). LLMs unterstützen auch die Datenverarbeitung im Hintergrund zB durch die **Erstellung komplexer Makros** in Excel (Dhana, 2023) oder als **Text-To-Code-Generator für SQL-Abfragen und Python-Codes**, wodurch die Abhängigkeit von der IT-Abteilung verringert wird (Wegenstein & Waniczek, 2023). Ein weiterer bedeutender Anwendungsbereich von LLMs, die in der Literatur diskutiert wird, ist die Analyse und Identifikation von Mustern und Anomalien in großen Datenmengen sowie daraus abgeleitete Empfehlungen (Chen et al., 2023; Hüsch et al., 2023; Rane, 2023; Singh & Singh, 2023). Die Bereitstellung der von LLMs generierten Empfehlungen können das Management bei der Entscheidungsfindung unterstützen. Dies kann insbesondere in komplexen und dynamischen Geschäftsumfeldern von großem Vorteil sein, um unterschiedliche Aspekte, die so für das Management nicht ersichtlich sind, aufzuzeigen. Die in der Literatur diskutierten Einsatzmöglichkeiten sind damit vielfältig. Rund 1,5 Jahre nach Veröffentlichung von ChatGPT gibt es jedoch noch uWn keine empirischen Studien wie ChatGPT von Mitarbeiter:innen im Finanzbereich genutzt wird.¹

¹ Eine allgemeine Studie über Einstellungen, Wünsche und Nutzungsverhalten von Österreicher:innen wurde von im März 2023 von PwC (Hrsg.) (2023) veröffentlicht.

Um die aktuelle Nutzung von LLMs im Finanzbereich besser zu verstehen, wurde eine quantitative Online-Befragung durchgeführt, an der 203 Personen aus dem Finanzbereich teilnahmen. Diese Untersuchung erfasste die Nutzungshäufigkeit und Anwendungsbereiche von ChatGPT & Co. im beruflichen Kontext und analysierte die größten wahrgenommenen Herausforderungen sowie wie und inwieweit Mitarbeitende aus dem Finanzbereich Prompt Engineering zur Optimierung des Outputs von LLMs heranziehen.

Der Beitrag gliedert sich in mehrere Abschnitte. In der Einführung in Kapitel eins werden die grundlegenden Funktionen von LLMs sowie den im Finanzbereich diskutierten Anwendungsmöglichkeiten vorgestellt. Das zweite Kapitel erläutert Prompt Engineering zur Optimierung des Outputs von LLMs, einschließlich Techniken wie Zero-Shot- und Few-Shot-Prompting. Im dritten Kapitel wird die empirische Studie dargelegt, die die Nutzungshäufigkeit und Anwendungsbereiche von LLMs als auch die größten wahrgenommenen Herausforderungen sowie die Anwendung von Prompt Engineering im Finanzbereich untersucht. Ebenso werden alters- und geschlechterspezifische Diskrepanzen in der Nutzung identifiziert werden. Das abschließende Kapitel vier fasst die wichtigsten Ergebnisse zusammen und gibt einen Ausblick auf zukünftige Forschungsfelder.

2 Prompt Engineering zur Optimierung des Outputs von Large Language Models

Um die Leistung des Outputs eines LLMs zu steigern, stehen eine Vielzahl von Ansätzen zur Verfügung. Diese umfassen die gezielte Formulierung von Prompts, die Erweiterung der Trainingsdaten und die Anpassung des Modells selbst. Im folgenden Kapitel wird erläutert wie Nutzer:innen, und damit Mitarbeiter:innen des Finanzbereichs das sogenannte Prompt Engineering nutzen können, um die Leistung von LLMs weiter zu steigern und präzisere Ergebnisse zu erzielen. Ein Prompt umfasst die Anweisungen, die eine Benutzer:in an ein LLM richtet, um den gewünschten Output zu generieren (White et al., 2023). Im Zuge von ChatGPT erfolgt diese über ein Textfeld in der für die Nutzer:in zugänglichen Benutzeroberfläche. Mittels Prompt werden die Informationen, die vom LLM angefordert werden, spezifiziert. Der Prompt bietet somit Kontext sowie Richtlinien für die Generierung des Outputs. Im Kern stellen Prompts den von der Benutzer:in eingegebenen Text in einem Chatbot dar. Prompt Engineering umfasst somit die Kunst der Erstellung von Prompts, die es dem LLM ermöglichen, benutzer:innendefinierte Outputs zu generieren (Ekin, 2023). Die Herausforderung besteht in der Feinabstimmung der Prompts, da selbst geringfügige Änderungen signifikante Auswirkungen auf den generierten Output haben können (Taulli, 2023). Beim sog. **Zero-Shot-Prompting** ist es von entscheidender Bedeutung, dem LLM einen adäquaten Kontext bereitzustellen, der es befähigt, einen präzisen und situationsabhängigen Output zu generieren (Taulli, 2023). In Bezug auf den Finanzbereich könnte dies zB die Angabe des Rechnungslegungsstandards oder einer spezifischen Information im Unternehmenskontext sein (Street et al., 2023). Die Verwendung eindeutiger Begriffe und Ausdrücke ist von besonderer Relevanz, um potenzielle Missverständnisse zu vermeiden. Dies ist insbesondere bei der Nutzung von Akronymen von Relevanz, da diese potenziell mehrdeutig sein können (Street et al., 2023; Taulli, 2023). Die Literatur empfiehlt außerdem dem Large Language Model eine spezifische Rolle zuzuweisen, zB „Du bist Bilanzierer eines österreichischen KMUs in der Baubranche“. Diese Vorgabe ermöglicht die Konzentration auf einen spezifischen Bereich. Dieser Ansatz kann auch bei kreativen Prozessen von Vorteil sein, wenn beispielsweise ein bestimmter Schreibstil in E-Mails gewünscht wird (Gao, 2023). Aufgrund der begrenzten Kapazität des Kontextfensters und der Tendenz des LLMs, Informationen am Anfang einer Konversation im Laufe der Zeit zu vergessen, besteht die Möglichkeit die bereits erwähnten Informationen dem LLM mithilfe der sogenannten System Prompts zu übermitteln. **System Prompts**, auch bekannt als Meta- oder Base-Prompts, ermöglichen es, dem LLM Informationen bereitzustellen, die in allen Konversationen berücksichtigt werden sollen. Umfangreiche System Prompts zu verschiedenen Themen können direkt über die API oder spezielle Benutzeroberflächen erstellt werden (McTear & Ashurkina, 2024). In Abgrenzung dazu sind die sogenannten Prompt Settings zu nennen, bei denen Parameter der Wahrscheinlichkeitsrechnung modifiziert werden können. Exemplarisch sei hier der Parameter „Temperature“ genannt, der die Zufälligkeit des Outputs steuert. Um die genannten Parameter anzupassen, muss direkt über APIs oder spezielle Benutzeroberflächen, wie etwa den „Playground“ von OpenAI, auf das LLM zugegriffen werden (McTear & Ashurkina, 2024). Beim **Few-Shot Prompting** werden dem LLM Beispiellösungen vorgelegt. Das Ziel besteht darin, dem Modell durch die Vorlage von Input-Output-Varianten aufzuzeigen, welche

Art von Output erwartet wird. Auf diese Weise erfolgt eine sowohl inhaltliche als auch formale Einschränkung möglicher Antworten (Gao, 2023; Sahoo et al., 2024). Bei komplexen Aufgaben, die mehrere Lösungsschritte erfordern, sind LLMs fehleranfälliger. Daher empfiehlt es sich, umfangreiche Aufgaben in Subaufgaben aufzuteilen und jede dieser Aufgaben dem LLM einzeln zu stellen. Dadurch kann die Nutzer:in die einzelnen Schritte auf ihre Korrektheit leichter überprüfen (Street et al., 2023). Ist eine Aufgliederung der Aufgabe nicht möglich oder nicht gewünscht, kann nach dem Prinzip des Few-Shot-Prompting eine „**Chain-of-Thought**“ als Muster zur Verfügung gestellt werden. Dabei wird dem LLM anhand einer Beispielaufgabe der Lösungsweg aufgezeigt. Dadurch lernt das System, wie es bei der Lösung der Aufgabe vorgehen und welche logischen Schritte vorzunehmen sind (Sahoo et al., 2024). Studien haben gezeigt, dass allein die Anweisung „Denke Schritt für Schritt“ ausreichend ist, um die Genauigkeit der Ausgabe des LLMs zu erhöhen (Gao, 2023). Wenn es nicht möglich ist, einen Lösungsweg beizufügen, kann die Verwendung eines „Fact-Check“ eine Option sein. Dabei wird das LLM aufgefordert, eine Liste aller von ihm verwendeten Fakten zu erstellen, ev. einschließlich der Quellen. Auf diese Weise können die Benutzer:innen die Richtigkeit dieser Fakten überprüfen (White et al., 2023).

3 Empirische Studie

3.1 Forschungsfragen

Folgende Forschungsfragen gilt es im Zuge der empirischen Untersuchung zu klären.

- Inwieweit nutzen Mitarbeiter:innen im Finanzbereich im privaten oder beruflichen Kontext ChatGPT & Co. und können Alters- und Geschlechterdiskrepanzen identifiziert werden?
- Welche Anwendungsbereiche und Herausforderungen werden in der Anwendung von ChatGPT & Co. Mitarbeiter:innen im Finanzbereich ChatGPT & Co. gesehen?
- Inwieweit setzen Personen im Finanzbereich „Prompt Engineering“ zur Optimierung des Outputs von ChatGPT & Co. ein?

3.2 Methodik

Die empirische Studie wurde als quantitative Befragung mittels Online-Fragebogen des Softwaretools „LimeSurvey“ umgesetzt. Der Fragebogen wurde insgesamt an 768 Absolvent:innen und 178 aktuell Studierende des Bachelor- und Masterstudiengangs „Controlling, Rechnungswesen und Finanzmanagement“ (kurz: CRF) versandt, wobei der Link zum Fragebogen auch über persönliche Kontakte via WhatsApp verbreitet wurde. Nur 2 % der CRF Studierenden sind nicht berufstätig bzw. haben keine Berufserfahrung. Die Beantwortung des Online-Fragebogens war von 17. bis 28. April 2024 möglich. Den Fragebogen haben insgesamt 203 Teilnehmer:innen vollständig beantwortet. Abbildung 1 zeigt eine nahezu ausgeglichene Geschlechterverteilung mit einem leichten Überhang weiblicher Teilnehmer:innen (51 %) im Vergleich zu männlichen Teilnehmern (48 %), während 1 % der Befragten keine Angabe zum Geschlecht gemacht haben. Dies deutet darauf hin, dass die Ergebnisse der Studie weitgehend auf eine gleichmäßige Repräsentation beider Geschlechter basieren können, was die Aussagekraft und Generalisierbarkeit der Erkenntnisse erhöht.

Geschlechterverteilung

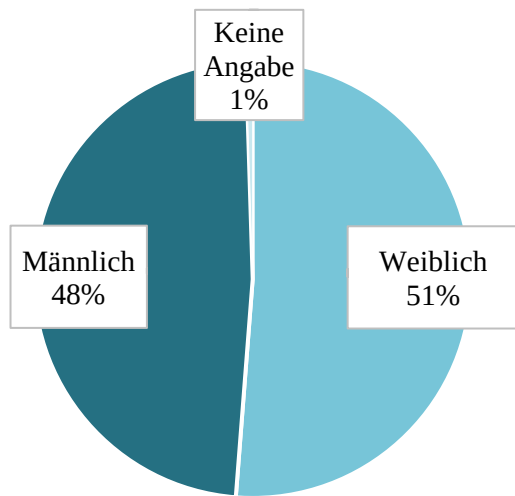


Abbildung 1: Geschlechterverteilung (eigene Darstellung)

Berufserfahrung (im Finanzbereich)

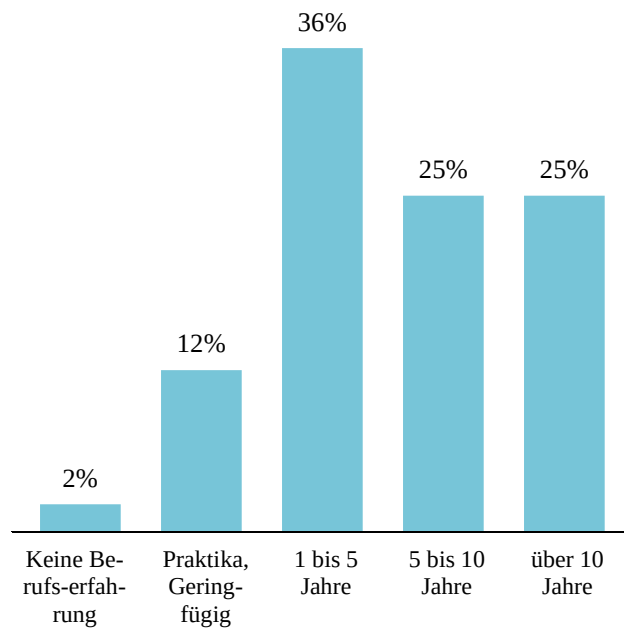


Abbildung 2: Berufserfahrung (eigene Darstellung)

Altersverteilung

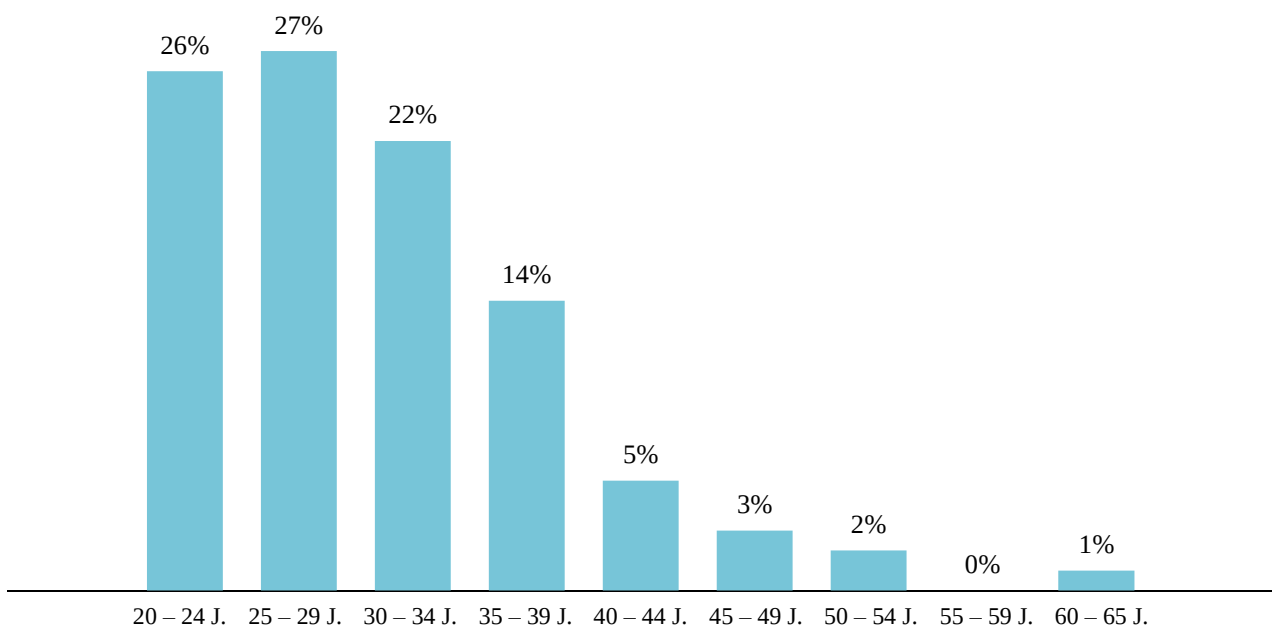


Abbildung 3: Altersverteilung (eigene Darstellung)

In Abbildung 3 ist eine deutliche rechtsschiefe Altersverteilung der Teilnehmer:innen zu erkennen. Das durchschnittliche Alter beläuft sich auf 30,3 Jahre. Diese Verteilung spiegelt die Altersstruktur der aktiv Studierenden des Studiengangs CRF als auch der Absolvent:innen des Studiengangs „CRF“ wieder. Aufgrund des relativ niedrigen Durchschnittsalters ist auch die Berufserfahrung der Teilnehmer:innen mehrheitlich gering (vgl. Abbildung 2). Vier Befragte (2 %) geben an, keine Berufserfahrung zu besitzen, während 25 Personen (12 %) lediglich über Erfahrungen aus Praktika oder geringfügiger Beschäftigung verfügen. Die größte Gruppe bildet mit 73 Personen (36 %) jene mit 1 bis 5 Jahren Berufserfahrung. Die Berufserfahrung von 5 bis 10 Jahren sowie jene von über 10 Jahren sind mit jeweils 25 % (50 bzw. 51 Personen) gleich verteilt. In Bezug auf das berufliche Tätigkeitsfeld der Befragten zeigt sich, dass über 40 % der Teilnehmer:innen im Bereich Controlling tätig sind, gefolgt von etwa 17 %

im externen Rechnungswesen, 12 % in der Steuerberatung und jeweils rund 7 % als Wirtschaftsprüfer:innen und Finanzberater:innen tätig sind (vgl. Abbildung 4). Zum sonstigen Bereich gehört zB der Tätigkeitsbereich der Unternehmensberatung.

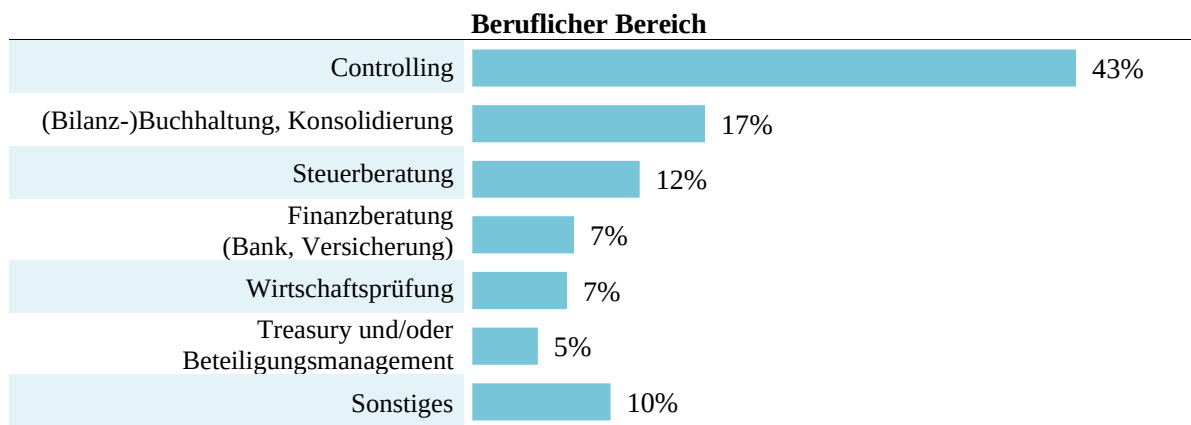


Abbildung 4: Beruflicher Bereich (eigene Darstellung)

3.3 Ergebnisse der Befragung

3.3.1 Nutzungsverhalten von ChatGPT & Co. im privaten und beruflichen Kontext und der Einfluss von Alter und Geschlecht

Zu Beginn des Fragebogens werden die Teilnehmer:innen zu ihrem Nutzungsverhalten bezüglich ChatGPT & Co. befragt. Bei der Nutzungshäufigkeit im privaten Kontext geben lediglich 19 % der Befragten an, ChatGPT & Co. nicht privat zu nutzen und 38 % ChatGPT & Co. nicht beruflich nutzen (vgl. Abbildung 5). Während 16 % der Teilnehmer:innen ChatGPT & Co. wöchentlich oder mehrmals pro Woche privat nutzen, setzen nur 15 % diese Technologie wöchentlich und 11 % mehrmals pro Woche beruflich ein. Eine tägliche Nutzung von ChatGPT & Co. sowohl im privaten (4 %) als auch im beruflichen Kontext (3 %) bleibt die Ausnahme. Es stellt sich die Frage, inwieweit ein Zusammenhang zwischen der privaten Nutzungshäufigkeit und der beruflichen Nutzungshäufigkeit besteht. Daraus lässt sich folgende Null-Hypothese formulieren:

NH1: Es gibt keinen signifikanten Zusammenhang zwischen der Nutzung im privaten und beruflichen Kontext.

Das Ergebnis der Spearman Korrelation zeigt einen statistisch signifikanten mittleren bis hohen positiven Zusammenhang ($r(200) = 0,52$, $p < 0,001$) (Kuckartz et al., 2013) zwischen der Nutzungshäufigkeit im privaten und dem beruflichen Kontext gibt. Das bedeutet, dass mit steigender privater Nutzungshäufigkeit auch die berufliche Nutzungshäufigkeit tendenziell steigt und umgekehrt.

Wie oft nutzen sie ChatGPT & Co. im ...? (in % von 203)

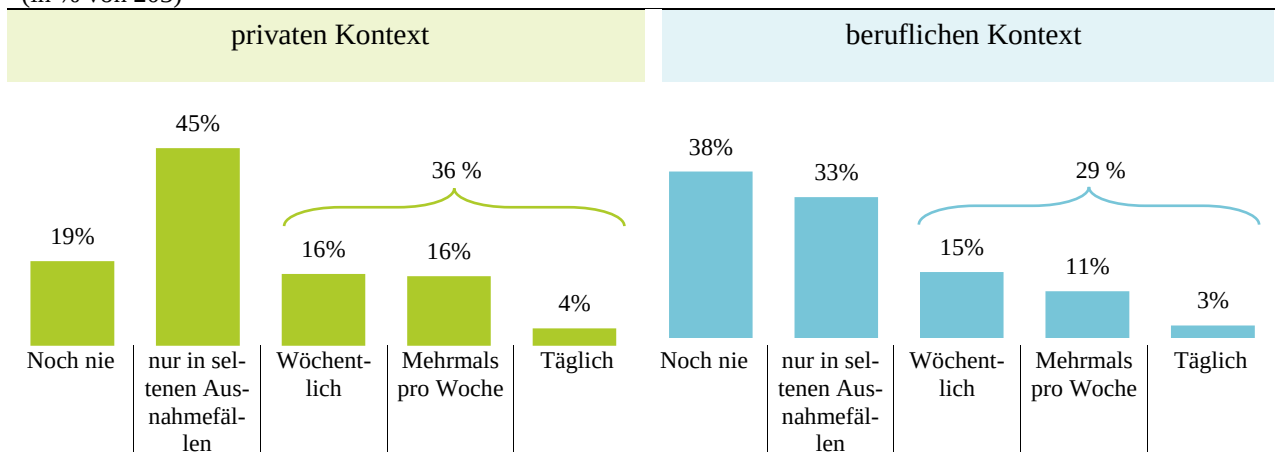


Abbildung 5: Nutzungshäufigkeit ChatGPT und Co. (eigene Darstellung)

Studien zeigen, dass Alter² und Geschlecht (West et al., 2019) die Akzeptanz neuer Technologien beeinflussen. Jüngere Menschen nehmen sie schneller an, während ältere Menschen sie als schwieriger empfinden (Draxler, 2023, Morris et al., 2006). Männer nutzen neue Technologien tendenziell früher als Frauen, was auf Unterschiede im Selbstvertrauen und technischer Ausbildung zurückzuführen ist (West et al., 2019). Ein besseres Verständnis der demografischen Faktoren ist wichtig, um den Zugang zu dieser neuen Technologie, die in unglaublich raschem Tempo die Nutzer:innen erreicht, gerechter zu gestalten. Intergenerationelle Kommunikation und Technologieausbildung können helfen, Alters- und Geschlechterbarrieren zu überwinden, unterstützt durch die UN-Nachhaltigkeitsziele für hochwertige Bildung und Geschlechtergleichstellung.

Um zu erheben, inwieweit Gender- als auch Altersdiskrepanzen bezüglich der Nutzungshäufigkeit bei Mitarbeiter:innen im Finanzbereich vorliegen, werden folgende vier Nullhypothesen formuliert.

NH2: Es gibt keinen signifikanten Zusammenhang zwischen dem Geschlecht und der Nutzungshäufigkeit von ChatGPT & Co. bei den befragten Mitarbeiter:innen im Finanzbereich im privaten Kontext.

NH3: Es gibt keinen signifikanten Zusammenhang zwischen dem Geschlecht und der Nutzungshäufigkeit von ChatGPT & Co. bei den befragten Mitarbeiter:innen im Finanzbereich im beruflichen Kontext.

NH4: Es gibt keinen signifikanten Zusammenhang zwischen dem Alter (im Sinne von Altersgruppe) und der Nutzungshäufigkeit von ChatGPT & Co. bei den befragten Mitarbeiter:innen im Finanzbereich im privaten Kontext.

NH5: Es gibt keinen signifikanten Zusammenhang zwischen dem Alter (im Sinne von Altersgruppe) und der Nutzungshäufigkeit von ChatGPT & Co. bei den befragten Mitarbeiter:innen im Finanzbereich im beruflichen Kontext.

Die deskriptive Analyse des Nutzungsverhaltens nach Altersgruppen offenbart eine deutliche Tendenz. Demnach nutzen jüngere Teilnehmer:innen häufiger ChatGPT als Ältere. Wie in Abbildung 6 illustriert, nutzen 48 % der unter 30-Jährigen regelmäßig und damit zumindest wöchentlich ChatGPT & Co. Es zeigt, sich das mit zunehmendem Alter die Bereitschaft sinkt ChatGPT & Co. für private Zwecke zu nutzen. Nur mehr 9 % der über 40ig-Jährigen nutzen ChatGPT & Co. regelmäßig. Dies deutet darauf hin, dass jüngere Generationen im privaten Kontext offener und vertrauter im Umgang mit neuen Technologien wie ChatGPT & Co. sind, während ältere Generationen zurückhaltender sind und seltener auf solche Tools im Alltag zurückgreifen. Eine tägliche Nutzung bleibt in allen Altersgruppen die Ausnahme. Eine Studie der IU Internationalen Hochschule in Erfurt aus dem April 2023 ergab, dass

² Im Kontext mit LLMs siehe Draxler et al. (2023), im Allgemeinen zu Altersdiskrepanz und Technologieanwendung siehe Morris et al. (2005), S. 375 ff.

die Hauptgründe für die Nutzung von ChatGPT und Co. Unterhaltung und Neugier sind. Die tatsächliche Nutzung für das Studium oder zum Erlernen neuer Fähigkeiten spielt demgegenüber eine weniger wichtige Rolle. Es ist bemerkenswert, dass im Rahmen der vorliegenden Studie dennoch festgestellt wurde, dass 66 % der Studierenden³ berichten, ChatGPT und Co. regelmäßig im privaten Rahmen zu nutzen, im Gegensatz zu lediglich 36 % der gesamten Befragtengruppe. Ein Vergleich der privaten Nutzung unter den Nicht-Studierenden zeigt, dass nur 24 % der unter 30-Jährigen regelmäßig ChatGPT verwenden. Diese Daten legen nahe, dass insbesondere Studierende eine höhere Affinität zur regelmäßigen Nutzung von ChatGPT aufweisen.

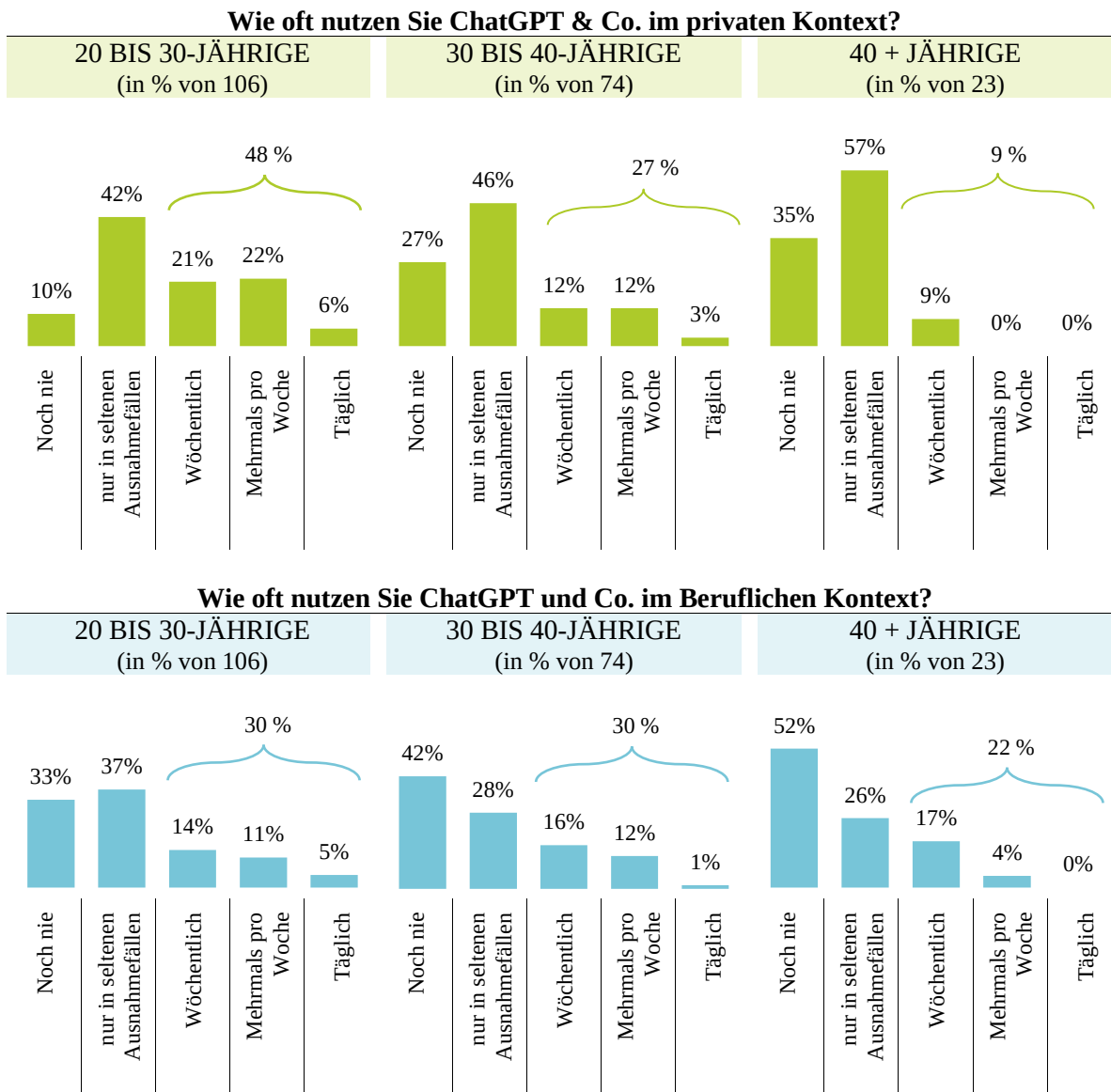


Abbildung 6: Berufliche und Private Nutzung nach Altersgruppe (eigene Darstellung)

Basierend auf einem Chi-Quadrat-Test zeigt sich, dass kein statistisch signifikanter Zusammenhang zwischen dem Geschlecht der Proband:innen als auch der Häufigkeit der Nutzung im beruflichen Bereich sowie zwischen der Altersgruppe und der beruflichen Nutzung festgestellt werden konnte (vgl. Tabelle 1). Basierend auf einer Spearman

³ Studierende sind in dieser Studie zu 90 % unter 30 Jahre.

Korrelationsanalyse lässt sich ein leichter statistisch signifikanter positiver Zusammenhang nur zwischen der Altersgruppe und der Häufigkeit der privaten Nutzung feststellen ($r(200) = 0,25$, $p < 0,001$). Demnach nimmt mit zunehmendem Alter die Nutzungshäufigkeit von ChatGPT & Co. im privaten Kontext ab und vice versa. Eine ähnliche Tendenz lässt sich jedoch nicht für den beruflichen Kontext und das Alter nachweisen (vgl. Tabelle 2).

Tabelle 1 Chi-Quadrat Hypothesentest (eigene Darstellung)

	Chi2 (χ^2)	df	p
Zusammenhang Geschlecht und private Nutzung	5,5	4	0,24
Zusammenhang Geschlecht und berufliche Nutzung	6,07	4	0,194

Tabelle 2 Spearman Korrelationsanalyse (eigene Darstellung)

	r(200)	p
Zusammenhang Altersgruppe und private Nutzung	0,25	< 0,001
Zusammenhang Altersgruppe und berufliche Nutzung	0,07	0,328

3.3.2 Aufgaben, die mit ChatGPT & Co. im Berufsalltag, bewältigt werden

Abbildung 7 illustriert, für **welche Aufgaben** die 59 Personen, die ChatGPT beruflich regelmäßig nutzen, ein LLM einsetzen. Die häufigste Nutzung findet mit **64 % im Bereich des Wissenserwerbs** statt. Diese Praxis ist jedoch kritisch zu betrachten, da der Output von LLMs von Halluzinationen⁴ und Ungenauigkeiten geprägt sein kann. Ein weiterer signifikanter Anwendungsbereich von LLMs unter den Teilnehmer:innen ist die **Beantwortung von E-Mails**. Hierbei berichten 56 % der Nutzer:innen, dass sie zumindest zeitweise nicht mehr selbst E-Mail-Antworten formulieren, sondern lediglich Stichworte vorgeben und diese dann mit Hilfe von LLMs in einen situationsangepassten E-Mail-Text umwandeln lassen. Des Weiteren berichten 24 % der Nutzer:innen, dass sie teilweise ihre Antwort-E-Mails vollständig von LLMs formulieren lassen, ohne zuvor den Inhalt selbst zu erarbeiten. Nicht nur kurze Texte, wie E-Mails, werden mithilfe von LLMs verfasst; 37 % der Befragten nutzen LLMs auch zur **Erstellung ganzer Berichtstexte** anhand von Stichworten. Hier stellt sich die Frage, ob diese durch LLMs formulierten Berichtstexte dieselben sind, die später von den Berichtsempfänger:innen mithilfe von LLMs wiederum zusammengefasst werden. Denn 54 % der Teilnehmer:innen geben an, LLMs zur Erstellung von Zusammenfassungen zu nutzen. Die Mehrsprachigkeit von LLMs wird von 42 % der Nutzer:innen zur **Text-to-Text Übersetzung** genutzt. Zudem nutzen 32 % der Befragten die Fähigkeit von LLMs der Beherrschung von Programmiersprachen. Dabei werden LLMs sowohl in der **Text-to-Code-Übersetzung** als auch für die Verbesserung von Programmiercodes eingesetzt. Diese relativ hohe Nutzungsdichte unterstreicht, dass auch im Finanzbereich gewisses IT-Wissen und technische Fähigkeiten partiell zur Anwendung kommen. Darüber hinaus unterstützen LLMs 31 % der Teilnehmer:innen bei der **Entscheidungsfindung**. Unter den sonstigen Tätigkeiten fallen die Datenanalyse und die Erstellung von Lehrmaterialien. Diese Nutzungsvielfalt zeigt, dass LLMs ein vielseitiges Werkzeug im professionellen Umfeld darstellen, die über sprachbasierte Aufgaben in unterschiedlichen Phasen des Schreib- und Leseprozesses hinaus auch in technischen und entscheidungsunterstützenden Funktionen zum Einsatz kommen.

⁴ Im Kontext von LLMs bezeichnet der Begriff „Halluzinieren“ das Phänomen, wenn der generierte Output des Modells falsche oder unsinnige Aussagen enthält (vgl. Kulkarni et al., 2023, S. 146.).

Für welche der folgenden Tätigkeiten nutzen sie LLMs im Berufsalltag?

(in % von 59)

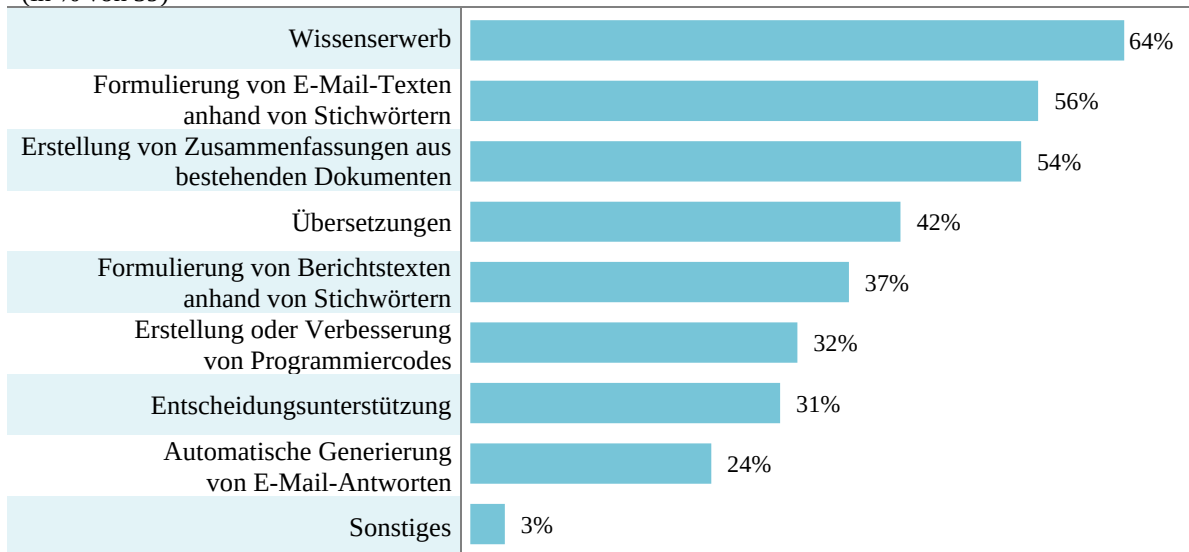


Abbildung 7: Anwendungen von LLMs im beruflichen Kontext des Finanzbereichs (eigene Darstellung)

3.3.3 Wahrgenommene Herausforderungen

Die letzte Frage der Fragengruppe „Nutzung von ChatGPT und Co.“ untersucht die **Herausforderungen** bei der Nutzung von LLMs im beruflichen Kontext. Diesbezüglich werden die Befragten gebeten, die aus Ihrer Sicht drei größten Herausforderungen zu identifizieren. Wie aus Abbildung 8 ersichtlich, identifizieren **77 % der Teilnehmer:innen die Halluzinationen von LLMs** im Zusammenhang mit der daraus resultierenden Schwierigkeit, diese Fehlinformationen als solche zu erkennen, als eine der größten Herausforderungen. Umso erstaunlicher, dass 64 % der Befragten, LLMs als Werkzeug für den Wissenserwerb einsetzen. An zweiter Stelle der größten Herausforderungen steht mit **60 % das Problem der Datensicherheit** bei der Nutzung von LLMs. Vor diesem Hintergrund ist zu hoffen, dass Personen, die ChatGPT zur Generierung von Berichten und E-Mails nutzen, keine sensiblen Daten zur Textgenerierung in ChatGPT eingeben. An dritter Stelle mit **41 % liegt die Herausforderung der ungewissen Rechtslage bezüglich der Nutzung von LLMs**. Per dato noch offen ist, inwieweit mit Umsetzung und Anwendung der KI-Verordnung der EU (2024/1689) diese Herausforderung verringert werden kann. Zudem sehen **32 % der Befragten in der „mangelnden Transparenz“** eine Herausforderung, insbesondere bezüglich der Trainingsdaten vieler LLM-basierter Chatbots und der Tatsache, dass der Output von LLMs auf Wahrscheinlichkeitsberechnungen basiert und somit nicht nachvollziehbar ist. In der Untersuchung betrachten **29 % der Befragten das mangelnde Verständnis der Funktionsweise von LLMs** als eine Herausforderung. Dieser relativ geringe Prozentsatz sollte jedoch nicht als Indikator dafür gesehen werden, dass ein umfangreiches Wissen über die Technologie hinter LLMs in der Bevölkerung vorhanden ist. Die Ergebnisse der Umfrage zeigen vielmehr, dass bereits grundlegende Begriffe im Zusammenhang mit LLMs, wie „Prompt Engineering“, einen geringen Bekanntheitsgrad aufweisen (vgl. Abbildung 9). Daraus lässt sich ableiten, dass umfassendes Wissen über die Technologie oft nicht als Voraussetzung angesehen wird, um LLMs zu nutzen. Die Vorlaufkosten und laufenden Kosten für den Einsatz von LLMs werden lediglich von 22 % der Befragten als eine der drei größten Herausforderungen identifiziert. Die Gründe hierfür können vielfältig sein. Einerseits wird möglicherweise der Schulungsbedarf für Mitarbeiter:innen unterschätzt, eine weitere mögliche Auswirkung der geringen Bekanntheit der Methode des Prompt Engineerings (vgl. Abbildung 9). Andererseits könnten die gering wahrgenommenen Kosten dadurch bedingt sein, dass öffentlich-zugängliche kostenlose LLMs oder kostengünstige Komplettpakete eingesetzt werden, die kein unternehmensspezifisches Training erfordern. Um die genauen Kosten von LLMs in Unternehmen zu evaluieren, sind weiterführende Studien notwendig. Zusätzlich geben nur 15 % der Befragten an, dass der Bedarf an internen Trainingsdaten eine große Herausforderung darstellt. Dies lässt keine eindeutigen Schlüsse zu, ob die Digitalisierung in Unternehmen bereits so weit

fortgeschritten ist, dass ausreichend Daten vorhanden sind, oder ob LLMs mit unternehmensinternen Daten nicht zum Einsatz kommen. Es könnte auch sein, dass diese Problematik zwar erkannt wird, aber andere Punkte als wichtiger erachtet werden. Die beschriebenen Aspekte unterstreichen die Notwendigkeit, den Einsatz von LLMs in Unternehmenskontexten detaillierter zu untersuchen, um die Ursachen und Auswirkungen der spezifischen Herausforderungen besser zu verstehen.

Was sind Ihrer Meinung nach die 3 größten Herausforderungen in Bezug auf die Nutzung von LLMs in Bezug auf berufliche Aufgaben?

(in % von 203)

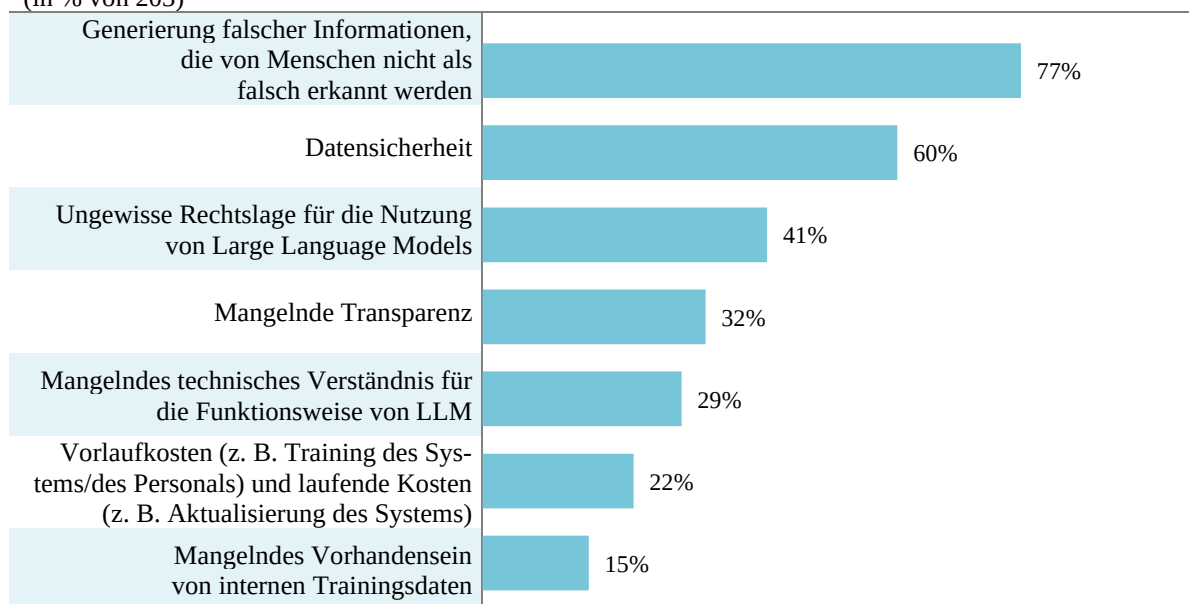


Abbildung 8: Wahrgenommene Herausforderungen (eigene Darstellung)

Im Anschluss an die Benennung der drei größten Herausforderungen haben die Teilnehmer:innen die Möglichkeit in einer offenen Frage weitere Schwierigkeiten im Umgang mit LLMs anzuführen. Hierbei wird wiederholt angeführt, dass unter den Kolleg:innen ein deutliches Defizit hinsichtlich des Wissens über die Einsatzmöglichkeiten von LLMs besteht. Diese Problematik ist eng verknüpft mit einer mangelnden Akzeptanz seitens einzelner Personen sowie einem unzureichenden Verständnis darüber, wie Schulungen zur Reduktion dieser Problematik gestaltet sein sollten. Zusätzlich wird die Qualität der Trainingsdaten von öffentlich zugänglichen LLMs als problematisch beschrieben. Es wird sowohl die fehlende Aktualität der Trainingsdaten als auch ein Mangel an fachlich hochwertigen Texten in den Trainingsdaten hervorgehoben. Insbesondere rechtliche Vorgaben sind dem LLM nicht bekannt, was bei einem unsachgemäßen Einsatz zu erheblichen Problemen führen kann. Des Weiteren wird die Integration von LLMs in bestehende Systemarchitekturen als herausfordernd beschrieben.

3.3.4 Einsatz von Prompt Engineering im Finanzbereich

Im Zuge der Studie wurde auch erhoben, inwieweit Prompt Engineering (vgl. Kapitel 2) den Befragten bekannt ist und zum Einsatz kommt, um den Output zu optimieren. Die Ergebnisse zeigen, dass 73 % der Teilnehmer:innen, obwohl sie ChatGPT verwenden, den Begriff „Prompt Engineering“ nicht kennen. Gemäß den in Abbildung 10 dargestellten Daten geben nur 9 % an, mit Prompt Engineering vertraut zu sein und es aktiv anzuwenden. Des Weiteren geben 18 % der Befragten an, den Begriff zu kennen, die Methode jedoch nicht zu nutzen.

Im nächsten Schritt werden den Teilnehmer:innen 10 aus der Literatur abgeleiteten Nutzungsempfehlungen des Prompt Engineering aufgezeigt. Auf die Frage, welche der 10 vorgeschlagenen Nutzungsempfehlungen von den Teilnehmer:innen eingesetzt werden, geben 138 Personen (ca. 70 %) an, mindestens eine der im Fragebogen beschriebenen Nutzungsempfehlungen zu nutzen (vgl. Abbildung 9). Im Durchschnitt werden von den Befragten 2,77 von 10 der Nutzungsempfehlungen angewendet. Diese Diskrepanz zwischen dem Wissen über Prompt Engineering und das Anwenden von Prompt Engineering-Methoden lässt darauf schließen, dass vielen Nutzer:innen nur der Begriff bzw. die vollständigen Facetten des Prompt Engineering unbekannt sind. Die Ergebnisse zeigen, dass ein **beträchtlicher Anteil der Nutzer:innen unbewusst Techniken des Prompt Engineering anwendet**, um die Interaktionseffizienz mit ChatGPT zu steigern, um dadurch potenziell genauere und relevantere Antworten zu generieren.

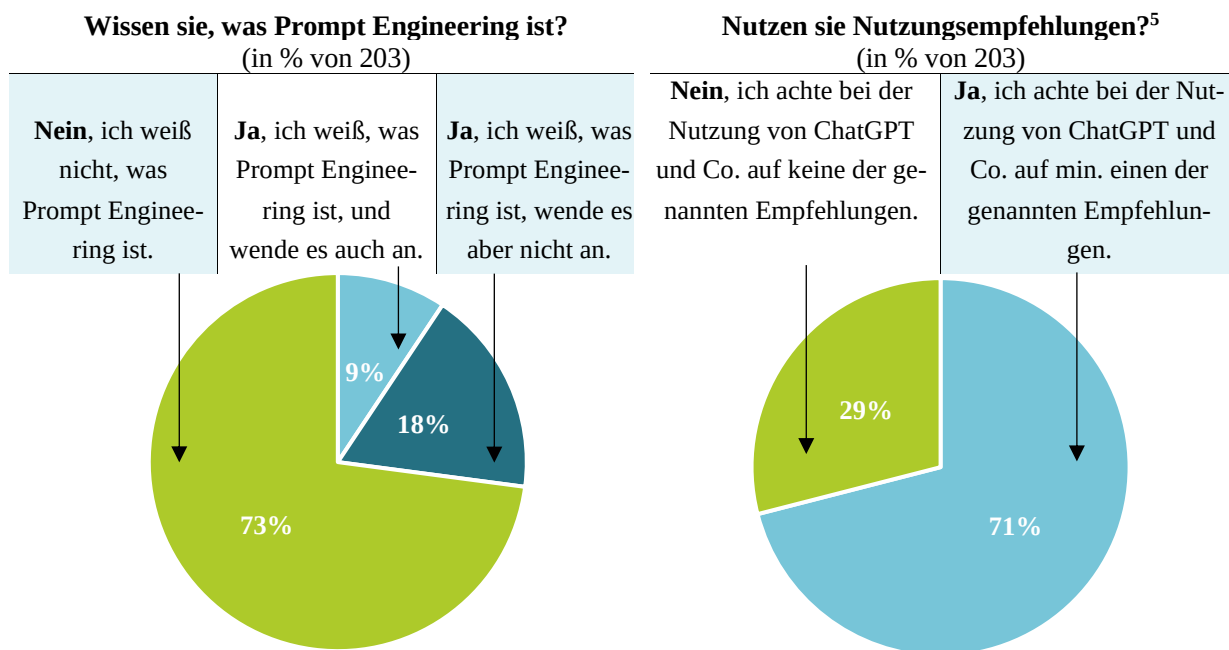


Abbildung 9: Kenntnis und Nutzung von Prompt Engineering im Finanzbereich (eigene Darstellung)

In Abbildung 10 wird dargestellt, welche in der Literatur beschriebenen Prompt Engineering Methoden von den 138 Nutzer:innen, die Nutzungsempfehlungen anwenden, verwendet werden. 77 % der Teilnehmer:innen geben an, dass sie im Prompt explizit spezifizieren, in welchem Format der Output erfolgen soll. Hierzu zählen beispielsweise Vorgaben bezüglich der Anzahl der Wörter oder Sätze sowie des Tonfalls. Des Weiteren stellen 65 % der Personen gezielt Informationen über den Kontext, in dem die Antwort benötigt wird, ChatGPT zur Verfügung. Zudem geben 38 % der Teilnehmer:innen an, welche Rolle das LLM übernehmen soll, etwa durch eine Formulierung wie „Du bist Bilanzierer eines österreichischen, produzierenden KMUs“. Diese drei am häufigsten genannten Aspekte unterstützen ChatGPT dabei, einen Output zu generieren, der den Wünschen der Nutzer:innen entspricht. Weiterhin nutzen 23 % der Teilnehmer:innen die „Fact-Check“-Methode, und damit die Angabe einer Quelle, um die Nachvollziehbarkeit von Outputs zu verbessern. Ferner setzen 80 % der Befragten darauf, dass ChatGPT mit ihren Akronymen vertraut ist, da sie deren explizite Vermeidung in den Prompts nicht praktizieren.

Die in Kapitel zwei vorgestellte **Few-Shot-Methode findet trotz ihrer Empfehlung in der Literatur unter den Befragten nur geringe Anwendung**. Lediglich 10 % der Befragten geben an, einen Musteroutput manchmal anzugeben, und nur 15 % legen einen expliziten Lösungsweg im Prompt fest. Es ist jedoch zu berücksichtigen, dass

⁵ Evaluierung im Rahmen der Frage: „Welcher dieser Nutzungsempfehlungen nutzen Sie?“.

die Few-Shot-Methode nicht für jede Tätigkeit geeignet ist. Die Option der benutzerdefinierten Systemeinstellung, bei der einmalig ChatGPT Output-Formatvorgaben oder Rollen zugewiesen werden können, wird von nur 11 % der Teilnehmer:innen genutzt. Die Gründe hierfür könnten vielfältig sein. Möglicherweise sind den Nutzer:innen diese Optionen nicht bekannt, da sie in der Benutzeroberfläche weniger sichtbar sind, oder die sehr unterschiedlichen Einsatzmöglichkeiten von ChatGPT machen diese standardisierten Vorgaben unpraktikabel. Änderungen der Systemparameter (=Prompt Settings) – wie in Kapitel zwei dargelegt – werden lediglich von 2 % der Teilnehmer:innen vorgenommen. Dieser geringe Prozentsatz könnte darauf zurückzuführen sein, dass über die Standardoberfläche von ChatGPT keine Anpassung der Systemparameter möglich ist. Zudem erfordert das sinnvolle Anwenden dieser Funktionen ein tieferes Verständnis der Funktionsweise von LLMs.

Welche dieser Nutzungsempfehlungen nutzen sie?

(in % von 138)

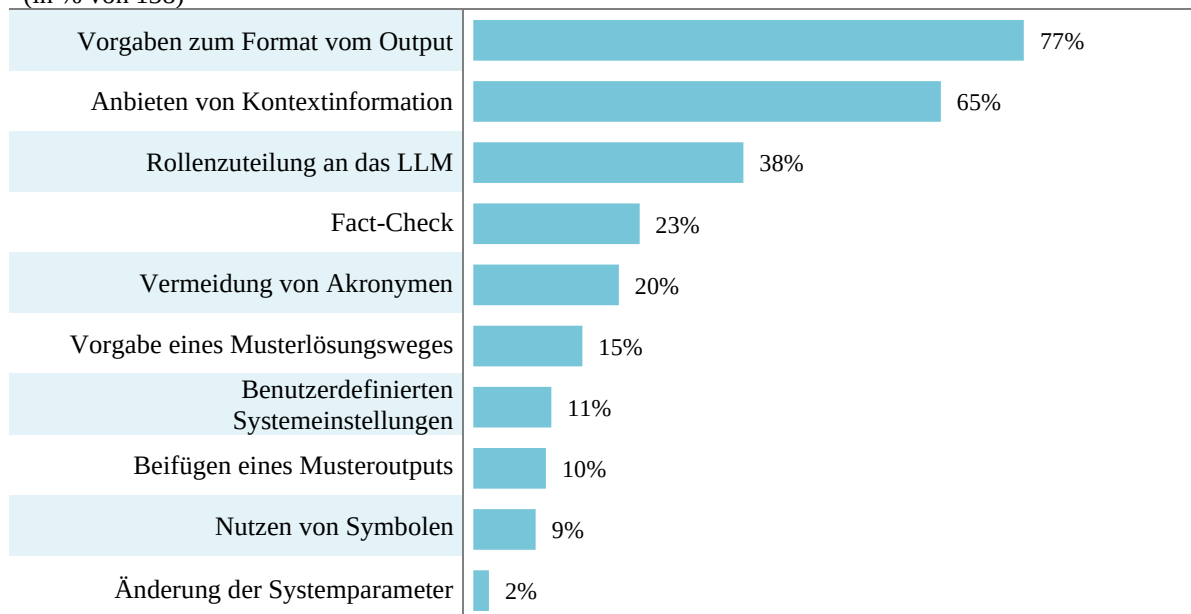


Abbildung 10: Nutzungsempfehlung (eigene Darstellung)

Das ausgeprägte Unwissen über den Begriff „Prompt Engineering“ unterstreicht die dringende Notwendigkeit, durch gezielte Informationsbereitstellung das Bewusstsein und das Verständnis für das Konzept „Prompt Engineering“ zu fördern. Es ist jedoch zu berücksichtigen, dass viele **Nutzer:innen bereits intuitiv grundlegende Methoden des Prompt Engineering** im Zuge des Verfassens der Anweisung **anwenden**, ohne dies explizit zu benennen. Eine vertiefte Kenntnis über Prompt Engineering könnte jedoch nicht nur die Nutzungserfahrung und die Qualität der Ergebnisse verbessern, sondern auch dazu beitragen, die Akzeptanz von ChatGPT im beruflichen Kontext zu steigern und auf ein Niveau zu bringen, das mit dem im privaten Kontext vergleichbar ist.

4 Fazit

Die Studie zeigt, dass Large Language Models (LLMs) wie ChatGPT im beruflichen und privaten Kontext unterschiedlich intensiv genutzt werden. Im Finanzbereich setzen 64 % der befragten Personen LLMs primär für den Wissenserwerb ein, obwohl die Genauigkeit der generierten Informationen oft kritisch betrachtet wird. Ein weiterer signifikanter Anwendungsbereich ist das E-Mail-Management, wobei 56 % der Nutzer:innen E-Mails mit Hilfe von LLMs formulieren. Zudem nutzen 24 % der Befragten LLMs, um komplette E-Mails ohne eigene Vorarbeit schrei-

ben zu lassen. Die Vielseitigkeit von LLMs zeigt sich in weiteren Anwendungsbereichen: 37 % der Teilnehmer:innen erstellen Berichtstexte anhand von Stichworten, 54 % nutzen LLMs zur Erstellung von Zusammenfassungen und 42 % setzen sie für Übersetzungen ein. Im technischen Bereich unterstützen LLMs 32 % der Befragten bei Programmieraufgaben, sowohl in der Text-to-Code-Übersetzung als auch in der Verbesserung von Programmiercodes. Außerdem nutzen 31 % der Teilnehmer:innen LLMs zur Entscheidungsunterstützung. Diese breite Anwendungspalette verdeutlicht das Potenzial von LLMs, weit über reine Textverarbeitung hinaus, auch in technischen und strategischen Aufgaben hilfreich zu sein.

Die Studie identifiziert jedoch auch mehrere Herausforderungen bei der Nutzung von LLMs. Am häufigsten genannt wurden Halluzinationen (77 %), die zu Fehlinformationen führen können, und die Datensicherheit (60 %), insbesondere im Umgang mit sensiblen Informationen. Die unklare Rechtslage (41 %) und mangelnde Transparenz der Modelle (32 %) wurden ebenfalls als erhebliche Hürden wahrgenommen. Zusätzlich gaben 29 % der Teilnehmer:innen an, dass das fehlende Verständnis der Funktionsweise von LLMs eine Herausforderung darstellt. Die Vorlauf- und laufenden Kosten wurden von 22 % der Teilnehmer:innen als Hürde gesehen, während nur 15 % den Bedarf an internen Trainingsdaten als große Herausforderung einschätzten.

Die Studie weist einige Limitationen auf, die bei der Interpretation der Ergebnisse berücksichtigt werden sollten. Erstens ist die Stichprobe auf Personen im Finanzbereich beschränkt, was die Generalisierbarkeit der Ergebnisse auf andere Branchen einschränkt. Zweitens könnte die relativ geringe Teilnehmer:innenzahl die Aussagekraft der Ergebnisse beeinflussen. Zukünftige Studien sollten größere und vielfältigere Stichproben umfassen, um umfassendere und repräsentativere Erkenntnisse zu gewinnen. Trotz dieser Einschränkungen liefert die Studie wertvolle Einblicke in die Nutzung und Herausforderungen von LLMs im beruflichen Umfeld und legt die Grundlage für weiterführende Untersuchungen.

Die Ergebnisse der Studie legen nahe, dass LLMs ein hohes Potenzial zur Effizienzsteigerung und Unterstützung bei verschiedenen beruflichen Aufgaben haben – wie zB im vorliegenden Fall beim Verfassen eines Fazits. Es ist jedoch notwendig, die Herausforderungen zu adressieren, um eine breitere und sicherere Anwendung zu ermöglichen. Zukünftige Forschung sollte sich darauf konzentrieren, Methoden zur Reduzierung von Halluzinationen zu entwickeln und die Datensicherheit zu verbessern. Auch die Klärung der Rechtslage und die Erhöhung der Transparenz der Modelle sind wesentliche Schritte, um das Vertrauen und die Akzeptanz von LLMs zu erhöhen. Darüber hinaus können Schulungen und Aufklärung über die Funktionsweise und den optimalen Einsatz von LLMs dazu beitragen, die Nutzungseffizienz zu steigern.

Literaturverzeichnis

- Chen, B., Wu, Z., & Zhao, R. (2023). From fiction to fact: the growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, 21(4), 471–496. <https://doi.org/10.1080/14765284.2023.2245279>
- Dhana, F. (2023). Writing advanced Excel macros with GPT-4. *Journal of Accountancy*. <https://www.journalofaccountancy.com/news/2023/jul/writing-advanced-excel-macros-with-chatgpt-4.html>. Abgerufen am 21.08.2024.
- Draxler, F., Buschek, D., Tavast, M., Hämäläinen, P., Schmidt, A., Kulshrestha, J., & Welsch, R. (2023). Gender, age, and technology education influence the adoption and appropriation of LLMs. arXiv preprint arXiv:2310.06556.
- Ekin, S. (2023). Prompt Engineering For ChatGPT: A Quick Guide To Techniques, Tips, And Best Practices. Authorea Preprints. <https://doi.org/10.36227/techrxiv.22683919.v2>
- Gao, A. (2023). Prompt Engineering for Large Language Models. SSRN. <http://dx.doi.org/10.2139/ssrn.4504303>
- Hacker, B. (2023). Wird ChatGPT das Rechnungswesen revolutionieren? Der Nutzen von Künstlicher Intelligenz (KI) im Rechnungswesen. *IRZ Zeitschrift für Internationale Rechnungslegung*, 18(3), 117–121.
- Hüsch, A., Distelrath, D., & Hüsch, T. (2023). Einsatzmöglichkeiten von GPT in Finance, Compliance und Audit. Springer Gabler, Wiesbaden. <https://doi.org/10.1007/978-3-658-41419-1>
- IU Internationale Hochschule (Hrsg.) (2023). ChatGPT & Co. persönliche genutzt. Wie werden KI-Bots eingesetzt? Welchen Einfluss haben sie? <https://static.iu.de/studies/Studie-Einfluss-von-KI-Bots-auf-unser-Privatleben.pdf>. Abgerufen am 10.07.2024.
- Kuckartz, U., Rädiker, S., Ebert, T., & Schehl, J. (2013). *Statistik: eine verständliche Einführung*. Springer-Verlag.
- Kulkarni, A., Shivananda, A., Kulkarni, A., & Gudivada, D. (2023). The ChatGPT Architecture: An In-Depth Exploration of OpenAI's Conversational Language Model. In: A. Kulkarni et al. (Hrsg.), *Applied Generative AI for Beginners. Practical Knowledge on Diffusion Models, ChatGPT, and Other LLMs*, Berkeley, CA. 55–77.
- Leitner-Hanetseder, S., Losbichler, H. & Altendorfer, F. (2023). Generative KI im Controlling – Wie können Large Language Modelle das Controlling verändern?. *CFOaktuell*, 17(5), 168-172.
- McTear, M., & Ashurkina, M. (2024). Transforming Conversational AI. Exploring the Power of Large Language Models in Interactive Conversational Agents. Apress Berkeley, CA. <https://doi.org/10.1007/979-8-8688-0110-5>
- Morris, M. G., & Venkatesh, V. (2006). Age differences in technology adoption decisions: Implications for a changing work force. *Personnel psychology*, 53(2), 375-403. <https://doi.org/10.1111/j.1744-6570.2000.tb00206.x>
- Pröhl, T., Baumgartner, S., & Zarnekow, R. (2023): Cloud-basierte KI-Anwendungen für Natural Language Processing: Marktüberblick und Vorstellung von typischen Anwendungsfällen für den Mittelstand, *HMD Praxis der Wirtschaftsinformatik*, 60(5), 1093–1107. <https://doi.org/10.1365/s40702-023-00994-w>
- PwC (Hrsg.) (2023): PwC ChatGPT & KI-Studie: Fast zwei Drittel der Österreicher:innen sind für ein Verbot von ChatGPT an Schulen – jede:r Dritte fürchtet um Arbeitsplatz, download unter: <https://www.pwc.at/de/presse/2023/chatgpt-ki.html>.
- Rane, N. (2023). Role and Challenges of ChatGPT and Similar Generative Artificial Intelligence in Finance and Accounting. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.4603206>
- Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications. <http://arxiv.org/pdf/2402.07927>
- Schulze, M., Broßmann, F., & Wallraff, B. (2023). Generative KI im Controlling: Was bleibt nach dem Hype?. *bilanz + buchhaltung*, 11, 39–42.
- Singh, H., & Singh, A. (2023). ChatGPT: Systematic Review, Applications, and Agenda for Multidisciplinary Research. *Journal of Chinese Economic and Business Studies*, 21(2), 193–212. <https://doi.org/10.1080/14765284.2023.2210482>.

- Street, D., Wilck, J., & Chism, Z. (2023). Six Principles for the Effective Use of Artificial Intelligence Large Language Models: How to Leverage ChatGPT, Bard, and Bing Chat in Accounting Work. *The CPA journal*, 93(11-12), 50-56.
- Taulli, T. (2023). ChatGPT and Bard for Business Automation: Achieving AI-Driven Growth. Apress. <https://doi.org/10.1007/978-1-4842-9852-7>
- Teubner, T., Flath, C. M., Weinhardt, C., van der Aalst, W., & Hinz, O. (2023). Welcome to the era of chatgpt et al. the prospects of large language models. *Business & Information Systems Engineering*, 65(2), 95-101. <https://doi.org/10.1007/s12599-023-00795-x>
- Wegenstein, K., & Waniczek, M. (2023). KI als Game-Changer in der Finanz- und ControllingOrganisation. Über die Reife der KI-Technologien und der Organisation für eine erfolgreiche KI-Integration. *CFOaktuell*, 17(6), 200-202.
- Wennker, P. (2020). Künstliche Intelligenz in der Praxis. Springer Gabler, Wiesbaden. <https://doi.org/10.1007/978-3-658-30480-5>
- West, M., Kraut, R., & Ei Chew, H. (2019). I'd blush if I could: closing gender divides in digital skills through education, download unter: <https://unesdoc.unesco.org/ark:/48223/pf0000367416>.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2302.11382>
- Zadorozhnyi, ZM., Muravskiy, V., Pochynok, N., Muravskiy, V., Shevchuk, A., & Majda, M. (2023). Application of chatbots with artificial intelligence in accounting. In *2023 13th International Conference on Advanced Computer Information Technologies (ACIT) IEEE*, 196-200.
- Zhao, J., & Wang, X. (2023). Unleashing efficiency and insights: Exploring the potential applications and challenges of ChatGPT in accounting. *Journal of Corporate Accounting & Finance*, 35(1), 269–276. <https://doi.org/10.1002/jcaf.22663>

Gesetzesverzeichnis

VERORDNUNG (EU) 2024/1689 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz)

The influence of culture on the capital structure of SME – A pan European quantitative research

Research Paper

Dr. Markus Wienbreyer

L&W GmbH, Langeoogerstr. 44, 27755 Delmenhorst, E-Mail: wienbreyer@hv-lw.de

Prof. Dr. Carola Spiecker-Lampe

Hochschule Bremen, Fakultät 1, Werderstr. 73, 28199 Bremen, E-Mail: carola.spiecker-lampe@hs-bremen.de

Prof. Dr. Maria Iborra

Universitat de València, Facultat d'Economia, Valencia, E-Mail: Maria.Iborra-Juan@uv.es

Abstract

Different finance theories have tried to explain the capital structure of companies, considering it a rational choice with limited results. In contrast, the upper echelons theory highlights the influence of top managers' values and beliefs on firms' decisions. To fill the gap between the rational finance theories and the upper echelons' proposition, our research examines the role of cultural values in explaining why SMEs' differ in their financial preferences. Our study uses the cultural framework from Hofstede (2021) to analyze the impact on the capital structure of SMEs'. Through a multilevel analysis of the financial information of 27,778 SMEs from 24 EU countries between 2015 and 2019 we find evidence that culture affects SMEs' choice of capital structure. However, in contrast to previous studies that found a strong relationship between culture and firms' financial behavior, our findings demonstrate a very small effect size of culture on financial behavior.

1 Introduction

The choice of a company's capital structure is probably a CEO's most fundamental financial management decision to minimize capital costs and maximize the firm's value. Whether there is an optimal capital structure, and if so, which factors influence it, has been a central research question for more than 60 years. When determining how to optimize a company's capital structure, with an ideal proportion of debt and equity capital, there are two main factors to consider: minimizing the cost of capital and maximizing the company's value (Hermanns, 2006). Modigliani & Miller's (1958) theorem of the irrelevance of capital structure laid the foundation for further research into the optimal capital structure in the field of finance¹. According to the traditional view, capital structure choice is entirely rational (Hens & Wang, 2007). After decades of studying the subject, financial economist Myers (2001) concluded: "There is no universal theory of the debt-equity choice and no reason to expect one." (p. 81).

The historical and nearly exclusively theoretical discussion has given way in recent years to a more pragmatic and realistic approach in which individual aspects relevant to the capital structure are increasingly being theoretically studied and empirically tested. Surprisingly, all these financial theories fail to analyze the role of CEOs and top management teams (TMTs) and their impact on firms' capital structure, even though upper echelons have an undeniable influence on corporate decisions and actions (Hambrick & Mason, 1984) and most of them have a predefined position how to structure their company's capital (Spiecker-Lampe, 2018). The CEO and top managers influence the organization, its actions, and its behavior in a certain way. Hambrick and Manson (1984) consider socio-economic background to be a significant factor in linking the CEO's decisions to the firm's strategy. However, the *upper echelons theory* (from now on, UET) is frequently used in conjunction with other theories to increase its explanatory power. For instance, Nielsen and Nielsen (2013) combine North's (1990) institutional economic theory with the UET to show that the individual, organizational, and institutional levels are tightly connected. Following this approach, this research combines the UET with institutional theory to reveal the strong link between the institutional, corporate, and individual levels in explaining firms' financial decision-making.

The proposals of both theories are even more significant for small and medium-sized enterprises (from now on SMEs) (Dolz, Iborra, & Safon, 2019). SMEs are essential to the local economy since they are linked to local customers, suppliers, financiers, and stakeholders. In this sense, SMEs' main actors, including CEOs and TMTs, are usually deeply rooted in the community, with all the privileges and obligations it entails.

Returning to the capital structure of companies and their organizational structure, it would seem that their managers have an enormous influence on these structures and that, in turn, institutional values and norms influence managers and their surroundings. Therefore, certain overlooked but nonetheless, relevant factors need to be taken into account other than those normally considered from the traditional viewpoint, according to which financial decisions are made using rational economic calculations and a narrow focus (Hens & Wang, 2007). Thus, the effect of individual and institutional values on financial decisions demands researchers' attention.

We argue that culture is one of the key factors influencing this decision-making since it is an essential dimension of institutional theory and the UET. In this regard, the first literature relating financial decision-making with culture reaches back to Smith (1776) and Weber (1905). Since then, a growing body of literature, connecting financial research and culture has emerged in different areas of economic research. Aggarwal and colleagues, for instance, present in various studies that culture, trust, and transaction costs are deeply linked to financial choices (Aggarwal & Goodell (2010), (2014)a, (2014)b, and Aggarwal et al. (2016)).

The probably most well-known and cited cultural model is from Hofstede (2001), wherefore the present study uses this culture classification. From Hofstede's six essential dimensions, the latest literature presents mainly three cultural variables from Hofstede, that have an essential influence on financial decision-making in general and more closely on a firm's capital structure, for instance, Haq et al. (2018) find that cultures with high individualism hold more leverage, Wang and Esqueda (2014) state that firms in countries with high long-term orientation tend to have

¹ To date, various theories have emerged, the most well-known probably being the trade-off theory by Kraus and Litzenger (1973), the pecking order theory by Myers and Majluf (1984), and more recently, dynamic approaches such as Baker and Wurgler's (2002) market timing theory.

less leverage, and Ashraf et al. (2016) have shown that bank-risk-taking is negatively related to a high uncertainty-avoidance.

Thus, the main objective of this work is to analyze the impact of culture on the capital structure of SMEs. Using multilevel analysis, we examine a database containing financial information from 27,778 SMEs in 24 EU countries between 2015 and 2019. We found support that indicates that culture has an impact on the capital structure of SMEs, although with a small effect size, suggesting that culture may no longer be as influential as it once was.

This study makes several valuable contributions. First, it contributes to finance research by providing a theoretical explanation for how other non-rational variables, like cultural background, affect SMEs' capital structure. This focus on values and beliefs may enrich traditional financial models. Thus, by adding these new features that influence the financial behavior of SMEs, we complement the rational view and advance the field of finance research. Second, this study contributes to the body of knowledge about culture's different levels of influence. Specifically, our research demonstrates that culture plays a role in SMEs' capital structure, although not as significantly as in the past. Third, we provide evidence that multilevel analysis can be helpful in understanding complex phenomena that involve at least two different levels of data.

2 Literature Review

2.1 Capital structure as a rational choice

The irrelevance theorem of Modigliani and Miller (1958) is considered the basis of many theories devoted to optimal capital structure. This theory is still valid under Modigliani and Miller's conditions. These imply, among other things, that "the total value of a company is given through the expected value of the future profit, [...] regardless of the capital structure. Therefore, the cost of capital cannot be optimized" (Hermanns, 2006, pp. 15-16). Unfortunately, this condition contradicts reality, in which there is no perfect market, and other factors, such as taxes, information asymmetries, bankruptcy costs, etc., must be taken into account.

As a result, new capital structure theories were developed from the original theory, which analyzed and included the market's imperfections. This way, models were created that emphasized the relevance of capital structure. First, the trade-off theory was developed, which considered the factors of an imperfect market, the existence of taxes, and the possibility of bankruptcy. The trade-off theory of Kraus and Litzenberger (1973) is characterized by the attempt to find the perfect balance between the tax advantages resulting from tax-deductible interest payments and the increased likelihood of insolvency and related bankruptcy costs brought on by taking on additional debt.

Since the trade-off theory does not consider other crucial factors for a company, such as information asymmetries, the pecking-order theory, proposed by Donaldson (1961), and refined by Myers and Majluf (1984), emerged as a further development of the capital structure theories. In addition to taking new influencing factors into account, the pecking-order theory rejects the assumption of an optimal target capital structure. It assumes that a specific ranking will be used to inform financial decisions. This ranking results from the reverse order of the expected information costs of a specific financial instrument. Accordingly, a company first prefers financing from its own funds, then by debt financing, and finally, through the issuance of equity.

Contrary to the pecking order theory, Baker and Wurgler's (2002) market timing hypothesis holds that firms' current capital structures are strongly correlated with historical market values and that firms generally do not care to finance their assets with debt or equity. They instead consider market timing and choose the most suitable option at the time. This approach has its roots in the analysis of a dynamically changing capital structure using the so-called "window-of-opportunities approach," which laid the foundation for market timing theory. Although this notion is probably the closest to how modern firms really choose their capital structures, and all these capital structure theories and approaches still have their place in academia and are essential when looking at an issue from a general perspective, they all rely on rational decision-making. Moreover, only external factors like taxes, information asymmetries,

bankruptcy, etc., are considered. These rational models do not take into account the influence of upper echelons' values and perceptions which are part of the "black box" of capital structure choice.

2.2 The role of values and beliefs on capital structure: the upper echelons and institutional theories

It has been shown that top managers influence the success of a company through their behavior and decisions (Finkelstein *et al.*, 2009; Hambrick and Mason, 1984). Hambrick and Mason (1984) stated that "organizational outcomes both strategies and effectiveness—are viewed as reflections of the values and cognitive bases of powerful actors in the organization." (p. 193). Along these lines, the UET holds that "organizational outcomes—strategic choices and performance levels—are partially predicted by managerial background characteristics" (Hambrick & Mason, 1984, p. 193). These range from socioeconomic background to formal education, age, or gender. Accordingly, it could be concluded that the organization reflects its top managers and their origins.

Historically, Weber (1905) was the first to suggest that an organization's surrounding institutions influence its members, which seems consistent with concepts long after developed by institutional theory, specifically by Meyer and Rowan (1977), Zucker (1977), DiMaggio and Powell (1983), and North (1990). The central premise of institutional theory is that everyday interactions inside organizations and between individuals are shaped by institutional norms and values. For example, institutional norms and values have an impact on organizational rules, norms, structures, and beliefs that govern individuals' interactions and behaviors (Lammers & Barbour (2006) & Lammers & Garcia (2014)). This premise may explain why researchers like Nielsen and Nielsen (2013) use the UET and institutional theory jointly.

Consequently, we argue that CEOs' and TMTs' beliefs affect how they make financial decisions for their companies; specifically, they influence how capital is structured and the level of risk they assume. We argue that, in particular, cultural values may influence the main financial choices of SMEs, their capital structure, and their level of risk.

3 Model and Hypothesis

3.1 Development of the model

Prior to modern research, Adam Smith (1776) and Max Weber (1905) recognized the potential influence of culture on finance and discussed this topic in their writings. Aggarwal and colleagues, for instance, present in various studies that culture, trust, and transaction costs are deeply linked to financial choices (Aggarwal & Goodell (2010), (2014)a, (2014)b, and Aggarwal *et al.* (2016)).

However, culture is a dynamic concept in a flowing status caused by social development and transformation (Kühnen, 2015). For example, Maridal (2013) defines Culture as "society's beliefs and value system", Licht (2001) argues that Culture is the "mother of all path dependencies", and Hofstede and Bond (1988) describe Culture as the way people think, feel, and act, including values, beliefs, behaviors, education, manners, and arts.

The most well-known and cited is the cultural model from Hofstede (2001), next to Trompenaars (2005), Schwartz (1999), GLOBE Study (Global Leadership & Organizational Behavior Effectiveness) Foundation (2018), and World Values Survey (2018), wherefore the present study uses these Culture classifications. From Hofstede's six essential dimensions (power distance, uncertainty avoidance, collectivism versus individualism, femininity versus masculinity, long-term orientation, and indulgence vs. restraint), some of these dimensions are specifically connected with financial decision making. The latest literature presents mainly three cultural variables from Hofstede and variables with similar meanings from the other theories, that have an essential influence on financial decision-making in general and more closely on a firm's capital structure. Tested together, they do not show any statistically significant correlation, implying that each variable captures a different dimension of national Culture (Mourouzi-dou-Damtsa *et al.* 2019).

In that sense, the dimensions of individualism, uncertain avoidance, and long-term orientation have been connected by various researchers with financial research. Individualism pertains to societies where there are loose ties between individuals, in contrast to collectivism, which belongs to Cultures where people look after themselves, have healthy relationships and behave cohesively (Haq et al. 2018). Haq et al. (2018) find that cultures with high individualism hold more leverage. Uncertainty-avoidance measures to which degree individuals feel comfortable with ambiguity as members of society (Hofstede, 1984). Commonly scientific research has shown that bank-risk-taking is negatively related to a high uncertainty-avoidance (Ashraf et al. 2016). Long-term orientation, on the contrary, is mainly based on the “Confucian thinking”, which implies to preferably have long-term success, including goals and thrift (Hofstede, 2001), and researchers state that firms in countries with high long-term orientation tend to have less leverage (Wang & Esqueda, 2014). However, for a holistic approach, the cultural dimensions of power distance, masculinity, and indulgence are included as well.

3.2 The impact of Culture on the capital structure of SMEs

The impact of culture can be grouped into three different channels. Firstly, on the firm directly. Secondly, through the institutional level and the influence from the institution on the organization (Institutional Theory by North (1992)). Furthermore, the institutional level influences the individual level as well (Crossan et al. 1999), which leads to, thirdly, the impact of the individual level on organizations (UET by Hambrick and Mason (1984)).

On the institutional level, Jordaan et al. (2016) find that national wealth favors a significant development in stock markets, and therefore equity financing is increasing. Furthermore, they connect legal, political, and democratic institutions to promote of a developed stock market (Jordaan et al. 2016). Stone et al. (2014) argue that the presence and intensity of an equity culture for companies within Europe has external institutional factors and internal managerial ones. This approach concludes that institutional and cultural values play a significant role in the change of equity and the organizational and individual preference within the firm (Stone et al. 2014). Karolyi (2016) describes in “The gravity of culture for finance” that cultural distance explains foreign bias in international portfolio holdings.

Next to culture in general, individual cultural dimensions are also linked to certain values. More specifically, Gaganis et al. (2019) find evidence that culture matters in SME’s profitability. In a large study with around 40.000 European firms from 25 EU countries, their results show that national culture plays an essential role in the terms and conditions of getting credit and enhancing profitability. In specific, individualism, masculinity, and long-term orientation positively impact the firm’s profitability, whereas power distance and uncertainty avoidance have the opposite effect. Chui et al. (2016) investigate how culture affects the cost of debt within companies using the cultural dimensions of Schwartz (1994). Using data from 33 countries, they find evidence that embeddedness and mastery have a strong negative relation to debt cost. Furthermore, they find that embeddedness is negative, and mastery is positively related to bankruptcy risk. Schwartz’s embeddedness is comparable to Hofstede’s individualism, which would lead to the statement that a high level of individualism leads to a lower cost of debt, and Schwartz’s mastery to Hofstede’s masculinity, which leads to the statement, that high level of masculinity leads to the high cost of debt.

Continuing with the individual cultural values, the role of individualism is one of the most used ones in the literature. Antonczyk and Salzmann (2014) provide evidence that a high degree of individualism exhibits strong optimism and overconfidence, causing higher debt ratios; also, Gaganis et al. (2019), in a large study with around 40.000 European firms from 25 EU countries, found a positive relationship between individualism and debt through its essential role in the terms and conditions of getting credit and enhancing profitability. Similarly, Haq et al. (2018) and Wang & Esqueda (2014) find that high individualism cultures hold more leverage.

Wand and Esqueda (2014) also find evidence that a high level of uncertainty avoidance leads to lower levels of debt. These findings are in line with Haq et al. (2018), who also find similar results regarding the level of leverage and the level of uncertainty avoidance. Both findings are in contrast with Aggarwal and Goodell (2010), who find evidence for the opposite, that uncertainty avoidance leads to less equity and, therefore, to more leverage.

For the cultural value of long-term orientation, Haq et al. (2018) and Wang and Esqueda (2014) find similar evidence, namely that high long-term orientation rather leads to lower amounts of leverage.

Similar statements are for the value of power distance. Here, Haq et al. (2018) and Wang and Esqueda (2014) also find evidence that high level of power distance rather led to lower amounts of leverage. This is again in contrast with Aggarwal and Goodell (2010), who find evidence for the opposite that a high level of power distance rather leads to more equity.

Masculinity and its association with financial architecture has not been studied in depth. However, due to the findings of Chui et al. (2016), with similar values from Schwartz, this gives a pointing direction.

Indulgence has only been studied by Wang and Esqueda (2014) in depth. They find that a high level of indulgence leads to more leverage.

Therefore, and based on the literature, we hypothesize that there is a general influence of various cultural values on the capital structure and their use of either equity or debt capital within companies. The individual hypotheses state that:

Hypothesis 1: Culture, measured by the following six Variables of Hofstede, has an influence on the capital structure of a company, measured by the level of debt:

Hypothesis 1a: Individualism has a negative influence on the level of debt, meaning the higher the level of Individualism, the higher the level of debt.

Hypothesis 1b: Uncertainty Avoidance has a positive influence on the level of debt, meaning the higher the level of Uncertainty Avoidance, the lower the level of debt.

Hypothesis 1c: Long Term Orientation has a negative influence on the level of debt, meaning the higher the level of Long-Term Orientation, the higher the level of debt.

Hypothesis 1d: Power Distance has a negative influence on the level of debt, meaning the higher the level of Power Distance, the higher the level of debt.

Hypothesis 1e: Masculinity has a negative influence on the level of debt, meaning the higher the level of Masculinity, the higher the level of debt.

Hypothesis 1f: Indulgence has a negative influence on the level of debt, meaning the higher the level of Indulgence, the higher the level of debt.

4 Research Methodology

4.1 Definition of the sample

The firm data were collected in March 2021 from the ORBIS® database² to select a very large but manageable dataset of small- and medium-sized enterprises.

In contrast to prior studies that only used US data (Hilary & Hui, 2009; Adhikari & Agrawal, 2016; He & Hu, 2016; and Cai & Shi, 2017), worldwide data (Baxamusa & Jalal, 2014; Jiang *et al.*, 2015; and Diez-Esteban *et al.*, 2019) or from continental European countries (Gaganis *et al.* 2019), the main focus of this study is on firms located in one of the EU 27 Member States. The United Kingdom was not an EU Member State at the time of data collection (2021) and, therefore, was excluded. We use an unbalanced sample, like McGuinness *et al.* (2018) and Diez-Esteban

² Orbis® is an online database compiled by Bureau Van Dijk, a Moody's Analytics Company

et al. (2019), because European countries differ in terms of size and the number of firms they host. The unbalanced samples of these authors range, respectively, from 27% in Spain to 0.02% in Latvia and from 26.25% in the United States to 0.25% in Luxembourg.

As Iborra *et al.* (2020) have shown, the selection of only medium-size enterprises, rather than small and medium-sized enterprises, has the advantage that the data are more reliable since their accounts are more likely to be externally audited. The size of an SME can be defined by its sales (Hilary & Hui, 2009; Jiang *et al.*, 2015), its total assets (Baxamusa & Jalal, 2014, Adhikari & Agrawal, 2016), or by the number of its employees (Dasí *et al.*, 2015, Iborra *et al.*, 2022). The present study uses the last option and focuses on firms with 50 to 250 employees.

The capital structure of SMEs varies depending on the industry, with clear differences between the manufacturing and service industries. Following Dasí *et al.* (2015), Dolz *et al.* (2019), and Iborra *et al.* (2022), this study selected firms that belonged to the manufacturing sector, specifically NAICS 2017 (Primary code): Ranking – 31-33. The justification for this is that firms in the manufacturing industry tend to have more capital-intensive assets, such as land, buildings, equipment, machines, etc., and a higher amount of current assets that can affect their capital structure choices.

We used data from 2015 to 2019 to generate the dataset, excluding firms that were subsidiaries of larger companies, such as multinational enterprises. This search produced a total of 32,197 companies from 24 countries since there were no data available for firms in Cyprus, Luxembourg, or Malta due to the aforementioned restrictions. In a subsequent phase, we eliminated firms with missing information that was necessary for the analysis. The final dataset comprised 27,778 medium-sized firms from 24 EU Member States.

4.2 Definition of variables

Dependent variables: The dependent variables represent the two aspects examined in this study: firms' capital structure.

Capital structure: When it comes to studying the capital structure, researchers typically examine equity and leverage. It is relevant because companies rely on a mixture of equity and debt to finance their operations, and knowing the leverage -through in our case the debt to assets ratio and debt to revenue ratio- allows evaluating whether the firm can pay off its debts as they come due. Leverage can be measured by using the total debt to assets ratio and debt to revenue ratio³.

The use of ratios in financial research is widespread. Titman and Wessels (1988), for instance, used different ratios to assess short-term and long-term debt. The same measurement is used by Mogha and Williams (2021), while Rashid *et al.* (2021) even use ratios to measure short-term, long-term, and total debt. Yin *et al.* (2020) relied on various ratios to evaluate firm-specific credit risks, including the debt to assets ratio; and Rokhanyati *et al.* (2019) used a debt ratio as a measure of the total leverage of SMEs. By way of example, recent studies that have used leverage in connection with religious influences include Chui *et al.* (2002), Jacobson *et al.* (2005), Baxamusa and Jalal (2014), Jiang *et al.* (2015), Adhikari and Agrawal (2016), Chen *et al.* (2016), Chui *et al.* (2016), He and Hu (2016), Mourousidou-Damtsa *et al.* (2019), Diez-Esteban *et al.* (2019), Li *et al.* (2019), Papageorgiou *et al.* (2020) and Czerwionka and Jaworski (2021). Due to the uniqueness of our data, we follow several of the cited studies and take into account one ratio as dependent variables⁴.

³ Hybrid capital or the differentiation of certain types of equity or debt are not considered in this study.

⁴ However, Verduch Arose *et al.* note with regard to the usage of ratios as variables that, "Our dependent variable is the debt-to-asset ratio, and not all researchers agree on how this should be measured." (Verduch Arosa, Richie, & Schumann, 2014, p. 182). The use of a debt to asset ratio is therefore not without controversy. Studies with titles such as "Divided We Fall: How Ratios Undermine Research in Strategic Management" (Certo *et al.*, (2018)) discuss the adverse effects of ratios as dependent, independent, or control variables, although their use in financial and management research is common. As Jasienski and Bazzaz state: "Empirical researchers love ratios; statisticians loathe them" (Jasienski & Bazzaz, 1999, p. 321). Wisemann (2009) reported that 74% of all empirical articles published in the *Strategic Management Journal* in 2007 included at least one ratio. In 2015, the use of ratios rose to 79% of all empirical articles published in that journal (Certo, Budenbark, Kalm, & LePine, 2018).

Our assessment of leverage and capital structure is the total debt to asset ratio, one of the most commonly used variables in the literature. In addition, since we analyze SME financial decision-making, we also used the total debt to revenue ratio, as robustness check. We argue that for small and medium-sized firms, the debt to revenue ratio has greater predictive power since it links the firm's capital structure with its ability to repay leverage through its revenue. Measurements for the capital structure are calculated in the following way:

$$\text{Total Debt to Assets Ratio} = \frac{\text{Short Term Debt} + \text{Long Term Debt}}{\text{Total Assets}} \quad (1)$$

$$\text{Total Debt to Revenue Ratio} = \frac{\text{Short Term Debt} + \text{Long Term Debt}}{\text{Revenue}} \quad (2)$$

Independent variable: Our independent variable in the present model is *Culture*. Within Culture, we use the cultural variables of Hofstede, which is one of the mainly used cultural models and was used in similar studies, for example, if culture can influence corporate cash holdings (Chen et al. 2015), if culture influences the profitability (Gaganis et al. 2019), if national culture can predict the cost of debt – (Chui et al. (2016) & Chen et al. (2016)), if culture can determine corporate risk-taking (Diez-Esteban et al. 2019), if culture can influence the kind of debt in firms (Mogha & Williams, 2021), and if culture has an influence on the capital structure (Ramirez and Kwok (2009) and Rashid et al. (2020)).

Following them, the six individual independent variables, which have been partially or completely used, are:

1. Power Distance
2. Individualism
3. Masculinity
4. Uncertainty Avoidance
5. Long Term Orientation
6. Indulgence

Control variables: Here, we employ various control variables for our regression model. The collection of the variables follows the argumentation of Hilary and Hu (2009)⁵ and Bernerth and Aguinis (2016) and are divided into two categories—on the one hand, control variables at the unit (firm) level, and on the other hand, control variables at the group (country) level.

At the firm level, we used the following control variables:

Return on assets: A ratio calculated from net income divided by total assets. Various studies analyzing leverage and risk have used this ratio as a control variable, e.g., Adhikari and Agrawal (2016), Chen *et al.* (2016), Chui *et al.*

⁵ The study of Hilary and Hu (2009) laid the argumentative basis for later studies and the use of organizational and institutional variables. Citing studies are, among others: Adhikari and Agrawal (2016), Cai and Shi (2017) and Diez-Esteban, Farinha, and Garcia-Gomez (2019).

(2016), He and Hu (2016), Cai and Shi (2017), Mourousidou-Damtsa *et al.* (2019), Gaganis *et al.* (2019), and Yin *et al.* (2020).

Size: This is one of the key control variables in most similar studies (e.g., Michaelas *et al.*, 1999; Chui *et al.*, 2002; Baxamusa and Jalal, 2014; Verduch Arosa *et al.*, 2014; Jiang *et al.*, 2015; Adhikari and Agrawal, 2016; Chui *et al.*, 2016; Mourousidou-Damtsa *et al.*, 2019; Diez-Esteban *et al.*, 2019; Li *et al.*, 2019; Papageorgiou *et al.*, 2020; Czerwonka and Jaworski, 2021, and Rashid *et al.*, 2021). The size of a company can be measured by three different variables: first, the firm's total revenue; second, the number of total assets; and third, the number of employees, as in this study.

The second category of the control variable is that of the country-level variables. These variables, due to the small size at the country level (N=24), include only one variable, which is:

Median income: In several studies, the per capita income has been used as a variable related to the financial needs of business (Baxamusa and Jalal, 2014; Adhikari and Agrawal, 2016; He and Hu, 2016; and Cai and Shi 2017). For this reason, we included the country's median income as a control variable.

5 RESULTS ANALYSIS AND DISCUSSION

5.1 Results analysis

The analysis strategy was based on the uniqueness of our dataset, representing variables at both the organizational and institutional levels.

Prior research used three different statistical methods: (i) the ordinary least squares method (OLS) (Hilary and Hui, 2009; Jiang *et al.*, 2015; Adhikari and Agrawal, 2016; He and Hu, 2016), (ii) country-weighted least squares regression to control for different country representations in the sample (Froot, 1989; Edwards, 1992; Williams, 2000; Dragolov & Boehnke, 2015; and Chen *et al.* 2016), and (iii), multilevel regression based on the nature of the data, i.e., organizational data (level 1) and institutional data (level 2) (Hayward & Kemmelmeier, 2011; Donati *et al.*, 2016; Chen *et al.*, 2018; and Molina-Azorin *et al.*, 2000).

Therefore, and given the similarity of our data to other studies, particularly a multilevel structure with firms nested within different countries, we used hierarchical linear modeling to test the proposed hypotheses, as this technique can simultaneously account for the variances among the different levels. Also, the cross-level effects are more accurately estimated using this method (Hofmann & Gavin, 1998).

Independent and control variables were centered before incorporating them into the models. The decision to use grand mean centering (GMC) or centering within cluster (CWC) depends heavily on one's substantive research questions (Enders & Tofighi, 2007, pp. 134-135). Thus, in accordance with our hypotheses, we used GMC for the level 2 variables (country) and CWC for the level 1 variables (firm).

Due to the small sample size at the country level (N=24), which limits the testing of models with many predictors, we tested the hypotheses with simple models, including only the control variables and those of the investigated hypothesis. We tested the three different dependent variables individually against the various control variables. The descriptive statistics and correlations of the variables at the individual level are presented in Table 1.

Table 1: Descriptive statistics and correlations

Variable	N	Mean	S.D.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Y1_Debt_to_Assets	30081	0.566	0.289										
2. Y2_Debt_to_Revenue	25981	0.480	0.425	.448**									
3. Cj3_Individualism	32196	61.307	16.687	.060**	.009								
4. Cj4_Uncertainty_Avoidance	32196	74.583	15.783	-.023**	.070**	-.507**							
5. Cj5_Median_Income_Countrylevel	32196	15692.998	7401.690	.101**	.015*	.609**	-.637**						
6. Cj6_Power_Distance	32196	50.401	17.145	-.042**	.035**	-.580**	.769**	-.723**					
7. Cj7_Masculinity	32196	53.522	20.077	.040**	.074**	.411**	.175**	.065**	-.205**				
8. Cj8_Long_Term_Orientation	32196	62.703	16.202	.019**	-.068**	.404**	-.310**	.375**	-.381**	.348**			
9. Cj9_Indulgence	32196	37.998	14.240	.034**	-.037**	.315**	-.524**	.754**	-.512**	-.281**	-.062**		
10. X1_In_Amount_Employees	32196	4.592	0.410	-.018**	-.005	-.027**	.021**	-.018**	.024**	.006	.013*	.005	
11. X2_ROA	27783	5.052	10.556	-.401**	-.316**	-.040**	-.054**	-.047**	-.013*	-.049**	.008	.002	-.010

** $p < .01$

Table 2 shows the test of the first hypothesis 1a-1f, which studies cultural variables of Hofstede. The dependent variable studied is *Debt to Assets*. The model 1 introduces the country and firm-level control variables. The regression is significant with an adjusted R^2 of 17.7%. The variables *Median Income* at country level and *ROA* at firm level have a significant influence, while *Firm's size* is not statistically significant. The significances of the control variables -*Median Income* and *ROA*- continue in models 2 to 7.

Model 2 to 7 introduce the six Hofstede Values, one by one. From the 6 predictors analyzed, the result shows that *Uncertainty Avoidance*, is statistically significant at the 5% level, with an adjusted R^2 of 18,1%.

Table 3 reproduces the analyses in Table 2 with the dependent variable *Debt to Revenue*. Model 8 introduces again the country and firm-level control variables. This regression has an adjusted R^2 of 11,1%. Similar to the first model, the control variables *ROA* and *Median Income* are significant, whereas the *Amount of Employees* continues to be non-significant throughout the other six models, too. Differently to the first seven models, *Median Income* does not continue to be significant in Model 9 to 14.

Masculinity, *Uncertainty Avoidance* are statistically significant at a level of 0.05, whereas *Individualism* is supported but in the opposite direction. However, the three predictors, *Power Distance*, *Long Term Orientation*, and *Indulgence* are not significant. Again, the effect sizes are low and reach from an R^2 of 12,1% with *Individualism* and an adjusted R^2 of 11,8% for *Masculinity* and *Uncertainty Avoidance*.

Table 2: Test of Hypotheses 1-6 – Debt to Assets

Variable	Y1_Debt_to_Assets													
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	Beta	<i>P</i>	Beta	<i>p</i>	Beta	<i>p</i>	Beta	<i>p</i>	Beta	<i>p</i>	Beta	<i>p</i>	Beta	<i>p</i>
Intersection	0.547	<.001	0.549	<.001	0.545	<.001	0.548	<.001	0.552	<.001	0.545	<.001	0.548	<.001
Country Level														
Cj3_Individualism					-0.001	0.160								
Cj4_Uncertainty_Avoidance									0.001	0.023				
Cj5_Median_Income_Countrylevel	0.000	0.008	0.000	0.002	0.000	0.003	0.000	0.008	0.000	<.001	0.000	0.009	0.000	0.11
Cj6_Power_Distance			0.001	0.076										
Cj7_Masculinity							0.000	0.904						
Cj8_Long_Term_Orientation											0.000	0.455		
Cj9_Indulgence													0.000	0.849
Firm Level														
X1_In_Amount_Employees	-0.008	0.196	-0.008	0.198	-0.008	0.196	-0.008	0.196	-0.008	0.197	-0.008	0.196	-0.008	0.196
X2_ROA	-0.012	<.001	-0.012	<.001	-0.012	<.001	-0.012	<.001	-0.012	<.001	-0.012	<.001	-0.012	<.001
Model deviance	3780.223		3777.015		3778.212		3780.209		3775.001		3779.657		3780.187	
Adjusted R ²	0.177		0.179		0.178		0.178		0.181		0.177		0.178	

Table 3: Test of Hypotheses 1-6 – Debt to Revenue

Variable	Y2_Debt_to_Revenue													
	Model 8		Model 9		Model 10		Model 11		Model 12		Model 13		Model 14	
	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p
Intersection	0.447	<.001	0.451	<.001	0.440	<.001	0.465	<.001	0.459	<.001	0.435	<.001	0.429	<.001
Country Level														
Cj3_Individualism					-0.003	0.02								
Cj4_Uncertainty_Avoidance									0.003	0.02				
Cj5_Median_Income_Countrylevel	0.000	8.47E-01	0.000	0.185	0.000	0.128	0.000	0.641	0.000	0.113	0.000	0.987	0.000	0.244
Cj6_Power_Distance			0.002	0.075										
Cj7_Masculinity							0.002	0.041						
Cj8_Long_Term_Orientation											-0.002	0.055		
Cj9_Indulgence													0.003	0.143
Firm Level														
X1_In_Amount_Employees	0.007	0.31	0.008	0.306	0.007	0.31	0.007	0.311	0.008	0.306	0.007	0.308	0.007	0.309
X2_ROA	-0.014	<.001	-0.014	<.001	-0.014	<.001	-0.014	<.001	-0.014	<.001	-0.014	<.001	-0.014	<.001
Model deviance	24050.774		24047.485		24045.250		24046.427		24045.488		24047.054		24048.564	
Adjusted R ²	0.111		0.115		0.121		0.118		0.118		0.116		0.115	

To provide some additional support for the role of culture over financial decision-making in SMEs the authors did additional tests taking into account the frameworks of Schwartz and GLOBE. The results are equivalent.

In summary, and with the underlying data it can be concluded that culture, in this case measured by the variables *Uncertainty Avoidance*, *in relation to Debt to Assets*, and *Individualism*, with the opposite direction, *Masculinity*, *Uncertainty Avoidance*, *in relation to Debt to Revenue* have an impact on the capital structure of firms.

6 DISCUSSION AND CONCLUSIONS

Does culture matter in financial decisions? To answer this question, we used a large database of 27,778 medium-sized firms from the EU27 over the 2015-2019 period. Our study provides some support for the impact of culture on financial decisions, but compared to previous studies, this impact appears to be smaller.

The differential impact of culture revealed by our dataset can be explained by our unique research design characteristics, which shed new light on the aforementioned research question. Our research design differs from previous studies in four ways.

First, due to the timeframe of our dataset, which includes the most recent firm-level data, with the closest period in the literature being 2007-2015, by Diez-Esteban *et al.* (2019). This implies that, in addition, our study has considered the most recent balance sheet and accounting standards.

Second, compared to prior studies, we used one of the largest firm-level datasets, with up to N=27,778, the average sample ranging from N=5,000 to N=10,000 (Baxamusa & Jalal, 2014; Adhikari & Agrawal, 2016; and He & Hu, 2016). Also, aside from being a large dataset, it specializes in privately held manufacturing SMEs, unlike other databases including publicly traded companies, as used by Baxamusa and Jalal (2014), among others.

As a third aspect, we created a full European dataset based on the EU-27 countries, similar to Gaganis *et al.*'s (2019), and not a worldwide one like Diez-Esteban *et al.*'s (2019) or a US dataset (Adhikari & Agrawal 2016; He & Hu 2016). In this regard, our results suggest that European societies have evolved and that nowadays, culture only slightly affects financial behavior.

Finally, most previous studies used an inadequate methodology by applying ordinary least square (OLS) regression on a multilevel dataset with multilevel variables, such as variables at the national and organizational levels (Hilary and Hui, 2009; Jiang *et al.*, 2015; Adhikari & Agrawal, 2016; He & Hu, 2016). Therefore, we focused on a multilevel research question and employed hierarchical linear modeling to explore the uniqueness of our multilevel dataset, where organizational data is nested within institutional data. According to Molina-Azorin *et al.* (2020), "multilevel research includes the development of multilevel theory (e.g., combining different theoretical approaches at different levels and establishing relationships between constructs at different levels), as well as the main elements of methods for empirical studies (sampling, data collection, variables and their measures, and analysis techniques..." (p. 309). Multilevel research can help overcome the problems associated with classical data aggregation and disaggregation methods that can lead to irrelevant conclusions when using data measured and analyzed at a different level because the hierarchical and nested structure of the data is not considered (Molina-Azorin *et al.* 2020).

With respect to our hypothesis of the impact of culture on firms' capital structure, our two measures of capital structure produced different outcomes. Even though the debt to asset ratio has been widely used and accepted for measuring the capital structure of firms, the second variable, debt to revenue, seems to have a higher predictive power in SMEs. In this sense, the context of SMEs may be relevant for highlighting the role of borrowing in relation to revenue and the possibility of paying back the debts with it.

Our findings indicate that a significant percentage of the capital structure of European SMEs can be explained by country and firm-level variables. Specifically, the *ROA* and the country variable *Median income* account for up to 17,7%.

Based on the debt to assets and debt to revenue ratios, the explanatory value of the independent cultural variables only adds up to between 0,1% and 1,0% of the capital structure of firms. Both ratios are significant, but the effect size is small. So, returning to the initial question: does culture have an impact on SMEs' capital structure across Europe? The answer is definitely yes, but lower and in a different way than previous studies in the literature have found.

With the variable *Debt to Assets*, significant results for the variable *Uncertainty Avoidance* were found, at the 5% level. All the other Hofstede variables were not significant. The reproduction of the test with the dependent variable *Debt to Revenue*, presented a significant influence, at the 5% Level, of *Individualism*, in the other direction, *Masculinity*, and *Uncertainty Avoidance*. The other ones were not significant. These results, from the three different cultural theories, are in line with each other.

The results for *Uncertainty Avoidance* with both dependent variables were in line with the hypothesis that a higher level of uncertainty avoidance leads to lower level of debt. These findings support earlier findings, for example, and among others, Wang and Esqueda (2014) and Haq et al. (2018), and stay in contrast with Aggarwal and Goodell (2010), who find evidence for the opposite, that uncertainty avoidance leads to less equity and, therefore, to more leverage.

The findings for *Individualism* in relation to *Debt to Revenue* was not in line with the hypothesis that a higher level of individualism leads to higher level of debt, which was based on the studies of Antonczyk and Salzmann (2014), Gaganis et al. (2019), Haq et al. (2018), and Wang and Esqueda (2014). However, the insignificance of the first dependent variable accompanies the findings of Aggarwal and Goodell (2010), who also did not find any significance for *Individualism*.

The evidence of higher *Masculinity*, which leads to higher amount of debt, is also in line with the predictions, and the study of Chui et al. (2016). The rest was not significant. *Long Term Orientation*, hypothesis 1c, which should have had an influence, did not show any significance.

So, why are our results inconsistent with previous studies, pointing to the non-significance of culture and its influence on financial decisions? On the one hand, it could mean that culture is losing influence on the financial structure of firms, and other factors are becoming more influential. On the other hand, it remains unclear whether this is an overall trend or whether the results of our study of European SMEs between 2015 and 2019 are so specific that the effect of culture is only denied within this timeframe and context, as opposed to other studies that have found a strong correlation between culture influence and financial decisions (Aggarwal & Goodell (2010); Antonczyk & Salzmann (2014); Wang & Esqueda (2014); Chui et al. (2016); Haq et al. (2021); Gaganis et al. (2019)). Table 4 provides a summary of our findings.

Table 4: Overview of Models 1-12 – The impact of culture on Debt to Assets & Debt to Revenue

Hypotheses	Model	Dependent Variable	Independent Variable	Result
1a	3	Debt to Assets	Individualism	Not supported
1b	5	Debt to Assets	Uncertainty Avoidance	Supported at 5% Level
1c	6	Debt to Assets	Long Term Orientation	Not supported
1d	2	Debt to Assets	Power Distance	Not supported
1e	4	Debt to Assets	Masculinity	Not supported
1f	7	Debt to Assets	Indulgence	Not supported
1a	10	Debt to Revenue	Individualism	Rejected
1b	12	Debt to Revenue	Uncertainty Avoidance	Supported at 5% Level
1c	13	Debt to Revenue	Long Term Orientation	Not supported
1d	9	Debt to Revenue	Power Distance	Not supported
1e	11	Debt to Revenue	Masculinity	Supported at 5% Level
1f	14	Debt to Revenue	Indulgence	Not supported

The question of optimal capital structure has attracted several researchers throughout the last sixty years. Over time, theories that departed from a perfect market scenario where the structure of capital was irrelevant have incorporated more and more assumptions and factors. However, the influence of CEOs, their socioeconomic background, and the firm's institutional surrounding were not among them. In this paper, the UET and institutional theory are used to contend that firms' capital structure is affected by their managers' values, which are shaped and influenced by cultural values and those of the country in which the firm is based. From an organizational perspective, we proposed that certain cultural values had a specific impact on the capital structure of firms. Here, we present evidence that indicates that culture has some bearing on firms' capital structure. Our results are only partially significant, with very small effect sizes overall. This leads us to the question of whether culture has reduced its impact on decision-making and on financial decisions.

Our study, like any other, has limitations. First, we only used average data from each country to measure the culture and country variables. We are aware that there are local variations within countries. Second, in this study, we have attempted to explain the individual influences of decision-makers, but only a personal qualitative or quantitative interview could provide a more precise picture of each decision-maker's cultural values.

Future lines of research could start exactly where the study's limitations lie. CEOs could be questioned on the role that their cultural values plays in their financial decisions, similar to Zitelmann's study (2017). Also, the proposed model could be applied to smaller samples, such as local companies at the state level (Dasi et al. 2015), or even at the county level.

References

- Adhikari, B. K., & Agrawal, A. (2016). Does local religiosity matter for bank risk-taking? *Journal of Corporate Finance*, 272-293.
- Aggarwal, R., & Goodell, J. W. (2010). Financial markets versus institutions in European countries: Influence of culture and other national characteristics. *International Business Review*(19), 502-520.
- Aggarwal, R., & Goodell, J. W. (2014). Cross-national differences in access to finance: Influence of culture and institutional environments. *Research in International Business and Finance*, 193-211.
- Aggarwal, R., & Goodell, J. W. (2014). Culture, institutions, and financing choices: How and why are they related? *Research in International Business and Finance*, 101-111.
- Aggarwal, R., Faccio, M., Guedhami, O., & Kwok, C. C. (2016). Culture and finance: An introduction. *Journal of Corporate Finance*(41), 466-474.
- Antonczyk, R. C., & Salzmänn, A. J. (2014). Overconfidence and optimism: The effect of national culture on capital structure. *Research in International Business and Finance*(31), 132-151.
- Ashraf, B. N., Zheng, C., & Arshad, S. (2016). Effects of national culture on bank risk-taking behavior. *Research in International Business and Finance*(37), 309-326.
- Baker, M., & Wurgler, J. (2002, Feb). Market Timing and Capital Structure. *The Journal of Finance*, LVII(1), 1-32.
- Baxamusa, M., & Jalal, A. (2014). Does religion affect capital structure? *Research in International Business and Finance*, 112-131.
- Bernerth, J. B., & Aguinis, H. (2016). A Critical Review and Best-Practice Recommendations for Control Variable Usage. *Personal Psychology*, 229-283.
- Cai, J., & Shi, G. (2017). Do Religious Norms Influence Corporate Debt Financing? *Journal of Business Ethics*.
- Chen, H., Huang, H. H., Lobo, G. J., & Wang, C. (2016). Religiosity and the cost of debt. *Journal of Banking and Finance*(70), 70-85.
- Chen, Y., Dou, P. Y., Rhee, S. G., Truong, C., & Veeraraghaven, M. (2015). National culture and corporate cash holdings around the world. *Journal of Banking & Finance*, 50, 1-18.
- Chui, A. C., Kwok, C. C., & Zhou, G. S. (2016). National culture and the cost of debt. *Journal of Banking & Finance*(69), 1-19.
- Chui, A. C., Lloyd, A. E., & Kwok, C. C. (2002, 1st Qtr.). The Determination of Capital Structure: Is National Culture a Missing Piece to the Puzzle? *Journal of International Business Studies*, 33(1), 99-127.
- Crossan, M. M., Lane, H. W., & White, R. E. (July 1999). An Organizational Learning Framework: From Intuition to Institution. *The Academy of Management Review*, 3(24), 522-537.
- Czerwonka, L., & Jaworski, J. (2021). Capital structure determinants of small and medium-sized enterprises: evidence from Central and Eastern Europe. *Journal of Small Business and Enterprise Development*, 28(1), 277-297.
- Dalwai, T., Singh, D., & S., A. (2021). Intellectual capital, bank stability and risk-taking: evidence from Asian emerging markets. *Competitiveness Review: An International Business Journal*(1059-5422), DOI 10.1108/CR-03-2021-0031.
- Dasi, A., Iborra, M., & Safon, V. (2015). Beyond path dependence: Explorative orientation, slack resources, and managerial intentionality to internationalize in SMEs. *International Business Review*, 24, 77-88.
- Dastmalchian, A. (12. 06 2018). *Globe Project*. Von <http://globeproject.com> abgerufen
- Diez-Esteban, J. M., Farinha, J. B., & Garcia-Gomez, C. D. (January-March 2019). Are religion and culture relevant for corporate risk-taking? International evidence. *Business Research Quarterly*, 22(1), 36-55.
- DiMaggio, P. J., & Powell, W. W. (1983, Apr.). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2), 147-160.

- Dolz, C., Iborra, M., & Safon, V. (1. 4 2019). Improving the likelihood of SME survival during financial and economic crises: The importance of TMTs and family ownership for ambidexterity. *Business Research Quarterly*, 22(2), 119-136.
- Donaldson, G. (1961). *Corporate Debt Capacity - A Study of Corporate Debt Policy and the Determination of Corporate Debt Capacity*. Boston: Graduate School of Business Administration, Harvard.
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods*, 12, 121-138.
- Gaganis, C., Pasuiuras, F., & Voulgari, F. (2019). Culture, business environment and SME's profitability - Evidence from European Countries. *Economic Modelling*, 78, 275-292.
- Haerpfer, C. W. (12. 06 2018). *World Value Survey*. Von <http://www.worldvaluessurvey.org/wvs.jsp> abgerufen
- Hambrick, D. C., & Mason, P. A. (1984, April). Upper Echelons: The Organization as a Reflection of Its Top Managers. *The Academy of Management Review*, 9(2), 193-206.
- Haq, M., Hu, D., Faff, R., & Pathan, S. (2018, September). New evidence on national culture and bank capital structure. *Pacific-Basin Finance Journal*, 50, 41-64.
- He, W., & Hu, M. (2016). Religion and bank loan terms. *Journal of Banking & Finance*, 205-215.
- Hens, T., & Wang, M. (2007). Does Finance have a cultural Dimension? *National Centre of Competence in Research Financial Valuation and Risk Management*. Working Paper No. 377.
- Hermanns, J. (2006). *Optimale Kapitalstruktur und Market Timing - Empirische Analyse börsennotierter deutsche Unternehmen*. Wiesbaden: Deutscher Universitäts-Verlag.
- Hilary, G., & Hui, K. (2009). Does religion matter in corporate decision making in America? *Journal of Financial Economics*, 455-473.
- Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Theoretical and methodological implications for organizational science. *Journal of Management*, 24, 623-641.
- Hofstede. (2021). *Hofstede Insights*. Von <https://www.hofstede-insights.com/product/compare-countries/> abgerufen
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations* (Vol. 2nd). Thousand Oaks, CA: Sage Publications Inc.
- Hofstede, G., & Bond, M. H. (1988). The Confucius connection: From cultural roots to economic growth. *Organizations Dynamics*, 1(15), 4-21.
- Iborra, M., Safon, V., & Dolz, C. (2020). What explains resilience of SMEs? Ambidexterity capability and strategic consistency. *Long Range Planning*, <https://doi.org/10.1016/j.lrp.2019.101947>.
- Iborra, M., Safon, V., & Dolz, C. (2022). Does ambidexterity consistency benefit small and medium-sized enterprises' resilience? *Journal of Small Business Management*, 1-44.
- Jacobson, T., Linde, J., & Roszbach, K. (2005). Credit Risk Versus Capital Requirements under Basel II: Are SME Loans and Retail Credit Really Different? *Journal of Financial Services Research*, 28(1), 43-75.
- Jiang, F., Jiang, Z., Kim, K. A., & Zhang, M. (2015). Family-firm risk-taking: Does religion matter? *Journal of Corporate Finance*(33), 260-278.
- Jordaan, J. A., Dima, B., & Golet, I. (2016). Do societal values influence financial development? New evidence on the effects of post materialism and institutions on stock markets. *Journal of Economic Behaviour & Organization*, 197-216.
- Kühnen, U. (2015). *Tierisch kultiviert - Menschliches Verhalten zwischen Kultur und Evolution*. Berlin Heidelberg: Springer Spektrum.
- Karolyi, A. G. (2016). The gravity of culture for finance. *Journal of Corporate Finance*(41), 610-625.
- Kraus, A., & Litzenberger, R. H. (1973, Sep.). A State-Preference Model of Optimal Financial Leverage. *The Journal of Finance*, 28(4), 911-922.

- Lammers, J., & Garcia, M. (2014). Institutional Theory. In L. Putnam, & D. Mumby, *The SAGE handbook of organizational communication: Advances in Theory, Research, and Methods* (S. 195-216). Thousand Oaks, CA: SAGE.
- Lammers, L., & Barbour, J. B. (2006, July). An institutional theory of organizational communication. *16*, 356-377.
- Li, K., Niskanen, J., & Niskanen, M. (2019). Capital structure and firm performance in European SMEs – Does credit risk make a difference? *Managerial Finance*, *45*(5), 582-601.
- Licht, A. N. (2001). The mother of all path dependencies: toward a cross-culture theory of corporate governance systems. *Delaware Journal of Corporate Law*, *26*, 147 - 205.
- Maridal, H. J. (2013). Cultural impact on national economic growth. *The Journal of Socio-Economics*(47), 136-146.
- Meyer, J. W., & Rowan, B. (Sep. 1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology*, *2*(83), 340-363.
- Modigliani, F., & Miller, M. H. (1958, Jun.). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, *48*(3), 261-297.
- Mogha, V., & Williams, B. (2021). Culture and capital structure: What else to the puzzle? *International Review of Financial Analysis*, *73*, 1-19.
- Molina-Azorin, J. F., Pereira-Moliner, J., Lopez-Gamero, M. D., Pertusa-Ortega, E. M., & Tari, J. J. (2020). Multilevel research: Foundations and opportunities in management. *Business Research Quarterly*, *23*(4), 319-333.
- Mourouzidou-Damtsa, S., Milidonis, A., & Stathopoulos, K. (February 2019). National culture and bank risk-taking. *Journal of Financial Stability*, *40*, 132-143.
- Myers, S. C. (2001). Capital Structure. *Journal of Economic Perspectives*(15), 81-102.
- Myers, S. C., & Majluf, N. S. (1984). Corporate Financing and Investment Decisions when Firms have Information that Investors do not have. *Journal of Financial Economics*, *13*, 187-221.
- Nielsen, B. B., & Nielsen, S. (2013). Top Management Team Nationality Diversity and Firm Performance: A Multilevel Study. *Strategic Management Journal*(34), 373-382.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press.
- North, D. C. (1992). *Institutionen, institutioneller Wandel und Wirtschaftsleistung*. Tübingen: Mohr Siebeck.
- Papageorgiou, T., Savvakis, G., & Kenourgios, D. (2020). Is political risk a driver of listed SMEs leverage? *Applied Economic Letters*, 1-4.
- Ramirez, A., & Kwok, C. (2009). Multinationality as a moderator of national institutions: the case of culture and capital structure decisions. *Multinational Business Review*, *17*(3), 1-27.
- Rashid, M., Johari, D. K., & Izadi, S. (2020). National culture and capital structure of the Shariah compliant firms: Evidence from Malaysia, Saudi Arabia and Pakistan. *International Review of Economics and Finance*(<https://doi.org/10.1016/j.iref.2020.10.006>).
- Rashid, M., Johari, D. K., & Izadi, S. (2021). National culture and capital structure of the Shariah compliant firms: Evidence from Malaysia, Saudi Arabia and Pakistan. *International Review of Economics and Finance*(<https://doi.org/10.1016/j.iref.2020.10.006>).
- Rokhanyati, I., Pramuka, B. A., & Sudarto. (2019, November). Optimal Financial Leverage Determinants for SMEs Capital Structure Decision Making: Empirical Evidence From Indonesia. *International Journal of Scientific & Technology Research*, *8*(11), 1155-1161.
- Schwartz, S. (1994). Beyond individualism/collectivism: new cultural dimensions of values. In U. Kim, H. C. Triandis, C. Kagitcibasi, S.-C. Choi, & G. Yoon, *Individualism and Collectivism: Theory, Method, and Applications* (pp. 85-119). Thousand Oaks, CA: Sage.
- Schwartz, S. H. (1999). A Theory of Cultural Values and Some Implications for Work. *Applied Psychology: An International Review*, *1*(48), 23-47.

- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of the Nations*. London: W. Strahan and T. Cadell.
- Spiecker-Lampe, C. (2018). Nachhaltige Optimierung von Kapitalkosten im Mittelstand mit Supply Chain Finance. In I. Dovbischuk, G. Siestrup, & A. Tuma, *Nachhaltige Impulse für Produktion und Logistikmanagement - Festschrift zum 60. Geburtstag von Prof. Dr. Hans-Dietrich Haasis* (pp. 177-187). Wiesbaden: Springer Gabler.
- Stone, Z., Filippaios, F., & Stoian, C. (2014). Equity culture development in Central and Eastern Europe: The role of institutional and managerial factors. *Research in International Business and Finance*(31), 234-263.
- Titman, S., & Wessels, R. (1988). The Determinants of Capital Structure Choice. *The Journal of Finance*, 43(1), 1-19.
- Trompenaars, F., & Hampden-Turner, C. (2005). *Riding the Waves of Culture: Understanding Cultural Diversity in Global Business* (Bd. 2nd ed.). London: Brealey.
- Verduch Arosa, C. M., Richie, N., & Schumann, P. W. (2014). The impact of culture on market timing in capital structure choices. *Research in International Business and Finance*(31), 178-192.
- Wang, D., & Esqueda, O. A. (2014). National cultural effects on leverage decisions: Evidence from emerging-market ADRs. *Research in International Business and Finance*(31), 152-177.
- Weber, M. (1905). *Die protestantische Ethik und der Geist des Kapitalismus*. Tübingen: Mohr.
- Yin, C., Jiang, C., Jain, H. K., & Wang, Z. (2020). Evaluating the credit risk of SMEs using legal judgements. *Decision Support Systems*, 136, 1-13.
- Zitelmann, R. (2017). *Psychologie der Superreichen - Das verborgene Wissen der Vermögenselite*. München: FinanzBuch Verlag.
- Zucker, L. (1977). The Role of Institutionalization in Cultural Persistence. *American Sociological Review*, 42(5), 726-743.

Interne Ratings – Einordnung und Ableitung weicher Rating-Faktoren

Extended Abstract

Dipl.-Kfm. Matthias Kretschmar

Hochschule Fulda, Fachbereich Wirtschaft, Fulda, E-Mail: matthias.kretschmar@w.hs-fulda.de

Prof. Dr. Kai Wiltinger

Hochschule Mainz, FB Wirtschaft – LS ABWL, Rechnungswesen und Controlling, E-Mail: kai.wiltinger@hs-mainz.de

Abstract

Neben externen Ratings durch Rating Agenturen wie Standard & Poor's gibt es interne Ratings, die von den Kreditinstituten selbst durchgeführt werden und im Prozess der Kreditvergabe eine zentrale Rolle spielen. Sie erfüllen Funktionen in folgenden Bereichen: Risikomanagement, Erfüllung aufsichtsrechtlicher Anforderungen, Preisgestaltung der Kredite, Kreditgenehmigungsprozess und Verwaltung des Kreditportfolios. Interne und externe Ratings haben grundsätzlich einen ähnlichen Prozess, unterscheiden sich aber hinsichtlich der verwendeten Quellen, der Methodik, der Datenbasis und der Funktionen.

In internen Ratings werden neben harten und meist quantifizierbaren Rating-Faktoren weiche Faktoren verwendet. Diese basieren auf Erfahrungen der Firmenkundenbetreuer aus zum Teil langjährigen Geschäftsbeziehung und nutzen so eine Informationsquelle, die externen Ratingagenturen in aller Regel nicht zur Verfügung steht. Für eine spätere empirische Analyse der Einflussfaktoren und Wirkungen wird auf Basis einer Literaturanalyse ein Set von elf weichen, internen Rating-Faktoren hergeleitet.

1 Bedeutung und Funktionen interner Ratings in Banken

Interne Ratings - auch interne Bank-, Kredit- oder Risikoratings genannt - sind Bewertungen, die von Kreditinstituten selbst durchgeführt werden, um die Kreditwürdigkeit von potenziellen und gegenwärtigen Kreditnehmern zu beurteilen (Achleitner et al., 2007, S. V). Sie quantifizieren die Einschätzung des Kreditinstituts, ob ein Kreditnehmer die Fähigkeit und den Willen hat, seinen Zahlungsverpflichtungen aus einem Kreditvertrag vollständig und pünktlich nachzukommen, mittels Scoring-Werten und ordnen die Kreditnehmer so Risikoklassen mit Ausfallwahrscheinlichkeiten zu (Brusov et al., 2021, S. 5ff.; Sparkassen Rating und Risikosysteme GmbH, 2024).

Ziel des Beitrags ist es, die weichen Faktoren zu identifizieren, die bei internen Ratings von den Kundenberatern bewertet werden. In einer nachfolgenden empirischen Studie soll untersucht werden, inwiefern die Kundenberater über die notwendigen Kenntnisse bezüglich dieser Faktoren bei ihren Kunden verfügen.

Interne Bankratings sind für ein effektives Risikomanagement, die Kapitalallokation, die Preisgestaltung bei Krediten, die Kreditentscheidung und das Portfoliomanagement in Kreditinstituten unerlässlich. Durch Basel II und III sind Anforderungen an den Aufbau und Qualität der Systeme vorgeschrieben. Interne Ratings ermöglichen es den Banken, durch eine strukturierte und objektive Bewertung des Kreditrisikos sicherer, effizienter und rentabler zu arbeiten (Reichling et al., 2007, S. 87). Dabei übernehmen interne Ratings die in Abbildung 1 dargestellten konkreten Funktionen.

Risikomanagement	– Berechnung des Probability of Default Ratings (PD-Rating) für die Ertrags- und Verlustrisikosteuerung sowie die Festlegung von Kreditbedingungen
Erfüllung aufsichtsrechtlicher Anforderungen	– Berechnung der gesetzlich notwendigen Eigenkapitaldeckung – Erfüllung der aufsichtsrechtlichen Berichtspflichten
Basis der Preisgestaltung der Kredite	– Berechnung des Zinsspreads bei der Kalkulation von neuen Krediten – Tracking des Zinsspreads über das gesamte Kreditportfolio
Bestandteil des Kreditgenehmigungsprozesses	– Standardisierung und Qualitätssicherung des Kreditgenehmigungsprozesses – Sicherstellung der Konsistenz und Willkürfreiheit der Kreditentscheidungen
Verwaltung des Kreditportfolios	– Steuerung der Mischung von Krediten aus unterschiedlichen Risikoklassen – Tracking von Veränderungen des Risikos einzelner Kredite bzw. des Portfolios

Abbildung 1: Funktionen interner Ratings
(Quellen: Michaelis & Schmeisser, 2016, S. 59; Reichling et al., 2007, S. 94)

2 Abgrenzung interner Ratings von externen Ratings

Interne und externe Ratings sind beide Instrumente zur Beurteilung der Ausfallwahrscheinlichkeit, sie unterscheiden sich aber wie in Abbildung 2 dargestellt in ihren Erstellern, den verwendeten Methoden, den Zugriff auf Daten und in ihren Funktionen.

Externe und interne Ratings sind für unterschiedliche Akteure im Finanzsystem von entscheidender Bedeutung, da sie Erkenntnisse liefern, die zu einer wirksamen Steuerung und Minderung des Ausfallrisikos von Finanzanlagen wie dem Kreditportfolio beitragen (Behr & Güttler, 2004, S. 63).

Gemeinsam ist beiden die ähnliche Struktur und der ähnliche Prozess der Ermittlung. Zunächst muss eine ausreichende Anzahl an Ratingobjekten, hier Kreditnehmer, vorhanden sein und Ratingkriterien identifiziert werden, die mittels eines geeigneten stochastischen Verfahrens die Berechnung von Score-Werten ermöglichen. Auf Basis der Score-Werte werden Risikoklassen hinsichtlich der Ausfallwahrscheinlichkeit gebildet (Daldrup, 2006, S. 27ff.).

Die Qualität von Ratingverfahren hängt von der Trennschärfe ab, also der Eignung, neue Kreditnehmer den Risikoklassen eindeutig und richtig zuzuordnen (Michaelis & Schmeisser, 2016, S. 65; Wimmer, 2023, S. 446).

Methodik	Daten
Rating-Agenturen verwenden standardisierte, veröffentlichte Methoden, um Konsistenz und Vergleichbarkeit zwischen verschiedenen Unternehmen und Branchen zu gewährleisten. Bei internen Ratings entwickeln Banken ihre eigenen Modelle, die auf ihre spezifischen Bedürfnisse, ihre Risikobereitschaft und ihren Kundenstamm zugeschnitten sind. Grundsätzlich ist die Grundmethodik allerdings ähnlich.	Externe Ratings stützen primär auf öffentlich verfügbare Finanzdaten, Branchen-Benchmarks und Ratinggespräche mit dem Top-Management. Interne Ratings verwenden zusätzlich vertrauliche, bankinterne Daten, die externen Agenturen nicht zugänglich sind, wie z. B. Transaktionshistorien oder Kontoverhalten. Weiche Faktoren, die die Erfahrung der Kundenberater in der Kundenbeziehung abfragen, spielen eine wichtige Rolle.
Ersteller	Funktionen
Externe Ratings werden von Analysten unabhängiger Kredit-Rating-Agenturen wie Standard & Poor's, Moody's und Fitch erstellt, interne Ratings von internen Kreditanalysten der Banken selbst unter Verwendung eigener Modelle, u. a. auf Basis von Erfahrungen der Kundenberater.	Externe Ratings sollen Investoren über das Kreditrisiko verschiedener Unternehmen und Finanzinstrumente informieren. Sie erhöhen Transparenz und steigern das Vertrauen der Teilnehmer am Kapitalmarkt. Die Funktionen interner Ratings ergeben sich aus Abb. 1.

Abbildung 2: Abgrenzung externer und interner Ratings hinsichtlich Methodik, Daten, Ersteller und Funktionen (Quellen: Bösch, 2019, S. 206; Everling 2007, S. 7; Reichling et al., 2007, S. 94; Wimmer, 2023, S. 363)

3 Bedeutung weicher Kriterien in internen Ratings

3.1 Strukturierung der Rating-Faktoren

Nach der Messbarkeit und Bestimmbarkeit lassen sich Rating-Faktoren in harte und weiche Faktoren unterteilen (vgl. Abbildung 3). Zu den harten Faktoren zählen neben extern publizierten Größen wie dem Umsatz oder Jahresüberschuss auch interne Informationen wie etwa Volumen der Kontoüberziehungen (quantitativ) oder Rechtsform (qualitativ). Den harten Faktoren ist gemein ist, dass hier für den Bewertenden kein Ermessensspielraum besteht. Weiche Faktoren haben aufgrund der subjektiven Einschätzung durch den Firmenkundenbetreuer einen Ermessensspielraum. Beispiele sind die Qualität des Managements oder der Unternehmensstrategie.

Messbarkeit/ Bestimmbarkeit	harte Faktoren		weiche Faktoren	
Beobachtungs- objekt	Bilanz-/ Finanzkennzahlen	Kontoführung	Unternehmens-/ Umweltmerkmale	
			ohne Ermessensspielraum	mit Ermessensspielraum
Beispiel	Eigenkapitalquote, Schuldendeckungs- grad	Umfang Kontoüberziehungen	Dauer der Kundenbeziehung, Rechtsform	Managementqualität, Unternehmens- strategie

Abbildung 3: Systematisierung Rating-Faktoren

3.2 Exklusive und weiche Informationen in internen Ratings

Interne Ratings von Kreditinstituten integrieren etliche exklusiven Informationen, die auf Erfahrungen der Firmenkundenbetreuer aus einer engen und meist langjährigen Zusammenarbeit mit dem Kreditnehmer und dem tatsächlichen Verhalten basieren. Externe Ratings beruhen hingegen auf größtenteils öffentlich zugänglichen Informationen und bedürfen keiner Erfahrungshistorie (Reichling et al., 2007, S. 88). Im Bereich der weichen Faktoren kann der Firmenkundenbetreuer durch gute Kenntnis und Einschätzung des Kunden zur Qualität des Ratingergebnis seines

Kunden beitragen. Diese exklusiven Kenntnisse lassen sich nicht oder nur mit großem Aufwand in gleicher Qualität von externen Ratingagenturen generieren.

Aus der Analyse des Schrifttums zu internen Ratings lassen sich relevante interne Rating-Faktoren ableiten, die der oben ausgeführten Definition von weichen Faktoren entsprechen.

Abbildung 4 zeigt häufig in internen Ratings genutzte weiche Rating-Faktoren und die primären Quellen, in denen diese genannt wurden. Welche der weichen Rating-Faktoren zur Trennschärfe beitragen und damit Eingang in die internen Ratings finden, wird i.d.R. nicht allein auf Basis von theoretischen Überlegungen oder erfahrungsbasiert entschieden, sondern auf Basis statistischer Verfahren wie der multivariaten Diskriminanzanalyse (Wimmer, 2023, S. 447). Die Verwendung der weichen Faktoren in der Praxis zeigt aber, dass diese signifikant zur Trennschärfe beigetragen haben.

Rating Faktoren (Einschätzung des Kundenberaters für seinen Firmenkunden)	Literaturquelle*								
	Bö	Br	Gu	KE	MS	Cb	VR	Sp	Wi
Entwicklungsperspektiven der Branche	x			x				x	x
Marktstellung und Wettbewerbssituation			x	x	x	x		x	x
Geschäftsmodell und der Wertschöpfungskette			x			x			
Qualität d. Management/Organisationsstruktur	x			x	x	x	x	x	
Nachfolgeregelung			x	x	x		x		
Unternehmens- und Marktstrategie	x		x	x				x	
Prognose- und Planungsqualität	x		x		x		x	x	
Reifegrad des Risikomanagements			x	x				x	
Kostenstruktur und Kostenmanagement	x		x					x	
Informationspolitik in Krisensituationen			x		x		x	x	x
Nachhaltigkeit der Produkte und Produktion		x	x						

* Abkürzungen: Bö = Bösch, 2019, S. 207; Br = Barthruff, 2014, S. 129ff; Gu = Gusserl, 2014, S. 82; KE = Kreuz & Everling, 2013, S. 362 ff.; MS = Michaelis & Schmeisser, 2016, S. 63; Cb = Commerzbank in Reichling et al., 2007, S. 94f.; VR = Volks- und Raiffeisenbanken in Reichling et al., 2007, S. 88f.; Sp = Sparkassen in Reichling et al., 2007, S. 91f.; Wi = Wimmer, 2023, S. 442f.)

Abbildung 4: Systematisierung und Herleitung der Ratingfaktoren.

In einer Befragung von Firmenkundenberatern soll in Hinblick auf eine konkrete Geschäftsbeziehung mit einem Unternehmen erhoben werden, wie gut sie die elf internen Rating-Faktoren für einen konkreten Kunden einschätzen können. Diese Einschätzung dient als abhängige Variable für die Fragestellung, welche Aspekte der Kundenbeziehung wie z. B. Dauer, Transparenz der Kommunikation oder Beständigkeit der Ansprechpartner einen Einfluss auf die Fähigkeit der Firmenkundenberater haben, die weichen Rating-Faktoren einzuschätzen. Zudem soll untersucht werden, welche Auswirkung die Einschätzbarkeit der internen Ratingfaktoren durch die Firmenkundenbetreuer auf abhängige empirische Konstrukte wie z. B. auf das Vertrauen des Firmenkundenberaters in den Firmenkunden haben.

4 Zusammenfassung

Weiche Faktoren wie die Qualität des Management oder die Nachhaltigkeit der Produkte und der Produktion sind für die Einschätzung des Kreditausfallrisikos von hoher Bedeutung. Während diese in externen Ratings durch Rating Agenturen nur in Ausnahmefällen vorliegen, kann bei internen Ratings die exklusiven Erfahrungen der Firmenkundenberater aus den zum Teil langjährigen Geschäftsbeziehungen mit Kreditnehmern zurückgegriffen werden.

Der vorliegende Beitrag umfasst die Definition und Abgrenzung internen Ratings und hat auf Basis einer Literaturanalyse ein Set von elf häufig verwendeten internen Rating-Faktoren abgeleitet. In zukünftigen empirischen Studien sollen zunächst die Fähigkeit der Firmenkundenberater, die Kunden hinsichtlich der internen Rating-Faktoren zu bewerten, analysiert werden, ebenso wie Vorraussetzungen und Auswirkungen eines guten Bewertungsvermögens.

5 Literaturverzeichnis

- Achleitner, A.-P., Everling, O. & Niggemann, K. A. (2007): Vorwort. In: Achleitner, A.-P., Everling, O. & Niggemann, K. A. (Hrsg.), Finanzrating – Gestaltungsmöglichkeiten zur Verbesserung der Bonität. Gabler, Wiesbaden: V-VI.
- Barthruff, C. (2014): Nachhaltigkeitsinduzierte Kreditrisiken. Springer Gabler, Wiesbaden.
- Behr, P. & Güttler, A. (2004): Interne und externe Ratings. Bankakademie Verlag, Frankfurt.
- Bösch, M. (2019): Finanzwirtschaft: Investition, Finanzierung, Finanzmärkte und Steuerung, 4. Auflage. Vahlen, München.
- Brusov, P., Filatova, T. & Orekhova, N. (2021). The Importance of Rating and the Disadvantages of Existing Rating Systems. Springer, Cham.
- Daltrup, A. (2006): Rating, Ratingsysteme und ratingbasierte Kreditrisikoquantifizierung. Arbeitspapier Nr. 17/2006. Institut für Wirtschaftsinformatik der Universität Göttingen.
- Everling, O. (2007): Wesen und Bedeutung der Finanzratings. In: Achleitner, A.-P., Everling, O. & Niggemann, K. A. (Hrsg.), Finanzrating – Gestaltungsmöglichkeiten zur Verbesserung der Bonität. Gabler, Wiesbaden: 3-14.
- Guserl, R. (2014): Mit Hard Facts und guter Kommunikation zu besseren Ratings. In: Controlling & Management Review, 58(Sonderheft 2): 78–83.
- Kreutz, R. & Everling, O. (2013): Rating von Debt Produkten und die Bedeutung von Rating Advisory. In: Hasler, P. T., Launer, M. A. & Wilhelm, M. K. (Hrsg.), Praxishandbuch Debt Relation. Springer Fachmedien, Wiesbaden: 361-378.
- Michaelis, T. & Schmeisser, W. (2016): Rating und Basel III, UVK, Konstanz.
- Reichling, P., Bietke, D. & Henne, A. (2007): Praxishandbuch Risikomanagement und Rating. |Gabler, Wiesbaden, 2007
- Sparkassen Rating und Risikosysteme GmbH (2024): Risikoklassifizierung. <https://www.s-rating-risikosysteme.de/unsere-expertise/Risikoklassifizierung/>. Abgerufen am 03.07.2024.
- Wimmer, K. (2023): Bankkalkulation und Risikomanagement, 4. Auflage. Erich Schmidt, Berlin.

5

Lehre



Digifit Steuerlehre: Ein Pilotprojekt zur Umsetzung von Blended Learning Konzepten in Modulen mit den Schwerpunkten Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung

Umgesetztes Lehrprojekt

Prof. Dr. Katharina Dillkötter, StBin

Technische Hochschule Mittelhessen, Fachbereich Wirtschaft, Professur für Allgemeine Betriebswirtschaftslehre, insb. Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung, Gießen,
E-Mail: katharina.dillkoetter@w.thm.de

Prof. i.V. Dr. Melanie Frieling, StBin

Hochschule Bielefeld, Fachbereich Wirtschaft, Lehrgebiet Controlling, Finanzen und Rechnungswesen, Bielefeld,
E-Mail: melanie.frieling@hsbi.de

Abstract

„Digifit Steuerlehre“ ist ein Pilotprojekt zur Umsetzung von Blended Learning Konzepten in Modulen mit den Schwerpunkten Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung im Studiengang Betriebswirtschaft an der Technischen Hochschule Mittelhessen. Ziel dieses Projekts ist die Steigerung der studentischen Lernmotivation und die Förderung der Fähigkeit zur Einschätzung des persönlichen Lernerfolgs. Dazu sollen Materialien erarbeitet werden, welche dazu dienen den Lernerfolg zu erhöhen und als schwierig empfundene Themen nachzuarbeiten. Der Einsatz der Materialien soll in ein ganzheitliches Blended Learning Konzept eingebettet werden. Das Projekt wird in mehreren Zyklen durchgeführt, die methodisch dem Vorgehen der Aktionsforschung folgen. Es wird jeweils das Feedback von unterschiedlichen Akteuren einbezogen. Die Besonderheit des Projekts besteht darin, dass die Studierenden selbst Vorschläge für passende E-Learning Materialien und zu bearbeitenden Themen unterbreiten und wesentlich an der Umsetzung beteiligt sind.

1 Ausgangssituation

In den Grundlagenmodulen der betriebswirtschaftlichen Studiengänge an der Technischen Hochschule Mittelhessen ist die Lerngruppe regelmäßig heterogen. Dies zeigt sich bspw. in der Vorbildung der Studierenden. Ein Teil der Studierenden verfügt über fachliche Vorkenntnisse. Studierende dieser Gruppe haben bereits eine Ausbildung absolviert, teilweise bereits wenige Jahre im Beruf gearbeitet oder einige Semester in einem anderen Studiengang studiert (meist Recht/Betriebswirtschaftslehre/Volkswirtschaftslehre an einer Universität). Diese Studierenden können den Lehrinhalten weit besser folgen als diejenige Studiengruppe, die ohne fachliche Vorbildung ist. Diese Ausgangssituation spiegelt sich in den Evaluationen wider, in denen das Vorlesungstempo von den Studierenden sehr unterschiedlich wahrgenommen wird.

Hinzu kommt, dass vor allem zu Beginn des Studiums bei den Studierenden Schwierigkeiten bestehen, den eigenen Lernerfolg einzuschätzen und Verantwortung für diesen zu übernehmen. Projektauslöser ist der Gedanke, die unterschiedlichen Lernbedürfnisse der Studierenden über asynchrone Elemente des E-Learnings abzudecken und somit ein Blended Learning Szenario zu gestalten.

2 Projekthinhalte und Zielsetzung

Die Idee für das Lehrprojekt „Digifit Steuerlehre“ ist aus der vorgenannten Problemstellung zu Beginn des Sommersemesters 2023 entstanden. Digifit ist die Abkürzung für digitale Fitness. Durch die digitale Unterstützung sollen die Studierenden „fitter“ in ausgewählten Modulen werden. Sie sollen in die Lage versetzt werden, ihren eigenen Lernstand einzuschätzen und eigenverantwortlich ihren Lernbedarf und ihr Lerntempo zu steuern. Das Projekt „Digifit Steuerlehre“ soll anhand seiner organisatorischen, fachlichen, didaktischen und wissenschaftlichen Dimension vorgestellt werden:

- *Organisatorisch* bezieht sich das Lehrprojekt auf Module in den Bereichen Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung. Diese Module sind an der Technischen Hochschule Mittelhessen am Fachbereich Wirtschaft im Studiengang Betriebswirtschaft angesiedelt. Dabei handelt es sich um einen 7-semestrigen Bachelor-Studiengang. Das angestoßene Lehrprojekt wurde im Rahmen des Lehr-Innovations-Fonds der Technischen Hochschule Mittelhessen ausgewählt. Die Projekte des Lehr-Innovations-Fonds sind wiederum Bestandteil des von der Stiftung Innovation in der Hochschullehre geförderten Verbundprojekts NIDIT (Network for Impactful Digital International Teachingskills). Dies hat auch zur Folge, dass das Projekt durch weitere Einheiten der Hochschule, insbesondere das Zentrum für kooperatives Lernen und Lehren (ZekoLL) und das Zentrum für Qualitätsentwicklung (ZQE), begleitet wird. Im Rahmen des Projekts werden zwei studentische Onlinetutor*innen beschäftigt, die selbst Vorschläge für passende E-Learning Materialien und zu bearbeitenden Themen machen und wesentlich an der Umsetzung beteiligt sind.
- In *fachlicher* Hinsicht erfolgt die Umsetzung zunächst im Rahmen des Moduls Steuerlehre. Beim Modul Steuerlehre handelt es sich um eine Lehrveranstaltung im 3. Semester des Bachelors Betriebswirtschaft am Fachbereich Wirtschaft der Technischen Hochschule Mittelhessen. Das Modul Steuerlehre ist ein Pflichtfach für alle Studierenden des vorgenannten Studiengangs mit einem Umfang von 4 SWS bzw. 5 ECTS. Prüfungsform ist eine 90-minütige Klausur. Studierende des Moduls Steuerlehre werden in die Lage versetzt die Hintergründe der Steuererhebung zu verstehen und das Steuerrecht anzuwenden. Für wesentliche Ertragsteuerfälle lernen sie mit Hilfe des Gesetzes die persönliche wie sachliche Steuerpflicht abzuleiten, Bemessungsgrundlagen zu ermitteln und Steuerbelastungen zu bestimmen. Sie beherrschen das Zusammenspiel der Ertragsteuern, in Form des Trennungs- und des Transparenzprinzips, sowie die Bedeutung und die Erhebungsform der Umsatz- wie auch der Erbschaftsteuer.
- Für dieses Modul liegt ein regelmäßig aktualisierter und mit vielen Übungsaufgaben und Beispielen angereicherter Foliensatz vor, der in den Präsenzveranstaltungen eingesetzt wird. Der Foliensatz stellt zugleich

das Lehrskript dar. Aufgrund der Gruppengröße werden die in Präsenz durchgeführten Lehrveranstaltungen an zwei Terminen gelesen. Die Lehrinhalte werden von studentischen Tutor*innen anhand weiterer Übungsaufgaben in den angebotenen Tutorien wiederholt. Das Modul Steuerlehre wird bisher trotz starker Gruppengröße im Lehrformat des seminaristischen Unterrichts durchgeführt. Hintergrund dafür ist, dass der Anwendung des Wissens, z. B. die Ermittlung einer Steuer sowie die Wirkungen verschiedener Einflussgrößen auf die Steuer, eine wesentliche Bedeutung zukommt. Neben der Entwicklung von digitalen Lehrmaterialien ist deren Einsatz sowie Gewichtung im Rahmen eines Gesamt-Lehrkonzepts zu erarbeiten. Hierbei sind die individuellen Lernstände und Lernpräferenzen zu berücksichtigen. Aus *didaktischer* Perspektive fokussiert das Projekt damit auf die Themen „heterogene Lerngruppen“, „E-Learning“, „Blended Learning“ und „Motivation“.

- Aus *wissenschaftlicher* Sicht erfolgt die Einbettung in die aktionsforschungs-basierte Analyse. Dies impliziert, dass das Projekt in mehreren Teilschritten durchgeführt wird, die jeweils wieder reflektiert werden. Das NIDIT und das ZekoLL beziehen sich in abgeschlossenen Projekten bereits auf die Aktionsforschung (Bock & Heiny, 2019; Heiny & Weyers, 2017).

Die einzelnen Sichtweisen stehen dabei nicht isoliert, sondern haben Überschneidungen und bedingen sich. So haben beispielsweise die komplexen fachlichen Inhalte Auswirkungen auf die didaktische Gestaltung. Diese Zusammenhänge werden auch aus Abbildung 1 deutlich.

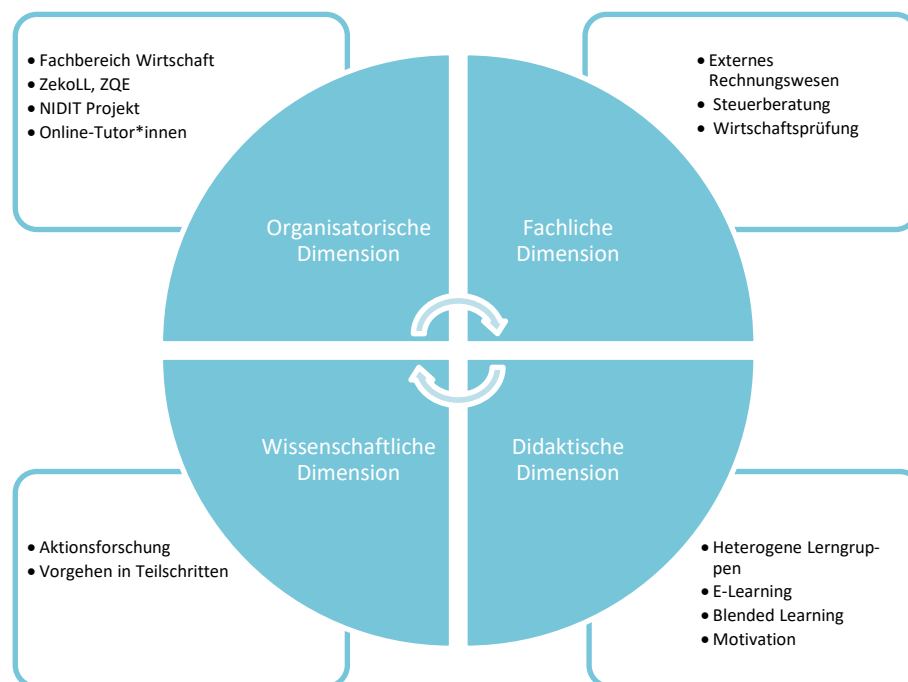


Abbildung 1: Dimensionen des Projekts

Das Projekt beinhaltet die Konzeption, die Erstellung und den Einsatz von digitalen Materialien. In einem ersten Schritt werden semesterbegleitende Moodle-Quizzes eingesetzt, um den Studierenden die Möglichkeit zu geben, ihre Leistung einzuschätzen. Im Rahmen des Projekts soll insbesondere die Unterstützung des studentischen Lernens durch Videos hinzukommen. Die neuen digitalen Materialien sollen gemeinsam mit dem bisherigen Präsenzunterricht ein stimmiges Blended Learning Konzept bilden. Dabei wird das sog. Anreicherungsmodell (KIT, 2024) gewählt.

Das Projekt verfolgt ausdrücklich den Ansatz in einen offeneren Kontakt mit den Studierenden zu kommen. Hierzu wird, in Ergänzung zu den weiterhin beschäftigten Präsenztutor*innen, verstärkt auf studentische Onlinetutor*innen zurückgegriffen. Diese unterstützen beim Austausch mit den Studierenden, liefern selbst Ideen für Themen/Medien und erstellen Online-Materialien wie Videos, die die Lehrinhalte in spannendere und lebensnahe Situationen von Studierenden projizieren. Zudem geben sie wertvolles Feedback.

Dem Lehrprojekt liegen demnach folgende Kernzielsetzungen zugrunde:

- Die Lernmotivation der Studierenden in den Modulen aus den Bereichen Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung soll erhöht werden.
- Zudem soll die Fähigkeit der Studierenden den eigenen Lernerfolg einzuschätzen gefördert werden. Ziel ist es, dass die Studierenden in die Lage versetzt werden, eigenverantwortlich das Lerntempo und den Lernbedarf zu bestimmen. Den Studierenden soll damit eine besondere Verantwortung, über ihre reguläre Rolle als Studierende hinaus, zukommen: Sie sollen beispielsweise anhand von sich selbst auswertenden Quizzes ihre individuellen Lernschwächen bzw. -schwierigkeiten identifizieren.
- Mit Unterstützung von zu erstellenden E-Learning-Materialien soll den individuellen Einzelinteressen der heterogenen Lerngruppe begegnet werden. So erhalten die Studierenden einen Bauchladen an digitalen Medien, der letztlich von ihnen initiiert wurde. Sie können aus diesem Bauchladen diejenigen Bausteine wählen, die für ihren – durch die Quizzes selbst ermittelten – persönlichen Lernstand passend sind.

Eine ergänzende Zielsetzung ist die Qualifikation der Tutor*innen.

3 Didaktischer Hintergrund

Der Begriff „blended“ hat in der Übersetzung die Bedeutung „vermischt“ oder „vermengt“. In der Literatur wird unter *Blended Learning* eine Mischung von Präsenzlehre und Onlinelehre verstanden (Zumbach & Astleitner, 2016, S. 148). Die Begriffe Onlinelehre, E-Learning und digitale Lehre werden im Beitrag synonym verwendet. Sie werden „verstanden als Einsatz von digitalen Technologien, wie z.B. Computern, Tablets und Smartphones, zur Unterstützung und Optimierung von Bildungsprozessen“ (Eisl et al., 2023, S. 364). Der Anteil von Präsenzlehre und Onlinelehre ist bei Blended Learning Konzepten nicht festgelegt. Das KIT (Karlsruher Institut für Technologie) hat unterschiedliche Formen des Blended Learnings systematisiert (Anreicherungsmodell, Wechselmedienmodell, Rotationsmodell, Wechselphasenmodell, Hyflex Modell, Inverted Classroom Modell – KIT (2024)).

An Hochschulen wird Blended Learning schon länger genutzt (Reinmann, 2008, S. 11). Ein weiterer Vorteil dieses Ansatzes ist, dass er insbesondere bei Absolvent:innen der Betriebswirtschaft auch der späteren beruflichen Lebensrealität bei Fort- und Weiterbildungen entspricht (mmb Institut - Gesellschaft für Medien- und Kompetenzforschung mbH, 2019).

Bisherige Metastudien zeigen, dass Blended Learning einen Vorteil gegenüber reinen Präsenz- bzw. reinen Online-Lehrveranstaltungen hat (Means et al., 2013; Bernard et al., 2014; Vo et al., 2017). Insbesondere ist den Studien zu entnehmen, dass Blended Learning, sinnvoll eingesetzt, die Lernsituation verbessert. Einen positiven Einfluss hat es bspw., wenn die digitale Lehre durch eine stärkere Interaktion der Studierenden ergänzt wird (siehe dazu Hattie, 2023, S. 399 m. w. N.).

Dies zeigen auch bereits Studien aus Lehrprojekten an Hochschulen für angewandte Wissenschaften (Dressler et al. 2016; Dressler & Rachfall, 2016; Rachfall et al. 2019; Dressler & Rachfall, 2020; Eisl et al., 2023). So bevorzugen Studierende eine Mischung aus Präsenzlehre und digitaler Lehre (Dressler et al., 2017, S. 254).

Blended Learning geht zudem über die bloße Kombination von Präsenzlehre und Onlinelehre hinaus (Kerres, 2018, S. 24). In aufbauenden Definitionen wird betont, dass es darauf ankommt, die zwei Teilbereiche der Präsenz- und

der Onlinelehre sinnvoll aufeinander abzustimmen (Gardner & Thielen, 2015, S. 19-20). Hier bedarf es einer Analyse, welche Inhalte zweckmäßigerweise in der Präsenzphase und welche in der Onlinephase platziert werden sollten (Garrison & Kanuka, 2004, S. 96). Zudem ist es erforderlich, nicht nur einzelne Elemente hinzuzufügen oder wegzunehmen. Es muss sich im Gesamten ein Modulkonzept ergeben, das mit den Lernzielen stimmig ist (McCarthy, 2016, S. 333). Dressler & Rachfall beschreiben Erfolgsfaktoren der Einführung von Blended Learning Konzepten an Hochschulen für angewandte Wissenschaften. Dabei wird insbesondere die Bedeutung der Kursstruktur für die Akzeptanz des neuen Kursformats hervorgehoben. Es muss eine Kooperation zwischen Studierenden und Lehrenden bestehen, strukturelle Aspekte sind notwendig, aber allein unzureichend. So muss die Überführung der Struktur in eine inspirierende und motivierende Lehre erfolgen (Dressler & Rachfall, 2020, S. 187 basierend auf Zierer 2014, S. 6 und Hattie, 2012).

Eisl et al. 2023 nutzten im Rahmen eines Videoprojektes die Erfolgsfaktoren *für heterogene Lerngruppen* nach Spector 2014. Demnach sind die (1) Personalisierung des Lernprozesses, die (2) Förderung der Selbstregulation, z. B. über eine Möglichkeit der Fortschrittsüberwachung im Hinblick auf die Lernziele, die (3) Möglichkeit Feedback im Hinblick auf die Lernziele zu erlangen und (4) eine Lernumgebung für den Austausch besonders wichtig. Im Rahmen des Lehrprojekts von Eisl et al. wurden diese Kriterien beachtet (Eisl et al., 2023, S. 365). Dies führte zu einer hohen allgemeinen Zufriedenheit mit dem Kursformat, einer hohen wahrgenommenen Qualität und einem hohen wahrgenommenen Nutzen der Lernvideos sowie der Veranstaltung an sich. Es zeigte sich ein positives Zusammenspiel von strukturierenden synchronen Veranstaltungen und Arbeit in eigenem Lerntempo (Eisl et al., 2023, S. 373-374). Auch wenn es sich bei diesem Projekt um einen Kurs im rein digitalen Format handelt, lassen sich hieraus wertvolle Erkenntnisse für die Videoplanung herleiten.

4 Vorgehen auf Basis der Aktionsforschung

Die Aktionsforschung ist eine Forschungsrichtung, zu der eine Vielzahl von Definitionen zu finden sind und die ihren Ursprung im angloamerikanischen Raum hat (Warneke, 2007, S. 66 ff.). Anliegen des Forschungsansatzes ist es, die Forschung praxis- und anwendungsorientiert zu gestalten (Cendon, 2015, S. 4 mit Verweis auf Lewin, 1951, S. 169). Ziel der Forschung ist die Lösung eines konkreten praxisnahen Problems (Cendon, 2015, S. 5), und zwar gegebenenfalls über mehrere Stufen. Den Forschungsansatz der Aktionsforschung kennzeichnend, ist die Lösung von konkreten Problemen (Fox et al., 2007, S. 51) durch Aktion und Reflexion. In diesem Zuge kann sich Wissen, insbesondere in Form konkreter Anwendungsergebnisse bzw. Methoden, ergeben. Die Aktionsforschung selbst ist keine eigene Methode; sie verfügt auch nicht über die „eigenen“ Methoden. Vielmehr lassen sich im Rahmen der Aktionsforschung eine Vielzahl von Methoden verwenden, die sich insbesondere in der Datenerhebung (z. B. Fragebögen oder qualitative Interviews), in der Auswertung sowie der Entwicklung von Strategien und den zugehörigen Maßnahmen zeigen.

Die allgemeinen Merkmale zur Aktionsforschung wurden durch Altrichter et al. auf den Fall der Unterrichtsforschung angepasst (Altrichter et al., 2018, S. 13). Auf Basis dieser Merkmale werden im Folgenden die wesentlichen Charakteristika des Lehrprojekts zusammengefasst.

Fragestellung aus der Praxis (Altrichter et al., 2018, S. 13)

In der Aktionsforschung geht es um eine konkrete Problemstellung aus der Praxis. Aus der Problemstellung lassen sich Forschungsfragen ableiten, die nachstehend zu folgenden Kernprojektzielen -mit Blick auf die Studierenden- formuliert wurden:

- Erhöhung der Lernmotivation
- Förderung der Fähigkeit, den eigenen Lernerfolg einzuschätzen
- Eigenverantwortliche Bestimmung von Lerntempo und Lernbedarf

- Identifizierung von Lernschwierigkeiten und passgenaue optionale Nutzung von noch zu erstellenden E-Learning-Inhalten zur Angleichung der Wissensstände

Mit Blick auf die Tutor*innen ist die Qualifikation dieser als zusätzliches Ziel zu benennen.

Forschung der Betroffenen, Konfrontation unterschiedlicher Perspektiven und Einbettung in die professionelle Gemeinschaft (Altrichter et al., 2018, S. 13, 15 f.)

Der Forschungsansatz zeichnet sich dadurch aus, dass Forschende und Beforschte Betroffene der Forschung sind (Altrichter et al., 2018, S. 16). Im Unterrichtsszenario ist zunächst die Lehrperson selbst betroffen und gibt zugleich den Anstoß für die Forschung. Um die spezifische Situation zu verbessern, ist es nach der Aktionsforschung essenziell auch die weiteren Betroffenen, wie im betrachteten Fall die Studierenden, einzubeziehen. Diese stellen die unmittelbar beforschte Gruppe dar. Es können aber auch weitere Personen, wie in diesem Lehrprojekt die Tutor*innen Betroffene sein. Diese sind lediglich als mittelbar Beforschte betroffen, weil sie zwar aktuell nicht das konkrete Modul belegen, aber mittelfristig ein anderes Modul aus den Bereichen Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung in ihre Studienplanung aufnehmen könnten, das gegebenenfalls vom Projekt profitiert.

In der Aktionsforschung soll im Idealfall eine Einbettung der individuellen Forschung in eine professionelle Gemeinschaft erfolgen (Altrichter et al., 2018, S. 16). Es wurde bereits in der Literatur beschrieben, dass insbesondere eine hochschuldidaktische Beratung zielführend für die Lehrentwicklung ist (Merkt, 2014, S. 32-33).

Die Projekte sind bewusst darauf angelegt, unterschiedliche Perspektiven einzuholen. Die Herausforderung der Lehrenden ist dabei, durch die Konfrontation mit anderen Meinungen eine Distanz zu finden (Altrichter et al., 2018, S. 15 f.). Zusammengefasst können die beteiligten Personen und Organisationen des Lehrprojektes -je nach Betroffenheit- zu drei Akteursgruppen zusammengefasst werden, wie nachstehende Abbildung 2 zeigt:

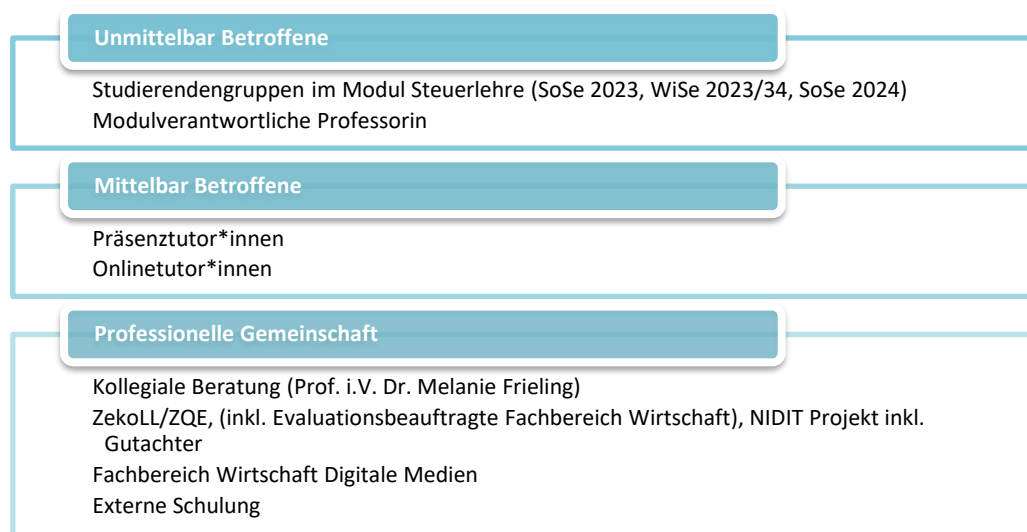


Abbildung 2: Akteure im Lehrprojekt (eigene Darstellung)

Längerfristige Forschungs- und Entwicklungszyklen (Aktionsspirale nach Altrichter et al., 2018, S. 15)

Die Forschungsprojekte sind insgesamt längerfristig angelegt und umfassen mehrere Zyklen. Das konkrete Lehrprojekt umfasst derzeit drei zyklisch ablaufende Prozesse (Phasen), die jeweils vier Teilphasen beinhalten. Auf die drei Zyklen wird in den Kapiteln 5.1 bis 5.3 genauer eingegangen. Das letztlich in einem revolvierenden Zyklus verlaufende Lehrprojekt kann durch folgende Darstellung erfasst werden (Abbildung 3):

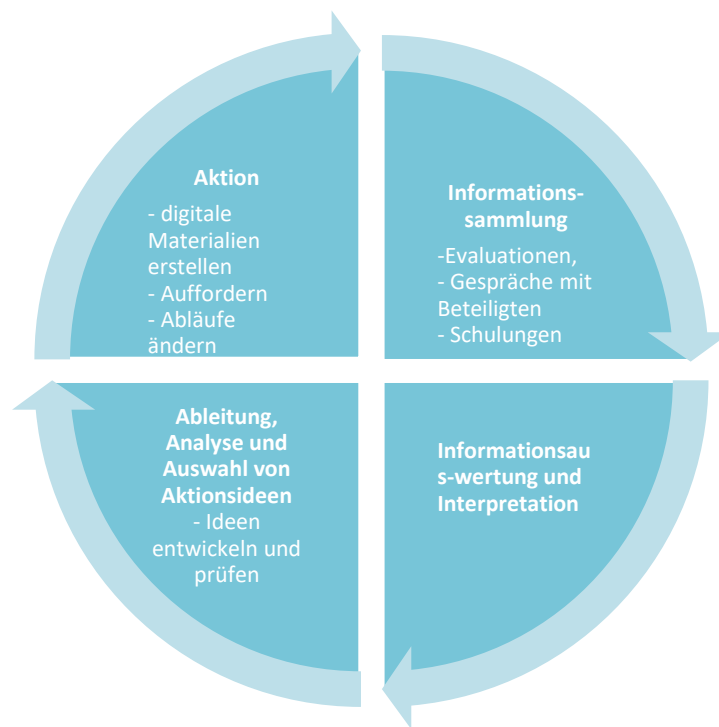


Abbildung 3: Zyklischer Prozess des Lehrprojekts (Grundform nach Altrichter et al., 2018, S. 88)

Dieser viergliedrige Zyklus kann in einen Aktions- und einen Reflexionsbereich geteilt werden. Der Aktionsteil umfasst die reine Durchführung der ausgewählten Handlungsalternativen (Viertelkreis oben links in der Abbildung 3); der Reflexionsteil umfasst die verbleibenden drei Teilphasen Informationssammlung, Auswertung und Interpretation sowie Ableitung von Handlungsalternativen. Die Erhebung von Daten erfolgt nicht am Ende des Projekts, sondern zeigt sich durch viele kleinere Zwischenerhebungen. Durch die Aneinanderreihung der Zyklen soll bildlich gesehen -im Idealfall- eine Aufwärtsspirale entstehen (Abbildung 4).

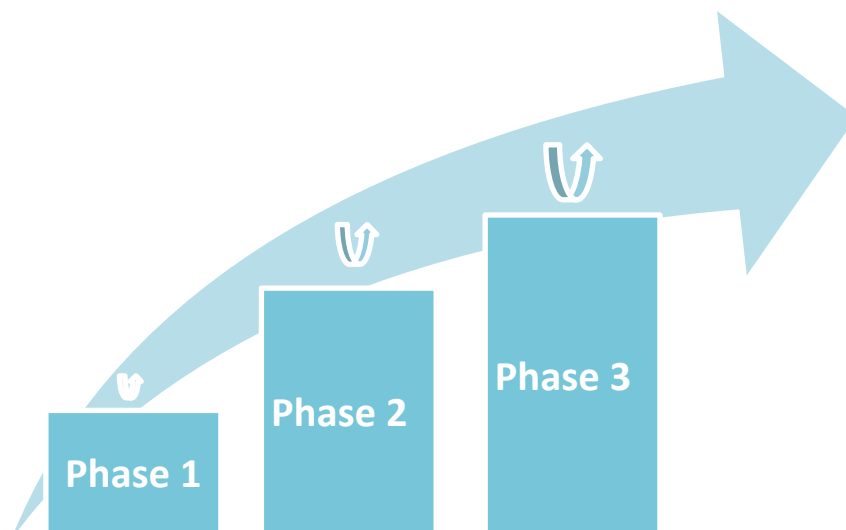


Abbildung 4: Aufwärtsspirale

Vereinbarung ethischer Regelungen für die Zusammenarbeit und Wertaspekte pädagogischer Tätigkeit sowie Veröffentlichung von Praktikerwissen (Altrichter et al., 2018, S. 16-17)

Im konkreten Projekt wurde den Tutor*innen ausdrücklich kommuniziert wurde, dass alle Themen eingebracht werden sollen. Bei den Tutor*innen handelt es sich um Personen, die in dem Modul keine Prüfungsleistung mehr erbringen müssen. Die Evaluationen am Fachbereich Wirtschaft werden zudem anonym durchgeführt. Die eigenen Wertvorstellungen der Forschenden haben Auswirkungen auf die ausgewählten und durchgeführten Handlungen. Damit ist die Aktionsforschung -im Kontext des konkreten Lehrprojekts- stark geprägt von den individuellen Wertvorstellungen der forschenden Professorin. Das gewonnene Praktikerwissen soll einer professionellen Community zur Verfügung gestellt werden.

Ziel (Altrichter et al., 2018, S. 17)

Das Ziel der Aktionsforschung ist die Erkenntnis als Ergebnis der Reflexion. Die darauf aufbauende Entwicklung ist das Ergebnis der Aktion. Aus Sicht der Vertreter*innen der praxisorientierten Aktionsforschung wird aber bereits durch das bloße Einnehmen der reflektierenden Haltung zu den beforschten Situationen, in diesem Fall des Unterrichtsgeschehens, ein Veränderungsprozess bewirkt (Grassl, 2016, S. 55). In Bezug auf das konkrete Projekt sollen die eingangs genannten Ziele erfüllt werden.

5 Ergebnisse der Projektphasen

5.1 Forschungsausgangspunkt

In einem ersten Schritt wurden bereits im Sommersemester 2023 Informationen zum Modul gesammelt, um den Status Quo zu klären. Die ersten Einschätzungen beruhten dabei zunächst auf eigenen Eindrücken und den ad hoc Rückmeldungen der Studierenden. Insbesondere entstand der Eindruck einer heterogenen Lerngruppe. Dieser wurden im Projektantrag dokumentiert. Wichtig zur Klärung des Ausgangspunktes ist es, die eigenen ersten Eindrücke nicht als korrekt hinzunehmen, sondern diese mit anderen Personen zu spiegeln (Altrichter et al., 2018, S. 60-63). Im Folgenden werden die drei Phasen des Projektes getrennt nach Reflexions- und Aktionsteil vorgestellt.

Phase 1 (April 2023 – Juli 2023)

Informationssammlung, Auswertung und Ableitung von Handlungsalternativen (Reflexionsteil)

Unmittelbar Betroffene

Lehrende: Von Bedeutung sind hier die eigenen Eindrücke, die im Projektantrag beschrieben wurden. Zudem wurde Literatur zur Thematik der heterogenen Lerngruppen und zum Blended Learning ausgewertet (siehe dazu Kapitel 3 „Didaktischer Hintergrund“).

Gespräche mit Studierenden des Moduls: Aus dem Feedback im Kurs wurde deutlich, dass einige Studierende das Vorlesungstempo entweder als zu schnell oder als zu langsam empfanden. Zudem wurde in diesen Gesprächen klar, dass im 3. Semester noch strukturgebende Aktivitäten sinnvoll sind.

Mittelbar Betroffene

Präsenztutor*innen und Onlinetutor*innen: In diesen Austauschgesprächen wurde zunächst ein positiver Gesamteindruck, aber auch der Eindruck des ungleichen Lernstandes bestätigt. Zugleich wurden häufige angesprochene Fehler bzw. Verständnisprobleme der Studierenden identifiziert und der Einsatz von digitalen Medien zur Lösung andiskutiert.

Einbettung in die Professionelle Gemeinschaft:

Hingewiesen wurde im Rahmen der Rückmeldung des NIDIT auf die notwendige Qualifikation der Onlinetutor*innen. Den Kernzielsetzungen wurde als Folge die Qualifikation der Onlinetutor*innen hinzugefügt. Weiterhin wurde angemerkt, dass eine Maßnahme zu finden ist, die dazu führt, dass die E-Learning-Materialien auch wirklich genutzt werden. Es erfolgte eine hochschuldidaktische Beratung zum Thema Video in der Lehre mit dem ZekoLL. Mit Prof. i. V. Dr. Melanie Frieling wurde ein Gespräch zur Situation geführt.

Aktionsdurchführung (Aktionsteil)

Auf Basis der gewonnenen Informationen wurde sich für ein Blended Learning Szenario entschieden. Es wurde das Anreicherungsmodell gewählt. Dabei handelt sich um eine „digitale Erweiterung der Präsenzlehre“. Diese Entscheidung wurde getroffen, weil es sich um ein Modul im 3. Semester, also zu Beginn, eines Bachelorstudiengangs handelt. Da die Lernerfahrung im Studium zu diesem Zeitpunkt noch nicht sehr hoch ist, soll durch den hohen Präsenzanteil eine Struktur gegeben werden, die durch die Lehrperson umgesetzt wird (siehe dazu auch Dressler & Rachfall 2020 zur Bedeutung der Struktur in Blended Learning Konzepten).

Die genannten Ideen wurden umgesetzt: Insbesondere wurde der Moodle-Kurs stärker strukturiert; indem dieser in die sechs Kapitel des Moduls untergliedert wurde. Zudem wurden jeweils die Lernziele pro Kapitel konkretisiert und auf der Moodle Seite unterhalb des fachlichen Themas sichtbar gemacht. Die wesentlichen Inhalte sowie der typische Lernablauf wurden grafisch im Moodlekursraum aufbereitet. So wurde im Rahmen des Constructive Alignment die Zielrichtung für die Prüfungsleistung deutlich. Beim Constructive Alignment werden Lernziele, Lernaktivitäten und Prüfung miteinander abgestimmt (Biggs & Tang, 2011, S. 11 ff.). Diese Lernziele wurden ebenfalls in Moodle veröffentlicht.

Zudem wurden für jedes der sechs Kapitel Quizfragen erarbeitet. Diese finden sich im Sinne eines Selbsttestes zu jedem Kapitel im Moodlekursraum. Die sechs Quizzes konnten ab Erstellung von jedem Studierenden anonym und auch mehrfach genutzt werden; das Ergebnis wurde unmittelbar nach Testende dem Studierenden angezeigt. Die Tests dienen dabei ausdrücklich als formatives Feedback, um den Studierenden Rückmeldung zu ihrem Lernfortschritt zu geben. Diese soll als Lernchance gesehen werden (Müller & Schmidt-Hertha, 2009, S. 27).

Die Studierenden hatten die Möglichkeit innerhalb des Moodlekursraums „vorzuarbeiten“ bzw. Themen zu wiederholen. Der Kursraum umfasst neben dem Skript, Übungsaufgaben sowie ergänzende Unterlagen. Durch die Kapitelstrukturierung verliert der Leser nicht den Überblick. Das Finden der eigenen Struktur wird für die Studierenden durch eine Fortschrittsanzeige erleichtert. Zudem findet dieser im allgemeinen Teil „häufige Fragen und Antworten“, „Informationen zu Klausur und Hilfsmitteln“ sowie „Tipps zum Lernen“. In einer persönlichen wöchentlichen E-Mail werden alle Studierenden „abgeholt“: Inhalte dieser E-Mail sind der aktuelle Lernstand, die nachzuarbeitenden Themen und Aufgaben sowie ein Ausblick auf die nächste Veranstaltung. Es wird jeweils ausdrücklich auf die E-Learning Materialien hingewiesen, sodass der Zusammenhang von Präsenzlehre und digitaler Lehre deutlich wird.

5.2 Phase 2 (August 2023 bis Dezember 2023)

Informationssammlung, Auswertung und Ableitung von Handlungsalternativen (Reflexionsteil)

Nach Abschluss des Sommersemesters 2023 wurden weitere Informationen gesammelt. Diese wurden wieder aus unterschiedlichen Quellen erhoben.

Unmittelbare Betroffene:

Lehrende: Fortbestand hatte auch in dieser Phase der persönliche Eindruck dahingehend, dass das Tempo von den Studierenden sehr unterschiedlich wahrgenommen wurde.

Evaluation Sommersemester 2023: In der Evaluation des Sommersemesters 2023 bestätigte sich die Erwartung, dass jeweils gleich viele Studierende, die Veranstaltung als zu schnell oder zu langsam empfanden. Im Hinblick auf die Frage nach dem Schwierigkeitsgrad und dem Stoffumfang wurde diese Beobachtung nicht bestätigt. Im Rahmen der Evaluation wurden zur besseren Übersicht weitere Schaubilder angeregt, da die Modulzusammenhänge nicht immer eindeutig waren. In der Klausur zeigte sich, dass insbesondere Themen aus dem Bereich Umsatzsteuer als schwierig empfunden wurden.

Auswertung Moodle: Im Sommersemester 2023 nutzten 69 von 147 Studierenden mindestens eines der sechs Quizzes. 15 Studierende haben jedes der sechs Quizzes mindestens einmal genutzt. Anzumerken ist, dass die Datengrundlage dadurch beeinflusst wird, dass der Moodlekurs eine freie Einschreibung hat. Es sind also gegebenenfalls Studierende eingeschrieben, die den Moodlekurs nicht nutzen, um das Modul abzuschließen, sondern beispielsweise, weil sie ein aufbauendes Modul absolvieren und etwas nachlesen möchten.

Mittelbare Betroffene:

Feedback Tutor*innen: Als Feedbackgeber wurden zusätzlich zu den vorhandenen Präsenztutor*innen zwei Onlinetutor*innen gefunden. Die Arbeit mit den Tutor*innen wurde gewählt, damit ausdrücklich in jedem Schritt die Sicht der Studierenden berücksichtigt wird. Eine der Onlinetutorinnen hat eine erste Themenliste erstellt, die dann von einer weiteren Onlinetutorin und den zwei Präsenztutor*innen ergänzt wurde. Anschließend sollten diese Themen in Videos gefasst werden. Dem Konzept nach sollen die Videos als Unterstützung beim Studieren im eigenen Lerntempo dienen und ausdrücklich nicht den Besuch der Vorlesung ersetzen. Anschließend erfolgte die Wahl von Technik und didaktischem Szenario.

Einbettung in die Professionelle Gemeinschaft:

Im Rahmen der Einbettung durch die professionelle Gemeinschaft wurden Ansätze zur Erhebung weiterer Daten gesammelt. Dazu wurden die Teaching Analysis Polls sowie eine Erweiterung der semesterweisen Evaluation in Betracht gezogen. Teaching Analysis Polls sind Feedbackverfahren, die dadurch gekennzeichnet sind, dass es sich um Gruppendiskussionen handelt, die durch Dritte moderiert werden und die sich auf konkrete Lehrveranstaltungen beziehen (Weitzel et al., 2022, S. 108 mit Verweis auf Frank et al., 2011 und Franz-Özdemir et al., 2019). Im Hochschulkontext kann die Rolle der Moderation beispielsweise durch Mitarbeiter*innen aus der Hochschuldidaktik oder aus dem Qualitätsmanagement ausgefüllt werden. Die Diskussion dient dazu lernförderliche bzw. lernhinderliche Aspekte innerhalb der Lehrveranstaltung zu identifizieren (Weitzel et al., 2022, S. 108).

Aktionsdurchführung (Aktionsteil)

Vor dem Hintergrund der noch nicht zufriedenstellenden Nutzung der Quizzes, wurde im Wintersemester ausdrücklicher, sowohl in der Vorlesung als auch über die wöchentlichen E-Mails per Moodle, auf die Nutzung der Quizfragen hingewiesen.

Die von den Tutor*innen erstellten Video-Themenlisten wurden ausgewertet und erste Themen zur Umsetzung identifiziert. Die Videos sollten möglichst kurz sein und ein erklärungsbedürftiges Thema aufgreifen. Unter dieser Maßgabe wurden einige der gesammelten Fachthemen nochmal neu strukturiert und z. B. in mehrere Unterthemen zerlegt. Die Themenliste wurde in diesem Zuge um die bei der Klausurauswertung festgestellten Schwächen zu Themen des Umsatzsteuerrechts ergänzt. Für das Setting (also die dargestellte Lebenssituation) wurde jeweils ein Vorschlag von den Onlinetutor*innen erarbeitet und als Entwurf umgesetzt.

Für die Videos wurden auf Basis der Beratungen drei verschiedene Techniken/Szenarien erarbeitet:

- Lehrvideo animierte Powerpoint: Der Aufbau sollte anhand der Hauptmerkmale: Lernziele, Thema, Zusammenfassung erfolgen. Eine Überarbeitung soll anschließend durch Camtasia erfolgen.
- Video Storytelling: Es wurden Scribble Videos mit kleiner humorvoller Geschichte in Betracht gezogen.

- Mashup Video: Das Video startet mit einer Scribble Sequenz und anschließend erfolgt die Erklärung mit einer Live Aufnahme.

Der initiale Vorschlag und die Entwürfe zu den konkreten Videos wurden hauptsächlich von den Onlinetutor*innen erarbeitet. In mehreren Feedback-Schleifen wurden diese immer weiter verfeinert. Dies führte einerseits dazu, dass die Themen näher an den Studierenden sind und andererseits wurden die Onlinetutor*innen durch dieses Vorgehen weiterqualifiziert.

Weiterhin wurden Übersichten in Form von Mindmaps zum Gesamtmodul und zu den einzelnen Kapiteln erstellt, die nun im Moodlekursraum sichtbar sind und die Studierenden durch das Modul leiten. Eine Tutorin hat die Erstellung selbstständig vorgenommen. Durch die Mindmaps ergibt sich eine bessere Übersicht, aber auch eine ansprechendere Ansicht. Um den Studierenden noch mehr Struktur auf dem Weg durch das Modul zu gehen, wurde zusätzlich eine Fortschrittsanzeige in Moodle eingefügt. Bei pdf-Dateien und Videos können die Studierenden den Abschluss manuell bestätigen. Bei den Quizfragen muss für die Fortschrittsanzeige das Quiz zumindest mit einer Bewertung abgeschlossen werden.

Bezüglich der Datenerhebung wurde entschieden, zunächst auf die Teaching Analysis Polls zu verzichten und eine Erweiterung der semesterweisen Evaluation zur Datenerhebung zu nutzen.

5.3 Phase 3 (Januar 2024 – Juni 2024)

Informationssammlung, Auswertung und Ableitung von Handlungsalternativen (Reflexionsteil)

Unmittelbare Betroffene:

Lehrende: Der persönliche Eindruck der Lehrenden war, dass das Semester strukturierter ablief. Insbesondere die Unterteilung der Präsenzveranstaltung in fachlichen Input und anschließendes Üben mit der Möglichkeit, in dieser Zeit individuelle Fragen zu klären und Feedback zu Lösungsansätzen zu erhalten, wurde als positiv wahrgenommen. Eine Auswertung der Klausur zeigte, dass den Studierenden insbesondere das sorgfältige, schrittweise Durchprüfen aller Tatbestandsmerkmale einer Rechtsnorm Schwierigkeiten bereitet hat.

Studierende: Der geschilderte Eindruck spiegelte sich auch in den Aussagen der Studierenden über das Semester hinweg wider. Es gab zudem Anfragen zu Karteikartentools.

Evaluation Wintersemester 2023/24: In der Evaluation zeigte sich weiterhin, dass das Veranstaltungstempo entweder als zu hoch oder zu niedrig eingeschätzt wurde. Die Evaluation wurde in dem betrachteten Semester um Fragen zu Lehrprojekt erweitert. Bei der Auswertung der Fragen zum aktuellen Stand des E-Learnings bzw. dem digitalen Angebot lassen sich folgende Erkenntnisse festhalten:

- Das Statement „Insgesamt bin ich mit dem Einsatz der digitalen Systeme in der Veranstaltung zufrieden.“ Wurde von allen Studierenden mit „trifft voll zu“ und „trifft zu“ bewertet (Skala von trifft voll (1) zu bis trifft nicht zu (5)).
- Auf das Statement „Insgesamt ist das aktuelle E-Learning Angebot im Modul Steuerlehre hilfreich.“ antworteten 85,7 % der Studierenden mit „trifft voll zu“ und „trifft zu“.
- 59,5 % der Studierenden gaben zum Zeitpunkt der Evaluation an, dass sie die Quizfragen ausprobiert hatten. Diese Gruppe schätzt in einer Folgefrage die Quizzes als hilfreich zur Ermittlung des eigenen Lernstandes ein. Dies zeigt Abbildung 5.

Die Online-Quizfragen sind hilfreich, um
meinen aktuellen Lernstand
einzuschätzen.

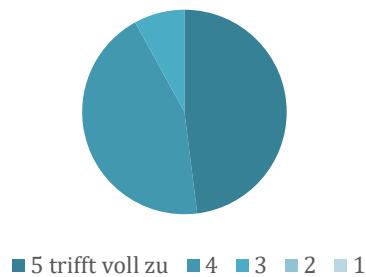


Abbildung 5: Auswertung Online-Quizfragen «Lernstand und Quizfragen»

Insgesamt zeigt sich somit eine starke Akzeptanz des E-Learning Angebotes. Um perspektivisch die Art der digitalen Inhalte passgenauer vorzubereiten, wurden auch entsprechende Fragen gestellt. Dabei wurde mit Blick auf die Erstellung von Videos abgefragt, welche Punkte den Studierenden bei Lehrvideos wichtig sind. Die Ergebnisse zeigt Abbildung 6.

WELCHE PUNKTE SIND IHNEN BEI LEHRVIDEOS WICHTIG?

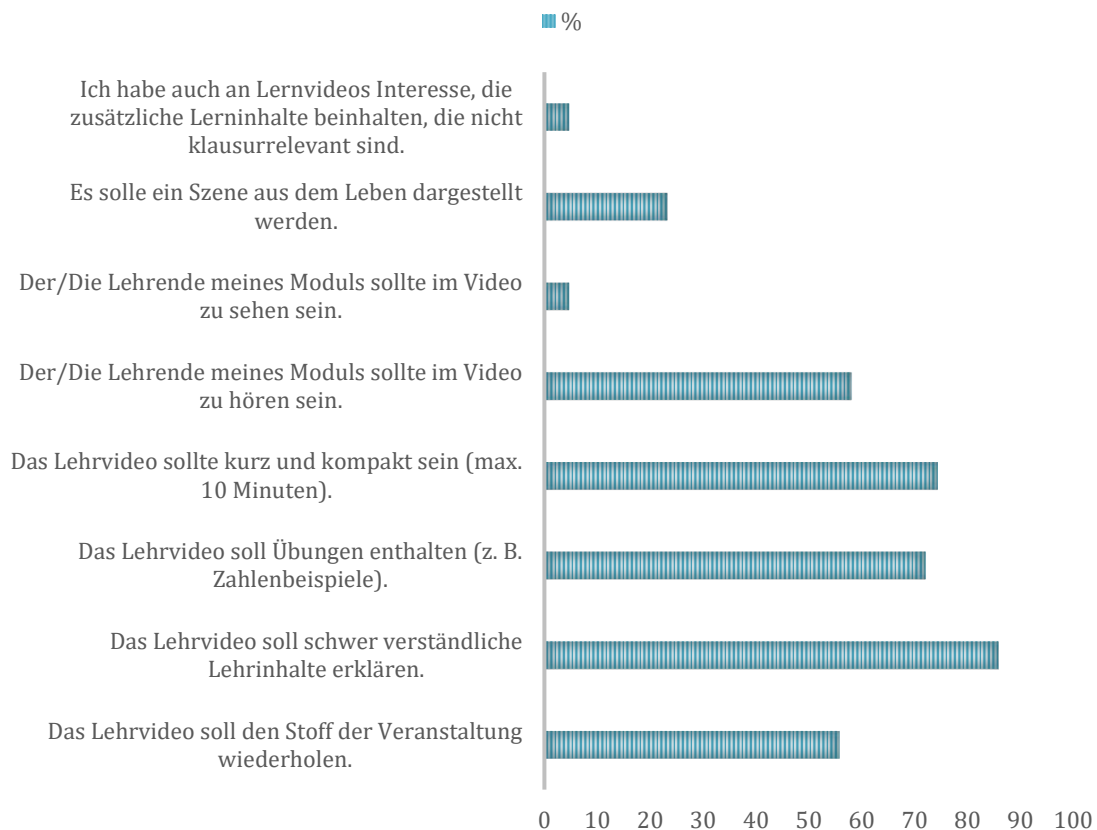


Abbildung 6: Wichtige Punkte bei Lehrvideos

Weiterhin wurde abgefragt, wie viel Zeit die Studierenden wöchentlich für ergänzende E-Learning Inhalte (Videos bzw. Quizfragen) aufwenden möchten. Abbildung 7 zeigt, dass zwischen 16 und 30 Minuten eine realistische Zeitspanne bilden.

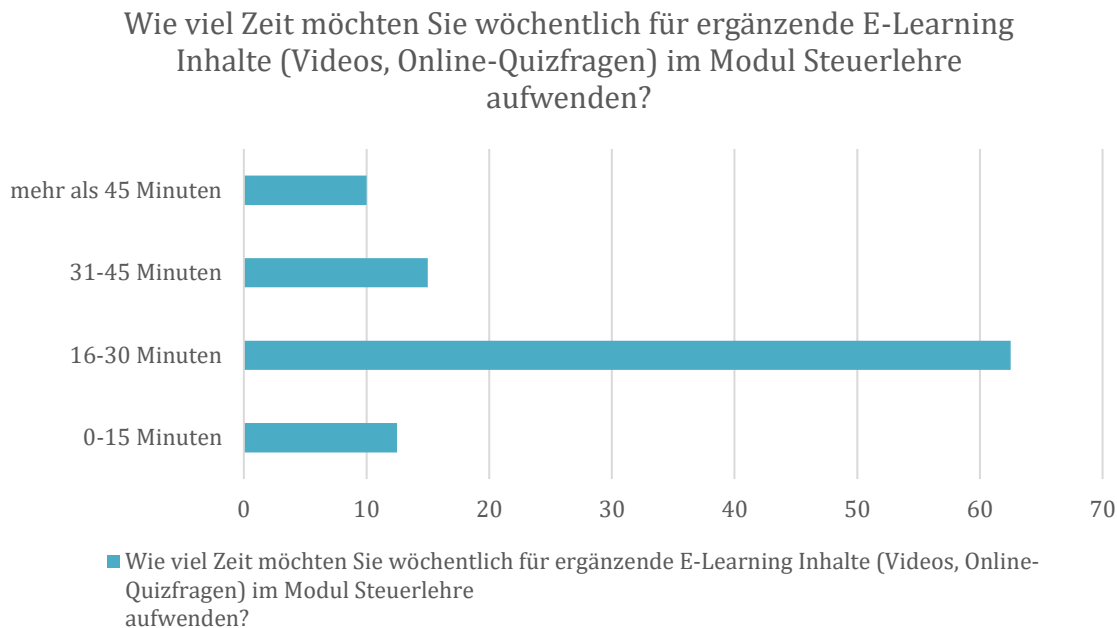


Abbildung 7: Zeitfaktor

Bei den offenen Fragen wurden die bisherigen Angebote als positiv bewertet. Die Studierenden wünschten sich, dass die Moodle Quizzes den Klausuraufgaben ähnlicher sind. Darüber hinaus wurde nach Videos bzw. nach Videos mit Rechenbeispielen gefragt. Zudem wurde das Tool Quizlet angesprochen.

Es wurde zudem erhoben, ob die Studierenden durch ein umfassenderes E-Learning stärker motiviert wären, sich auch außerhalb der Präsenzvorlesung mit den Inhalten zu beschäftigen. Dieser Aussage stimmen 69 % der Studierenden zu (Abbildung 8).

Durch ein umfassenderes E-Learning Angebot wäre ich stärker motiviert, mich auch außerhalb der Präsenzvorlesung mit den Inhalten zu beschäftigen.

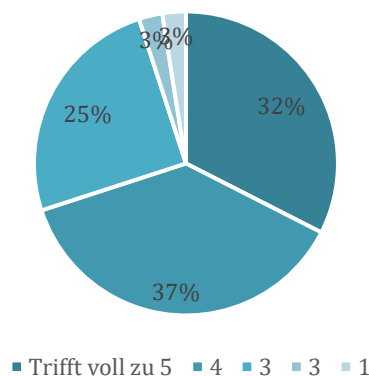


Abbildung 8: Motivation durch E-Learning

Auswertung Moodle: Im Wintersemester 2023/24 nutzten 92 von 165 Studierenden mindestens eines der sechs Quizzes. 37 Studierende haben jedes der sechs Quizzes einmal genutzt. Hinzuweisen ist auch hier darauf, dass die

Datengrundlage dadurch beeinflusst sein könnte, dass der Moodlekurs eine freie Einschreibung hat. Die Nutzung hat sich also deutlich erhöht.

Mittelbare Betroffene:

Feedback Tutor*innen: Die Onlinetutor*innen haben im Rahmen des laufenden Projekts jeweils Feedback zum Vorgehen und den einzelnen Ergebnissen (beispielsweise bereits erstellte Videos) gegeben.

Einbettung in die Professionelle Gemeinschaft:

Im Rahmen der kollegialen Beratung wurde empfohlen, Rechenaufgaben hinzuzunehmen sowie Aufgaben mit Begründungen zu integrieren; also über die Niveaustufen 1-2 nach Bloom 1971 bzw. erweitert nach Anderson & Krathwohl 2001 (kennen, verstehen) hinauszugehen und zum Niveau 3 (Anwendung) zu kommen.

Aktionsdurchführung (Aktionsteil)

Die Videoplanung wurde auf die angegebenen Charakteristika angepasst. Insbesondere wurde der Fokus auf die Wiederholung/Erklärung von schwierigen Sachverhalten und nicht auf ergänzende Inhalte gelegt. Weiterhin war in der Evaluation die Vermutung bestätigt wurde, dass die Lehrende selbst zu hören und idealerweise auch zu sehen sein sollte. In den Varianten „animierte Powerpoint“ und „Video Storytelling“ ist die Lehrende zu hören. In der Variante Mashup Video soll die Lehrende auch zu sehen sein.

Abbildung 9 zeigt die aktuelle Videoliste. Dabei sind die gewünschten Videos zur Berechnung kursiv gesetzt, weil diese für eine angedachte weitere Projektphase 4 berücksichtigt werden sollen.

Kapitel 1:	• Welche Rechtsquellen gibt es im Steuerrecht?
Kapitel 2:	• Wie werden Vorsorgeaufwendungen im Steuerrecht berücksichtigt? • Wie funktioniert der Verlustabzug bei der Einkommensteuer? • Welche Werbungskosten gibt es bei Einkünften aus Vermietung und Verpachtung? • Welche Werbungskosten gibt es bei Einkünften aus nichtselbstständiger Arbeit? • <i>Wie wird die Einkommensteuer berechnet?</i>
Kapitel 3:	• Wer ist unbeschränkt körperschaftsteuerpflichtig? • Wie funktioniert der Verlustabzug bei der Körperschaftsteuer? • <i>Wie wird die Körperschaftsteuer berechnet?</i>
Kapitel 4:	• Wie funktioniert der Verlustabzug bei der Gewerbesteuer? • Wie ist der Zusammenhang zwischen der Einkommensteuer und der Gewerbesteuer? • <i>Wie wird die Gewerbesteuer berechnet?</i>
Kapitel 5:	• Wann ist ein Umsatz steuerbar? • Unentgeltliche Lieferungen und sonstige Leistungen in der Umsatzsteuer • Welche Steuerbefreiungen gibt es im Umsatzsteuerrecht? • Steuerfreie Umsätze mit und ohne Optionsrecht
Kapitel 6:	• Wie werden die persönlichen Freibeträge und Steuerklassen bei der Erbschaft- und Schenkungsteuer bestimmt? • Wie funktioniert die Härtefallregelung?

Abbildung 9: Videoliste

In der Evaluation und in persönlichen Anfragen wurde nach einem Tool ähnlich Quizlet, Anki etc. gefragt. Auch in der kollegialen Beratung wurde auf die Möglichkeit der Nutzung von Karteikartentools hingewiesen. Daraufhin wurde gemeinsam nach einem Tool recherchiert, das einerseits funktional ist, aber andererseits die Anforderungen im Hinblick auf den Datenschutz erfüllt. Gefunden wurde das System THM.Cards. Das Programm ermöglicht, neben einigen anderen Funktionen, insbesondere das Lernen mit Karteikarten (siehe dazu Adam et al., 2019). Das Tool wurde bereits mit den Inhalten des Moduls gefüllt und kann voraussichtlich im Wintersemester 2024/25 an den Start gehen. Um den Einstieg zu erleichtern, haben die Onlinetutor*innen eine Anleitung für die Studierenden erstellt.

In der Evaluation wurde gewünscht, dass das Quiz anspruchsvoller bzw. klausurnäher ausgestaltet werden sollte. Dieser Wunsch konnte aktuell aus Kapazitätsgründen noch nicht umgesetzt werden. Aktuell sind die Quizfragen hauptsächlich für das Lernzielniveau 1-2 (kennen, verstehen) konzipiert. Das Lernzielniveau 3 „anwenden“ ist zwar klausurrelevant, wird derzeit aber lediglich vereinzelt in den Quizfragen berücksichtigt. Zu berücksichtigen ist, dass Fragen mit dem Lernzielniveau „anwenden“ häufig Berechnungen beinhalten und/oder Begründungen erfordern, die nicht oder nur eingeschränkt direkt über Moodle ausgewertet werden können.

5.4 Diskussion der Ergebnisse und Ausblick

Im Hinblick auf die eingangs genannten Ziele weisen Indikatoren darauf hin, dass die Lernmotivation durch die E-Learning Angebote steigt. Dies zeigt sich zum einen in den Evaluationen, in denen sich das Interesse der Studierenden an E-Learning-Angeboten weiter erhöht hat. Zudem kann bereits nach wenigen Semestern beobachtet werden, dass die E-Learning-Angebote zunehmend stärker genutzt werden. Dies hat auch Auswirkungen auf die Rückmeldungen zum Modul im Gesamten. Unabhängig von den projektspezifischen Fragen wird aus den Evaluationen deutlich, dass sich auch die weiteren Rückmeldungen positiv entwickeln. Dies geht einher mit Ergebnissen aus weiteren Studien zu Blended Learning Formaten an Hochschulen für Angewandte Wissenschaften (siehe dazu bspw. Federau, 2023, S. 395).

Aus Basis der Evaluation der Studierenden lässt sich zudem ein Bezug zu den Studien von Dressler et al. und Dressler & Rachfall herleiten. Diese haben herausgearbeitet, dass ein Erfolgsfaktor von Blended Learning Veranstaltungen in einer klaren Struktur liegt, die aber durch die Lehrperson motivierend umgesetzt werden muss. Anhand der übergeordneten Fragen zur Lehrveranstaltung wurde sichtbar, dass die Studierenden die Struktur im Wintersemester 23/24 bereits als deutlich klarer empfanden als im Sommersemester 2023. So wurde die Aussage „Die Vorlesung ist gut organisiert“ im Wintersemester von 81,4 % der Studierenden mit „trifft voll zu“ und von 18,6% der Studierenden mit „trifft zu“ bewertet. Im Sommersemester 2023 gaben nur 64,7 % „trifft voll zu“ und 17,6 % „trifft zu“ an. Der inhaltliche Aufbau der Vorlesung ist logisch und gut nachvollziehbar. Im Wintersemester bestätigten bereits 65,1 % der Studierenden, dass dies „voll zutrifft“ sowie 31,3 %, dass dies „zutrifft“. Im Sommersemester gaben nur 50 % der Studierenden „trifft voll zu“ an; 29,34% der Studierenden gaben „trifft zu“ an.

Die Ergebnisse zu den Videopräferenzen gehen einher mit der Literatur. Es wird betont, dass kurze Lehrvideos verwendet werden sollten (Hsin & Cigas, 2013, S. 253-259; Alksne 2016, S. 513), um ein höheres Engagement in der Beschäftigung mit den Videos zu erzielen. Die Literatur empfiehlt sogar eine maximale Dauer der Videos von 6 Minuten (Guo et al., 2014, S. 47). Nach Ergebnissen von Handke werden Videos mit einer Dauer von maximal 2 Minuten zu 98 % vollständig geschaut. Bei Videos mit einer Dauer von maximal 20 Minuten werden durchschnittlich nur 60 % des Videos abgeschlossen (Handke, 2023, S. 189). Der Sichtbarkeit der erklärenden Personen wird eine geringe Rolle zugesprochen. Lediglich das in einer der Studien betrachtete Alter der erklärenden Person war wichtig. So führte die Videoerklärung von einer aus Perspektive der Studierenden älteren Person zu einer höheren wahrgenommenen Qualität der Instruktion und zu einem höheren Lernerfolg als eine Videoerklärung durch Personen im ähnlichen Alter, z. B. andere Studierende (siehe Übersicht in Findeisen et al. 2019 mit Verweis auf weitere Studien). Im Rahmen des Videoprojekt wurden zwar wesentliche Arbeiten an den Videos von den Onlinetutorinnen vorgenommen, das Einsprechen erfolgt aber jeweils durch die, im Vergleich zu den Studierenden ältere, Dozentin.

Studien zeigen darüber hinaus, dass ein positiver Effekt zwischen der Interaktivität der Videos (z. B. anschließende Wiedergabe, Transfer, Übungsaufgabe) und dem Lernerfolg besteht (Findeisen et al., 2019 mit Verweis auf weitere Studien). Interaktive Videos führen zu höherer Aufmerksamkeit, begünstigen das selbstregulierte Lernen, fördern die Selbstständigkeit sowie die Reflexion (Cattaneo & Sauli 2017, S. 11 ff.). Eine weiterführende Studie sollte sich demnach aus damit beschäftigen, ob auch im entwickelten Szenario die Videos, die interaktive Elemente enthalten, als hilfreicher empfunden werden.

Die von Eisl et al. genutzten Kriterien (1) Personalisierung des Lernprozesses, die (2) Förderung der Selbstregulation, z. B. über eine Möglichkeit der Fortschrittsüberwachung im Hinblick auf die Lernziele, die (3) Möglichkeit Feedback im Hinblick auf die Lernziele zu erlangen und (4) eine Lernumgebung für den Austausch (Eisl et al., 2023, S. 365), wurden auch im Rahmen dieses Kurses betrachtet und führten ähnlich der Ergebnisse von Eisl et al. nicht nur zu einer hohen Zufriedenheit mit den Materialien, sondern zu einer höheren Zufriedenheit mit dem Modul insgesamt.

Mit Blick auf die durchgeführten Quizzes, dürften die Studierenden hinsichtlich ihres Lernbedarfs sensibilisiert und in der Lage sein eigenverantwortlich ihren Lernbedarf und das zugehörige Lerntempo zu bestimmen.

Eine passende Erweiterung der Nutzung von Quizfragen als formatives Feedback zur Unterstützung des Lernprozesses könnten E-Klausuren sein (siehe dazu Dressler et al. 2018, S. 352). Zudem wäre es möglich, die bisher freizunutzbaren Quizfragen um klausurähnliche Fragen zu ergänzen, die nur zu einem bestimmten Zeitpunkt freigeschaltet werden, zum Beispiel während einer der letzten Veranstaltungen des Semesters. Letzteres würde dazu führen, dass der Stressfaktor einer Prüfung simuliert werden könnte.

Auf Basis der letzten Reflexionsphase liegen bereits weitere Aktionsideen vor. Insbesondere sollen

- zusätzlich Videos zu Berechnungen erstellt werden,
- Quizfragen auf höherem Niveau und mit Rechenbeispielen entwickelt werden und
- die vorbereiteten Karteikarten im Wintersemester 2024/25 freigeschaltet werden und in die Erprobung gehen.

Zudem soll der im Sommersemester gestartete Videoeinsatz evaluiert werden. Sobald nach dem Pilotprojekt ein weiteres Ausrollen der Blended Learning Konzepte auf die weiteren Module erfolgt, sollte insbesondere konkreter erhoben werden, welche Videoformate für die Studierenden besonders motivierend und hilfreich sind. Dabei soll es einerseits um die unterschiedlichen Arten der technischen Umsetzung (animierte Powerpoint, Scribble Video, Mashup Video) und andererseits um die unterschiedlichen didaktischen Elemente (Interaktivität, Betonung von Lernzielen, Nutzung von Storytelling) gehen.

Das Gesamtergebnis der ersten drei Phasen kann durch folgende Aufwärtsspirale (Abbildung 10) abgebildet werden:

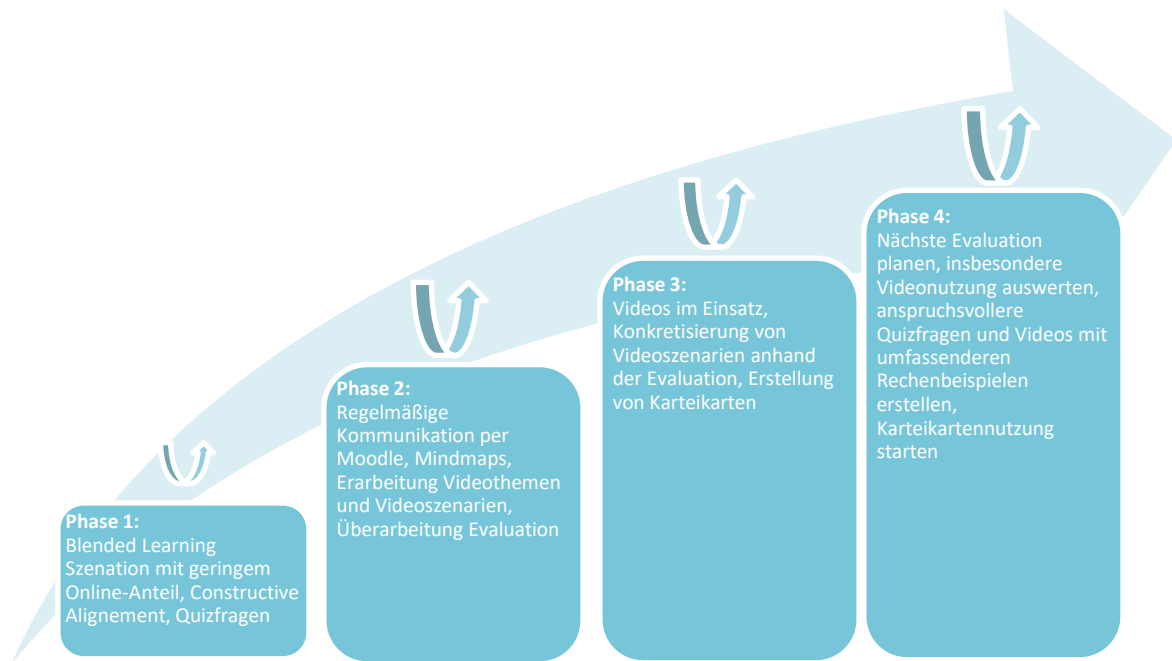


Abbildung 10: Aufwärtsspirale

6 Handlungsempfehlungen zur Umsetzung

Im Rahmen des Projekts haben sich praktische Handlungsempfehlungen zur Umsetzung herauskristallisiert:

Modulkonzept verdeutlichen: In der Präsenzveranstaltung selbst Zeit für die Erläuterung des Modulkonzepts einplanen. Das Konzept und die Erwartungshaltung sollten regelmäßig kommuniziert werden, idealerweise in der Veranstaltung und per zusätzlicher Moodle Nachricht. Die Erwartungshaltung zur Nutzung der Inhalte klar kommunizieren (z. B. Kategorie notwendig für die nächste Vorlesung oder individuelle Wiederholung und Vertiefung).

Zeit einplanen: Die Abstimmung der Vielzahl von Akteuren bindet Zeitressourcen. Das Vorgehen beinhaltet viele Schleifen. Dem Konzept ist zudem immanent, dass andere Ideen aufkommen, als vorab erwartet wurden, die aber ausgewertet und gegebenenfalls umgesetzt werden.

Offenheit und Vertrauensbasis: Da alle Akteure Anmerkungen einbringen können, ist es wichtig, eine gewisse Offenheit für neue Ideen mitzubringen. Im Konzept wird jede aufgekommene Idee im Rahmen der Informationssammlung ausgewertet, interpretiert und gegebenenfalls getestet. Die aktive Ideeneinbringung kann allerdings nicht erzwungen werden. Vielmehr ist es wichtig eine Vertrauensbasis aufzubauen, durch die Ideen niedrigschwellig eingebracht werden können. Hilfreich ist es, ausdrücklich und wiederholt zu kommunizieren, dass Freiräume bestehen und Vorschläge willkommen sind.

Rolle der Tutor*innen stärken: Die Onlinetutor*innen werden laufend über das Review der Projekte qualifiziert. Zudem haben sie an weiteren Schulungen teilgenommen. Die Präsenztutor*innen durchlaufen an der Technischen Hochschule Mittelhessen verpflichtend eine Didaktik-Schulung. Es wird ausdrücklich an die aktuelle Studierendengruppe und die Tutor*innen kommuniziert, dass Feedback erwünscht ist und auch über die Tutor*innen als Vertrauenspersonen anonym weitergegeben werden darf.

Weiterbildungsbedarf bedenken: Da nicht alle Projektergebnisse vorab feststehen, sollte ebenfalls ein Puffer für den Weiterbildungsbedarf eingeplant werden. Ein solcher liegt bei der Projektleitung. Aber insbesondere auch die in das Projekt eingebundenen Tutor*innen müssen in die Weiterbildung soweit wie möglich einbezogen werden.

Langfristige Vorteile des Projekts sehen: Der Einbezug der vielfältigen Akteure ist aufwendig. Die Aufbereitung des Materials und die Abstimmung zwischen Präsenz- und Onlinelehre kostet viel Zeit. Bei allen Restriktionen sollten aber die Vorteile und der hohe Wertbeitrag gesehen werden. Durch das Arbeiten mit den Tutor*innen ergeben sich beispielsweise viele Erkenntnisse zum Modul, die dem Verantwortlichen andernfalls nicht offenbart werden. Projekte mit einer solchen Vielzahl von Akteuren sind in der Praxis eher selten durchführbar. Vor diesem Hintergrund sollte das Augenmerk darauf gelegt werden, wie die Erkenntnisse auf andere Situationen (hier Module) übertragen werden können. Zweckmäßig ist diese Berücksichtigung durch langfristige Projektpläne.

7 Zusammenfassung

„Digifit Steuerlehre“ ist ein Blended Learning Pilotprojekt für Module aus den Bereichen Externes Rechnungswesen, Steuerberatung und Wirtschaftsprüfung, angesiedelt im Fachbereich Wirtschaft der Technischen Hochschule Mittelhessen. Das Lehrprojekt wurde zunächst für das Modul Steuerlehre durchgeführt und beinhaltet drei Kern-Zielsetzungen, damit die Studierenden eigenständig für ihren Erfolg einstehen und der unterschiedliche Lernstand der heterogenen Lerngruppen auf einer höheren Ebene angeglichen wird: 1. Erhöhung der Lernmotivation, 2. Förderung der Fähigkeit, den eigenen Lernerfolg einzuschätzen und 3. Eigenverantwortliche Bestimmung des individuellen Lerntempos sowie Lernbedarfs. Mit der Zielerreichung geht die Erstellung von fakultativen E-Learning-Materialien einher. Eine ergänzende Zielsetzung ist die Qualifikation der Tutor*innen.

Der Einsatz der E-Learning-Materialien wurde positiv beurteilt. Damit wurden weitere Anreize geschaffen, sich auch außerhalb der Veranstaltung mit dem Modul zu beschäftigen. Die Selbststudienzeit wurde durch die wöchentlichen E-Mails strukturiert und durch die Erweiterung der Lernmöglichkeiten attraktiver.

Ein großer Teil der Studierenden hat die E-Learning Angebote genutzt. Dies wurde zunächst anhand der Quizfragen evaluiert. Sie fühlten sich zudem besser in der Lage den eigenen Lernerfolg einzuschätzen.

Zum Stand Juni 2024 wurden die ersten drei Phasen des Projekts abgeschlossen. Auf Basis des zyklischen Prozesses der Aktionsforschung wurde aber deutlich, dass mindestens ein weiterer vierter Zyklus notwendig ist.

Die erzielten Erkenntnisse lassen sich auf andere Module der genannten Schwerpunkte übertragen. Da nicht immer gewährleistet werden kann, dass für Projekte entsprechende Tutor*innen zur Verfügung stehen, sollen alternative Möglichkeiten zur zwischenzeitlichen Datensammlung getestet werden. In Moodle ist es beispielsweise möglich, anonym Fragen einzustellen, die für Videos genutzt werden können. Eine weitere Möglichkeit wäre die Nutzung von Teaching Analysis Polls oder die Nutzung von Audience Response Systems.

Wir bedanken uns bei den Tutor*innen Zara Bayir und Carina Kraft, sowie Anja Baumann und Andreas Deublein vom Fachbereich Wirtschaft der Technischen Hochschule Mittelhessen.

Literaturverzeichnis

- Adam, C.; Kamutzki, D.; Quibeldey-Cirkel, K. (2019): ARSnova.cards: How to Sustainably Promote Mobile Learning with a Flashcard App. In: 11th International Conference on Education and New Learning Technologies. EDULEARN19 Proceedings 2019. Palma: 2067-2077.
- Alksne, L (2016): How to produce video lectures to engage students and deliver the maximum amount of information. In: Society Integration Education Proceedings of the International Scientific Conference (Band II). Rezekne: 503-516.
- Altrichter, H.; Posch, P.; Spann, H. (2018): Lehrerinnen und Lehrer erforschen ihren Unterricht. Unterrichtsentwicklung und Unterrichtsevaluation durch Aktionsforschung. 5. Aufl. Verlag Julius Klinkhardt, Bad Heilbrunn.
- Anderson L.W.; Krathwohl D. R. (2001): A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives. Pearson Education Group, Boston.
- Bernard, R. M.; Borokhovski, E.; Schmid, R. F.; Tamin, R. M.; Abrami, P. (2014): A meta-analysis of blended learning and technology use in higher education: From the general to the applied. Journal of Computing in Higher Education, 26 (1): 87-122.
- Biggs J.; Tang C. (2011): Teaching for quality learning at university. What the student does. 4th ed. Open University Press, Maidenhead.
- Bloom, B. S. (1972): Taxonomie von Lernzielen im kognitiven Bereich. 4. Aufl. Beltz Verlag, Weinheim und Basel.
- Bock, S.; Heiny J. (2019): Lehrentwicklung im Zusammenspiel von Hochschuldidaktik, Lehrenden, Studierenden und Aktionsforschung. In: Jenert, T.; Reinmann, G.; Schmohl, T. (Hrsg.), Hochschulbildungsforschung. Springer, Wiesbaden: 185-205.
- Cattaneo, A; Sauli F (2017): Die Integration interaktiver Videos in didaktische Szenarien. <https://www.sfuvet.swiss/sites/default/files/2022-07/linee%20guida-tedesco-digitale.pdf>. Abgerufen am 02.07.2024.
- Cendon, E. (2015): Praxisforschung. In: Thematische Berichte der wissenschaftlichen Begleitung des Bund-Länder-Wettbewerbs "Aufstieg durch Bildung: offene Hochschulen") - URN: urn:nbn:de:0111-pedocs-129838 - DOI: 10.25656/01:12983.
- Dressler, S.; Förster-Trallo, D.; Rachfall, T. (2016): Improved Learning Performance Based on a Flipped Classroom Concept – A Case Study about improved learning performance due to a digitally enhanced course structure. In: EDULEARN16 Proceedings 2016. Barcelona: 570-575.
- Dressler, S.; Rachfall, T. (2016): Improved learning performance based on a flipped classroom concept—A case study. In: Proceedings of the 2nd International Conference on Higher Education Advances 2016. Valencia: 219-227.
- Dressler, S., Rachfall, T.; Förster-Tallo, D.; Kapanen, A. (2017): Improved Learning Performance based on a flipped classroom concept — Results after three consecutive courses. In Proceedings of IAC-TLEI 2017. Budapest: 249-258.
- Dressler, S.; Rachfall, T.; Dressler, S.; Förster-Trallo, D. (2018): E-Tests als moderne Prüfungsform in den Wirtschaftswissenschaften: Wie wirken sich E-Tests auf die psychische Belastung und die akademische Unehrlichkeit von Studierenden aus? In: CARF Luzern 2018, Controlling.Accounting.Risiko.Finzen, Konferenzband. Zug: 352-361.
- Dressler, S.; Rachfall, T. (2020): Improved Learning Performance Based on a Flipped Classroom Concept – A Case Study based on the Course Introduction to Management Accounting for Business Engineers, in: Walker, Z.; Tan, D.; Koh, N. K. (Hrsg.), Flipped Classroom with Diverse Learners. Springer, Singapur: 183-202.
- Eisl, C.; Freudenthaler-Mayrhofer, D.; Wagner, G. (2023): Controlling-Lehre im Spannungsfeld zwischen Selbststeuerung und Anleitung – ein Fallbeispiel aus dem Digital Learning mittels Lernvideos, In: CARF Luzern 2023, Controlling.Accounting.Risiko.Finzen, Konferenzband. Zug: 361-378.

- Federau, M. (2023): Grundlagen der Wirtschaft und finanziellen Unternehmensführung: Ein feedbackbasierter Flipped Classroom-Ansatz. In: CARF Luzern 2023, Controlling.Accounting.Risiko.Finanzen, Konferenzband. Zug: 2023: 385-396.
- Findeisen, S.; Horn, S.; Seifried, J. (2019): Lernen durch Videos – Empirische Befunde zur Gestaltung von Erklärungsvideos, Medienpädagogik. Zeitschrift für Theorie und Praxis der Medienbildung, 19 (10): 16-36.
- Frank, A.; Fröhlich, M.; Lahm, S. (2011): Zwischenauswertung im Semester: Lehrveranstaltungen gemeinsam verändern. Zeitschrift für Hochschulentwicklung, 6(3): 310-318.
- Franz-Özdemir, M.; Reimann, J.; Wessel, K. (2019): Teaching Analysis Poll (TAP) – Konzept und Umsetzung einer aktuellen Methode an der Schnittstelle von Evaluation und Lehrentwicklung. In Berendt, B; Szczyrba, A. Fleischmann, N. Schaper & J. Wildt (Hrsg.), Neues Handbuch Hochschullehre. DUZ, Berlin: 37-64.
- Fox, M.; Martin, P.; Green, G. (2007): Doing practitioner research. Sage Publications, London.
- Gardner, C. & Thielen, S. (2015). Didaktische Prinzipien für E-Learning. wvb Wiss. Verlag, Berlin.
- Garrison, D. R.; Kanuka, H. (2004): Blended learning: Uncovering its transformative potential in higher education. Internet and Higher Education, 7 (2): 95-105.
- Grassl, R. (2016): Veränderung durch Fragen. Die sokratische Methode als Instrument der Aktionsforschung. In Cendon, E.; Mörtz, A.; Pellert, A. (Hrsg.), Theorie und Praxis verzahnen. Lebenslanges Lernen an Hochschulen. Münster, Waxmann: 51-66.
- Guo, P., J.; Kim, J.; Rubin, R. (2014): How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. In: Proceedings of the first ACM conference on Learning @ scale conference. Atlanta: 41-50.
- Handke, J. (2023): Handbuch Hochschullehre digital, 4. Aufl. Nomos Verlag, Baden-Baden.
- Hattie, J. (2012): Visible learning for teachers. Maximizing impact on learning. 1. publ. London u.a.: Routledge.
- Hattie, J. (2023): Visible Learning: The Sequel: A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement. London, Routledge.
- Heiny, J.; Weyers, S. (2017): Wie (effektiv) nutzen Studierende die Selbstlern- und Präsenzphasen einer Flipped Classroom Veranstaltung?. Das Hochschulwesen, 65 (3): 79-87.
- Hsin, W. J.; Cigas, J (2013): Short videos improve student learning in online education. Journal of Computing Sciences in Colleges, 28(5): 253-259.
- Kerres, M. (2018). Mediendidaktik. Konzeption und Entwicklung digitaler Lernangebote De Gruyter Studium, 5. Aufl. De Gruyter, Berlin, Boston.
- KIT (2024): Blended Learning-Modelle. <https://www.zml.kit.edu/blendedlearning-modelle-methoden.php>. Abgerufen am 22.4.2024.
- Lewin, K. (1951): Field theory in social science. Selected theoretical papers. Harper, New York.
- McCarthy, J. (2016): Reflections on a flipped classroom in first year higher education, in: Issues in Educational Research, 26(2): 332-350.
- Means, B. M.; Toyama, Y.; Murphy, R.; Baki, M. (2013): The Effectiveness of Online and Blended Learning: A Meta-Analysis of the Empirical Literature. Teachers College Record 115(3): 1-47.
- Merkt, M. (2014): Hochschuldidaktik, Organisationsentwicklung und Begleitforschung an der Hochschule Magdeburg-Stendal – ein integrativer Ansatz. In Egger, R.; Kiendl-Wendner, D.; Pöllinger, M. (Hrsg.), Hochschuldidaktische Weiterbildung an Fachhochschulen. Durchführung – Ergebnisse – Perspektiven. Springer Fachmedien, Wiesbaden: 27-48.
- Müller, A.; Schmidt-Hertha, B. (2009): Prüfungen als Lernchance: Sinn, Ziele und Formen von Hochschulprüfungen. Zeitschrift für Hochschulentwicklung. 4 (1): 23-35.
- Rachfall, T.; Förster-Trallo, D.; Zimbelmann, E. E. (2019): Students and their requirements – a survey about using e-learning elements in part-time studies; In: EDULEARN19 Proceedings 2019. Palma: 5531-5537.
- Reinmann, G. (2008). Blended Learning in der Lehrerbildung. Grundlagen für die Konzeption innovativer Lernumgebungen, 3. Aufl. Pabst, Lengerich.

- Spector, J.M. (2014): Conceptualizing the emerging field of smart learning environments. Smart Learn. Environ. SpringerOpen, <https://doi.org/10.1186/s40561-014-0002-7> . Abgerufen am 03.07.2024.
- Vo, M. H.; Zhu, C.; Diep, A. N. (2017): The effect of blended learning on student performance at course-level in higher education: A meta-analysis. *Studies In Educational Evaluation*, 53(5):17-28.
- Warneke, D. (2007): Aktionsforschung und Praxisbezug in der DaF-Lehrerbildung. University Press, Kassel.
- Weitzel, J.; Timmann, A.; Franz-Özdemir, M.; Grunert, C.; Reimann, J.; Sachse, A., Salzmann, S.; Weiss, P.; Wessel, K. (2022): Dialogische Feedback- und Evaluationsverfahren für die Hochschulentwicklung. In Leben, N.; Reinecke, K. & Sonntag, U. (Hrsg.), *Hochschullehre als Gemeinschaftsaufgabe. Akteur:innen und Fachkulturen in der lernenden Organisation* wbv Publikation, Bielefeld: 107-119.
- Zierer, K. (2014): Kernbotschaften aus John Hatties “Visible Learning”. Konrad-Adenauer Stiftung e.V., Sankt Augustin.
- Zumbach, J.; Astleitner, H. (2016): *Effektives Lehren an der Hochschule*. Kohlhammer, Stuttgart.

Transfer anders gedacht – Entrepreneurship Education als Schul-Hochschul-Kooperation in der 3rd Mission

Geplantes Lehrprojekt

Prof. Dr. Tim Kampe

Hochschule Bielefeld – University of Applied Sciences and Arts, Center for Entrepreneurship, Bielefeld, E-Mail: tim.kampe@hsbi.de

Prof. Dr. Ingo Ballschmieter

Hochschule Bielefeld – University of Applied Sciences and Arts, Center for Entrepreneurship, Bielefeld, E-Mail: ingo.ballschmieter@hsbi.de

Abstract

Der Beitrag beschreibt ein intendiertes kooperatives Lehrprojekt zwischen der Marienschule der Ursulinen, einem Bielefelder Gymnasium, und dem Center for Entrepreneurship der Hochschule Bielefeld. Schülerinnen und Schüler der Jahrgangsstufe 10 werden im Rahmen eines neu einzuführenden Wahlpflichtfachs Wirtschaft für den Bereich Unternehmensgründung sensibilisiert, indem sie über einen Zeitraum von etwa 8 bis 12 Wochen in Hochschule und Schule die Grundlagen des Lean-Startup-Ansatzes vermittelt bekommen und diesen parallel auf eine eigene Gründungsidee anwenden. Das Modul schließt mit dem Pitch des in Gruppen entwickelten eigenen Geschäftsmodells vor einer Expertenjury in der Hochschule. Neben dem Einblick in selbständige Tätigkeit durch eigene Erfahrungen im Sinne des Action Learning gewinnen die Schülerinnen und Schüler auch einen ersten Einblick in eine Hochschule, die sich hier zugleich auch als möglicher Studienort vorstellt.

1 Einführung und Grundlagen

Neben Forschung und Lehre hat sich der Transfer in den letzten beiden Dekaden als wesentliche und zunehmend gleichberechtigte Säule in vielen Universitäten und Hochschulen etabliert, weshalb dieser auch als „3rd Mission“ bezeichnet wird. Dabei hat sich für den Begriff des Transfers und die hierunter zu subsumierenden Aktivitäten von Hochschulen noch kein einheitliches Verständnis herausgebildet. Gleichwohl kann unter Rückgriff auf das Centrum für Hochschulentwicklung (CHE) Transfer verstanden werden als „Aktivitäten, die einen direkten wechselseitigen Austausch mit Wirtschaft, Gesellschaft und (lokaler) Politik ermöglichen.“ (Roessler, 2024) Darunter fallen unter anderem Aspekte der Wissenschaftskommunikation, Kooperationsprojekte mit etablierten Unternehmen aber auch der Wissens- und Technologietransfer durch Ausgründungen bzw. die Gründungsförderung von Hochschulen.

Wesentliche Grundlage der Gründungsförderung an Hochschulen ist Entrepreneurship bzw. Entrepreneurship Education. Entrepreneurship bezeichnet im Allgemeinen die Lehre vom Suchen, Finden und Nutzen unternehmerischer Gelegenheiten und kann im weiteren Sinne als Unternehmertum umschrieben werden (Kampe & Wach, 2021, S. 108). Die Entrepreneurship Education, eine Übersetzung in die deutsche Sprache hat sich trotz mehrfacher Versuche nicht durchsetzen können (Halbfas & Liszt-Rohlf, 2019, S. 9), zielt dabei nicht nur auf die Vermittlung gründungsrelevanter Kompetenzen, sondern auch auf die dafür erforderlichen Bildungsstrukturen und die Auswirkungen des gründungsbezogenen Handelns auf die Gesellschaft (Halbfas & Liszt-Rohlf, 2019, S. 18).

Ein Ziel der Entrepreneurship Education ist es, für die selbständige bzw. unternehmerische Tätigkeit zu sensibilisieren und sie damit als Alternative zur rein abhängigen Beschäftigung sichtbar zu machen (Kampe & Wach, 2021, S. 109). Denkt man die o.g. Definition des Transfers in diesem Kontext nun weiter, bedeutet dies, dass Entrepreneurship Education im Sinne der 3rd Mission einer Hochschule nicht an den Grenzen der Hochschule haltmachen muss, sondern vielmehr auch weitere gesellschaftliche Gruppen umfassen kann. Eine solche – aus Sicht der Hochschule wesentliche – Gruppe ist die der Schülerinnen und Schüler weiterführender Schulen im unmittelbaren Umfeld der Hochschule. Vor dem Hintergrund der demographischen Entwicklung sowie einer Ausdifferenzierung auch des tertiären Bildungssystems stehen Hochschulen zunehmend im Wettbewerb um (potentielle) Studierende.

Schul- und Hochschulbildung sind in Deutschland überwiegend Ländersache (BMBF - Bundesministerium für Bildung und Forschung, 2024). In Nordrhein-Westfalen, dem mit rd. 18 Mio. Einwohnern bevölkerungsreichstem Bundesland Deutschlands, verfolgt die Landesregierung das Ziel, die ökonomische Bildung an allen Schulen als Bestandteil der Allgemeinbildung gezielt zu stärken (MSB NRW - Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2024). Dazu zählt auch die sukzessive Einführung bzw. Neuausrichtung eines Wahlpflichtfaches Wirtschaft in den Jahrgangsstufen 9 und 10 der Gymnasien des Landes.

Der vorliegende Beitrag beschreibt eine vereinbarte Zusammenarbeit zwischen der Fachkonferenz bzw. Fachgruppe Wirtschaft-Politik/Sozialwissenschaften der Marienschule der Ursulinen, staatlich anerkanntes Gymnasium für Mädchen und Jungen in Bielefeld-Schildesche, und dem Center for Entrepreneurship (CFE) der Hochschule Bielefeld (HSBI).

2 Zielsetzungen des Lehrprojekts und didaktische Konzeption

Die Marienschule der Ursulinen versteht sich in ihrem Leitbild als eine „katholische Schule mit ökumenischem Profil“ (Marienschule der Ursulinen, 2024). Auf der Grundlage des Leitbilds der Marienschule hat die Fachgruppe Wirtschaft-Politik/Sozialwissenschaften das Ziel, Schülerinnen und Schüler darin zu unterstützen, verantwortlich für sich und die Gemeinschaft zu handeln sowie Toleranz gegenüber Vielfalt zu entwickeln. Im Wahlpflichtfach Wirtschaft sollen insbesondere Mündigkeit und soziale Verantwortung in den Fokus der unterrichtlichen Arbeit gerückt werden. Dies geschieht durch den Erwerb von fachlicher und sozialer Kompetenz in verschiedenen mehrwöchigen Modulen bzw. thematischen Blöcken, u.a. zu Digitalisierung oder Internationalisierung von Unternehmen und Märkten und eben auch Unternehmensgründung.

Hinsichtlich der für eine mögliche Unternehmensgründung in der Entrepreneurship Education zu vermittelnden Kompetenzen existieren unterschiedliche Ansätze und Initiativen (eine Übersicht geben z.B. Grewe & Brahm, 2019, S. 139ff.). Für das Lehrprojekt bzw. allgemein im Bereich Qualifizierung des CFE wird der „Entrepreneurship Competence Framework (EntreComp)“ des Joint Research Center (JRC) der Europäischen Kommission herangezogen (Bacigalupo et al. 2016, S. 11), da dieser ein umfassendes, aktuelles und empirisch fundiertes Modell darstellt (Lindner 2018, S. 411; Grewe & Brahm 2019, S. 142). Wie in Abbildung 1 wiedergegeben, kann Entrepreneurial Competence nach dem EntreComp-Ansatz in drei Kompetenzbereiche mit jeweils fünf Kompetenzen eingeteilt werden, die auch zu der Zielerfüllung seitens der Marienschule beitragen.



Abbildung 1: Kompetenzen des Entrepreneurships (Kampe & Wach, 2021, S. 111, unter Rekurs auf Bacigalupo et al. 2016, S. 1).

Trotz vielversprechender Ansätze in einzelnen Bundesländern gilt die Entrepreneurship Education – im Gegensatz zur Situation im Hochschulsystem – im allgemeinbildenden Schulsystem in Deutschland im internationalen Vergleich nach wie vor als eher unterrepräsentiert (Bijedić 2019, S. 23). Hinsichtlich der konkreten Ausgestaltung der Entrepreneurship Education können nach Lackéus (2015, S. 8) zwei gegensätzliche Grundrichtungen definiert werden (Abbildung 2). In der praxisorientierten Ausbildung durch Entrepreneurship durchlaufen die Teilnehmenden einen tatsächlichen unternehmerischen Prozess, d.h. einen vornehmlich erfahrungsbasierten Ansatz. Die theorieorientierte Ausbildung über Entrepreneurship stellt einen inhaltsreichen und eher auf die Vermittlung von Wissen und Kompetenzen abzielenden Ansatz dar, wie er häufiger an Hochschulen gelehrt wird. Im internationalen Vergleich hat sich im schulischen Kontext eher ein Mittelweg etabliert, der Ausbildung für Entrepreneurship zum Inhalt hat und dem daher auch vorliegend gefolgt werden soll.

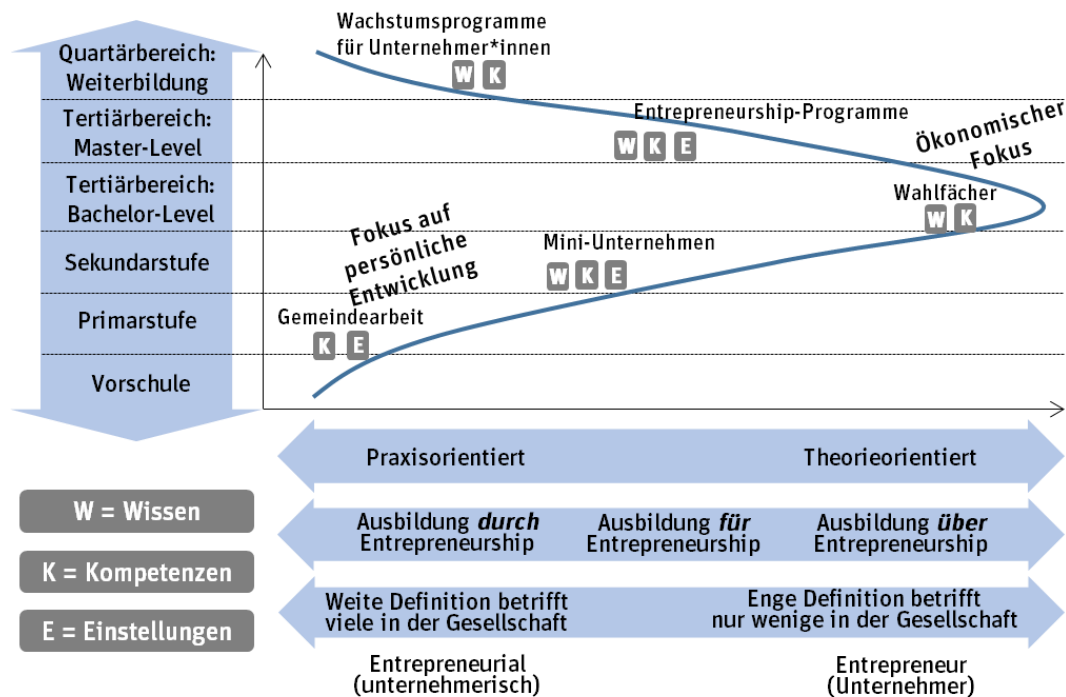


Abbildung 2: Ausgestaltung und Fokusse der Entrepreneurship Education im Bildungsbereich (Kampe & Wach, 2021, S. 112 unter Rekurs auf Lackéus, 2015, S. 8).

3 Geplante Ausgestaltung der Unterrichtseinheit

Die geplante Unterrichtseinheit soll über einen Zeitraum von 8 bis 12 Wochen laufen. Am Beginn steht eine Einführung im Sinne eines Kick-Offs an der Hochschule (3.1). Die eigentliche Entwicklung von Geschäftsidee und Geschäftsmodell, d.h. die Umsetzung des Lean-Startup-Ansatzes und des Business Model Canvas, erfolgt dann parallel zum Unterrichtsverlauf und unter Begleitung der Lehrkräfte (3.2). Die Unterrichtseinheit schließt mit einem finalen Pitch vor einer Expertenjury in der Hochschule (3.3).

3.1 Kick-Off an der Hochschule

Das Lehrprojekt beginnt mit einer Exkursion der Schülerinnen und Schüler in die Hochschule. Hier werden sie in dem typischen Lernsetting einer Hochschule für angewandte Wissenschaften begrüßt und erhalten zwei akademische Doppelstunden seminaristischen Unterricht im Co-Teaching durch die Verfasser, die gleichzeitig gemeinsam auch das Center for Entrepreneurship leiten. Diese Kick-Off Veranstaltung zielt auf die Vermittlung des grundlegenden Wissens und der Basiskompetenzen für die Auseinandersetzung mit der Unternehmensgründung und bildet somit den theoretischen Rahmen für die folgenden Wochen. Insbesondere folgende Aspekte sollen dabei thematisiert werden:

- Grundlagen: Begriff des Entrepreneurships, Entrepreneurship als Berufswahl, Mythen über Entrepreneurre
- Erkennen von Geschäftsmöglichkeiten: Arten, Entwicklung und Evaluation von Geschäftsideen
- Überleitung von Geschäftsideen in Geschäftsmodelle: Business Model Canvas als Goldstandard
- Lean Startup zur weiteren Vorgehensweise

Die Veranstaltung endet mit der Vorstellung eines konkreten Gründungsprojekts, das zu diesem Zeitpunkt am Inkubator-Programm des CFE teilnimmt. Das Gründungsteam soll als Role Model für die Schülerinnen und Schüler fungieren und Lust auf das eigene Gründungsprojekt machen.

3.2 Lehrbegleitende Geschäftsmodellentwicklung – Lean Startup & Business Model Canvas

In den nun folgenden 8 bis 12 Wochen sollen die Schülerinnen und Schüler in Gruppen von 3-4 Personen eine eigene Geschäftsidee entwickeln und testen. Dies muss nicht zwangsläufig eine innovative Idee sein, die in einem skalierbaren Geschäftsmodell resultiert, sondern kann auch eine weniger originelle Idee wie die Eröffnung eines veganen Restaurants in Bielefeld zum Gegenstand haben. Wichtig ist vielmehr, dass die Schülerinnen und Schüler sich grundsätzlich an die Vorgehensweise des Lean-Startup-Ansatzes nach Eric Ries (2011) halten.

Die grundsätzliche Idee des Lean-Startup-Ansatzes ist es, möglichst früh mit einer minimalen Version des Produkts oder der Dienstleistung („build“) an den Markt bzw. an potentielle Kunden zu gehen und deren Nutzungsverhalten zu erheben („measure“), um von dem dort erlangten Feedback zu lernen („learn“) und die Geschäftsidee bzw. das Geschäftsmodell kontinuierlich weiter zu verbessern (neuer Build-Measure-Learn-Zyklus). Diese minimale Version des Produkts bzw. der Dienstleistung („minimal viable product“ – MVP) kann dabei von der bloßen verbalen Beschreibung der Geschäftsidee bis hin zu wirklichen Prototypen reichen. Die grundlegende Vorgehensweise sowie mögliche Beispiele für MVP-Experimente sind in Abbildung 3 zusammengefasst.

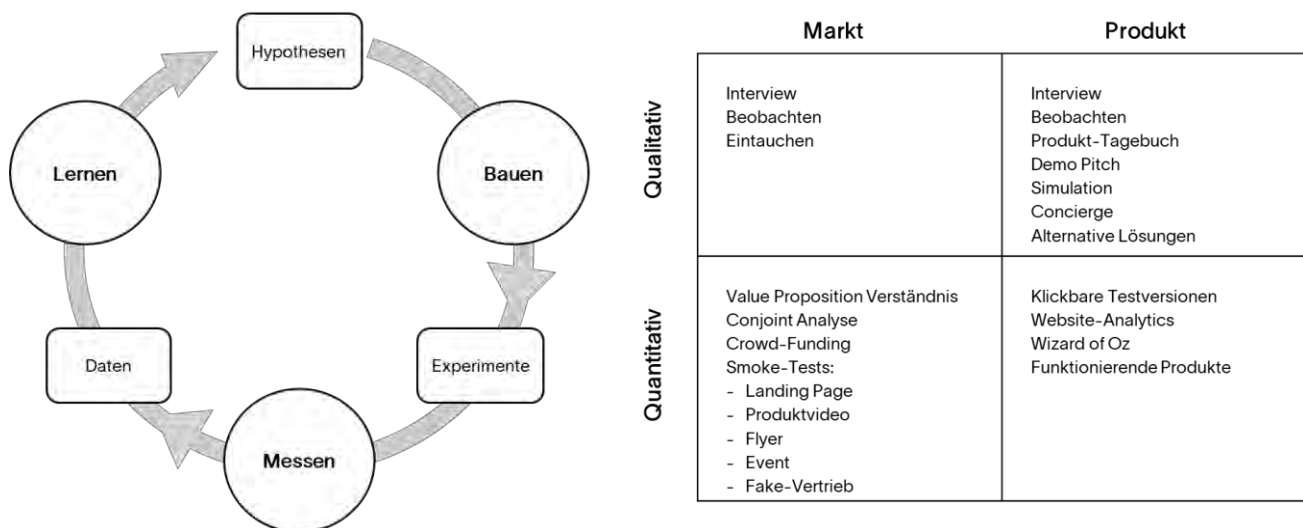


Abbildung 3: Vorgehensweise beim Lean-Startup-Ansatz und MVP-Experimente (Freiling & Harima, 2019, S. 196, S. 189)

Konkret bezogen auf das Beispiel des veganen Restaurants könnte die Vorgehensweise so aussehen, dass die erste zu testende Hypothese lauten könnte: „In Bielefeld besteht Bedarf für ein veganes Restaurant“. Für diese zunächst marktbezogene Fragestellung könnte sich eine eher qualitative Vorgehensweise wie die Befragung (Interview) von Passanten in der Fußgängerzone anbieten. Ist die Hypothese bestätigt, können erste Überlegungen zur Ausgestaltung des Geschäftsmodells angestellt werden. Diese Überlegungen können durch Beobachten von Kunden in der Gastronomie oder weitere Befragungen fundiert bzw. verifiziert werden. Das Konzept könnte bspw. einer Gruppe von Interessierten (Familie & Freunde) im Sinne eines Demo Pitch präsentiert werden usw.

Die Vorgehensweise des Lean-Startup-Ansatzes und insbesondere auch der MVP-Experimente wird parallel im Unterricht weiter vertieft. Hierzu werden den Lehrkräften entsprechende Materialien durch die Hochschule zur Verfügung gestellt. Diese beruhen zum Teil auf Standardlehrbüchern (wie z.B. Barringer & Ireland, 2019; Flueglister et al., 2019; Freiling & Harima, 2019; Grichnik et al, 2017; Kollmann, 2022) werden aber auch zum Teil hierfür extra erstellt.

Am Ende sollen die Schülerinnen und Schüler ein validiertes Geschäftsmodell zu einer vorher selbst gewählten Geschäftsidee vorweisen und präsentieren können. Die Erwartung ist auch, dass hierzu das Instrument des

Business Model Canvas nach Osterwalder und Pigneur (2011) genutzt wird, wie es in Abbildung 4 wiedergegeben ist.

Schlüsselpartner ▪ Wer sind die strategischen Partner? ▪ Wer sind die strategischen Lieferanten? ▪ Welche Ressourcen werden von Partnern eingebracht? ▪ Welche Aufgaben erbringen die Partner?	Schlüsselaktivitäten Welche Kernaufgaben erfordern Nutzenversprechen, Vertriebswege, Kundenbeziehungen u. Einnahmequellen? Schlüsselressourcen Welche Schlüsselressourcen erfordern Nutzenversprechen, Vertriebswege, Kundenbeziehungen u. Einnahmequellen?	Value Proposition ▪ Welcher Nutzen wird für die Kunden gestiftet? ▪ Welches Problem wird gelöst? ▪ Welche Produkt-/Dienstleistungsbündel werden angeboten? ▪ Welche Kundenbedürfnisse werden befriedigt?	Kundenbeziehung ▪ Welche Art der Beziehungen bestehen? ▪ Wie teuer sind sie? Kanäle ▪ Über welche Vertriebswege werden Kunden erreicht? ▪ Integration der Wege? ▪ Kosten/Funktion?	Kundensegmente ▪ Für wen wird Nutzen gestiftet? ▪ Wer sind die wichtigsten Kunden?
Kostenstruktur ▪ Was sind die bedeutendsten Kosten im Geschäftsmodell? ▪ Welche Schlüsselressourcen sind am teuersten? ▪ Welche Schlüsselaufgaben sind am teuersten?			Einnahmequellen ▪ Für welchen Nutzen sind die Kunden bereit zu zahlen? ▪ Wie bevorzugen Kunden zu zahlen? ▪ Wieviel trägt jede Einnahmequelle zu den Gesamteinnahmen bei?	

Abbildung 4: Business Model Canvas nach Osterwalder und Pigneur (2011)

3.3 „Höhle der Löwen“ – Final Pitch an der Hochschule

Die Geschäftsidee, die Vorgehensweise bei der Überführung in ein Geschäftsmodell sowie das Geschäftsmodell selbst in der oben dargestellten Struktur des Business Model Canvas sind durch die Teams zum Abschluss der Unterrichtseinheit zu präsentieren. Diese auch als „Final Pitch“ bezeichnete Präsentation von jeweils max. 10 Minuten erfolgt vor einer Expertenjury und findet wieder in der Hochschule statt. Das Setting erinnert dabei bewusst ein wenig an das vielen Schülerinnen und Schülern aus Fernsehen bzw. Streaming bekannte Format „Höhle der Löwen“ das deutschen Senders Vox, bei dem echte Gründungsteams um echtes Kapital pitchten.

Die Jury besteht aus Mitarbeitenden des CFE (Leitung und/oder wissenschaftliche Mitarbeitende), Gründerinnen und Gründern aus dem aktuellen Portfolio des Inkubator-Programms sowie den beteiligten Lehrkräften der Marienschule. Die Präsentation ist gleichzeitig die schulische Prüfungsleistung. Präsentationskompetenzen sind auskunftsgemäß bereits ausreichend bei den Schülerinnen und Schülern vorhanden.

Es wird darüber hinaus nachgedacht, den besten Pitch – wie bei Pitch-Veranstaltungen im Gründungsökosystem sonst auch üblich – mit einem Preisgeld zu versehen. Eine Voranfrage bei der Fördergesellschaft der Hochschule Bielefeld, das Preisgeld zu finanzieren, ist auf positive Resonanz gestoßen.

4 Fazit und Ausblick

Die beteiligten Lehrer und Professoren sind sich einig darüber, dass die Kooperation für beide Seiten eine Bereicherung darstellt. Ob diese Bereicherung auch von den Schülerinnen und Schülern so empfunden wird, wird sich dann bei erstmaliger Durchführung in der Jahrgangsstufe 10 im Schuljahr 2025/26 zeigen. Zumindest stößt die Kooperation bereits jetzt auf großes Interesse seitens der Schülerinnen und Schüler, die im Mai 2024 ihre Wahlfähcher festlegen mussten: von 126 Schülerinnen und Schülern haben sich mehr als die Hälfte für das neue Wahlpflichtfach Wirtschaft entschieden.

Sofern das Pilotprojekt mit der Marienschule aus Sicht aller Beteiligten erfolgreich verläuft, soll das Konzept bzw. die Zusammenarbeit auch anderen weiterführenden Schulen angeboten werden.

Literaturverzeichnis

- Bacigalupo, M.; Kampylis, P.; Punie, Y.; Van den Brande, G. (2016): EntreComp: The Entrepreneurship Competence Framework. Publication Office of the European Union, Luxembourg.
- Barringer, B. R.; Ireland, R. D. (2019): Entrepreneurship: Successfully Launching New Ventures. 6., globale Aufl., Pearson, Harlow.
- Bijedić, T. (2019): Begriffliche und disziplinäre Genese der Entrepreneurship Education. In: Bijedić, T; Ebbers, I.; Halbfas, B. (Hrsg.), Entrepreneurship Education: Begriff – Theorie – Verständnis. Springer Gabler, Wiesbaden, 21-40.
- BMBF (2024): Bildung: Zusammenarbeit von Bund und Ländern. https://www.bmbf.de/bmbf/de/bildung/zusammenarbeit-von-bund-und-laendern/zusammenarbeit-von-bund-und-laendern_node.html. Abgerufen am 10.07.2024.
- Flueglister, U.; Fust, A.; Müller, C.; Müller, S.; Zellweger, T. (2019): Entrepreneurship: Modelle - Umsetzung - Perspektiven. 5., überarb. Aufl., Springer Gabler, Wiesbaden.
- Freiling, J; Harima, J. (2019): Entrepreneurship: Gründung und Skalierung von Startups. Springer Gabler, Wiesbaden.
- Grewe, U; Brahm, T. (2019): Kompetenzorientierung in der Entrepreneurship Education. Wie kann die unternehmerische Kompetenz operationalisiert werden? In: Bijedić, T; Ebbers, I.; Halbfas, B. (Hrsg.), Entrepreneurship Education: Begriff – Theorie – Verständnis. Springer Gabler, Wiesbaden, 133–150.
- Grichnik, D.; Brettel, M.; Koropp, C.; Mauer, R. (2017): Entrepreneurship: Unternehmerisches Denken, Entscheiden und Handeln in innovativen und technologieorientierten Unternehmungen. 2., überarb. Aufl., Schäffer-Poeschel, Stuttgart.
- Halbfas, B.; Liszt-Rohlf, V. (2019): Entwicklungslinien und Perspektiven der Entrepreneurship Education – eine Analyse von Definitionen. In: Bijedić, T; Ebbers, I.; Halbfas, B. (Hrsg.), Entrepreneurship Education: Begriff – Theorie – Verständnis. Springer Gabler, Wiesbaden, 3-20.
- Kampe, T.; Wach, B. (2021): Entrepreneurship Education in dualen Studiengängen: Zielsetzungen, Praxisperspektiven und Studienformate. Duales Studium: Personal in Hochschule und Betrieb gemeinsam entwickeln, 3(2): 107-119.
- Kollmann, T. (2022): Digital Entrepreneurship: Grundlagen der Unternehmensgründung in der Digitalen Wirtschaft. 8., überarb. u. erw. Aufl., Springer Gabler, Wiesbaden
- Lackéus, M. (2015): Entrepreneurship in Education – what, why, when, how. Entrepreneurship360 Background Paper. LEED Division of the OECD.
- Lindner, J. (2018): Entrepreneurship Education. In: Faltin, G. (Hrsg.), Handbuch Entrepreneurship. Springer Gabler, Wiesbaden, 407-423
- Marienschule der Ursulinen (2024): Marienschule Bielefeld – eine katholische Schule mit ökumenischem Profil. <https://marienschule-bielefeld.de/schulleben/programm/>. Abgerufen am 10.07.2024.
- MSB NRW (2024): Schulfach Wirtschaft. <https://www.schulministerium.nrw/schulfach-wirtschaft>. Abgerufen am 10.07.2024.
- Osterwalder, A.; Pigneur, Y. (2011): Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus, Frankfurt a. M. und New York
- Ries, E. (2011): The Lean Startup: How Constant Innovation Creates Radically Successful Businesses. Portfolio Penguin, London.
- Roessler, I. (2024): DatenCHECK 3/2024: In der Third-Mission-Förderung erfolgreiche Hochschulen. <https://hochschuldaten.ch.de/third-mission-foerderung/>. Abgerufen am 10.07.2024.

Lern-App BWL Champion: Konzeption und Anwendung in der digitalen Lehre/Prüfung

Laufendes Lehrprojekt

Prof. Dr. Jörn Littkemann

FernUniversität in Hagen, Lehrstuhl für Betriebswirtschaftslehre, insbesondere Unternehmensrechnung und Controlling, Universitätsstr. 11, D-58084 Hagen, E-Mail: joern.littkemann@fernuni-hagen.de

Kristopher Pantani, M. Sc.

FernUniversität in Hagen, Lehrstuhl für Betriebswirtschaftslehre, insbesondere Unternehmensrechnung und Controlling, Universitätsstr. 11, D-58084 Hagen, E-Mail: kristopher.pantani@fernuni-hagen.de

Abstract

Lern-Apps sind in der wirtschaftswissenschaftlichen Hochschullehre noch wenig verbreitet. Dieses überrascht vor dem Hintergrund der zunehmenden Digitalisierung der Lehre und Prüfung in den Hochschulen. In diesem Beitrag wird mit der Lern-App BWL Champion ein Beispiel vorgestellt, welche von Lehrenden konzipiert wurde, hochschulübergreifend genutzt werden kann und laufend weiterentwickelt wird. Zudem werden praktische Anwendungen in der modulbezogenen Lehre und Prüfung aufgezeigt. Als theoretische Grundlage dient der Gamification-Ansatz, der zunehmend Akzeptanz in der Hochschullehre findet. Mittlerweile nutzen ca. ein Dutzend Hochschulen und ihre Lehrenden die Lern-App zur Unterstützung und Ergänzung ihrer Lehrveranstaltungen und Prüfungsaktivitäten.

1 Einführung

Im Jahr 2022 gab es rund zwei Millionen verschiedene Apps im Apple App Store, von denen 99 % von Drittanbietern stammen. Dies stellt nahezu eine Vervierfachung seit 2015 dar. Die Entwicklung ist umso bemerkenswerter, wenn sie mit der Anzahl der verkauften Geräte verglichen wird, die im selben Zeitraum nur moderat gewachsen ist (Apple Inc., 2022; Caminade & Wartburg, 2022, S. 8). Ein Blick auf die Top 10 der weltweit führenden Bildungs-Apps zeigt, dass sich diese mit dem Erlernen von Sprachen befassen und überwiegend für Kinder und Schüler konzipiert worden sind (Blacker, 2023). Doch neben Schülern arbeitet und lernt auch die junge Studierendengeneration zunehmend mobil mit Tablets und Smartphones. Auch Lehrende bieten, nicht zuletzt begünstigt durch die Corona-Pandemie, vermehrt digitale Formate in Lehre und Studium an. Diese Formate sind jedoch häufig web- und weniger app-basiert, obwohl sich Apps besonders gut für das mobile Lernen mit Smartphones und Tablets eignen, während webbasierte Lösungen eher für das Lernen vom Laptop oder Desktop gedacht sind. Zudem stehen viele Hochschulen app-basierten Lösungen aufgrund von Risiken bezüglich IT-Sicherheit und Datenschutz sowie der damit oftmals verbundenen Abhängigkeiten von den US-Mediengiganten wie Apple und Google skeptisch gegenüber. Aufgrund dieser Zögerlichkeit seitens der Hochschulen springen in jüngster Vergangenheit Schul-Apps in diese Lücke und erweitern ihre Lernangebote um Lehrinhalte für Studierende, vielfach anpassbar nach Hochschulen, Studiengängen und -jahren, jedoch nicht durchgängig in der erforderlichen Qualität.

Vor diesem Hintergrund ist die Lern-App *BWL Champion* eine der ersten Apps, die vollständig von Lehrenden selbst konzipiert und entwickelt wurde und die sich primär an potenzielle Studierende der Betriebswirtschaftslehre sowie angrenzender Fachdisziplinen richtet. Ferner können Lehrende und Dozierende, unabhängig von ihrer Hochschule, die App zur Vermittlung und Überprüfung von Lerninhalten einsetzen. Die App enthält das für ein betriebswirtschaftliches Studium notwendige Basiswissen. Sie ist sowohl im Apple App Store (iOS) als auch im Google Play Store (Android) verfügbar.

Im Folgenden werden zuerst die theoretischen Grundlagen zum digitalen Lernen vorgestellt. Dazu eignet sich besonders der Gamification-Ansatz, der das spielerische Lernen beschreibt. Zur Einordnung von *BWL Champion* wird darauffolgend der Stand der Forschung dargelegt. Anschließend wird die Konzeption von *BWL Champion* dargestellt und auf die Einsatzmöglichkeiten sowohl für Lehrende als auch für Studierende eingegangen und im Kontext des digitalen Lernens diskutiert. Darauffolgend wird ein Anwendungsbeispiel von *BWL Champion* im Rahmen eines Bachelorwahlpflichtmoduls illustriert. Der Beitrag schließt mit einem Ausblick auf zukünftige Entwicklungen.

2 Theoretische Grundlagen

2.1 Gamification

Gamification ist ein Konzept, das keine einheitliche Definition besitzt, da es in unterschiedlichen Bereichen wie im Bildungs- oder Arbeitswesen angewendet wird. Historisch betrachtet, argumentieren Experten, dass eine Art von Gamification bereits im frühen 20. Jahrhundert mit der Pfadfinderbewegung begann, bei der Abzeichen für Leistungen vergeben wurden (Johnston, 2015, S. 509; Zainuddin et al., 2024, S. 71). Andere sehen Gamification als ein Phänomen, das untrennbar mit IT-Elementen und der Bewältigung von Aufgaben durch technologische Unterstützung verbunden ist.

Faust (2021) unterscheidet zwischen zwei Definitionen von Gamification: der strukturierten und der interaktiven Sichtweise (Bueno Muñoz et al., 2023, S. 6). Die strukturierte Sichtweise basiert auf der Struktur der Aufgabe oder des Systems, während die interaktive Sichtweise das Ergebnis der Mensch-Aufgabe-Interaktion darstellt und daher von der individuellen Wahrnehmung abhängt (Faust, 2021, S. 9).

Zainuddin et al. (2024) unterscheiden neben der Gamification außerdem zwischen spielbasiertem Lernen und ernsthaften Spielen („Serious Games“). Das Hauptziel beim spielbasierten Lernen ist es, Problemlösungs- und kritisches

Denkvermögen durch unterhaltsame Aktivitäten zu fördern, während Serious Games für ernste Zwecke wie das Gesundheitswesen oder das Militär entwickelt wurden (Zainuddin et al., 2024, S. 69–70).

Grundsätzlich kann daher unter Gamification die Nutzung von Spielmechaniken, wie Punkte oder Bestenlisten, in nicht-spielerischen Kontexten verstanden werden, um das Engagement und die Motivation der Beteiligten zu steigern und deren Ziele zu erreichen (Deterding et al., 2011, S. 1; Warszta, 2023, S. 29; Zainuddin et al., 2024, S. 67). Es beschreibt daher die Anwendung von Spielmechanismen wie Punkte, Bestenlisten, und Aufgaben in nicht spielerischen Kontexten, um Einzelpersonen zu engagieren und zu motivieren, ihre Ziele zu erreichen. Im Bildungskontext bedeutet Gamification daher, Spieldesigns anzuwenden, um neue, häufig kognitive Fähigkeiten zu erwerben (Hou, 2023, S. 2). Die aktuellen Entwicklungen lassen erahnen, dass die Zukunft der Gamification im Bereich des Remote- oder Blended-Game-Based Learning liegen.

Daneben müssen Gamification-Inhalte sicherstellen, dass Lernaktivitäten nicht nur oberflächlich, sondern tiefgründig und nachhaltig sind. Dazu eignen sich die vier Säulen des Lernens, die wie folgt charakterisiert werden:

- Aktives Lernen: geistigste Beteiligung während des Lernprozesses,
- Engagement: Fähigkeit, während des Lernens aufmerksam zu bleiben,
- Bedeutsames Lernen: Anwendung des Gelernten,
- Soziale Interaktion: Interaktionen mit Lehrenden aber auch gleichgesinnten Lernenden (Hirsh-Pasek et al., 2015, S. 4; Tamis-LeMonda et al., 2014, S. 121–122).

Werden die vier Säulen des Lernens, von denen jedoch die ersten beiden überwiegend intrinsischer Art sind, erfolgreich mit Gamification-Inhalten unterstützt, so kann der Lernerfolg und insbesondere die Lernmotivation gesteigert werden (Hirsh-Pasek et al., 2015, S. 22). Die weitere Diskussion der vier Säulen und insbesondere die Umsetzung in BWL Champion erfolgt im nächsten Kapitel. Der nächste Abschnitt widmet sich der aktuellen Forschung im Bereich von Lern-Apps und insbesondere Lern-Apps in der Hochschulbildung.

2.2 Stand der Forschung

Der aktuelle Stand der Forschung zum mobilen Lernen zeigt, dass Lern-Apps häufig nicht ausgereift sind und vor allem im Bereich der Betriebswirtschaftslehre zu wenig Beachtung finden, trotz dem großen Potenzials, das sie den Studierenden zur effektiven Lernunterstützung bieten.

Eine Studie aus dem Jahre 2012 bestätigt diese Beobachtung, indem sie aufzeigt, dass zu diesem Zeitpunkt weniger als 2 % der im englischsprachigen Raum verfügbaren Bildungs-Apps auf die Betriebswirtschaftslehre ausgerichtet waren. Die Diskrepanz zwischen dem vorhandenen Potenzial und der tatsächlichen Nutzung von Lern-Apps in der Betriebswirtschaftslehre unterstreicht die Notwendigkeit für weitere Entwicklungen und Forschungen in diesem Bereich. Dabei stellt der Großteil der angebotenen Apps, heute wie damals, allgemeine Lernhilfen wie Karteikarten dar oder bietet mehrere Disziplinen im Rahmen der Drill-And-Practice-Methodik an (Highfield et al., 2013, S. 382; Papadakis et al., 2018, S. 145). Diese Methodik wird bereits seit 200 Jahren im amerikanischen Bildungssystem genutzt und kann auf heutige Bildungssysteme übertragen werden. Sie beschreibt eine Vorgehensweise, bei der durch wiederholte Übung, oft routinemäßig und mechanisch, Wissen und Fertigkeiten gefestigt werden sollen. Sie kann daher mit dem deutschen Begriff der Fleißarbeit verglichen werden. Besonders geeignet ist die Methode beim Erwerb grundlegender Fertigkeiten wie Lesen oder Fremdsprachen, sie kann jedoch auch in jedem anderen Bildungsbereich angewendet werden (Ornstein, 1990, S. 112–114).

Neben diesen allgemeinen Studien zu Bildungs-Apps, konnte eine Studie im deutschsprachigen Raum identifiziert werden, die betriebswirtschaftliche Inhalte innerhalb einer Bildungs-App abbildet. In dieser Studie konnten Voshaar et al. (2023) validieren, dass ihre App zu signifikant besseren Ergebnissen in der Endprüfung führt, wenn Studierende diese ernsthaft nutzen. Abgeleitet bestätigt dies außerdem, dass die zweite Säule des Lernens „Engagement“

zu Lernerfolgen führen kann. Da die App jedoch ausschließlich auf Studierende der Universität Bremen beschränkt ist, lässt sich hier sowohl eine Forschungs- als auch Marktlücke finden.

Daneben konnte eine Studie identifiziert werden, die Lern-Apps für Schüler untersuchte. Dabei kamen Papadakis et al. (2018) zu dem Ergebnis, dass der Großteil der Apps nicht entwicklungsgerecht ist und die Bedürfnisse der Kinder nicht erfüllt. Daraus ist abzuleiten, dass Apps der jeweiligen Benutzergruppe anzupassen sind.

Sowohl die Forschung als auch bestehende Lern-Apps konzentrieren sich hauptsächlich auf Schüler, während die Hochschulbildung nur wenig berücksichtigt wird und wenn, dann häufig nur hochschulintern. Insbesondere im wirtschaftswissenschaftlichen Bereich gibt es kaum Anwendungen. Diese Lücke soll durch die Einführung von **BWL Champion** geschlossen werden. Die App kann eine bedeutende Unterstützung für Studierende der Betriebswirtschaftslehre darstellen. Daher wird im Folgenden die Konzeption und der Anwendung der App dargestellt.

3 **BWL Champion**

3.1 Die Konzeption von **BWL Champion**

Die Lern-App **BWL Champion** wurde von dem von Prof. Dr. Jörn Littkemann mitgegründeten digitalen Bildungsunternehmen **APP Academic Product Partner GmbH** in Emsdetten entwickelt. Sie richtet sich vornehmlich an Studierende der Betriebswirtschaft sowie angrenzender Disziplinen und Studiengänge, aber auch an Schüler sowie an Weiterbildungsstudierende oder sonstige (Außer-)Berufstätige, die betriebswirtschaftliche Kenntnisse besitzen bzw. aufbauen möchten.

Die Lerninhalte stammen aus Zwecken der Qualitätssicherung ausschließlich von Professoren und können allgemein oder modulbezogen ausgestaltet sein. Die Basisversion der Lern-App ist kostenlos, die Premiumversion (ohne Werbung und mit ausführlichen Lösungshinweisen) ist zu studierendenfreundlichen Abonnentenpreisen¹ erhältlich.

Zur einmaligen Registrierung bei **BWL Champion** reichen die Angabe des frei wählbaren Nutzernamens und Passworts sowie der E-Mail-Adresse (empfohlen wird die Registrierung mit der Uni-Mailadresse) aus. Öffentlich sichtbar werden dabei nur der Nutzername, ein gewählter Avatar sowie ggf. eine gewählte Hochschule und der Nutzerstatus (z. B. Studierender).

Zum gemeinsamen Lernen können alle bekannten Personen (z. B. per Mail oder über soziale Medien) oder angefragte Nutzer der App eingeladen werden und bei Freundschaftsannahme der Freundesliste hinzugefügt werden (vgl. Abbildung 1).

¹ Ein Jahresabonnement beträgt derzeit 24 €.

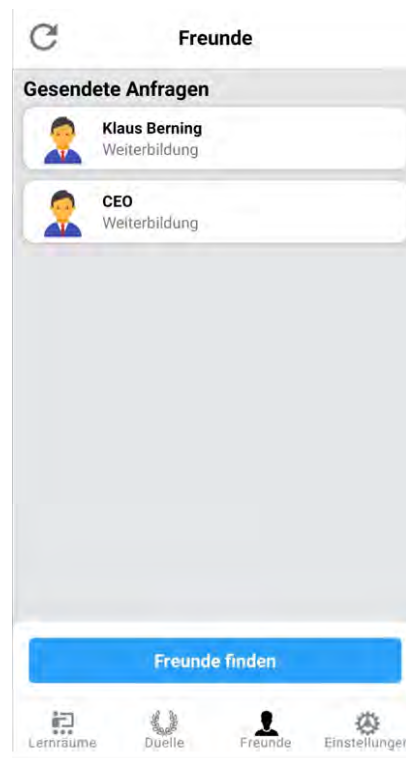


Abbildung 1: Freundschaftsanfrage

Nutzer können in Form von Duellen oder als Tests in sogenannten Lernräumen (= Modulen) lernen. Ein Duell beinhaltet 12 Fragen aus vier Modulen mit drei unterschiedlichen Schwierigkeitsgraden, wobei jede bzw. jeder der beiden am Duell Teilnehmenden zwei Module wählen darf. Dies fördert die soziale Interaktion, angelehnt an die vier Säulen des Lernens, was zu einer gesteigerten Lernmotivation und nachhaltigem Lernerfolg beitragen soll. Die App enthält derzeit über 4.000 Fragen und deckt das BWL-Basisstudium inhaltlich ab. Zu jeder Frage lassen sich ausführliche Lösungserläuterungen einsehen. Eine Liste der verwendeten Literatur befindet sich in den Einstellungen.

Innerhalb eines Lernraums sind Unterthemen zu finden, so dass sich ein Modul auch themenbezogen lernen lässt (vgl. Abbildung 2). Nutzer können alle Fragen, im Rahmen der in Kapitel 2 dargestellten Drill-And-Practice-Methodik, lernen oder einen Test simulieren, um so, wie von Hirsh-Pasek et al. (2015) vorgeschlagen, aktives Lernen zu fördern. Ein Probetest besteht aus 12 Zufallsfragen mit jeweils vier Fragen zu den drei Schwierigkeitsgraden. Freunde können in die einzelnen Lernräume eingeladen werden, um ein gemeinsames Lernen zu ermöglichen. Bei Bedarf können Lehrende hier den Lernfortschritt ihrer Studierenden überprüfen und ggf. Steuerungsmaßnahmen einleiten.

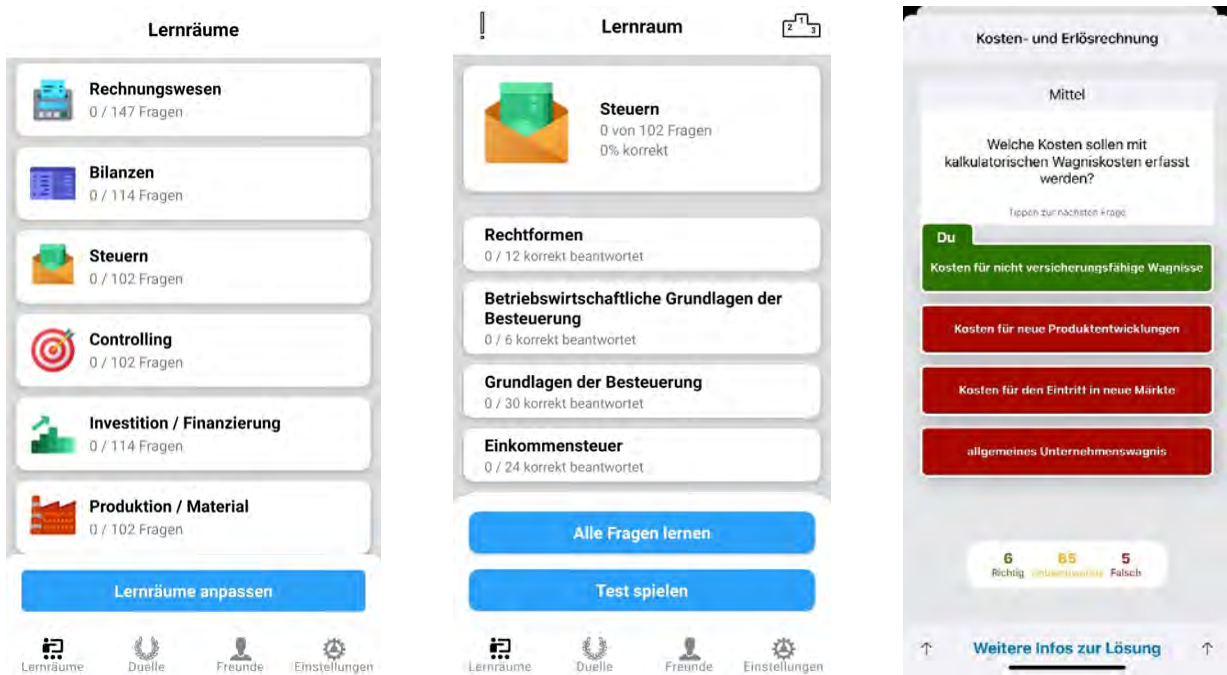


Abbildung 2: Lern-, Unterräume und Beispielfrage (Module)

Seit dem Sommersemester 2023 besteht für sich an der Lern-App beteiligende Hochschulen und deren Lehrenden die Möglichkeit, eigene Lernräume (Module) in unterschiedlichen Sprachen² in BWL Champion zu konzipieren und einzustellen. Dieses gebietet sich in den Fällen, in denen die Module der jeweiligen Hochschulen sich entweder formal oder inhaltlich stark von den Lehrinhalten der Standardlernräume in der Lern-App unterscheiden. Beispielsweise enthält der Lernraum Rechnungswesen die beiden Themengebiete Kostenrechnung und Buchhaltung. Bei einigen Hochschulen gehört die Kostenrechnung jedoch in das Modul Controlling und die Buchhaltung gemeinsam mit der Bilanzierung zum Modul externes Rechnungswesen. Unabhängig davon kann es zu Abweichungen aufgrund unterschiedlicher ECTS kommen, was dazu führt, dass bestimmte Lehrinhalte nicht oder in lediglich rudimentärer bzw. sehr ausführlicher Form behandelt werden.

Die von den Lehrenden selbst erstellten Lernräume erscheinen mit dem Logo der jeweiligen Hochschule in der App.³ Diese übernimmt im Regelfall die Abonnementgebühren für ihre Studierenden, die damit Zugriff auf die Premiumfunktionen aller Lernräume von BWL Champion haben.⁴

3.2 Anwendung und Evaluation von BWL Champion und Lehre und Prüfung

Im Wintersemester 2023/2024 wurde von dem Controlling-Lehrstuhl von Prof. Dr. Littkemann an der FernUniversität in Hagen eine Online-Übung zum Bachelorwahlpflichtmodul Instrumente des Controllings angeboten (Montag abends im ersten Teil des Wintersemesters). Die Übung bot die Möglichkeit, Kenntnisse in den Bereichen des Controllings zu vertiefen, ausgewählte Lerninhalte zu diskutieren und gezielt auf die bevorstehende Klausur vorzubereiten. Dies umfasste die Erörterung und Diskussion ausgewählter Lehrinhalte sowie die praktische Prüfungsvorbereitung durch das Üben und Lösen klausurähnlicher Aufgaben.

² Derzeit in deutscher und englischer Sprache.

³ Bislang haben Professoren folgender Hochschulen und Weiterbildungsinstitutionen eigene Lernräume konzipiert und in die Lern-App eingestellt: Bergische Universität Wuppertal, FernUniversität in Hagen, Hagener Institut für Managementstudien (HIMS), Hochschule für Polizei und öffentliche Verwaltung NRW, Institut für wirtschaftswissenschaftliche Forschung und Weiterbildung (IWW), Institut für wissenschaftliche Weiterbildung (FeUW) und Universität Trier.

⁴ Interessierte Lehrende, die sich an BWL Champion beteiligen und Teil der Community des digitalen/mobilen Lernens werden möchten, wenden sich bitte an: info@app-partner.com. <https://app-partner.com/>

Die Teilnahme an der Übung war freiwillig. Eine vorherige Anmeldung war nicht erforderlich. Um an den einzelnen Veranstaltungen teilzunehmen, brauchten die Studierenden lediglich die in der Moodle-Lernumgebung angegebenen Zoom-Links für die jeweiligen Termine anzuklicken. Zum Abschluss der Übung konnten die Studierenden durch das erfolgreiche Absolvieren eines freiwilligen Tests Bonuspunkte in Höhe von 5 % zur Verbesserung des Gesamtergebnisses der kommenden Klausur erreichen. Der Test wurde in der letzten Zoom-Veranstaltung mit dem Smartphone mittels QR-Code basierend auf Fragen der Controlling-Lernräume aus der Lern-App BWL Champion durchgeführt. Die relevanten Lernräume wurden den Studierenden zu Beginn der Eröffnungsveranstaltung bekanntgegeben. Zudem wurde der Ablauf des Tests erläutert und im Rahmen der Aufgaben der einzelnen Übungsveranstaltungen mehrmals geprobt.

Der Aufbau des Tests lehnte sich an den aus der Lern-App BWL Champion bekannten Probetest an:

- 12 Zufallsfragen aus den ausgewählten Lernräumen (jeweils drei leichte, mittlere und schwere);
- jeweils vier Antwortmöglichkeiten pro Frage, nur eine Antwort ist richtig;
- pro Frage 15 Sekunden Zeit für die Beantwortung;
- für jede richtig beantwortete Frage gibt es einen Punkt, die maximale Punktzahl ist 12;
- der Test ist bestanden, wenn mindestens sieben Punkte erreicht wurden.

Die Bonuspunkte für die kommende Klausur konnten die Studierenden erlangen, indem sie nacheinander zwei Tests mit unterschiedlichen Fragen aus den Lernräumen absolvierten. Wenn einer der beiden Tests bestanden wurde, erhielten sie 5 %-Bonuspunkte. Diese Punkte galten nur für die kommende Klausur und wurden dort automatisch berücksichtigt; sie waren nicht in kommende Semester übertragbar. Zur Teilnahme an den Tests mussten sich die Studierenden mit ihrer Matrikelnummer anmelden.

Nach der Durchführung der beiden Tests fand eine kurze Evaluation der Online-Übung und der Lern-App statt.⁵ Dabei beurteilten die knapp 60 Studierenden die Möglichkeit des Absolvierens von Prüfungen mittels der Lern-App BWL Champion mit großer Mehrheit als „sehr gut“.

Für zukünftige Online-Übungen gaben die Studierenden folgende Anregungen: Es sollten mehrere Termine mit kürzeren Zeiträumen angeboten werden, und die Anmeldung zu den Veranstaltungen sollte verpflichtend sein. Die Präsentationsfolien sollten bereitgestellt und die Veranstaltungen aufgezeichnet werden. Zudem sollte die Lern-App durch Rechenaufgaben ergänzt werden. Die Abonnentenpreise der Lern-App seien angemessen, allerdings äußerten die Studierenden den Wunsch, dass die Universität die Gebühren übernimmt. Abschließend konnten die Studierenden noch offene Fragen stellen.

4 Fazit und Ausblick

Die Vision von BWL Champion lautet „Studieren am Handy“! Sie zielt darauf ab, das Studium durch die Lern-App zu reformieren. Die App bietet zahlreiche Vorteile für Studierende und Lehrende. Lehrende können Lerninhalte individuell anpassen und gestalten, was eine effektivere Vermittlung des Stoffes ermöglicht. Gleichzeitig dient die App als Ergänzung oder Erweiterung der Lehrveranstaltungen, wodurch die Lehrenden entlastet werden. Studierende profitieren von der Möglichkeit sowohl individuell als auch in Gruppen zu lernen, was ihren unterschiedlichen Lernpräferenzen entgegenkommt. Außerdem können sie die App auch außerhalb der Hochschule nutzen, wodurch das außerhochschulische Lernen und die Vernetzung gefördert wird. Das hohe Weiterentwicklungspotenzial der App, etwa durch die Erweiterung um Prüfungsfunktionen oder die Kombination mit anderen digitalen Tools, macht die App zu einem innovativen Werkzeug zur Optimierung der Lehr- und Lernprozesse.

⁵ <https://evaluation-lernapp.limesurvey.net/994198>

Als Nachteile sind evtl. hohe Abhängigkeiten von Apple und Google, bspw. beim Datenschutz und zum Teil bei der Monetarisierung sowie zusätzliche Kosten für Studierende bzw. Lehrende/Hochschulen zu nennen. Letzter Punkt relativiert sich insofern, dass durch die Nutzung der Lern-App Kopierkosten und Ausgaben für Literatur gespart werden können. Zudem trägt das digitale/mobile Lernen zur ökologischen Nachhaltigkeit bei.

Bis das Ziel des „Studierens am Handy“ erreicht werden kann, ist es jedoch noch ein langer Weg. Folgende Weiterentwicklungen sind für die nahe Zukunft geplant:

- Neue Lernräume/Module mit Inhalten von Partnerhochschulen/-instituten,
- neue Lernräume/Module mit Inhalten betriebswirtschaftlicher Nebenfächer,
- neue Aufgabentypen: Einfachwahl- (nicht „nur“ 1 aus 4), Mehrfachwahlaufgaben, Begriffe, Zahlen, Freitext,
- englisch- (bzw. mehrsprachige) Versionen von BWL Champion,
- Lern-Apps für weitere wissenschaftliche Disziplinen,
- Weiterführung des Testens von Fragen aus BWL Champion in Lehrveranstaltungen und zu Prüfungszwecken (sowohl online als auch in Präsenz).

Zusätzlich gibt es die Möglichkeit, in der Zukunft digitale Abzeichen für Studierende oder Nutzer in die App zu integrieren. Laut Facey-Shaw et al. (2020) empfinden die Mehrheit der Studierenden Abzeichen als motivationsfördernd.

Insgesamt gesehen ist die App BWL Champion auf dem besten Wege, durch Gamification neue Lernanreize und Motivationsmöglichkeiten für Studierende der Betriebswirtschaftslehre und anderer Disziplinen zu schaffen. Der innovative Ansatz der App nutzt die Vorteile sowohl traditioneller als auch digitaler Lernmethoden, um das Lernen flexibler, interaktiver und ansprechender zu gestalten. Besonders hervorzuheben ist außerdem, dass BWL Champion eine Marktlücke schließt, indem es eine hochschulübergreifende und speziell auf betriebswirtschaftliche Inhalte fokussierte Lern-App anbietet. Die Kombination aus Individualisierung, Flexibilität und spielerischen Elementen stellt sicher, dass die App einen erheblichen Mehrwert für die Studierenden und Lehrenden darstellt und zur Optimierung der Lehr- und Lernprozesse beiträgt.

Literaturverzeichnis

- Apple Inc. (2022). *Studie zeigt weltweiten Erfolg von Drittanbieter-Apps im App Store*. <https://www.apple.com/de/newsroom/2022/04/report-finds-third-party-apps-see-global-success-on-the-app-store/>
- Blacker, A. (2023). *Worldwide and US Download Leaders 2022: 10 Most Downloaded Education Apps in 2022*. https://apptopia.com/blog/worldwide-and-us-download-leaders-2022/?mod=anlink&google_analytics_client_id=GA1.1.782931992.1694101996#Education
- Bueno Muñoz, C., Nanclares, N. H., Zamorano, Luis R. Murillo & Sánchez, J. Á. L. (2023). *Gamification and Design Thinking in Higher Education: Innovation in the Economics Classroom: Case Studies for Instructional Innovation in the Economics Classroom*. Routledge. <https://doi.org/10.4324/9781032675558>
- Caminade, J. & Wartburg, M. von. (2022). *The Success of Third-Party Apps on the App Store*.
- Deterding, S., Nacke, L. E., O'Hara, K., Sicart, M. & Dixon, D. (2011). Gamification. Using game-design elements in non-gaming contexts. In Association for Computing Machinery (Hrsg.), *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*. Association for Computing Machinery.
- Facey-Shaw, L., Specht, M., van Rosmalen, P. & Bartley-Bryan, J. (2020). Do Badges Affect Intrinsic Motivation in Introductory Programming Students? *Simulation & Gaming*, 51(1), 33–54.
- Faust, A. (2021). *The effects of gamification on motivation and performance*. Dissertation. Springer Gabler. <https://doi.org/10.1007/978-3-658-35195-3>
- Highfield, K., Goodwin, K., Steinle, V., Ball, L. & Bardini, C. (2013). Apps for Mathematics Learning: A Review of 'educational' Apps from the iTunes App Store. *Mathematics Education Research Group of Australasia*, 378–385.
- Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B. & Kaufman, J. (2015). Putting Education in "Educational" Apps: Lessons from the Science of Learning. *Psychological Science in the Public Interest*, 16(1), 3–34. <https://doi.org/10.1177/1529100615569721>
- Hou, H. T. (2023). Game-Based Learning and Gamification for Education. In H. T. Hou (Hrsg.), *Game-Based Learning and Gamification for Education* (S. 1–4). Multidisciplinary Digital Publishing Institute.
- Johnston, S. (2015). Courting public favour: The Boy Scout movement and the accident of internationalism, 1907–29. *Historical Research*, 88(241), 508–529. <https://doi.org/10.1111/1468-2281.12079>
- Ornstein, A. C. (1990). Practice and Drill: Implications for Instruction. *NASSP Bulletin*, 74(525), 112–117. <https://doi.org/10.1177/019263659007452523>
- Papadakis, S., Kalogiannakis, M. & Zaranis, N. (2018). Educational apps from the Android Google Play for Greek preschoolers: A systematic review. *Computers & Education*, 116, 139–160. <https://doi.org/10.1016/j.compedu.2017.09.007>
- Tamis-LeMonda, C. S., Kuchirko, Y. & Song, L. (2014). Why Is Infant Language Learning Facilitated by Parental Responsiveness? *Current Directions in Psychological Science*, 23(2). <https://doi.org/10.1177/0963721414522813>
- Voshaar, J., Knipp, M., Loy, T., Zimmermann, J. & Johannsen, F. (2023). The Impact of Using a Mobile App on Learning Success in Accounting Education. *Accounting Education*, 32(2), 222–247. <https://doi.org/10.1080/09639284.2022.2041057>
- Warszta, T. (2023). Personalauswahl 4.0 aus der Perspektive der Kandidat:innen. In K. P. Stulle & R. T. Justenhoven (Hrsg.), *Personalauswahl 4.0: KI, Machine Learning, Gamification und andere Innovationen in der Praxis* (S. 21–38). Springer Gabler.
- Zainuddin, Z., Chu, S. K. W. & Perera, C. J. (2024). *Gamification in the Flipped Classroom: Pedagogical Methods and Best Practices*. Springer Nature Singapore. https://doi.org/10.1007/978-981-97-2219-8_4

Digitalisierung in der Lehre – Projekt zur Neukonzeption von Prüfungsleistungen

Abgeschlossenes Lehrprojekt

Janina Matern, M.Sc.

FernUniversität in Hagen, Lehrstuhl für Betriebswirtschaftslehre, insbesondere Unternehmensrechnung und Controlling, Universitätsstr. 11, D-58084 Hagen, E-Mail: janina.matern@fernuni-hagen.de

Abstract

Mit dem Auftreten der Corona-Pandemie hat die Digitalisierung in der Hochschullehre einen bedeutenden Impuls erfahren. Seither sind die Hochschulen stärker bestrebt, die traditionelle Lehre um digitale Lernformate zu erweitern. In diesem Zuge kommt auch der Digitalisierung von Prüfungsformaten eine hohe Bedeutung zu. Während zu Zeiten der Corona-Pandemie viele Hochschulen digitale Prüfungsformaten angeboten haben, ist seit dem Ende der Pandemie allerdings wieder ein Trend zur Rückkehr zu den Präsenzprüfungen zu verzeichnen, obwohl Online-Prüfungen diverse Vorteile mit sich bringen können. Eine nachhaltige Digitalisierung des gesamten Prüfungsprozessen unter didaktischen Gesichtspunkten zeigt das vorliegende Projekt. Neben Implikationen für eine systematische Planung und Umsetzung von digitalen Prüfungsvorhaben, steht dabei die Überführung von offenen Frageformaten in ein geschlossenes Fragesystem für eine teilautomatisierte Klausurkorrektur im Vordergrund.

1 Einleitung

1.1 Problemstellung

Die Digitalisierung gewinnt auch in der Lehre immer weiter an Bedeutung und hat insbesondere durch die Corona-Pandemie einen deutlichen Impuls erfahren (Dittler & Kreidl, 2021, S. 5-9; Pellert, 2020a; Wollmann, 2020, S. 253, 274-278). Seither sind die Hochschulen stärker denn je bestrebt, die traditionelle Lehre um digitale Lern- und Prüfungsformate zu erweitern (Bandtel et al., 2021, S. 9-11). Hierbei besteht die Herausforderung, die Lehre nicht nur durch mediale Unterstützung abzubilden, sondern diese auch mit didaktisch überlegten Konzepten zu untermauern, um so einen Mehrwert für die Hochschulbildung im Gesamten zu erreichen (Pellert, 2020b).

Die FernUniversität hat an der Fakultät Wirtschaftswissenschaften derzeit über 18.000 eingeschriebene Studierende, die sich über das gesamte Bundesgebiet verteilen (FernUniversität Hagen, 2024).¹ Charakteristisch für die Studierenden ist dabei, dass sie überwiegend berufstätig sind und häufig in Teilzeit studieren. Eine hohe zeitliche und örtliche Flexibilität des Studiums ist daher von herausragender Bedeutung. Dieses beinhaltet nicht nur die digitale Bereitstellung von Lernangeboten, sondern auch eine Anpassung der bisherigen Prüfungsformate an die Bedürfnisse der Studierenden. Während in der Corona-Zeit viele Universitäten Online-Prüfungsformate umsetzten, ist wieder ein Trend zurück zur Präsenzklausur zu erkennen, auch wenn digitale Prüfungsszenarien insgesamt an Bedeutung gewonnen haben (Budde et al., 2023, S. 3, 10). Da der Lehrstuhl für Betriebswirtschaftslehre, insbesondere Unternehmensrechnung und Controlling bestrebt ist, eine nachhaltige Digitalisierung in der Lehre und bei den Prüfungsleistungen der Studierenden zu erreichen, wurde ein längerfristig geplantes Projekt zur Digitalisierung der zu erbringenden Prüfungsleistungen umgesetzt, dass ergänzend zu einem zentral angelegten Projekt zur Implementierung von Online-Prüfungen an der FernUniversität stattfand.

1.2 Ziel des Projektes

Ausgangslage der Prüfungsvorleistung und Prüfungsleistung:

Die Studierenden an der wirtschaftswissenschaftlichen Fakultät legen am Ende jedes belegten Moduls eine Modulabschlussprüfung ab, welche im Regelfall die Prüfungsform Klausur vorsieht. Der Lehrstuhl für Betriebswirtschaftslehre, insbesondere Unternehmensrechnung und Controlling bietet insgesamt vier Bachelor- und Mastermodule an. Voraussetzung für die Zulassung zu der Modulabschlussprüfung ist der Nachweis, dass die Studierenden Prüfungsvorleistungen,² die grundsätzlich dem Aufbau der späteren Klausur gleichen, erfolgreich bearbeitet haben (Prüfungsordnung Bachelor Wirtschaftswissenschaften, 2020, S. 4). Die Prüfungsvorleistungen bearbeiten die Studierenden während des Semesters von zuhause aus und senden diese per Post bis zu einem bestimmten Abgabezeitpunkt zur Korrektur an die FernUniversität zurück. In der Regel handelt es sich dabei um zwei Prüfungsvorleistungen, wobei mindestens eine bestanden werden muss.

Die nachfolgende Grafik skizziert den Prozess für die Prüfungsvorleistungen und die Prüfungsleistung Klausur für ein belegtes Modul:

¹ Darüber hinaus verfügt die FernUniversität auch über einzelne Studienzentren im Ausland.

² Für die Prüfungsvorleistungen wird an der FernUniversität auch der Begriff Einsendearbeiten verwendet.

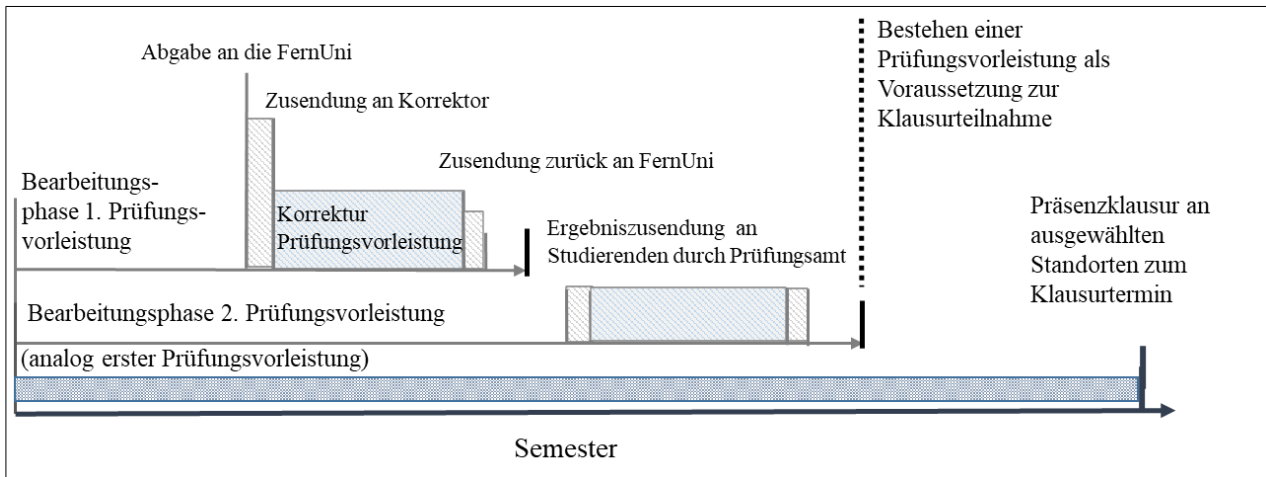


Abbildung 1: Prozess der Prüfungs(vor)leistungen vor der Umstellung

Hauptziel des Projektes:

Das Hauptziel des Projektes bestand in der vollständigen Digitalisierung des Prozesses der Bearbeitung der Prüfungsvorleistungen und der Prüfungsleistungen für die Module des Lehrstuhls.

Um eine zeitlich deutlich schnellere, studienfreundlichere Korrekturzeit zu ermöglichen, die Voraussetzungen für eine weitere Automatisierung der Prüfungsabfrage in der Zukunft zu schaffen sowie unter Berücksichtigung der Kapazitätssteuerung des Lehrstuhls wurde im Zuge der Umstellung beschlossen, eine Transformation der Aufgaben in ein geschlossenes Aufgabenprüfungsformat zu vollziehen.

Der Zielzustand beinhaltet unter anderem folgende Vorteile:

- Durch die automatische, systemseitige Vor-Korrektur entfallen Teile der Korrekturaufwendungen und die Studierenden erhalten nach Ablauf der Abgabefrist deutlich schneller ihre Bewertungsergebnisse.
- Mit Blick auf die zunehmende Digitalisierung der Prüfungsformen kann bereits durch die digitalen Prüfungsvorleistungen ein Beitrag dazu geleistet werden, dass sich die Studierenden stärker an digitale Aufgabenstellungen gewöhnen.
- Eine weitere Flexibilisierung in der örtlichen Struktur der Bearbeitung wird erreicht, da die Prüfungen Online am PC bearbeitet werden können.
- Die Prüfungsleistung im Online-Format ist kostengünstiger, da für die Mitarbeitenden der FernUni keine Reisekosten anfallen, keine Hörsäle gemietet werden müssen und die Klausuren papierlos bereitgestellt werden können.
- Für die Prüfungsvorleistungen entfallen verschiedene Verteilzeiten, die Einbindung des Prüfungsamtes als Koordinationsstelle sowie eine mitunter einige Wochen dauernde Korrekturzeit, in der die Studierenden auf das Ergebnis warten müssen. (Helpdesk Wiwi, 2020).

2 Planungsphase

Bevor die Aufgaben in ein auszuwählendes Online-System eingestellt werden können, muss eine Neukonzeption der Aufgaben durchgeführt werden. Für die Prüfungsvorleistungen, die als Aufgaben bereits bestanden, bedeutet dies eine Transformation der bestehenden Freitext-Aufgaben in die neuen, geschlossenen Aufgabentypen des Online-Systems. Die Klausuraufgaben sind hingegen erstmals direkt als geschlossene Aufgaben zu konzipieren. Nach

der Neukonzeption der Aufgaben kann die Einstellung in ein Online-System stattfinden. An der FernUniversität wird hierfür das Online-Übungssystem verwendet.³

Das mediengestützte Lernen und Prüfen sollte hierbei grundsätzlich dieselben Anforderungen erfüllen wie in der traditionellen Lehre. Dabei kommt es nicht auf ein möglichst umfassendes, sondern didaktisch zur Themenstellung passendes Angebot an (Ulrich, 2016, S.141-142). Bei der didaktischen Aufbereitung von Arbeitsmaterialien und Umgebungen sollten verschiedene Aspekte berücksichtigt werden, um das Verständnis von Lehrinhalten und die Informationsaufnahme zu fördern. Hierzu gehören u. a., dass die Medien intuitiv verständlich und einfach zu bedienen sind. Im besten Fall motivieren die eingesetzten Medien die Studierenden, sich selbstständig tiefergehend mit ihnen auseinanderzusetzen (Kerres, 2018, S. 351). Webbasierte Übungen und Lehr-/Informationsvideos helfen dementsprechend selbstgesteuerte Lernszenarien zu erarbeiten (Erpenbeck et al., 2015, S. 35-39).

Für eine erfolgreiche didaktische Umsetzung des Projektes umfasst das Vorgehen neben der reinen Digitalisierung der Prüfungsleistungen weitere Bestandteile. Neben der Transformation und Einstellung der Prüfungsvorleistungen und Erarbeitung der Klausuren im neuen Format sind weiterführende Informationsmedien geplant worden, um die Studierenden auf die Neuerungen adäquat vorzubereiten. Hierunter fällt zum einen eine neu entwickelte Übungs-umgebung, die es den Studierenden ermöglicht, die neuen Aufgabentypen vor Bearbeitungsbeginn selbstständig kennenzulernen sowie Informationsvideos (und weitere Informationshinweise) zur Umstellung der Prüfungsformate.

Der zukünftig geplante Zielzustand der vollständig digitalen Bearbeitung der Prüfungsvorleistungen und Prüfungsleistungen ist in der Darstellung zusammenfassend abgebildet:

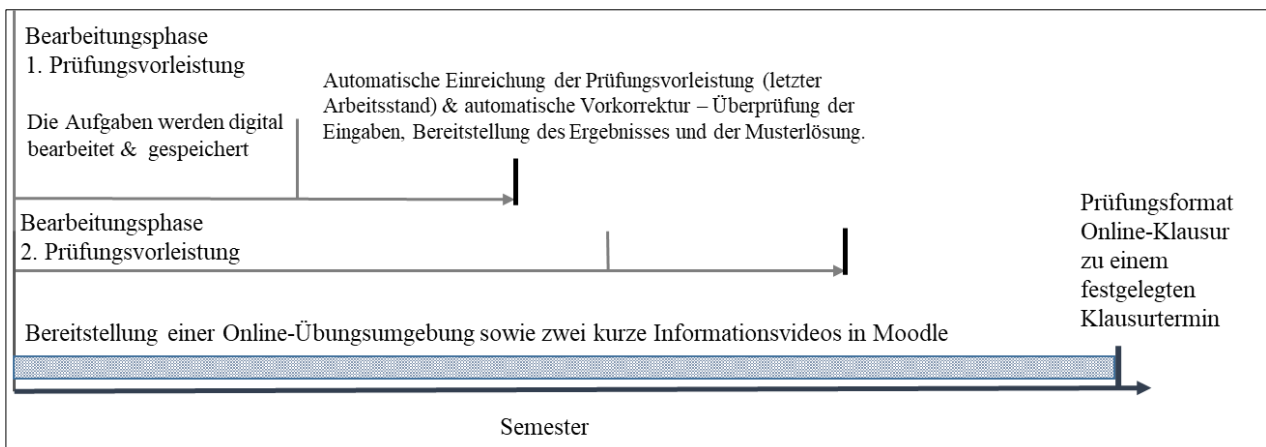


Abbildung. 2: Zukünftige Bearbeitung der Prüfungsleistungen im Zeitablauf⁴

Die technische Bereitstellung der Online-Prüfungsräume, der Aufsichten und Anforderungen an die Studierende in der Klausur wurden durch eine Projektgruppe der Fakultät geleitet und waren nicht Teil der didaktischen Umsetzung des Projektes.⁵ Für einen Überblick werden in Kap. 3.3 die wesentlichen Charakteristika der Prüfungssituation dennoch kurz dargestellt.

Eine umfassende Teilaufgabe ist zusätzlich in der Organisation und Kommunikation innerhalb des Lehrstuhls und in Bezug auf das Prüfungsamt zu sehen. Grund hierfür ist, dass die Konzeption und Einarbeitung der Aufgaben

³ Das *Online-Übungssystem* ist eine Eigenanwendung für verschiedene Prüfungsleistungen. Die im weiteren Verlauf genannten Art der Aufgabentypen bieten jedoch auch andere Systeme wie bspw. Moodle.

⁴ Die zweite Einsendearbeit weist wiederum den gleichen, lediglich zeitlich versetzten Ablauf der ersten EA auf.

⁵ Die digitale Umstellung der Prüfungen erfolgte dabei an einigen anderen Lehrstühlen in einem deutlich geringeren Umfang, so dass bspw. pdf-Dateien online bereitgestellt werden, woraufhin die Studierenden ihrerseits die Antworten als pdf-Dateien hochladen.

durch sechs Lehrstuhlmitarbeitende geleistet wird. Da der Lehrstuhl einer der ersten ist, der die Prüfungsvorleistungen und Prüfungen teilautomatisiert korrigieren lässt und ein breites Spektrum der möglichen Aufgaben nutzt, ist zusätzlicher Abstimmungs- und Entwicklungsaufwand mit dem Prüfungsamt und der IT-Abteilung einzuplanen.

Ausgehend von den aufgezeigten Bausteinen, die bei der Umstellung Berücksichtigung finden sollten, wurde ein Projektplan erarbeitet. Dieser sollte sowohl die zeitliche Planung als auch den ungefähr zu beanspruchenden Arbeitsanfall wiedergeben. Die nachfolgende Grafik beinhaltet den ursprünglichen Projektplan, ergänzt um die Erfahrungswerte und den tatsächlich angefallenen Arbeitsaufwand im Rahmen der Umstellung. Hierbei wurden Tätigkeiten in mehreren Zeilen ergänzt, die zu Beginn der Projektplanung noch nicht berücksichtigt wurden, diese sind in der Spalte der *Einschätzung des zeitlichen Aufwandes* grau hinterlegt. Nur in der im Nachgang hinzugefügten Spalte *geschätzter tatsächlich angefallener Arbeitsaufwand* ist folglich ein Wert enthalten. Außerdem wurde eine weitere Spalte eingefügt, welche farblich markiert, ob der ursprünglich angesetzte Zeitaufwand eingehalten (hellgrün), übererfüllt (dunkelgrün) oder überschritten (rot) wurde.

Insgesamt zeigt sich, dass die zunächst aufgestellte zeitliche Planung kaum eingehalten wurde. Insbesondere der Aufwand, der in Bezug auf Abstimmungen, Koordination und Qualitätskontrollen anfiel, wurde im Vorfeld deutlich unterschätzt. Zusätzlich haben das Kennenlernen und der tiefergreifende Umgang mit den unterschiedlichen Systemen nicht eingeplante zeitliche Ressourcen in Anspruch genommen. Sofern andere Lehrstühle ebenfalls eine Neukonzeption und Umstellung der Einsendearbeiten planen, kann hiermit ein Anhaltspunkt gegeben werden, welche zeitlichen Ressourcen zur Verfügung gestellt werden sollten.

Projektplan				
Zeitliche Einordnung	Aufgabe und Ziele	Einschätzung des zeitlichen Aufwandes	Geschätzter tatsächlich angefallener Arbeitsaufwand	Einhaltung Zeitplan
Vor dem Projektstart	Grundsätzliche Ideenfindung und Konzeptionierung	6,00	6,00	
	Konkretisierung, Erarbeitung der Projektskizze	5,00	10,00	
	Absprachen zum Vorgehen mit dem Prüfungsamt, Lehrstuhlinhaber und den Lehrstuhlmitarbeitenden	10,00	10,00	
Projektstart April				
Erstellung der Aufgaben in digitaler Form	Einrichten der Kursumgebung (Online-Übungsplattform)	4,00	2,00	
	Sichten der Prüfungsvorleistungen und Prüfung des Änderungsbedarfs	4,00	4,00	
	Einarbeitung der Prüfungs(vor)leistungen im Online-System je MA	50,00	50,00	
	Qualitätskontrolle der anderen Lehrstuhlmitarbeitenden der EA		10,00	
	Einarbeitung in das Online-System		10,00	
	Erstellung einer Übungsumgebung für die Studierenden	10,00	20,00	
	Nach dem Review-Prozess: Einarbeitung der Hinweise in die Übungsumgebung	2,00	4,00	
Anpassung von Informationsquellen	Sichtung und Identifikation von Informationsquellen (Moodle, FAQ, Homepage, Kurseinheiten, Rücksprache Prüfungsamt...)	8,00	8,00	
	Anpassung von Informationsquellen	4,00	6,00	
	Erstellung eines Informationsvideos zur geplanten Umstellung	12,00	20,00	
	Schulung und Kennenlernen des Programmes Camtasia zur Videoerstellung		10,00	
Absprachen und Freigabe	Im Rahmen der Umsetzung entstehende Absprachen mit Lehrstuhlinhaber, dem Prüfungsamt und den Lehrstuhlmitarbeitenden	5,00	23,00	
	Weiterentwicklung, Ideenaustausch, Problemebehebung im Rahmen der Einarbeitung mit ZMI und Prüfungsamt		10,00	
Projektende Oktober				
Projektbericht	Anfertigung des Projektberichts	12,00	25,00	
SUMME		132,00	228,00	96,00

Abbildung 3: Projektplan und tatsächlicher Umsetzungsaufwand

Der nachfolgende Abschnitt setzt sich zunächst mit der Transformation und Einstellung der Aufgaben in das Online-Übungssystem auseinander.

3 Umsetzungsphase

3.1 Transformation und Einstellung der Prüfungs(vor)leistungen

Vor der eigentlichen Transformation und Einstellung der Aufgaben mussten lehrstuhlinterne Absprachen, zentrale Koordinationsaufgaben sowie die Anlage und der Zugang für die Studierenden geklärt und eingerichtet werden.

Die geplante Umstellung der Prüfungen wurde im Rahmen einer lehrstuhlinternen Besprechung beschlossen. Für jeden Mitarbeitenden bedeutete dieses die Umstellung von zwei Prüfungsvorleistungen, die Klausur zum verantworteten Modul sowie die Qualitätskontrolle von eines anderen Mitarbeitenden. Der Zeitplan wurde so festgelegt, dass innerhalb der üblichen Abgabefristen der Prüfungsleistungen an das Prüfungsamt alle Aufgaben abgeschlossen sein sollten.

Im Rahmen der zentralen Koordination war es unter anderem die Aufgabe, sämtliche Informationsmedien zu den bestehenden Prüfungs(vor)leistungen des Lehrstuhls zu prüfen, da diese gegebenenfalls angepasst werden mussten. Die unterschiedlichen Informationsarten und -orte wurden tabellarisch notiert und die zu ändernden Texte vorgeschrieben, um diese zwei Wochen vor dem Semesterstart, in dem die Umstellung stattfinden sollte, auszutauschen. Dies betraf u. a. die FAQ-Seiten der Module, die Lehrstuhl-Homepage, die Moodle-Lernumgebung sowie allgemeine Informationsseiten der FernUniversität zu den Prüfungs(vor)leistungen der Module.⁶

Die Anlage der Prüfungsumgebungen wurde nach Absprache mit dem Prüfungsamt zentral durch die IT vorgenommen. Der Zugang zu den Prüfungs(vor)leistungen ist in die Moodle-Umgebung der Kurse integriert worden, so dass die Studierenden die Prüfungsräume ohne Komplikationen über das ihnen vertraute Portal erreichen können. Mit der angegebenen Verlinkung gelangen die Studierenden zu der Übersicht der einzelnen Aufgabenhefte der Prüfungs(vor)leistungen des jeweiligen Semesters.

Aus Sicht des Lehrstuhls bestehen für die Ausgestaltung der geschlossenen Prüfungsabfragen verschiedene Möglichkeiten. Es wurde sich für die folgenden Aufgabentypen entschieden, die sich von den bisherigen Aufgabentypen deutlich unterscheiden:⁷

- Einfachauswahl,
- Multiple-Choice-Aufgaben,
- Zahlenaufgaben zur Eingabe von Rechenergebnissen,
- Zuordnungsaufgaben,
- Lückentexte.

Nachfolgend ist beispielhaft eine umgestellte Aufgabe einer Prüfungsleistung abgebildet. Die vorher als Freitextaufgabe formulierte Frage nach den Vor- und Nachteilen eines lokal-unterstellten Controllers wurde hier in eine Zuordnungsaufgabe mit integrierter Abbildung umgewandelt.

⁶ Bei Letzterem erfolgte eine Änderung zentral durch das Prüfungsamt nach vorheriger Rücksprache.

⁷ Für weiterführende Informationen siehe Littkemann (2021).

b) Unterstellungsmöglichkeiten des lokalen Controllers (120 Punkte)

Es existieren verschiedene Unterstellungsmöglichkeiten für die lokalen Controller, mit denen unterschiedliche Vor- und Nachteile einhergehen.
Ordnen Sie die aufgeführten Aussagen (1.-6.) in der folgenden Abbildung den Stärken und Schwächen der verschiedenen Unterstellungsformen zutreffend zu.

	A: Stärken Dotted-Line-Prinzip	B: Stärken Unterstellung Zentralcontrolling	C: Stärken Unterstellung Teilbereichsleiter	D: Schwächen Dotted-Line-Prinzip	E: Schwächen Unterstellung Zentralcontrolling	F: Schwächen Unterstellung Teilbereichsleiter
1.	☐	☐	☐	☐	☐	☐
2.	☐	☐	☐	☐	☐	☐
3.	☐	☐	☐	☐	☐	☐
4.	☐	☐	☐	☐	☐	☐
5.	☐	☐	☐	☐	☐	☐
6.	☐	☐	☐	☐	☐	☐

Abbildung 4: Beispiel einer umgestellten Prüfungsaufgabe aus dem Online-System

3.2 Herausforderungen bei der Anpassung der Aufgabenerstellung

Die Integration von Multiple-Choice Aufgaben hatte zur Folge, dass die Bepunktung der Klausur angepasst werden musste, um prüfungsrechtlich Bestand haben. In einem bestehenden Gerichtsurteil wurde entschieden, dass es grundsätzlich zu keiner Negativbepunktung bei MC-Mehrfachauswahlfragen kommen darf. Der aus diesem Grund vom Prüfungsamt vorgeschlagene Bewertungstyp sieht eine gestaffelte, exponentielle Punktevergabe vor. Um diesen Bewertungstyp umsetzen zu können, musste die ursprüngliche Punktzahl um das 20-fache angehoben werden.⁸ Dies ist dem Umstand geschuldet, dass das verwendete Online-System keine Teilpunkte ausweisen kann und bei nur einer richtig gewählten Antwort bei dieser Variante 1 % der Aufgabenpunkte zugesprochen werden. Eine MC-Mehrfachauswahlaufgabe, für die nach der alten Logik etwa 5 Punkte (= 5 Minuten) vergeben worden wären, erhält nun 100 Punkte. Bei 5 Antwortmöglichkeiten und einer richtig bewerteten Aussage⁹ erhält der Studierende somit einen von 100 Punkten. Die intuitive Bepunktung für die Studierenden von „1 Minute Bearbeitungszeit = 1 Punkt“ entfällt durch die Änderung.

Die Einschränkung, dass innerhalb des Online-Systems nur Teilpunkte vergeben werden können, hat auch Auswirkungen auf andere Fragetypen wie die Zuordnungsaufgaben, bei denen beispielsweise nur ein Vielfaches der abgefragten Zuordnungen als Punktzahl verwendet werden kann. Derartige Restriktionen müssen im Rahmen der Aufgabenkonzeption berücksichtigt werden.

Weitere Herausforderung während des Umstellungsprozesses betrafen u. a. den Aufgabentyp Zahlen. Grund hierfür ist, dass innerhalb des Online-Systems noch keine Folgefehlerbewertung möglich ist. Da die Controlling-Module jedoch viele Rechenverfahren beinhalten, mussten hier andere Wege zur Abfrage von Zwischenlösungen gefunden

⁸ Bisher war die zu erreichende maximale Punktzahl einer Prüfungsvorleistung auf 60 Punkte und die einer Klausur auf 120 Punkte festgelegt. Für jeden Punkt wurde in etwa mit einer Minute Bearbeitungszeit gerechnet, so dass die Aufgaben grundsätzlich innerhalb von einer Stunde lösbar sein sollten.

⁹ Hierunter fällt sowohl das Ankreuzen einer richtigen Antwort als auch das Nicht-Ankreuzen einer als falsch erkannten Antwort.

werden, zum Beispiel über die Vorgabe von alternativen Zwischenergebniswerten. Dabei bestand insbesondere die Herausforderung, dass durch die Abfrage von geschlossenen Aufgaben das Schwierigkeitsniveau nicht herabgesetzt wird. Die Abfrage von Transferwissen gestaltet sich bei geschlossenen Aufgaben deutlich schwieriger und bedeutet einen höheren, zeitlichen Konzeptionsaufwand als die ursprüngliche Klausurerstellung.¹⁰ Zuletzt sind noch kleinere Administrationsfehler bei der Übertragung der Aufgaben und im Online-System zu nennen, die ebenfalls im Prüfprozess aufgefallen sind.

Letztendlich ist die exakte Formulierung der Aufgabenstellung und der zu erwartenden Antworten nicht zu unterschätzen. Anders als bei Freitextaufgaben, in welchen die Studierenden gegebenenfalls Annahmen treffen können, ist bei geschlossenen Fragestellungen diese Möglichkeit nicht gegeben. Für die Konzeption und Reflexion der Aufgabenerstellung boten sich unter anderem die Überlegungen zu den Taxonomiestufen nach Bloom an, um ein angemessenes Schwierigkeitsniveau zu erreichen (Reinmann, 2013, S. 6-7; Tremp & Eugster, 2020, S. 219-233).

An den Zulassungsvoraussetzungen zur Klausur oder der Bestehensgrenze der Prüfungs(vor)leistungen wird keine Veränderung vorgenommen.

3.3 Ablauf und Ergebnisbekanntgabe der Prüfungs(vor)leistungen

Ablauf der Klausurprüfung:

Während die Studierenden für die Prüfungsvorleistungen mehrere Wochen Zeit haben, diese bis zum Abgabedatum von zuhause aus zu bearbeiten, erfolgt die Abschlussprüfung der Klausur zentral an einem bestimmten Tag zu einem bestimmten Zeitpunkt. Als Prüfungssoftware zur Überwachung der Studierenden arbeitet die FernUniversität dabei dauerhaft mit Zoom. Im eingestellten Aufsichtsmodus werden im Schnitt zwischen 20-40 Studierende durch drei Aufsichten überwacht; die Studierenden können sich in diesem Modus gegenseitig nicht sehen. Die Studierenden müssen für die Teilnahme an der Prüfung über eine externe Kamera verfügen und ihren Bildschirm über die gesamte Bearbeitungszeit hinweg teilen. Die Legitimationskontrolle erfolgt separat in Breakout-Räumen. Um diese zeitliche Unterbrechung einzupreisen, haben die Studierenden bereits fünf Minuten vor dem offiziellen Bearbeitungsbeginn Zugriff auf die Klausur. Wie auch in Präsenz erfolgt zu Beginn eine zentrale Ansage an die Studierenden. Am Ende der Bearbeitungsfrist werden die zuletzt gespeicherten Ergebnisse automatisch eingereicht und den Korrektoren zur Prüfung vorgelegt.

Veröffentlichung der Ergebnisse der Prüfungsvorleistung:

Nachdem die Bearbeitungsfrist beendet ist, erhalten die Studierenden nach Kontrolle ihre Ergebnisse in einer Online-Lösungsübersicht sowie die ebenfalls online bereitgestellten Musterlösungen. Der Korrekturprozess lief insgesamt schnell und ohne größere Komplikationen ab. Nur kleinere Verbesserungen mussten vorgenommen werden. So ist z. B. innerhalb der Aufgabenstellungen vereinzelt nicht genau benannt worden, in welchen Einheiten, mit wie viel Kommastellen etc. die Lösungen anzugeben waren.

Veröffentlichung der Ergebnisse der Prüfungsleistung Klausur:

Die Klausurergebnisse und eine detaillierte Übersicht der erreichten Punkte je Aufgabe werden den Studierenden im Online-System nach etwa drei Wochen zur Verfügung gestellt. Separat erhalten die Studierenden wie bisher einen Brief mit dem offiziellen Ergebnis vom Prüfungsamt. Die Widerspruchs- und Nachkorrekturquote hat sich seit der transparenteren Darstellung der Ergebnisse im Zuge der Umstellung deutlich reduziert.

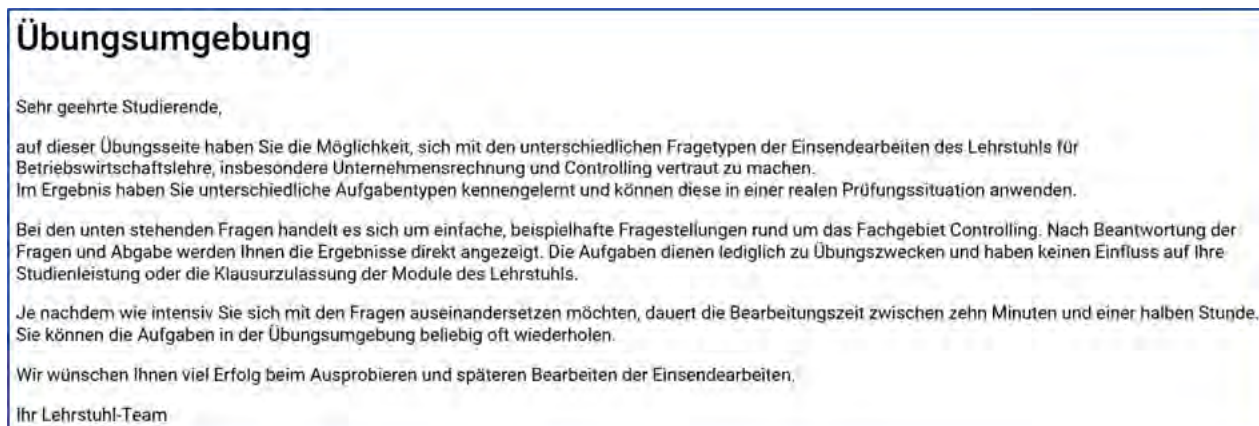
3.4 Die Übungsumgebung zum Kennenlernen der Aufgabentypen

Neben den Prüfungs(vor)leistungen im Online-Übungssystem sollte außerdem eine Übungsumgebung implementiert werden, in welcher die Studierenden die verschiedenen Aufgabentypen üben können. Dieses wird als sinnvoll

¹⁰ Darüber hinaus wurde der Abfragetyp Lückentexte erst im Laufe der Umstellung der Prüfungs(vor)leistungen gemeinsam mit dem Lehrstuhl entwickelt.

erachtet, da aufgrund der veränderten Prüfungsabfrage bei den Studierenden Unsicherheiten entstehen könnten und vermieden werden soll, dass aufgrund einer fehlerhaften Bedienung des Systems Punkte nicht erreicht werden.

Ein kurzer Einleitungstext mit den wesentlichen Hinweisen sollte den Studierenden das Ziel sowie weitere Informationen zur Übungsumgebung liefern:



Übungsumgebung

Sehr geehrte Studierende,

auf dieser Übungsseite haben Sie die Möglichkeit, sich mit den unterschiedlichen Fragetypen der Einsendearbeiten des Lehrstuhls für Betriebswirtschaftslehre, insbesondere Unternehmensrechnung und Controlling vertraut zu machen. Im Ergebnis haben Sie unterschiedliche Aufgabentypen kennengelernt und können diese in einer realen Prüfungssituation anwenden.

Bei den unten stehenden Fragen handelt es sich um einfache, beispielhafte Fragestellungen rund um das Fachgebiet Controlling. Nach Beantwortung der Fragen und Abgabe werden Ihnen die Ergebnisse direkt angezeigt. Die Aufgaben dienen lediglich zu Übungszwecken und haben keinen Einfluss auf Ihre Studienleistung oder die Klausurzulassung der Module des Lehrstuhls.

Je nachdem wie intensiv Sie sich mit den Fragen auseinandersetzen möchten, dauert die Bearbeitungszeit zwischen zehn Minuten und einer halben Stunde. Sie können die Aufgaben in der Übungsumgebung beliebig oft wiederholen.

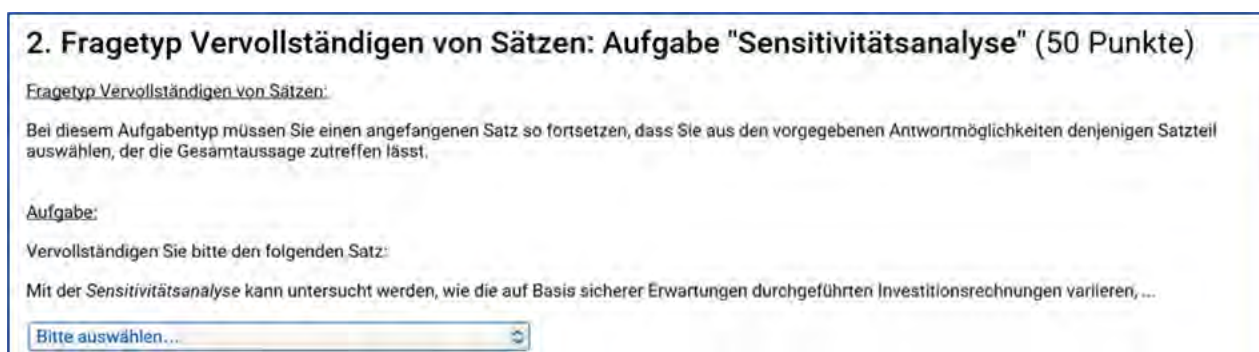
Wir wünschen Ihnen viel Erfolg beim Ausprobieren und späteren Bearbeiten der Einsendearbeiten.

Ihr Lehrstuhl-Team

Abbildung 5: Einleitungstext Übungsumgebung

Die Aufgaben wurden so konzipiert, dass eine Bearbeitung innerhalb von 10-30 Minuten möglich sein sollte. Hierzu wurde zu jedem Aufgabentyp eine einfache Aufgabe aus dem Themengebiet des Controllings entwickelt. Der Fokus lag nicht auf einer inhaltlichen Leistungsabfrage, sondern auf der Online-Sozialisation bzw. dem Kennenlernen des Online-Systems und der Art der Fragestellungen. Insgesamt existieren in der Übungsumgebung elf verschiedene Aufgaben, mit denen sich die Studierenden in Vorbereitung auf die Prüfungs(vor)leistungen auseinandersetzen können. In jeder Aufgabe wird zunächst der Fragetyp kurz benannt und anschließend erläutert, bevor die eigentliche Aufgabenstellung beginnt.

Nachfolgend ist beispielhaft eine Aufgabe der Übungsumgebung aufgezeigt, um einen Eindruck zu gewinnen, in welchem Stil diese verfasst sind.



2. Fragetyp Vervollständigen von Sätzen: Aufgabe "Sensitivitätsanalyse" (50 Punkte)

Fragetyp Vervollständigen von Sätzen:

Bei diesem Aufgabentyp müssen Sie einen angefangenen Satz so fortsetzen, dass Sie aus den vorgegebenen Antwortmöglichkeiten denjenigen Satzteil auswählen, der die Gesamtaussage zutreffen lässt.

Aufgabe:

Vervollständigen Sie bitte den folgenden Satz:

Mit der *Sensitivitätsanalyse* kann untersucht werden, wie die auf Basis sicherer Erwartungen durchgeführten Investitionsrechnungen variieren, ...

Bitte auswählen...

Abbildung 6: Beispielhafte Aufgabe aus der Übungsumgebung

Die Übungsumgebung wurde fachfremden Lehrstuhlmitarbeitenden zu Testzwecken vorgelegt, um noch vor der Veröffentlichung konstruktive Hinweise aufzunehmen und umzusetzen. Zur Validierung diente das Online-Tool des Verbundprojektes *wm³ - Weiterbildung Mittelhessen* zur Systematisierung und Validierung von Studienmaterialien (*wm³*, 2017, S. 1-76). Dieses betraf z. B. die bereits erwähnte Beschreibung des Aufgabentyps vor dem eigentlichen Aufgabentext, die Nennung der Bearbeitungsdauer, den Stand und die Versionsnummer sowie die Art der gestellten Aufgaben. Die gewonnenen Erkenntnisse wurden auch für die Verbesserung der Prüfungs(vor)leistungen genutzt.

Ähnlich wie bei den Prüfungsleistungen haben die Studierenden auch hier die Möglichkeit, sich ihre Eingabeergebnisse im Nachgang anzeigen zu lassen. Die Übersicht ist grundsätzlich analog der Ergebnisse der Prüfungs(vor)leistungen. Da sich die Studierenden jedoch zu Übungszwecken direkt ansehen sollen können, wie ihre Antworten bepunktet wurden, welche Lösungen zulässig waren und welche Fehler sie u. U. in der Bearbeitung gemacht haben, werden an dieser Stelle die Lösungen sofort mit Kommentierungen angezeigt. Einen Eindruck der Ergebnisdarstellung vermittelt die folgende Darstellung:

3. Fragetyp Zuordnung: Aufgabe "Risikoarten und -subtypen" Externe Risiken	Ihre Lösung: C Richtige Lösung: B Die richtige Zuordnung ist: Prozessrisiken; Steuerungsrisiken Risiken aus Personal; Managementrisiken Externe Risiken; Umweltstrisiken	0 von max. 20 P. 0 % (1 aus 5)
4. Fragetyp Wert/Zahl bestimmen: Aufgabe "Baumanalyse"	Ihre Lösung: 3500000 Richtige Lösung: [3500 / 3500] oder [3,5 / 3,5] oder [3500000 / 3500000] Bei dieser Aufgabe muss das graue Kästchen = zu erwartender Schaden für "die Mensa brennt" in der Darstellung berechnet werden. Wichtig bei dieser Art von Aufgaben ist, dass Sie auf die geforderte Rechengröße achten (z.B. Angabe in Tsd., Mio., ausgeschriebene Zahl...). Da alle Angaben gegeben sind, errechnet sich der Wert wie folgt aus den potentiellen Schäden der Ebene darüber: $0,25 \times 5 \text{ Mio.} + 0,75 \times 3 \text{ Mio.} = 3500 \text{ Tsd. Euro.}$	160 von max. 160 P. 100 % (Zahlen)
5. Fragetyp Lückentext: Aufgabe "Morgengenuß GmbH"	Ihre Lösung: funktionierendes Controllingsystem x, Beschaffung x, Finanzbuchhaltung -, x, x Richtige Lösung: Kosten- und Leistungsrechnung, funktionierendes Controllingsystem, Finanzbuchhaltung, Planungs- und Kontrolldaten, Beschaffung Die Morgengenuß GmbH ist eine traditionsreiche Firma, die bereits im 19. Jahrhundert gegründet wurde. Konrad Troller konnte sich in den vergangenen Wochen ein Bild über die bisher als Controlling bezeichneten Tätigkeiten bei der Morgengenuß GmbH verschaffen und stellt folgendes fest: Zwar existiert eine Kosten- und Leistungsrechnung, diese ist aber meilenweit davon entfernt, als Ausgangsbasis für ein funktionierendes Controllingsystem zu dienen, geschweige denn selbst als Controllingsystem bezeichnet zu werden. Die im Rahmen der sog. Kosten- und Leistungsrechnung verarbeiteten Daten entstammen ausnahmslos der Finanzbuchhaltung. In den Funktionsbereichen werden zwar vereinzelt Planungs- und Kontrolldaten erhoben, diese sind aber allenfalls als rudimentär zu bezeichnen. In keinem der primären Funktionsbereiche (Beschaffung, Produktion, Marketing, Logistik und Personal) kommt es zu einem zielorientierten und wirkungsvollen Controlling.	20 von max. 100 P. 20 % (Begriffsfolge)
6. Fragetyp wahr/falsch: Aufgabe "Innovationcontrolling in der Automobilindustrie"	Ihre Lösung: A Richtige Lösung: A Diese Aussage ist wahr. Besonders im Innovationsbereich ist das Controlling eine anspruchsvolle Aufgabe. Neben der Begleitung und Optimierung der Entwicklungsprozesse aus finanzieller Sicht geht es insbesondere auch darum, Fehlentwicklungen zu erkennen.	50 von max. 50 P. 100 % (1 aus 1)
7. Fragetyp MC Mehrfachauswahl: Aufgabe "Nutzwertanalyse Buskauf"	Ihre Lösung: A B C D E Richtige Lösung: A,B,E Richtig sind die folgenden Aussagen: A) Der alte Bus hat einen höheren Gesamtnutzwert als der neue Bus. B) Der errechnete Gesamtnutzwert des alten Buses beträgt 6,4. E) Die Kriterien "Charme" und "Preis" werden beim alten Bus besser bewertet als beim neuen Bus. Die anderen beiden Aussagen sind falsch. C) Der Gesamtnutzwert des neuen Buses beträgt $3,5 (= 0,2 \times 7 + 0,3 \times 2 + 0,5 \times 3)$. D) Die Zuverlässigkeit des neuen Buses fließt mit 1,4 in den Gesamtnutzwert ein.	30 von max. 100 P. 30 % (1 bis 3 aus 10 (2/2,0))

Abbildung 7: Auszug aus der Lösungsübersicht

Die Lösungsübersicht weist die erreichten Punkte absolut und relativ für jede Aufgabe aus. Bei Auswahl- bzw. Ankreuzaufgaben werden richtig bewertete Aussagen mit einem grünen Haken gekennzeichnet, falsche Aussagen erhalten ein rotes Kreuz. So können die Ergebnisse durch die Studierenden problemlos nachvollzogen werden. Am Ende jeder Zeile ist hinter der prozentual erreichten Punktzahl der Bewertungsmodus für die einzelnen Aufgaben hinterlegt. Durch Aufrufen des Links können sich die Studierenden informieren, wie der jeweilige Aufgabentyp bewertet wird. Dies ist z. B. bei den bereits erwähnten MC-Mehrfachauswahlaufgaben von Vorteil, welche in der Punkteverteilung nicht intuitiv verständlich sind.

Insgesamt kann festgehalten werden, dass den Studierenden vor den eigentlichen Prüfungsleistungen eine umfassende Übungsmöglichkeit bereitgestellt wird.

3.5 Informationsvideos zur Umstellung der Prüfungsformate

Neben den schriftlichen Hinweisen und der Übungsumgebung wurden zwei Lehr-/Informationsvideo zur Umstellung der Prüfungs(vor)leistungen verfasst, eines zu den Prüfungsvorleistungen, eines zu den Klausurprüfungen. Diese stehen den Studierenden als zusätzliches Informationsmedien in der Moodle Lehrumgebung zur Verfügung. Ziel war es hierbei, die wesentlichen Bearbeitungshinweise zu erläutern, um die Studierenden zu informieren und eine selbstständige Auseinandersetzung mit dem Online-Systems zu erreichen.

Die Videos beinhalten dabei Hinweise zum Einstieg und der Bearbeitung der Aufgaben des Online-Systems, zu den neuen Aufgabentypen, der Übungsumgebung und den Abgabefristen und dem Ablauf der Prüfungen. Außerdem

werden für inhaltliche und technische Fragen Ansprechpartner genannt, um unklare Sachverhaltsfragen gezielt an die richtigen Institutionen adressieren zu können.

Die Konzeption der Lehrvideos wurde vor der Aufnahme anhand eines Storyboards vorgenommen. Die eigentliche Umsetzung erfolgte mit der Software Camtasia, auf Basis einer Powerpoint-Präsentation und selbst gezeichneten Grafiken sowie Animationen.¹¹

Das Storyboard in seiner groben Struktur ist in der folgenden Grafik abgebildet:

Vorgehen:	Ziel	Inhalt	Methode	Dauer
Begrüßungsfolie	kurze Begrüßung	Vorstellung Lehrstuhl, Betreuung der Module ...	Powerpoint Folie, Einspann FernUni	20 Sek.
Folie Veränderung	Erläuterung der Digitalisierung der Prüfungs(vor)-leistungen	Digitalisierung der Prüfungsleistungen, Möglichkeit der flexiblen Bearbeitung ohne Postversand, Umfang der Prüfungsvorleistungen in Bearbeitungspunkten, verschiedene Frageformen (siehe beispielhafte Aufgabentypen)	Powerpoint, Animation in Camtasia	30 Sekunden
Wichtige Hinweise	relevante Bearbeitungshinweise erläutern	Wie läuft die Bearbeitung ab; bis wann muss das Ergebnis abgegeben sein, wie wird das Ergebnis abgegeben, wann wird das Bearbeitungsergebnis bereitgestellt, werden Lösungshinweise gegeben, Eingabe von Einzelwerten (Rechtschreibung, Recheneinheit), Änderung der Reihenfolge der Fragen	Animation Ipad, Screenshots Online-System, Animation Camtasia	2 Minuten
Beispielhafte Aufgabentypen	Darstellung der versch., geforderten Aufgabentypen	Übersicht der verschiedenen Aufgabentypen, (wahr/falsch, MC- Einfachauswahl, MC- Mehrfachauswahl, Zahl, Begriffe, Lückentext, Zuordnungsaufgaben, Vervollständigen von Sätzen)	Animation Ipad	1:30 Minuten
Beispielhafte Aufgabentypen	Darstellung der versch., geforderten Aufgabentypen	Vorstellung der einzelnen Aufgabentypen, Abgabe der Prüfungsleistungen	Powerpoint	2 Minuten
Schlusswort	Verabschiedung	"Hoffentlich waren die Hinweise für Sie hilfreich", Viel Erfolg für die Bearbeitung wünschen.	Animation und Logo der Fernuni	15 Sekunden

Abbildung 8: Drehbuch für eines der Informationsvideos

Zur Qualitätssicherung der Inhalte wurden die Informationsvideos zur Kontrolle an das Prüfungsamt gesendet, sowie Feedback von den Mitarbeitenden des Lehrstuhls eingeholt. Insgesamt ist festzuhalten, dass sich ausreichend Zeit für die Erstellung eines Lehrvideos genommen werden sollte. Hierbei ist zwischen der Erstellung der Präsentation, dem Einfügen der Animationen und der finalen Vertonung zu unterscheiden. Grafische Animationen führen insgesamt zu mehr Abwechslung des Videos und damit verbunden auch zu einer vermutlich größeren Aufmerksamkeitsspanne der Zuschauenden. Die perfekte Abstimmung zwischen Erscheinen und Ausblenden von animierten Informationen, den Präsentationsfolien und der sprachlichen Unterlegung ist entsprechend zeitaufwendig. Bei der Vertonung eines Videos bietet es sich an, sich den Text im Vorfeld zu notieren, um störende Zwischengeräusche oder Versprecher zu vermeiden.¹²

Nach der Einschätzung des Lehrstuhls ist es auch aufgrund der zahlreichen Informationsmedien und Hilfestellungen im Rahmen der Umstellung der Prüfungs(vor)leistungen kaum zu Rückfragen der Studierenden gekommen. Somit kann davon ausgegangen werden, dass die bereitgestellten Medien für die Vermittlung der Inhalte geeignet sind.

3.6 Evaluation innerhalb des Projektes

Ein weiterer Baustein innerhalb der Planung ist die Integration einer Evaluation der Umstellung. Die Lehre erreicht immer dann ein hohes Niveau, wenn diese auch reflektiert wird und Schwierigkeiten durch Innovationen gelöst

¹¹ Die Software Camtasia bietet bspw. umfangreichen Hilfsmaterialien für diverse Animationen, um eine ansprechende, abwechslungsreiche Gestaltung zu erreichen.

¹² Das Video zur Umstellung der Einsendearbeiten steht den Studierenden in der Moodle-Lernumgebung oder über den folgenden Link zur Verfügung: <https://www.youtube.com/watch?v=-GaMZrEklhY>.

werden. Das Einholen eines Feedbacks durch die Studierenden ist demnach von großer Bedeutung, um zukünftig eine noch ansprechendere Lehre anbieten zu können (Ulrich, 2016, S. 191). Insbesondere bei einer solch umfangreichen Umstellung ist das Feedback der Studierenden ein wichtiger Indikator, ob von einer didaktisch sinnvollen Umsetzung gesprochen werden kann bzw. an welchen Stellen grundsätzlich Verbesserungspotential existiert. Sofern die Ergebnisse eher negativ ausfallen würden, würde ein gezieltes Nachfragen über die Unzulänglichkeiten und Probleme über die Moodle-Foren der Kurse stattfinden. Die Ergebnisse der Auswertung werden kurz im nächsten Abschnitt vorgestellt.

Ein häufiges Problem bei Befragungen bzw. Evaluationen sind die niedrigen Rücklaufquoten (Berger-Grabner, 2016, S. 164). Um eine möglichst hohe Rücklaufquote der Studierenden zu erhalten, erfolgte daher eine Eingliederung der Evaluationsfragen direkt im Rahmen der Prüfungs(vor)leistungen. So wurde an das Ende jeder der Prüfungs(vor)leistungen der Module des Lehrstuhls im Rahmen der erstmaligen Umstellung der Bearbeitungsform der folgende Aufgabentext nebst Fragen implementiert, hier beispielhaft für die Prüfungsvorleistung:

Evaluation der veränderten Bearbeitung der Einsendearbeiten (0 Punkte)

Die Bearbeitung der Einsendearbeiten erfolgt seit diesem Semester erstmalig online und mit veränderten Aufgabentypen. Wir möchten Sie daher bitten, dem Lehrstuhl ein kurzes Feedback zu geben, wie Sie die neue Aufgabenbearbeitung empfunden haben.

Die Beurteilung hat selbstverständlich weder Einfluss auf das Bestehen Ihrer Einsendearbeit, noch erfolgt eine Einzelbetrachtung der Evaluationen der Studierenden. Diese dient lediglich dazu, in den kommenden Semestern ggf. Verbesserungen vornehmen zu können.

Vielen Dank für Ihr Feedback!

Ihr Lehrstuhl-Team für Betriebswirtschaftslehre,
insbes. Unternehmensrechnung und Controlling

Wie beurteilen Sie, ...

	A: sehr gut	B: gut	C: befriedigend	D: ausreichend	E: mangelhaft	F: keine Beurteilung möglich
... die Art der gestellten Aufgaben innerhalb der Einsendearbeit	☐	☐	☐	☐	☐	☐
... die Verständlichkeit der Aufgaben	☐	☐	☐	☐	☐	☐
... die Bedienungsfreundlichkeit	☐	☐	☐	☐	☐	☐
... die Übungsumgebung zum Kennenlernen der Aufgabentypen	☐	☐	☐	☐	☐	☐
... die weiterführenden Informationen zur Umstellung auf die neue Bearbeitungsform	☐	☐	☐	☐	☐	☐
... die neue Bearbeitungsform der Einsendearbeiten insgesamt	☐	☐	☐	☐	☐	☐

Aufgabe einsenden

Abbildung 9: Evaluation der Umstellung

Insgesamt werden den Studierenden sechs Fragen über die Art der Aufgabenstellung, deren Verständlich- und Bedienungsfreundlichkeit, die weiterführenden Informationen und die Beurteilung der Umstellung als Ganzes gestellt. Die Bewertung erfolgt hierbei anhand des klassischen Schulnotensystems von „sehr gut“ bis „mangelhaft“ sowie dem Auswahlfeld, dass keine Beurteilung möglich sei.¹³

4 Evaluationsergebnisse der Studierenden zur Umstellung

Insgesamt haben 733 Studierende an den verschiedenen Befragungen des Lehrstuhls teilgenommen. Hiervon haben 595 Studierende die Evaluationsfragen beantwortet.¹⁴ Dies entspricht einer Rücklaufquote von etwa 80 %. Die Ergebnisse wurden in Bezug auf die unterschiedlichen Module (Bachelor- und Master) sowie zwischen Prüfungsformaten ausgewertet.

Die ersten beiden Fragen zielen darauf ab, ob die Transformation der Freitextaufgaben von den Studierenden angenommen und als gelungen bewertet wurden. Die erste Frage bezog sich konkret auf die Art der gestellten Aufgaben

¹³ Aufgrund technischer Restriktionen sind systemseitig bei der Gesamtaufgabenerstellung leider keine Freitext-Aufgaben und somit weitere Stellungnahmen der Studierenden zu konkreten Optimierungen möglich.

¹⁴ Nicht jede Frage ist von allen Studierenden gleich oft beantwortet worden. Daher schwankt die Anzahl der Antworten zwischen 584 und 595 abgegeben Stimmen in den einzelnen Fragen.

innerhalb der neu konzipierten Einsendearbeiten. Eine große Mehrheit der Studierenden scheint die gestellten Aufgaben als angemessen empfunden zu haben, da 86 % hier ein *gut* oder *sehr gut* vergaben.

Ein sehr ähnliches Bild ist bei der Beurteilung der Verständlichkeit der Aufgaben zu sehen. Diese ist über alle drei Module mit einem Durchschnittswert von 1,91 positiv beurteilt worden. 36 % der Studierenden waren der Meinung, dass die Aufgaben *sehr gut* zu verstehen sind, die überwiegende Mehrheit hat die Verständlichkeit als *gut* bewertet. Lediglich 6 % der Antworten befinden sich im ausreichenden oder mangelhaften Antwortbereich.

Für die Transformation der Aufgabentypen von Freitextaufgaben in geschlossene Fragetypen lässt sich zusammenfassend für alle Module ein positives Resümee mit geringem Verbesserungsbedarf ziehen. Um in diesem Bereich ein noch positiveres Ergebnis erzielen zu können, erfolgte als Konsequenz u. U. noch einmal eine allgemeine Überprüfung der Klarheit, Stringenz und Verständlichkeit der Aufgabenstellungen durch die Lehrstuhlmitarbeitenden.

Die Bedienungsfreundlichkeit des Online-Systems ist ein vom Lehrstuhl unabhängiges Kriterium, das auf die allgemeine Handhabung des Systems abzielt und dem Lehrstuhl einen Eindruck vermitteln sollte, wie die Studierenden mit der Plattform zurechtkommen. Mit knapp 50 % an sehr guten Bewertungen und einem Durchschnittswert von 1,84 konnte hier ebenfalls ein zufriedenstellendes Ergebnis erreicht werden. Der Lehrstuhl war während der Umstellungszeit im stetigen Austausch mit den Administratoren des Online-Systems und tauschte sich fortwährend über Verbesserungsideen aus.

Aus didaktischer Sicht waren neben der Neukonzeption der Aufgaben als solches auch besonders die Ergebnisse aus dem Bereich der Hilfestellungen und Informationen zur Umstellung interessant. Zum einen betraf dieses die implementierte Übungsumgebung zum Kennenlernen der Aufgabentypen (Frage 4) sowie zum anderen die weiterführenden Informationen wie die Lehrvideos und die Hinweise in Moodle bzw. die Hinweise auf der Homepage des Lehrstuhls (Frage 5). Die Ergebnisse fallen mit Durchschnittswerten von 1,87 zur Übungsumgebung und 1,76 zu den weiteren Informationen ebenso positiv aus wie die zuvor aufgezeigten Resultate. Auffällig ist lediglich der höhere Anteil an Studierenden, die angegeben haben, dass keine Beantwortung möglich sei. Dies lässt sich damit begründen, dass nicht jeder die zusätzlich angebotenen Informationsmaterialien nutzte.

Als letztes wurde danach gefragt, wie die Umstellung der Prüfungsformate in die neue Bearbeitungsform insgesamt beurteilt wurde. Die Ergebnisse können zusammenfassend als Erfolg gewertet werden, da lediglich 5 % der Befragten geantwortet haben, dass sie die Umstellung als *ausreichend* oder *mangelhaft* bewerten; 80 % der Studierenden finden die vollzogene Umstellung *gut* oder *sehr gut*.

Die zusammenfassenden Ergebnisse der Beurteilung zeigt die folgende Abbildung 10:

	Art der gestellten Aufgaben innerhalb der Prüfungen		die Verständlichkeit der Aufgaben		die Bedienungs-freundlichkeit		die Übungsumgebung zum Kennenlernen der Aufgabentypen		die weiterführenden Infos zur Umstellung auf die neue Bearbeitungsform		die neue Bearbeitungsform der Prüfungen insgesamt	
	absolut	in %	absolut	in %	absolut	in %	absolut	in %	absolut	in %	absolut	in %
sehr gut	253	42,59%	214	35,97%	281	47,23%	195	33,39%	233	39,69%	279	47,29%
gut	255	42,93%	263	44,20%	179	30,08%	189	32,36%	211	35,95%	202	34,24%
befriedigend	65	10,94%	79	13,28%	81	13,61%	65	11,13%	51	8,69%	56	9,49%
ausreichend	10	1,68%	25	4,20%	36	6,05%	21	3,60%	16	2,73%	19	3,22%
mangelhaft	5	0,84%	10	1,68%	13	2,18%	8	1,37%	9	1,53%	12	2,03%
keine Beantwortung möglich	6	1,01%	4	0,67%	5	0,84%	106	18,15%	67	11,41%	22	3,73%
	594	100,00%	595	100,00%	595	100,00%	584	100,00%	587	100,00%	590	100,00%
Durchschnittsnote	1,74		1,91		1,84		1,87		1,76		1,74	

Abbildung 10: Gesamtauswertung der Evaluation für die Prüfungsleistungen für die Module

5 Fazit

Die im Rahmen des Projektes ursprünglich angestrebten Ziele konnten weit überwiegend erreicht werden. Die Umstellung der Prüfungsvorleistungen und den Prüfungsleistungen in den Modulen konnte fristgerecht und qualitätsgesichert umgesetzt werden. Die nachfolgenden Prüfungs(vor)leistungen sind mittlerweile mit wenigen Nachfragen und Korrekturen annähernd reibungslos verlaufen. Auch die positiv ausgefallene Evaluation der Studierenden war in diesem Zusammenhang für den Lehrstuhl eine Bestätigung des Vorgehens. Insbesondere die örtliche Flexibilität an der Klausurteilnahme von zuhause aus ist ein deutlicher Vorteil aus Sicht der Studierenden. Durch die ergriffenen Überwachungsmaßnahmen während der Klausur sind die Betrugsmöglichkeiten deutlich eingegrenzt. Auch die bequeme Online-Bearbeitung der Prüfungsvorleistungen sowie eine deutlich verkürzte Zeitspanne bis zur Ergebnisbekanntgabe sind als Vorteile zu nennen. Entsprechend ist auch der Arbeitsaufwand für die Korrekturen mit der Umstellung gesunken. Den Studierenden steht weiterhin die Übungsumgebung zum Kennenlernen der verschiedenen Aufgabentypen zur Verfügung. Gleiches gilt für die erstellten Informationsvideos.

Hingegen ist die Konzeption der Prüfungsaufgaben im neuen Format herausfordernd, wenn ein angemessener Schwierigkeitsgrad erreicht werden soll. Insbesondere die Transferleistungen sowie aufeinander aufbauende Aufgabenstellungen dauern in der Erstellung mitunter länger und können nicht an die Freitext-Aufgaben-Abfrage heranreichen.

Bezogen auf das Projekt ist in der Rückschau insbesondere die zeitliche Planung nichtzutreffend gewesen. Die einzelnen Teilelemente und der Abstimmungs- und Koordinationsaufwand waren deutlich über der ursprünglichen Ressourcenplanung. Für ein einheitliches Vorgehen der Umstellung der Prüfungs(vor)leistungen am Lehrstuhl und für die Klärung von offenen Fragen wurden in diesem Zusammenhang zum Beispiel Arbeitsanweisungen/-hilfen angelegt und regelmäßig aktualisiert. Die Funktion eines zentralen Koordinators und Ansprechpartners am Lehrstuhl hatte dabei den Vorteil, dass sich sehr tiefgreifend mit den verschiedenen Systemen auseinandergesetzt wurde und so umfassende Kenntnisse in verschiedenen Systemen gewonnen werden konnten.

Literaturverzeichnis

- Bandtel, M.; Baume, M.; Brinkmann, E.; Bedenlier, S.; Budde, J.; Eugster, B.; Ghoneim, A.; Halbherr, T.; Persike, M.; Rampelt, F.; Reinmann, G.; Sari, Z.; Schulz, A. (2021): Digitale Prüfungen in der Hochschule. Whitepaper einer Community Working Group aus Deutschland, Österreich und der Schweiz. Berlin.
- Budde, J.; Tobor, J.; Beyermann, J. (2023): Hochschulforum Digitalisierung – Blickpunkt digitale Prüfungen, Berlin.
- Berger-Grabner, D. (2016): Wissenschaftliches Arbeiten in den Wirtschafts- und Sozialwissenschaften, 3. Auflage, Wiesbaden.
- Dittler, U.; Kreidl, C. (2021): Einleitung- Covid wirbelt die Semestervorbereitung durcheinander. In: Dittler, U.; Kreidl, C. (Hrsg.), Wie Corona die Hochschullehre verändert – Erfahrungen und Gedanken aus der Krise zum zukünftigen Einsatz von eLearning, Wiesbaden.
- Erpenbeck, J.; Sauter, S.; Sauter, W. (2015): E-Learning und Blended Learning – Selbstgesteuerte Lernprozesse zum Wissensaufbau und zur Qualifizierung, Wiesbaden.
- Fernuniversität Hagen (2020): Prüfungsordnung für den Bachelorstudiengang Wirtschaftswissenschaft an der Fernuniversität Hagen vom 20.12.2006 (in der Fassung vom 20.05.2020). https://www.fernuni-hagen.de/wirtschaftswissenschaft/studium/bachelor_wiwi/pruefungsordnung.shtml. Abgerufen am 12.07.2024.
- FernUniversität Hagen (2024): Zahlen und Daten aus dem Wintersemester 2023/2024. <https://www.fernuni-hagen.de/universitaet/zahlen.shtml>. Abgerufen am 07.07.2024.
- Helpdesk Wiwi (2020): Online-Übungssystem für Wiwi-Studierende. https://wiki.fernuni-hagen.de/helpdesk/index.php/Online-%C3%9Cbungssystem_f%C3%BCr_Wiwi-Studierende. Abgerufen am 23.01.2023
- Kerres, M. (2018): Mediendidaktik – Konzeption und Entwicklung digitaler Lernangebote, 5. Auflage, Berlin, Boston.
- Littkemann, J. (2021): Hinweise zur Umstellung der Klausuraufgaben. <https://www.fernuni-hagen.de/controlling/lehre/module/index.shtml>. Abgerufen am 10.07.2024.
- Pellert, A. (2020a): Die Revolution des Lernens hat begonnen. <https://www.faz.net/aktuell/feuilleton/fernuni-rektorin-ueber-corona-und-die-revolution-des-lernens-16806312.html>, Abgerufen am 13.07.2024.
- Pellert, A. (2020b): So gelingt uns die digitale Transformation. <https://www.hrk-nexus.de/material/blog/so-gelingt-uns-die-digitale-transformation-prof-dr-ada-pellert/>. Abgerufen am 13.07.2024.
- Reinmann, G. (2013): Didaktisches Design, von der Lerntheorie zur Gestaltungstheorie. In: Ebner, M.; Schön, S. (Hrsg.), L3T - Lehrbuch für Lehren und Lernen mit Technologien, 2. Auflage, Berlin.
- Trempp, P.; Eugster, B. (2020): Klassiker der Hochschuldidaktik? Kartografie einer Landschaft, Wiesbaden.
- Ulrich, I. (2016): Gute Lehre in der Hochschule, Wiesbaden.
- Wm³ Weiterbildung Mittelhessen (2017): Systematisierung und Validierung von Studienmaterialien. https://www.wmhoch3.de/images/Systematisierung_und_Validierung_von_Studienmaterialien.pdf. Abgerufen am 13.07.2024.
- Wollmann, H. (2020): Entwicklungs- und Konfliktlinien des Bildungssektors im deutschen Föderalismus – Neue Dynamik durch Digitalisierung in Zeiten des Coronavirus? In: Roters, W.; Gräf, H.; Wollmann, H. (Hrsg.), Zukunft denken und verantworten, Wiesbaden.

Gamified financial education application for young adults

Lehrprojekt

Manuel Szecsenyi, MSc.

Innovation Process Technology AG, Zürich, E-Mail: manuel.szecsenyi@gmail.com

Prof. Dr. Max Monauni

Hochschule Luzern, Wirtschaft/IFZ, Zug, E-Mail: max.monauni@hslu.ch

Abstract

Teenagers and young adults do not feel prepared enough for their future when it comes to financial education. One of many topics which pop up in studies among Austrian students is the lack of knowledge about costs. Additionally, there is a push to increase financial literacy in the country and organizations are trying out different approaches. A gamified approach that teaches the target group is different from existing quizzes and presentations. Additionally, this topic has not been covered in that detail by other games.

The goal of this study was to create a prototype that teaches children about the costs and events in life, related to financial literacy. The following research question is asked: How can a gamified application increase the financial knowledge of teenagers between the ages of 13-18 and what could such a solution look like?

To answer these questions multiple methods were used. First a systematic literature review answered which gamification pattern can be used in the prototype. In addition to that the state of financial education in Austria and the knowledge gap were identified which set the base for an expert interview to answer which literacy topics should be included. After designing the prototype according to the gamification patterns, an expert on educational games reviewed the progress and gave further advice. Finally, after developing the game, three classrooms tested the application and gave feedback on it.

The Octalysis framework by Yu-kai Chou was used and provided an excellent base for the game. The topics introduced were costs of living, salaries, work-life balance, jobs, necessary education needed, and insurance. There was a noticeable difference in feedback between the classes, with each having different priorities. It is highly recommended for other games in that area to take the personality aspect into the application. Educational games should try to pack the learnings in a story with elements the target group can relate to. Competition between peers should not be understated and can have a very positive effect. In combination with educational games, a follow-up discussion on the topic with the classroom is advised.

Further research can build on the existing prototype or can explore teaching methods to use in combination with gamified approaches. Alternatively, the used framework cores can be further tested in that scenario to find the strongest core within each target group.

1 Introduction

1.1 The current problem

In a young adult survey by the “YEP” (Youth Empowerment Participation) in cooperation with the ERSTE Financial Life Park 50% of teenager aged 14 to 18 think they are not sufficiently prepared when it comes to financial literacy (YEP - Stimme der Jugend, 2021, p. 10). This also includes topics like taxes, banks, and contracts e.g., carriers or rentals (YEP - Stimme der Jugend, 2021, p. 16). This might lead to prioritizing different budgets wrong or underestimating the real cost of living before moving out.

There are existing companies and games on the Austrian market whose goal is to increase financial education. The mechanics of these are mostly in a quiz format and do not cover a full-life simulation in a gamified format (*FLiP Challenge*, 2023). Educated young adults, might make better financial decisions, and can make a positive impact on a country’s economy. For this reason among others, the government of Austria is working on implementing a national financial literacy strategy. (Vanessa Koch et al., 2022)

When it comes to the topic of gamification, there have been multiple studies on different types of Gamification systems (Koivisto & Hamari, 2019) as well as studies on Gamification in banking to increase customer satisfaction (Prasetyaningrum et al., 2022), other studies describe how gamification can be used in education areas (Nah et al., 2014).

1.2 Goal of this study

This paper addresses the need for financial education among teenagers by merging gamification with financial learning. It involves developing a prototype game with experts to illustrate the impact of financial decisions on teenagers' lives, promoting critical thinking about life choices and their outcomes. Initially, various gamification methods and organizational requirements will be explored. The prototype was then tested with teenagers, and their feedback shaped the final design. Additionally, the paper offers guidelines for creating gamified educational tools. Target audiences include government officials, banks, UI/UX experts, and educators, providing insights into gamification in education. The study was conducted as part of a master's thesis at Lucerne University of Applied Sciences and Arts, from January to June 2023. This summary is focusing primarily on testing and the results, while still giving an overview on the initial research and input that went into developing the application. Full access to the master's thesis including detailed description on research processes and game development is available upon request.

1.3 Research questions

The main research question of this thesis is:

How can a gamified application increase the financial knowledge of teenagers between the ages of 13-18 and what could such a solution look like?

Following this bigger question, the following sub-questions are extracted which the thesis will try to cover and answer:

- What gamified approach makes sense?
- How can selected financial knowledge be packaged and implemented to achieve the given goal?
- Is the application able to increase the knowledge of the target group?

1.4 Structure of the research

The structure of the work goes as follows. In a form of a literature review existing gamification patterns and work are explored and summarized. In addition to that, research is also done on the financial health topic. This includes the definition, the potential of it as well as the current problems and lacks teenagers have associated with it.

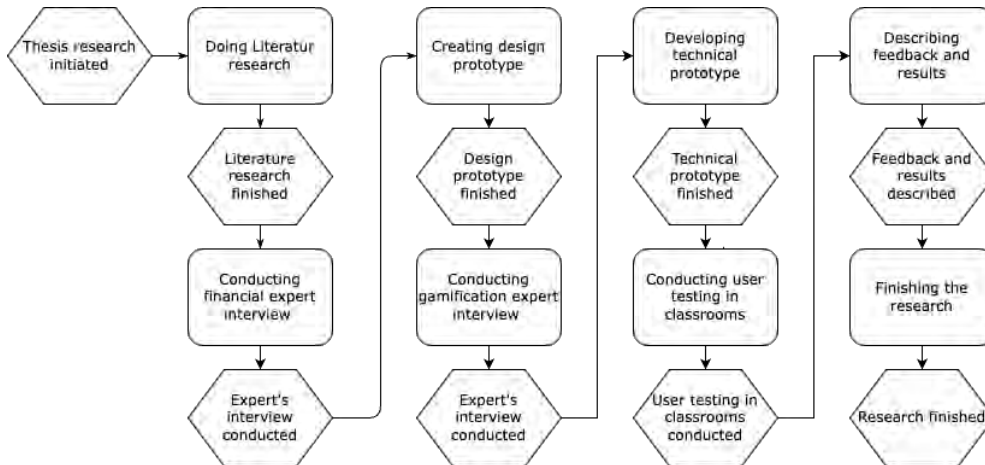


Figure 1 The process of this study is visualized as an event-driven process chain.

The interview with financial experts provides insights into the educational gaps within the target group, alongside lessons learned from existing financial games. These findings shaped the design of a prototype discussed in the following chapter. The prototype is digitally crafted and reviewed by another expert for feedback before the technical implementation. The application was then tested with real students in a classroom setting and their feedback collected and analyzed. A comparison between expected and actual outcomes was then conducted. The project concluded with a discussion encompassing a summary, prospects, and potential areas for further research.

1.5 Methods and expected outcome

This study used multiple scientific methods, beginning with a literature review to gain an overview of the current state of research on gamification and financial health, particularly from the perspective of teenagers. The outcome of this review was a comprehensive understanding of these topics and their current research landscape. The second method involved prototyping, where input from financial experts and the findings from the literature review resulted in the creation of a prototype for further development. Additionally, a qualitative interview with experts and quantitative surveys with pupils were conducted to test and review the app. Expert interviews focused on content and learning style, while user surveys gathered quantitative ratings from the target group. Overall, these methods contributed to the evaluation and refinement of the prototype.

All methods used including their expected outcome can be found in Table 1 Overview Method <> Outcome:

Table 1 Overview Method <> Outcome

Method	Outcome
Literature review	Overview and research state of Gamification patterns and financial health for the target group
Prototyping	A testable artifact
Interview	Expert's input on financial literacy. Expert's feedback on the design and UX of the prototype.
Survey	Targets group feedback on the prototype and game concept of the application

2 Research

2.1 Literature review

2.1.1 Procedure of the review

For each topic one specific goal was set. On the topic, of gamification, an overview of the usage and patterns should be found. The financial education part should highlight topics and knowledge gaps with a priority on Austria. A systematic literature review approach was used to get the current state of research and necessary information to lay a foundation for this study. Following previous approaches which recommended defining time period, keywords, and “Forward/Backward search”, the process was documented (Simons et al., 2009).

2.1.2 Source

For the research, the search engine Google Scholar was used. Only the first ten results (first page) & free to access studies were noted. The access to platforms was granted over a license of the Lucerne University of Applied Science and Arts. It was conducted in February 2023 and results were filtered to not be older than five years. In addition, only German and English articles were considered.

2.1.3 Terms & keywords

The keywords used were written in English and German. The following table shows the keywords to find gamification patterns and implementations correlated to the financial sector for further analysis. Additionally, the number of search results found.

Table 2 Keywords used in the research for gamification patterns and implementations.

Keyword in English	Results found	Keyword in German	Results found
Gamification	159.000	Gamifizierung	694
Gamification patterns	45.900	Gamification-Muster	2
Gamification frameworks	48.300	Gamifizierung Frameworks	430
Gamification in education	123.000	Gamifizierung im Bildungswesen	754
Gamification banking	18.000	Gamifizierung Banking	64
Gamification trading	21.100	Gamifizierung Trading	121
Total	415.300	Total	2.065

Next, papers, studies, and reports were researched related to financial education. The goal was to identify topics of financial education for young adults with a focus on Austria. As before, keywords both in English and German have been chosen. Next, the filter for languages and the five years age limitation was considered. After applying the filter, the first 10 studies for each keyword have been further analyzed. Leaving the amount of literature to 111 sources on the topic of gamification and 200 sources for financial education. As a next step, both languages have been combined and duplicate literature has been removed. Then the literature was scanned for relevant information. As a last step relevancy of the topic was defined. As for gamification research, relevant patterns or frameworks are to be searched for. Relevant literature had included those terms or were investigating the effect of those patterns. Regarding financial education, topics relevant to teach to the target group or giving an overview of financial literacy knowledge were filtered. Studies found, which also included knowledge about gamification in education, were also marked as relevant in both categories.

When reading the filtered sources, the references were also included and noted while studying. With the researched literature an overview of gamification and financial education in Austria was created, which served as a base and input for the prototype and further interviews.

Figure 2 The sourcing process in an overview, shows the process of filtering. The mentioned backward search is also listed there as a last point.

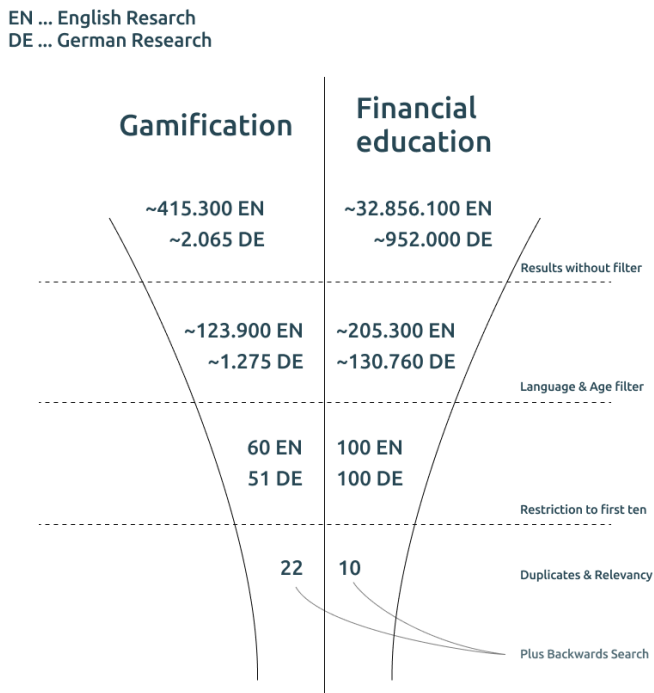


Figure 2 The sourcing process in an overview

2.1.4 Research outcome Gamification

Because of the significant mention in the literature, the Octalysis Framework by Yu-kai Chou was used when designing the prototype. Other motivators and research are also overlapping with this framework and therefore will also be partially applied. Nevertheless, the decision has been made, to focus on the gamification aspect by aligning with the framework. Later in the study, the different core drives were analyzed and applied to the game. Finally, each applied drive will be examined after testing, to see its impact.

2.1.5 Research outcome financial education

At the time of writing this thesis, Austria is working on implementing a national financial literacy strategy. Four main priorities are set in the agenda:

1. Develop base knowledge about financial education to avoid over-indebtedness by making sound financial decisions
2. Promote long-term financial planning
3. Raise awareness about financial education and ensure access to quality education for all
4. Increase the effectiveness of financial education through better coordination

(Vanessa Koch et al., 2022, p. 4)

Financial education surveys have shown that Austrian adults have a lack of understanding when it comes to withstanding “financial shocks”. Meaning being able to overcome negative financial. Teenagers between the age of 13 to 19 in Austria. 85% of which are between 14- and 18-years old feel that, financial safeness is one of the factors they see as essential to a good life. 47% say they are not well prepared for life. One of the reasons for that is a lack of financial health. Also, they are missing information about bureaucracy like how much tax needs to be paid. As far as life skills go, they say that handling finances, taxes, and income is one of the valuable skills. (YEP - Stimme der Jugend, 2022, pp. 1, 4, 6, 8)

2.1.6 Financial education interview

To include the most valuable topics and learning in the prototype, an interview was conducted with experts to find overlaps between research and expert opinions. One expert, Nina von Gayl, is a curator at the Erste Financial Life Park (FLiP), responsible for presenting topics and developing new ideas for financial education. The second expert, Philip List, has 25 years of experience with Erste Group and is the director of FLiP, responsible for infrastructure, budgeting, and resource allocation for new ideas.

Both experts agreed that budget planning is the main topic of financial education. Nina emphasized the importance of understanding one's monthly budget as a foundation for financial development and social life participation. Philip highlighted the need to incorporate real-life perspectives into financial education games, noting the challenge of applying game-learned steps to real life.

The experts identified several problems in current financial education. Students at “Handelsakademien” (HAKs) often lack practical knowledge about investments and real-life applications of financial concepts. There is also a significant lack of understanding of risk, with young adults often attracted to high-risk investments like cryptocurrencies. Additionally, Philip pointed out the issue of "Buy now, Pay later" schemes, which can lead to prolonged debt.

Regarding the target age group for financial education, the experts believe it should start with the first pocket money children receive and continue throughout life. FLiP mainly targets 14 to 17-year-olds, as this age group is still in compulsory schooling and can be significantly impacted.

The experts found it challenging to specify which financial knowledge is lacking, stating that students' knowledge is often inhomogeneous. They aim to provide a broad "toolset" of financial knowledge, including topics like surety bonds, to prepare teenagers for the future and prevent financial issues.

Frequently asked questions from students often relate to current media topics like cryptocurrencies. Other questions are selective and personal, such as inquiries about real estate loans.

In terms of teaching methods, the experts use a blended learning approach. While children use iPads to solve tasks, they can also seek help from an accompanying person. Nina noted that the methodology does not always need to involve electronic devices; it should be tailored to the target group. For young people, using devices and designs familiar from daily life can effectively convey content playfully.

Additional insights from the experts included the importance of addressing reading comprehension issues. FLiP has developed formats with more visual content to aid those who struggle with text assignments. They also found that incorporating competition into activities can be positively received by teenagers, contrary to their initial design of FLiP as a non-competitive safe space.

3 Prototyping

With the results of the literature and the interview, it was possible to design and plan the prototype.

3.1 General idea

The game aims to prepare young individuals for real-life decision-making by simulating daily choices that impact their finances and overall well-being. Players begin by personalizing their profiles and then face various decisions, each affecting their budget and quality of life. Real relatable values are used to enhance realism and relevance to players' lives. Success in the game is measured by two indicators: money balance and life points, indicating overall satisfaction. This two metrics encourages players to consider not only financial but also personal health when making decisions. The goal is to achieve a positive balance between money and life points, leading to an understanding of costs and budgeting through gameplay.

3.2 Connecting the financial knowledge

With implementing relatable values, the problem of missing knowledge about real-life costs, is included. Users can learn about real-life expenses and financial obligations. During the onboarding of the game, jobs and academic paths will be shown. Helping them explore potential careers and giving them an insight into the salaries associated with them.

Also, by implementing insurance, players might have first contact with that topic. Helping them make informed decisions later based on their experience during the game.

The prototype simulated unexpected events like dental emergencies or a broken washing machine. Teaching about the importance of being prepared for unforeseen expenses and again showing prices for real-life events. This way users should be taught the significance of setting aside the money, ensuring they have a financial safety net. The prototype shall not discourage the students from spending any money. With the implementation of the life points, it rather shall teach about making better financial choices and weighting their decisions.

3.3 User journey

For a better overview of the different parts of the game, the user journey is divided into three sections:

- 1) Onboarding
- 2) Game phase
- 3) Review

In the first phase, players should set the foundation for their digital life. After an initial personalization, they decide on an education. Depending on the training, certain professions are reserved for them. The next selection steps determine how much monthly budget is left, as well as the life points being deducted. The choice of insurance has an influence on certain events and could limit the costs for water damage, for example, to a certain deductible.

After selecting the basic parameters, the actual game begins. Users are confronted with various questions and events. After a defined number of events, a game year is completed. The players then have the possibility to invest unused assets in three different areas. The money can be saved to build up a financial buffer and to be better prepared for future events. Alternatively, some of it can be exchanged for life points. This aspect of the game is meant to simulate investing in activities and leisure time. In real life, this could be a vacation, for example. Finally, players could invest in themselves in the form of further education. In this game concept, it is assumed that investing in oneself has a positive impact on one's profession and therefore on one's salary. In summary, players can keep the money, exchange it for life points or invest it in a better salary.

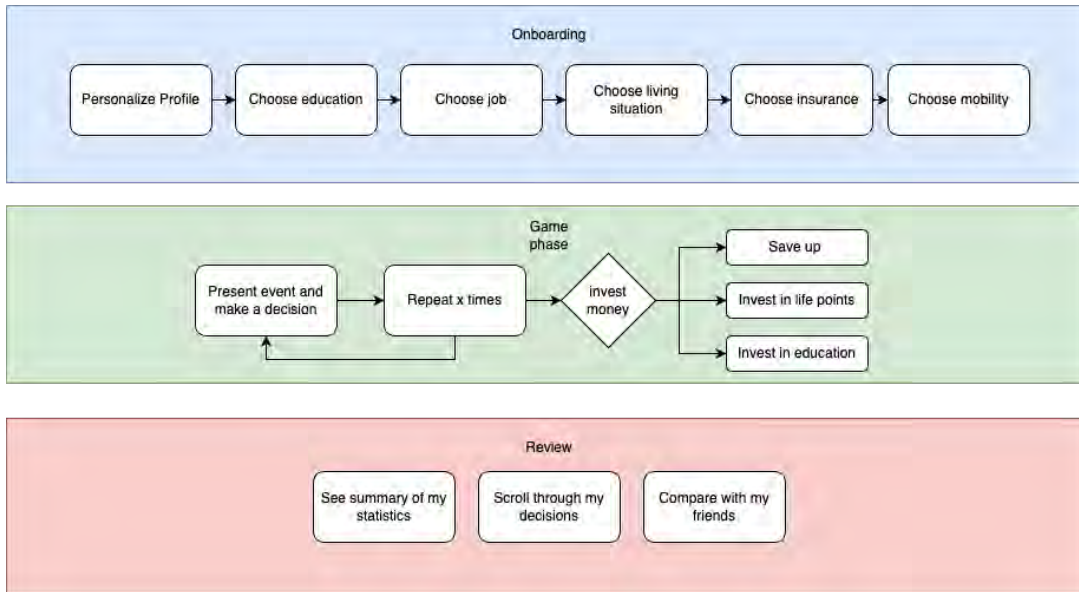


Figure 3 The three main processes in the game

Lastly, the prototype should display a summary of the game. This summary shall show the result of fortune and life points and present the recorded decisions. The goal is to allow students to reflect on their personal journey and share the outcome with their peers.

3.4 Applying the Gamification Pattern

The design of the game is based on the Octalysis framework and try's to implement as many core drives as possible.

With the help of the player statistics, it was tried to implement Core Drive 2. The users have the possibility to compare each other based on the assets or the life points.

Core Drive 3 is about the creative approach and how players can try out different combinations. By personally choosing certain starting criteria and then reviewing the decisions, students can try to better understand the impact of their actions.

Through personalization and the decisions that can be made individually, Core Drive 4 was attempted to be met. In a further development stage of the prototype, this drive can of course be fulfilled even more by further personalizing the avatar and having this personalization have an impact on the events.

Core Drive 5 is not strongly perceived in the prototype because the game can only be played alone in the first planned version. However, due to the points mentioned in Core Drive 2, a digital comparison with classmates could be incorporated in a later phase.

The themes of Core Drive 7 are implemented with unpredictable events for which the players can individually prepare themselves with insurance or cash set aside.

Lastly, users can of course lose the game if their expenses are greater than their income or their life points have fallen below a certain threshold. Core Drive 8 revolves around this avoidance and is attempted to implement it through this gameplay technique.

No game mechanics were found for Core Drives 6 and 1.

3.5 Wireframe

As a first step, the idea of the prototype was designed using the interface design tool “Figma” (Figma, 2023). In this first round of designing, the application was created as a wireframe. Meaning the flow of the game is built and not fully designed visual elements are placed on the canvas. “Figure 4 Wireframe of the player’s customization” shows the full process the players should walk through when starting the game. First, the player can customize the user with a name and generic character image. This was done to follow the “Ownership” Core Drive 4 from Chou Yu-kai. The goal of this, is that the player should feel more ownership towards the character (Chou, 2014, p. 26). In this first prototype, this is limited but could be extended in the future.



Figure 4 Wireframe of the player’s customization

For answering the questions, a “Tinder”-like approach has been developed. Laine and Lindberg list “familiar, safe, and comfortable controls” as one of their researched motivators (Laine & Lindberg, 2020, p. 8). It is taken as a hypothesis here, that young adults are familiar with the swipe UI. This approach makes sense, as the design is primarily designed for Boolean type of question: Yes or No. The user will have the opportunity to read more about the topic. However, it should not be the focus and should not interrupt the flow of the game.

With the prepared Wireframe it was fast and easy to understand and construct the game flow that the users will navigate through.

3.6 High fidelity

Before going into development, a high-fidelity prototype was designed within Figma. The primary goal of this stage was the option to conduct a second expert’s interview on the design of the application. In addition, it also helps when developing, as the focus can be on the development and not the visual part.

Figure 5 gives an overview about of visuals defined. For the user customization, a simple avatar image is used. The “Dicebear Adventurer”-API was chosen as it provides a great number of avatars that can be accessed over an API. Consistency can be insured by using a seed (Adventurer | DiceBear, n.d.). When choosing the images, it was considered that either the number of male characters equals the female ones or that they look neutral. Additionally, a great range of skin tones was important to include.

Each event should include a large colorful image to balance with the text-heavy content. Streamline Icons provides a great range of different packs. The “UX Colors” icon pack was chosen because it matches the style of the application. On the other hand, with around 9.350 different icons (the state of March 2023), the problem of finding a fitting image for an event seemed small. (Streamline UX Colors Illustrations Set, 2023)

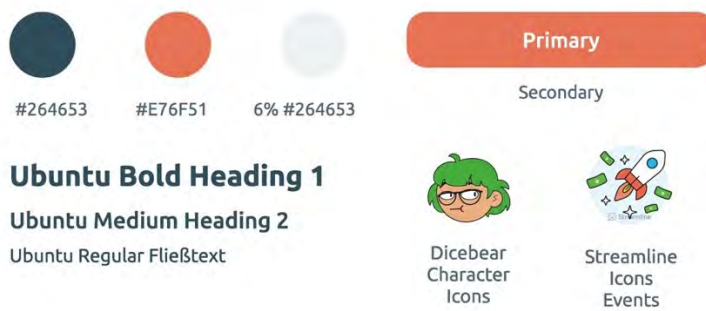


Figure 5 Design overview of the prototype

After applying the defined design guidelines and adding images a few further design modifications were being made. The finished design of the game, as well as the newly designed game summary, can be viewed below (Figure 6).



Figure 6 Two screenshots from the fully designed prototype. Left: The summary after onboarding. Right: The game view.

3.7 Expert review

Before going into developing the application, another expert interview was conducted. The goal was to get professional feedback on the design, find hidden flaws, compare the design process, and consult about the flow and idea of the game. The interview was with Joerg Hofstätter, owner of the agency "ovos,". "ovos" specializes in creating playful training tools and web solutions, and they have developed financial literacy apps for various clients, including Erste Bank and the Swiss National Bank. Besides giving insights about several best practices about developing and testing such application, feedback on the game was given.

Suggestions include adding an introductory splash screen, narrative sequences between decisions, and a coherent storyline to enhance engagement. Improved event descriptions and the use of videos or AI-generated audio could be more engaging than text.

Additionally it was mentioned that the target group of 14-18 years old is challenging due to high expectations from commercial games. Allowing failure in the game is crucial for safe experimentation and Post-game discussions enhance knowledge transfer. Overall, the expert emphasized the importance of making educational games engaging and relatable, using diverse approaches tailored to different player types.

3.8 Final adjustments

After the positive feedback from the expert interviewer, minor things were adjusted and added. First, before onboarding, a missing instruction screen was added (Figure 7). It shows the most important game parameters and is visually supported by again streamline icons. The navigation between these screens shall be implemented with a slider or button.

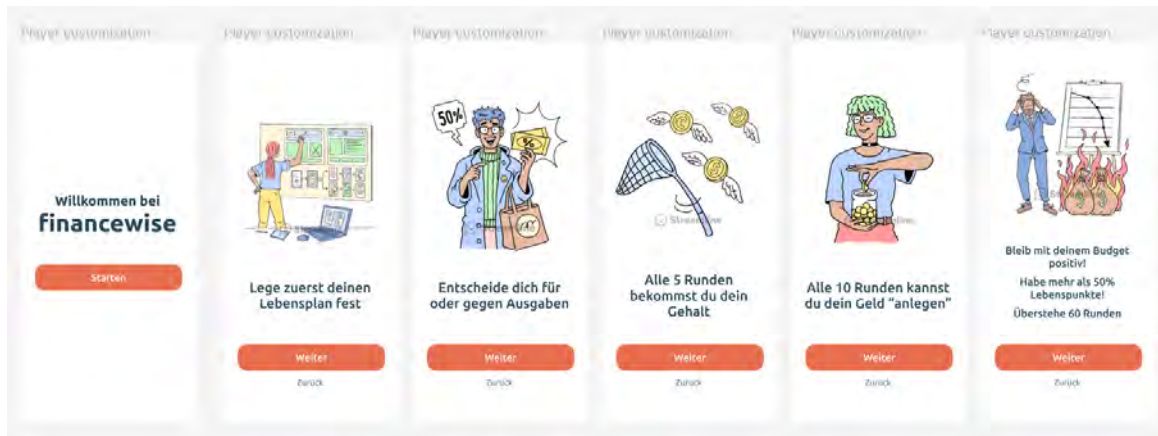


Figure 7 After the expert's interview a missing introduction screen was added.

Secondly, the decision was made that the events will not be random but in the same order for everybody. This will enable the users to exchange and compare themselves to their peers. In the hope of opening discussion about financial decisions in the classroom and between students.

Next, when designing the decisions, the users will encounter the number of unprecedented events shall be increased. This should make the game a bit harder and easier to fail. Giving the users the safe space to try out and fail to try out different paths.

Also, the expert gave feedback to might add more storytelling around the decisions and events. This idea will not be completely done in the first iteration of the prototype due time limitation. Nevertheless, the text of the events will be adjusted to better tell stories around the given decisions.

Finally, as the interviewee pointed out the difficulty and expectations the target group has, the goal of the testing will be directly communicated. At the beginning of each testing session, the classroom will be enlightened about the fact, that this will be an educational game.

4 Technical build

It was decided to create a web app. This type of application is easier for the testing sessions and gives almost no limitations in platforms as the users are just able to open a webpage to start the app. A native or cross-platform solution would require the users to download the application first, which may also be problematic when installing the prototype over school tablets.

As an underlying framework, Angular was chosen. Being among the biggest JavaScript frameworks it provides a great base for the prototype (Most Used Web Frameworks among Developers 2022, 2022). For simplicity reasons there is no additional backend. The whole application runs on the client side and publishes no data about the game.

A detailed description about the technical aspects of the project, including the process of constructing the technical prototype, can be found in the original thesis to this study. It outlines the technical stack used, and highlights the obstacles encountered during the development process.

5 Testing

The “FLiP” organization offered the opportunity to connect with teachers who are interested in trying out new games with their students. To do that, a tender had to be published which was sent via a newsletter. The target group for the testing was as follows: Students aged 13-16 in the Vienna area. The German tender used can be found as an Appendix.

The public tender unfortunately received no answers. Therefore, all three tests were conducted in one school on two different dates.

5.1 About the students

The three tests were conducted in the “Private Mittelschule St. Marien” in the Liniengasse 21, 1060 Vienna. Table 3 gives an overview of the students that tried out the game. The average age was calculated with a median and the number of students represent the number of children in the classroom, on that day, trying out the game. The classroom variable A, B, and C do not represent the name of the class in the school but rather was used to differentiate between the classes within this thesis.

Table 3 An overview of each classroom

Classroom	Date of test	Grade	Age Average (Median)	Number of students	Duration
A	3rd May	3rd	13	16	1h 30m
B	12th May	3rd	13	19	1h
C	12th May	4th	14	19	1h10m

5.2 Planning the testing

Before presenting to the class, the prototype was intensively tested multiple times to not stumble upon blockades when being in class.

The plan was to give a proper introduction to the game, give them time to test out the prototype in detail, get feedback and then discuss the learnings. A presentation was created with Google Slides. A playful design was selected, and gifs were added to the slides, to remove the formal aspect of the experiment. As the children will get to know the game and the presenter for the first time, it was important to earn their trust as fast as possible. It was assumed that, with a trusted presenter, the children would be open with their feedback and communicate their honest expectations. A screenshot of all slides can be found as an appendix to this thesis.

It was important to mention to the audience, that the game is about financial education and what goal the creation of the game was. This was planned due to the feedback from “ovos”, to better manage expectations. Afterwards, the four-game mechanics should be introduced: Character Creation, Event decision, Investment, and Reflection.

The mechanics were also shown with screenshots so that the users will know what to expect from each screen. The introduction of the game shall only take 2-5 minutes and was planned on purpose. This way it should be tested how intuitive the experience would be. To start the game, the children had to navigate to a web page. In classroom A, a few users had a problem typing in the URL and instead looked on Google, where the game could not be found. Because of this, a QR Code was added for classrooms B and C.

For the feedback a “mentimeter” was created. Five questions were chosen that the children could decide on a range of 1-10 with 1 meaning “Don’t agree” and 10 “I agree”. The questions are, translated from German to English:

- I liked the game
- The game was too difficult

- I learned something new
- I would recommend the game to a friend
- I would play the game again

After answering the four questions the students had the chance to answer the question “What was new?”, openly. The answers would be displayed as a “word cloud” on the screen, allowing discussions or clarification questions.

Finally, the children had the chance to openly express further feedback. It was expected that they would ask questions about the game, give new ideas or tell what they missed while playing. The students saw the questions of everybody and had the chance to upvote them to give them a higher priority. The whole feedback process was anonymous and could be performed without a child raising their voice. It was decided to choose this design as still after playing the game, there might not be the needed trust established, which is needed for honest feedback. Shy students could raise their feedback without being pointed out or having to speak in front of the class. Still, the whole session would be moderated and openly discussed, and speaking out was encouraged.

After getting the feedback an open-end discussion would take place. The exact topics of this phase would not be planned exactly as questions and topics might occur during the game phase. Still, two parts would be picked up:

“Did all of you choose insurances? What were they useful for and do all insurances make sense? How did you feel balancing life points and money? Was it easy?”

This last phase is meant to be the most educational one, as the teacher would have the chance to completely take over. As it was not clear at the time of planning how much the teacher would interfere with each class, it was expected to moderate all four phases alone.

For all three tests, a second person would join in class sitting offside and making notes when asked. This way the presenter was able to go to each student, help with problems and focus on presenting.

5.3 Conducting the test

The tests, lasted between one and one ½ hours, and involved a teacher who introduced the experiment. In the first classroom, some students initially struggled to access the game URL, but these issues were quickly resolved. Early finishers, generally above-average students, completed the game within 5-7 minutes. A common issue was failing with “game over” due to a fire event, as students lacked insurance, savings, or sufficient life points. After 20-30 minutes, all students managed to complete a full game. It was noticed that one student struggled to balance life points and money, leading to frustration. Post-game, students compared their scores and experiences with classmates.

In the second session, a QR code was added to slides for easier access to the game. Like the first class, the students had no prior finance lessons. Only two students had difficulty accessing the game, and most started without issues. Some students did not understand the term “Pendeldistance” (commute distance). The rest of the session was like the first, with all students completing the game at least once.

The third class consisted of some older students with prior lessons on money and finance. This class included many repeaters and students needing extra attention. Students showed less eagerness to play and experiment with the game. Some played only one round or gave up after a few attempts. Students spent more money in-game, leading to losses due to financial mismanagement rather than life points. One girl repeatedly lost the game by making the same choices each time. The noise level was higher, with audible reactions to game events. Students actively compared their decisions and outcomes with friends.

Overall, the tests revealed significant insights into student engagement and learning outcomes, highlighting the importance of balancing game elements to maintain interest and ensure educational value.

6 Test results

6.1 Feedback on the game

The five questions on the first slides of “mentimeter” could be answered via a range slider. The players could choose a number from one to ten. Each individual question and their answers will now be analyzed. To do that, it was counted how many students, for example, selected the number seven for the first question. This was done for every question and every number from zero to ten. After that, the counts have been normalized and then put next to each other. This way it can be compared how much percent of each classroom chose the same number. The charts contain the original question in German as a label.

6.1.1 Question 1: I liked the game

Generally, all classrooms liked the game. With all three being either above or slightly under 50% fully agreeing to that. The first classroom A was the happiest with 70% fully agreeing and the lowest number of seven. It is noticeable that in the 4th grade (classroom C), a few students choose a very low number and apparently disliked the experience. The statistics correspond to the observation in class mentioned in **Fehler! Verweisquelle konnte nicht gefunden werden.** that the older students did not seem less enthusiastic than the 3rd graders.

6.1.2 Question 2: The game was hard

Here it can be seen that the game wasn't too hard for most of the players. In comparison to the first question, again classroom A gave the most extreme answer with over 70% totally disagreeing with the statement. It is also noticed that there are also some students in classrooms B and C who thought that the game was slightly too hard and one student in classroom B (3rd grade), almost completely agreed that the game was too hard.

6.1.3 Question 3: I learned something new

With this statement being raised we can see to most Inhomogeneous distribution of answers so far. Because two different grades were present in the experiment, it should be looked at them individually. Around a third of the classroom, A totally agree that they have learned something new. Adding the numbers for seven to nine in addition it can be said that around 65% of the students think that they learned something new. For classroom B less than 50% agree with the statement and did not learn something new. This is interesting as the classroom median age and learning plan is the same.

Looking at the older class it can be seen that almost 40% chose to disagree completely with the statement meaning there apparently was nothing that they did not know before playing the game. Still, around a quarter of the class did learn something while trying out the prototype for the first time.

6.1.4 Question 4: I would recommend the game to a friend

More than 60% of the first class A would recommend the game. The other classes have been more critical with only around 50% saying they would show it to their friends. The average of the votes in classes B and C are the same.

6.1.5 Question 5: I would play the game again

More than 50% of the class would play the game again. This fits with the observation that many students retried after failing or replaying with different parameters multiple times in class. Also, a few students from classroom C were observed during the experiment. They played once and then did not have the motivation to try again. In classrooms A and B the question was raised if the game can be played outside of school as well.

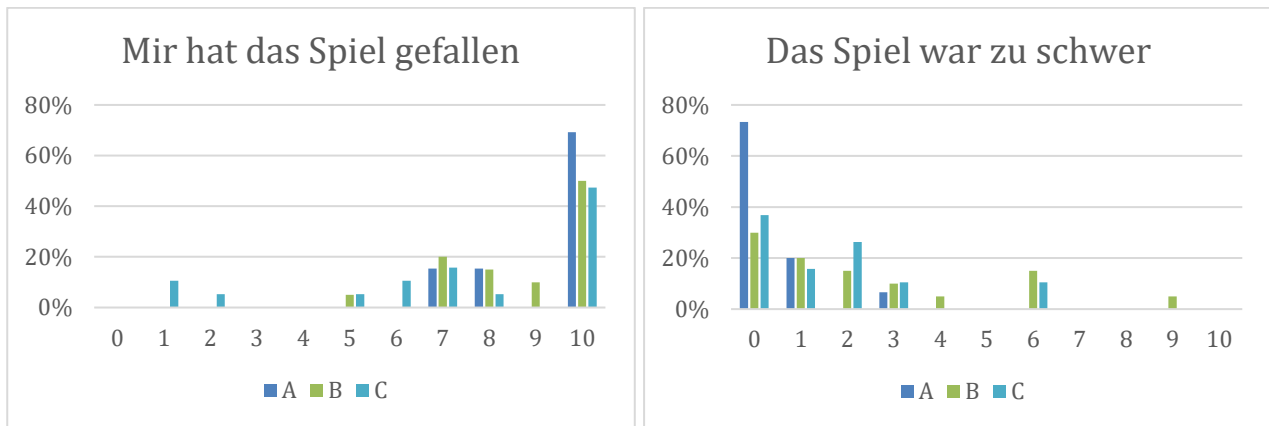


Chart 1 Showing the result of the first two questions

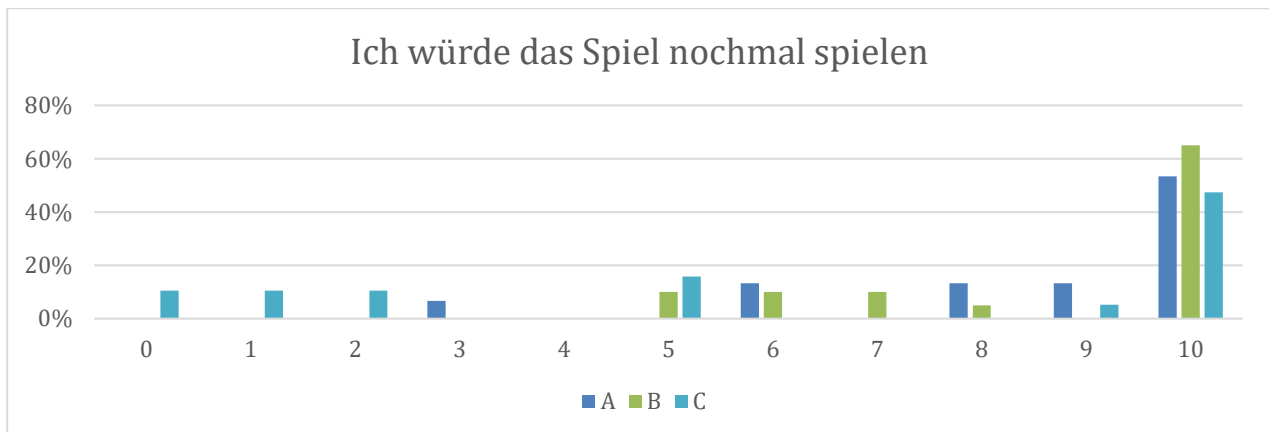
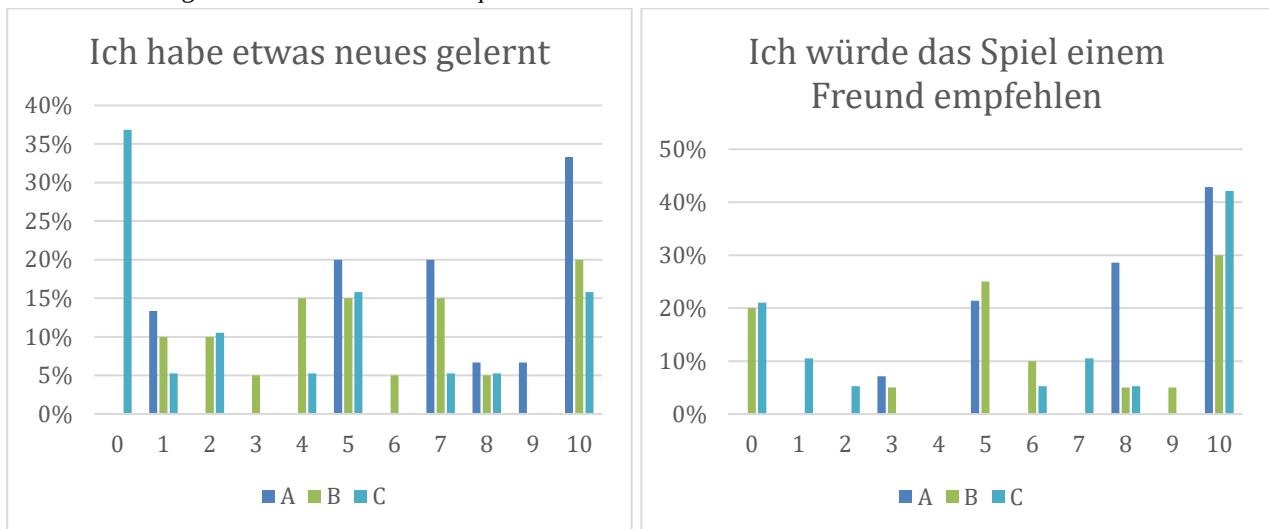


Chart 2 Overview about the last three questions and responses.

6.2 What did the students learn from the game?

The next question was open to answer in the classroom. Every student had the possibility to put what in their opinion was new to them and what they learned. The answer could be put in an open text field and was sent to the board where it was displayed as a “word cloud” for everybody to see.

6.2.1 Raw data

Looking at the raw answers of every class, group A gave the most answers. That although there are fewer students than in the other classes. The appendix shows the word cloud. In the bottom right corner, the number of people participating in the quiz can also be seen. It should be marked that for classroom C some students did not participate in the feedback round as their phone connection or battery gave up.

Looking at the words and data the students sent to the board, the feedback is significantly inhomogeneous. The first class seemed to give rather serious feedback and added many learnings. Class B added emojis but also provided similar learnings as A. The amount of feedback from class C is not only significantly smaller than that of the other classes but also indicates that some children did not encounter anything new and that they already knew all the mentioned terms shown in the game.

6.2.2 Breakdown

To further analyze the learning given by the students, each “significant learning” (SL) of them got categorized into one of seven classes. The definition for a SL is that it must be related to the game. The input “I learned nothing” or something similar is also not counted because the question is “What was new?”. Counting all the SL’s for each word cloud concludes in the following overview.

Table 4 Counting the SL's of each classrooms feedback

Classroom	No. of SL's
A	44
B	22
C	7

Next, the categories have been created. They have been derived from their feedback and are as followed:

Table 5 The seven categories of learning the students discovered.

Category ID	Category label	Examples (Quoted in German)
A	What does something cost?	„die kosten“, „zahnarztkosten“, „wenig geld“, „alles kostet“
B	How to balance life	„mit geld umgehen können“, „money is not everything“, „leben balancieren“, „über konsequenzen denken“
C	Jobs & Studies	„was ein studium ist“, „lehre“
D	Saving money & Cryptocurrencies	„sparen“, „schwer geld sparen“, „notgeld“
E	Taxes	„steuern“
F	Interest rates	„zinsen“
G	Insurances	„versicherungen“

With the categories and SL’s given, a chart has been created which gives an overview of which topics have been mentioned the most by the students.

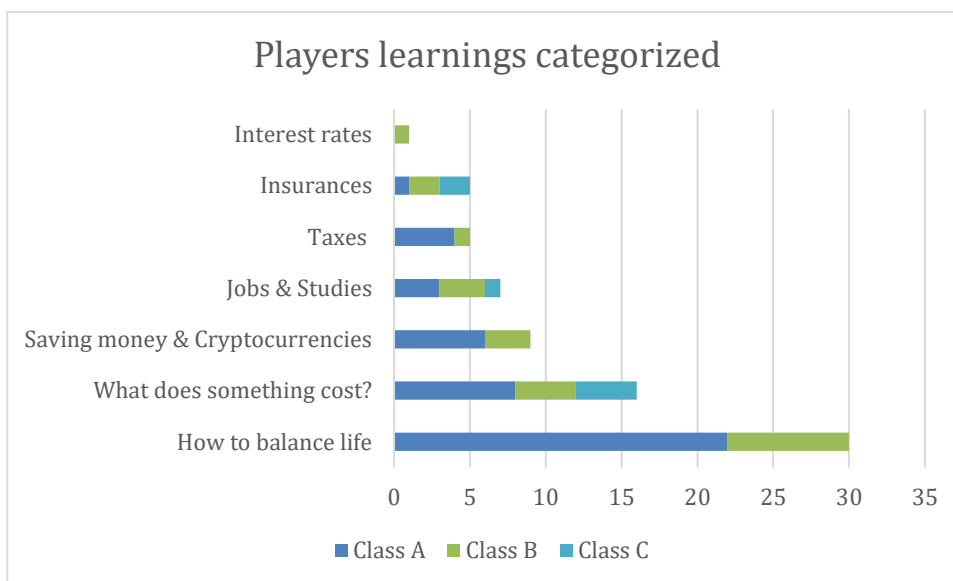


Chart 3 An overview of what the learnings of the players are.

Most learnings provided by the students are related to how to balance their lives. The participants mentioned that they learned how to organize their free time by making wiser choices when buying. Users from class A mentions that one learning was to not buy unnecessary stuff or to not “waste” money. Answers of class B are similar but independency is also mentioned twice. Nobody from classroom C mentions learning associated with balancing. Nevertheless, combining all user's answers, with 30 related learnings this category is almost twice as much mentioned as the next category.

Looking at the second most mentioned feedback is real-life costs. The next category, significantly less, was related to saving up money and cryptocurrencies. The 4th grade seems to be already educated on that topic. Jobs and studies have also been mentioned a few times. Especially the learning “what is a study”, mentioned by classroom A should be pointed out for this category. Learnings related to insurance and taxes are mentioned the same amount of time, but it can be noticed that only the 3rd grades mentioned them. Interest rates have also been mentioned once by classroom B. While presenting, the classroom has been asked where the interest rates popped up in the game. The event around the startup was correctly mentioned by the players.

6.3 Additional feedback

Lastly, there was an open question for feedback. It was noticed that all the classes wished for more jobs that they can choose from. In addition, they were interested in more events and adding pets to the game. The 4th grade additionally asked for more customization for their character. They wished to buy individual clothing, dress their character, and compare with their friends.

Another point that was mentioned was the ability for better comparisons. The wish was expressed that the students would like to have a high scoreboard. Also, without asking they mentioned the starting conditions everybody had. They would be ok with different starting conditions and randomized events. In their opinion “only this way, it would be like real life”.

A note was taken after the feedback from a teacher questioning how life points are deducted each month. Her concern is that work is badly associated as each month a new salary comes in and life points are deducted.

7 Discussion

7.1 The gamified approach

The decision was made to follow the eight core drives (Chou, 2014) when designing the game. As mentioned earlier only the two core drives one and six have not been implemented.

The game was giving the players the option to overcome different challenges throughout the game and offered statistics that they could compare themselves with. This element was openly used by the students in class. As mentioned in the result chapter, they stood up and compared their money with each other. Also, from the observation taken, the mood and excitement in the class were noticed when being faced with difficult decisions.

It can be said that this mechanic was positively impacting the impressions the players had from the game. Core drive two is proving itself to be making sense being implemented in such a scenario.

The positive implementation creativity and feedback which is connected to core drive three, was observed when the students tried out the game for multiple rounds with different parameters. Their motivation of trying again gave them the opportunity to try out different paths in life (getting a degree, choosing the job and apartment) being able to fail.

The personalization aspect of the application was also positively received. The additional feedback given by the students with the wish for more jobs and customization of their character shows, that there was more potential for the game in that area. The implementation of the prototype allows easy additional jobs to be added by just passing an object with defined properties. The wish for more player customization and change of outfits would need a bit of refactoring of the application but would fit the general concept well. This core drive can also be recommended for such games.

The fifth core drive was not strongly met but a wish for a stronger expression was mentioned. With the implementation of a multiplayer, this drive might increase the player's motivation for the game more.

Unpredictable events, which are aligning with the seventh drive, were also positively received by the audience. Especially events that were highly impacting their money were challenging the players, bringing them game over if they haven't saved up enough money or chose the wrong insurance. This avoidance of losing is also part of the eighth drive. The high number of players who replay the game shows that implementing those cores also positively impacted the experience.

It can be said that the positive feedback gathered by the audience, many times fitted with the implementation of the core drives. Meaning the approach of designing the game with these in mind can be recommended for similar applications. In addition to the drives the general game mechanics also worked well. The simple design and experience with just tapping a green or red button to decide also worked well and did not need more explanations for the students.

One mechanic which often was unclear to the users was the investment screen. Generally, the idea of the feature was clear, but while presenting, the students called for help asking what they can now do on that screen. A helpful tutorial or simulation of the investment phase could have resolved that problem. Little helpful texts and popups would be another method to resolve this issue. In a further developed prototype, this screen should be reworked and tested with the users. Generally, it can be said, that educational games should try to make use of such helpful hints to make the game easier to understand. In the best case, the UI should be self-explanatory, leaving the users focused on the content instead of the controls.

The feedback provided by the second interviewee, to add more storytelling to the game is implemented partially as the given timeframe was too short. The emotions expressed by the students while playing and encountering events,

as well as the wish for further character customization show how much the players identify themselves while playing. Adding a story part around that could potentially increase positive feedback and opens possibilities to show how different social backgrounds and starting conditions impact the gameplay and difficulty.

7.2 Packaging the financial knowledge

The financial knowledge which should be taught by the game was selected based on the expert interview as well as from selected studies and reports around Austrian schools. The main challenge the game was trying to solve was to show children real numbers of salaries and costs and how to balance life right. The topics used in the game were also derived from the reports mentioned in the research part of this thesis. With the onboarding phase, children can see how much salary certain jobs can bring them, and the aspect of selecting an apprenticeship or university gives them an insight into the requirements some careers have. Adding more parts to the onboarding phase like apartments and mobility can further educate on the pricing in different aspects of life.

In the interview with “ovos” a “fear-free space” was mentioned. Meaning giving the students the option to try and fail. This very important technique was added to the game on the game-over screen. Players were shown why they failed and can immediately restart. In the testing session, this feature was used by the children a lot. It was observed that they were motivated enough to retry on failure multiple times in a row. However, not all players can work with this strategy and need additional help. In class C a girl was frustrated, that she failed repeatedly. On being asked what she tried so far, she explained that she chooses the same events and strategy multiple times, but always went bankrupt. Educational games could prevent this frustration by adjusting the difficulty dynamically or helping the users with a notification. This way users can stay motivated and learn in their own peace. A connection to a statistical service or a game log can help educators to find which students need help and support them on their way.

The simply designed UI with the “tinder” user experience proved itself well. No further instructions were needed, and no student had questions about the game elements. The players started the game intuitively without a tutorial. Parts of the students finished the game within a few minutes. Except for the investment screen, there were no questions from the classes about buttons or the phase they are in right now. The modern UI and UX seemed to ease onboarding. Familiar button layouts and mechanics help users to start and play the game on their own.

Displaying one decision at a time gave the players the option to fully concentrate on the impact the single event has. There was no option to see what event will come next to prepare. This approach worked well in the given scenario and shifted the player's focus on each decision. Feedback, given by the 4th grade, was to add more pictures or videos to the game. This matches the feedback given in the “ovos” interview, where it was also mentioned that they remove large texts. A game might be more enjoyable for the target group if it would support animations, videos, or even work without any text. This would also increase the supported target group making the game more accessible for users who can't read or have difficulties with it.

Interactive elements and less text play a role and positively impact the user experience. A modern and clean design with a simple and intuitive experience lessens the entry barrier for trying out the game. Players can easily discover the applications on their own without the help of someone else.

The interview conducted with “FLiP” and the “YEP youth report” both mention interactive media or games in combination with traditional presentations or discussions. In the testing sessions, this was achieved by first letting the students play and then discussing the results. The prototype offers the possibility to open discussions about many different financial aspects. By giving the children, the possibility to try it out first on their own, the discussion can be lively as the users experienced the topics just a few minutes ago. Giving them the option to better relate what it means to have e.g., insurance in the case of a fire. The same feeling might be difficult to convey with just a presentation or video where there is one-way communication happening. A board game-like approach could also work out but increases the difficulty in handling. Physical elements would need to be stored, handed out, and limited in a classroom of unknown size.

7.3 Increasing the target groups knowledge

One of the questions this research is trying to cover is if the target group's knowledge could be increased with the application. To test this with the first prototype, the feedback was used in the end of the game. One of the two questions that are related to the research is a self-rating of the players if they learned something new. Here they could rate on a scale from one to ten if the statement "I learned something new" is correct (see 6.1.3). About half of class A says they have learned something when playing the game. Also, class B with around 40% agrees on that. With the 4th grade, only a ¼ learned something.

More tests with a greater number of classrooms have to be deducted to have a clear statement on this and to further limit the bias, mentioned by "ovos". Nevertheless, from the tests that have been done, it can be said that the prototype performed well for around 50% of the children in 3rd grade. 4th-grade classes have most of the topics already covered in the class and did not learn something new ($\approx 25\%$). However, it doesn't mean the game cannot be used in these age groups. To summarize, around 50% of the younger users declare that they have learned something new with this concept. By using the same approach in another age group, this number halves.

Next, with the additional question about the specific topics the children have learned, it can be evaluated if the learning goals have been achieved. The results mentioned in 6.2.2 show the user's feedback when they were being asked what they learned from the session. The majority of the users have stated that the topics around balancing life and money was new to them. Following this category, the second most mentioned learning was about costs in life. This matches the game's target of educating the students about this aspect in life. Saving up was displayed with unpredictable events like a broken washing machine and has also been mentioned by the children. However, the number of mentions halves. A reason for could be that many of them already know that they need to save money, but just didn't know how much because they had no numbers to relate to.

The topic of cryptocurrencies was also mentioned by a few users, but not like planned. The goal was to show the risk of these investments. However, the feedback to that was that they did not know what this term means. Resulting in clicking the decline button and not interacting with it. Introducing new topics and terms on the side might not achieve the best result. They seem to interrupt the game flow and users try to get past it as fast as possible. Topics that should be taught in the game might perform better if they are in the center of the experience or follow the player around. A way to implement learnings about cryptocurrencies in the existing prototype could be to add a portfolio that they can manage or the option to also invest in them on the investment screen. This way the users must interact with it and decide about the topic more than once.

Closely followed to that are mentions about the "Jobs & Studies". An observation was made here, that not all children chose to not go the "easiest" way of selecting a study and the highest earning job. This indicates that the users did not try to aim for the best score but tried to mimic their life best. Supporting this theory is the fact that one of the most wished improvements is more jobs in the game. Nevertheless, the goal was reached to teach about the earnings of different jobs and what other possibilities a university degree offers. Educational games might work better when they are relatable for the students, referencing to their wish of further customization noted in 6.3.

The topics of insurance and taxes have been mentioned by a handful of students. While giving feedback it was discussed with students, which insurance is needed. The prototype might have helped a few children to better understand the need and what it means to be insured. Although it is questionable how much the knowledge on insurance has been increased while playing. Similar to the topic of cryptocurrencies, insurances have been part of the game, but not the center.

The one statement about interest rates has been checked with the student giving the feedback. He clarified that the idea of interest rates when borrowing money was new to him. The majority however knew the topic and it was also not the goal of the test to teach about interest rates.

The result of the testing matches the expert's opinion on the lack of knowledge. The knowledge is inhomogeneous spread among the students and cannot be generalized. What can be said is that in this scenario there was some increase of knowledge for a few students in the general topic of financial education.

As mentioned in the additional feedback 6.3, a teacher gave feedback about how life points are deducted each month with the job. Her feedback was, that work might be associated in a bad way. The concern can be acknowledged but would need further research if this feature would really impact the user subconsciously. An adjustment of the prototype would be to separate the two messages. Additionally, the popups can be displayed in at least one event distance from one another. This would mean in the given implementation that the salary gets paid out in every fifth round. The stress would then be applied to every seventh event. This would separate the two topics from one another and would no longer connect the two topics together.

7.3.1 Generally

It can be summarized that children lack an understanding of costs in real life and what surprising events are happening. Terms like insurance, job opportunities, and how to save money are also lacking in general.

The Octalysis Framework by Yu-kai Chou is widely known in research about gamification and provided an excellent base for developing an educational game. The feedback provided by the testers overlaps with the expectations and guidelines the framework provides and supports this argument. Especially Core drive number four, which is about player customization, was observed to be very strongly developed with the target group. Instead of trying to win the prototype by choosing the best parameters, they tried to best reflect their future roadmap in life.

Educational games can make use of this by providing strong players customization and designing the gameplay with the core drives in mind. The results show that the players had fun playing the game, replayed, and would recommend it to friends. This although, the game only used one illustration per event and no animation or audio elements. From this, it can be concluded that complex graphics and animations are not needed to design a fun educational experience if the experience is well thought through.

When designing an educational gamified application, it should be emphasized that the gameplay should be designed around the content of the game. In this concept, costs and life balance were the main learning. Students had to constantly make decisions to balance these two factors. It was observed that other topics can be also added to the game but might not bring the same expected learning as with the main goals. As an example, cryptocurrencies have been mentioned but students who did not know the term, just skipped the event, as it was not necessary to deal with it in detail to continue.

Judging from the user's feedback, the target group's knowledge has increased. Financial knowledge is very heterogeneously distributed among students. Therefore, only parts of the classes, where the testing was conducted, had learned that they could mention. The targeted main learning about costs and balancing life was confirmed as the number one learning by the classrooms. Information about jobs, studies, and saving money follows as the next learning that could be taught to the students by letting them play the game.

It is recommended to further adjust the prototype with the feedback provided and test it in other classes and schools.

Educational games generally could achieve similar results by following the approach to design the application according to the core drives, identify the lack of knowledge it tries to fill, and iterative test and improve the game.

The objectives of the thesis have been achieved by developing a solid prototype that includes the knowledge gaps the target group has. The positive feedback and ideas the teenagers had after playing show, that an educational game in that area has great potential and should be further developed.

7.4 Limitations

One of the limitations of this thesis is that the testing and financial knowledge only cover the Austrian market. The testing has only been done in two 3rd and one 4th classrooms in Vienna. This was due first, to time limitations and secondly testing out the approach and prototype in a smaller setting.

7.5 Classification of the State of Research

The research connects to the surveys done in Austria about the lack of financial knowledge young people have. Additionally, it uses the Octalysis Framework by Yu-kai Chou, which is already known in the industry, and combines it in the educational setting. By documenting the process of game idea over topic, design, and technical research, including user testing, the study provides an overview and guideline of how an educational game is developed. The research outcome can be used for completely independent games with a different topic. However, it is also possible to further build on this study and derive additional educational applications about financial knowledge. Adapting the prototype to different knowledge areas can also bring new results.

7.6 Recommendation for further research

As mentioned, the study focus and testing group was in Austria. The study could be picked up in a different country and the prototype adapted. Possibly the results may differ. Future research can be conducted by taking the prototype and testing it with a larger set of classrooms. The study can also try to check if financial knowledge has been increased by performing a small quiz before and after letting the user play the game for the first time. Also adjusting events, pricing, and onboarding parameters, the game can be tested in different regions to conduct and compare if there are differences.

Another research could have a detailed look at the learnings and can conduct an additional survey while testing the prototype. With a quiz before and after playing, progress might be trackable. As mentioned in the second interview, it is recommended to conduct this research separately from the development, to decrease a certain bias.

Studies can further research, the differences in the knowledge that children in different classes might have. These results could help further improve the prototype by providing detailed gaps in financial knowledge. Other educational games would also profit from this research.

Additional research in the field of gamification could try to connect the core drives to educational games in other field areas. The outcome of these studies can be aggregated among different topics, providing information about which drives work for what educational field. Also, a comparison between age groups can be further researched. This can bring information about which motivators work best for different children. Combining all the mentioned research can provide valuable input for designing gamified experiences.

Further research can also be done using games in educational settings and which combination of teaching styles works best. A study can conduct if applications should be used with traditional presentations, open discussions, questionnaires, and more. This would provide educators with best practices for combining learning methods. Game developers would benefit from this study by knowing how the application will be used in an educational setting and can provide special features for educators.

It is possible to further adapt the prototype with new topics and game elements which can bring new learning to students. A study can pick up the existing state and iterative refactor the game with the help of the target group and feedback rounds.

7.7 Outlook for the prototype

In addition to an outlook and recommendation related to the research part of this study, suggestions for further adjustments of the prototype shall also be mentioned. As seen in the feedback of the students one of the first steps should be the introduction to more jobs that they can select from. This should make the game more relatable and will teach the children more about certain job opportunities including salary ranges.

A second low-hanging fruit is a better event variety. Introducing more activities and decisions to the application brings more diversity. Additionally, these events can also point out new concepts and terms to the players. More effort would be event dependencies which would bring even more variety. This way, decisions can have a longer impact on the game. If for example, players choose to select a car for mobility, then new events like an engine breakdown could show up. Players who chose the bike or public transport would not see those types of events. Another example would be to use the decisions for another type of investment. Players can select if they want to invest in Stock X. A few rounds later the result of that investment would show up. Event dependencies would open completely new event possibilities and teachings that can be implemented.

Lastly, as mentioned especially in the 4th grade, player customization can be added. While on first sight it might not directly impact the educational part. If users are willing to spend more time in the game and try out more aspects of financial decisions, there could be a greater learning effect in the long run. This statement however would need to be tested in a future study with an enhanced prototype.

Appendix 1: The public tender

Klassen mit Schüler:innen im Alter von 13-16 Jahren für neues Spielkonzept gesucht

Gesucht werden zwei Klassen mit Kindern im Alter von 13-14 Jahren welche gerne ein neues Spielkonzept ausprobieren möchten. Im Zuge einer Masterarbeit an der FH Technikum Wien (AT) und der HSLU Luzern (CH) wurde ein Prototyp entwickelt mit welchen die Schüler:innen finanzielle Entscheidungen in ihrem Leben durchspielen können. Nach dem setzen von einigen Grundparametern (Lehre oder Studium, Job, Wohnung etc.) werden die Kinder mit Ereignissen und Fragen konfrontiert. Mit simplen Ja/Nein Antworten spielen die Nutzer:innen das Spiel durch. Die Fragen behandeln Alltagsthemen und haben finanzielle als auch soziale Auswirkungen. Zwischen den Spielphasen können die Schüler:innen entscheiden ob sie ihr Geld sparen, in Weiterbildung investieren oder in Freizeitaktivitäten. Ziel des Spieles ist, ein ein Gleichgewicht zwischen Finanz und "Lebenspunkten" zu erreichen. Am Ende werden alle Entscheidungen und das Ergebnis auf einer Übersicht zusammengefasst. Dies ermöglicht spannende Einblicke und öffnet Diskussionsrunden im Klassenzimmer. "Warum habe ich mehr Geld als XY?", "Was wäre passiert hätte ich eine Versicherung abgeschlossen?".

Ablauf des Testes

Nach einer kurzen Erklärung können die Kinder selbständig das Spiel am Smartphone oder im Browser ausprobieren. Zusammen mit dem Lehrpersonal kann das Ergebnis mit der Klasse besprochen werden. Digital wird danach Feedback von den Schüler:innen für die Auswertung und Weiterentwicklung gesammelt.

Dauer

Die Spieldauer wird mit ca. 5-15 Minuten geschätzt. Das Einholen von Feedback wird in etwa 10 Minuten beanspruchen. Die Diskussion im Klassenraum kann vom Lehrpersonal geleitet und persönlich eingeteilt werden.

Mögliche Termine

2-4. Mai

10-12. Mai

15-19. Mai

Appendix 2: Prototype

The prototype can be accessed under the following link:

Code: <https://github.com/ManuelSzecsenyi/financial-education-game>

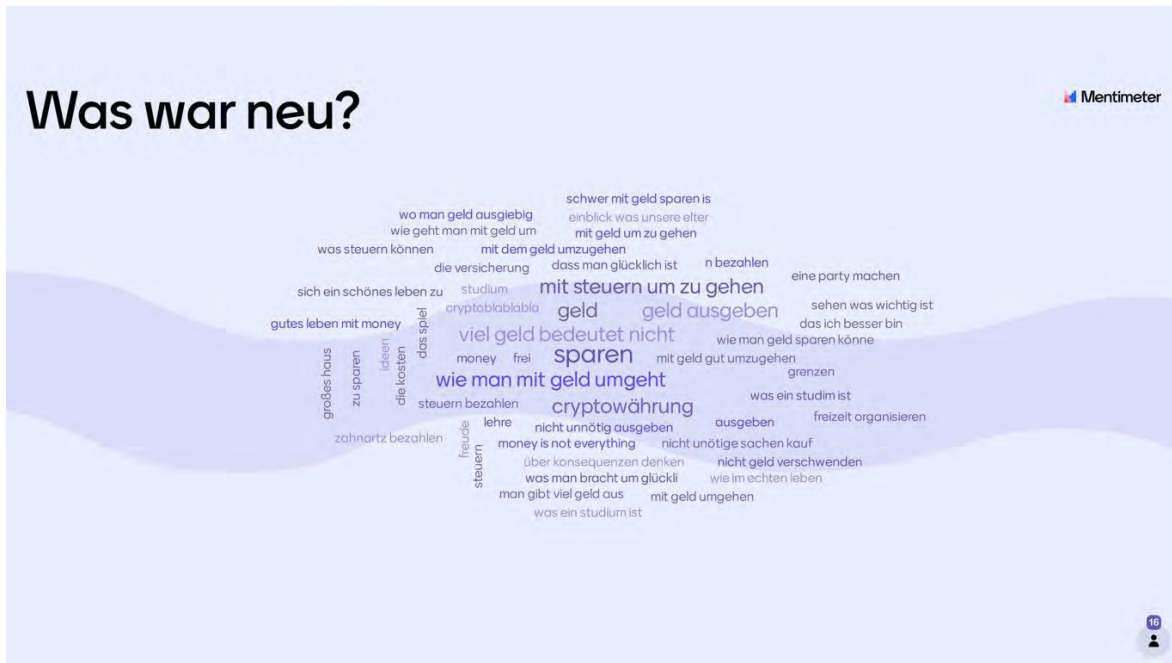
Game: <https://demo.marginleft.at>

Appendix 3: Master thesis

The full master thesis including the transcripts can be found here:

https://marginleft.at/data/Masterarbeit_Szecsenyi.pdf

Appendix 4: Word Clouds



Classroom A



Classroom B



Literaturverzeichnis

- Adventurer* | DiceBear. (n.d.). Retrieved May 24, 2023, from <https://www.dicebear.com/styles/adventurer>
- Chou, Y. (2014). *Actionable Gamification*.
- Figma. (2023, February 27). Figma. <https://www.figma.com/de/>
- FLiP Challenge. (2023, May 23). FLiP Challenge. <https://flipchallenge.at/>
- Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management*, 45, 191–210. <https://doi.org/10.1016/j.ijinfo-mgt.2018.10.013>
- Laine, T. H., & Lindberg, R. S. N. (2020). Designing Engaging Games for Education: A Systematic Literature Review on Game Motivators and Design Principles. *IEEE Transactions on Learning Technologies*, 13(4), 804–821. <https://doi.org/10.1109/TLT.2020.3018503>
- Most used web frameworks among developers 2022*. (2022, July). Statista. <https://www.statista.com/statistics/1124699/worldwide-developer-survey-most-used-frameworks-web/>
- Nah, F. F.-H., Zeng, Q., Telaprolu, V. R., Ayyappa, A. P., & Eschenbrenner, B. (2014). Gamification of Education: A Review of Literature. In F. F.-H. Nah (Ed.), *HCI in Business* (pp. 401–409). Springer International Publishing. https://doi.org/10.1007/978-3-319-07293-7_39
- Prasetyaningrum, P. T., Purwanto, P., & Rochim, A. F. (2022). Gamification On Mobile Banking Application: A Literature Review. *2022 1st International Conference on Information System & Information Technology (ICISIT)*, 222–227. <https://doi.org/10.1109/ICISIT54091.2022.9873034>
- Simons, A., Niehaves, B., Niehaves, B., & Reimer, K. (2009). *Reconstructing the giant: On the importance of rigour in documenting the literature search process*.
- Streamline UX Colors Illustrations Set*. (2023, May 24). <https://www.streamlinehq.com/illustrations/illustrations-multicolor>
- Vanessa Koch, Sebastian Swoboda, Bianca Schranz, Eva Arene, & Katharina Heindl. (2022). *Fortschrittsbericht 2021/22 Nationale Finanzbildungsstrategie*. 70.
- YEP - Stimme der Jugend. (2021). *YEP Jugendbericht*.
- YEP - Stimme der Jugend. (2022). *YEP Jugendbericht Wirtschaftsbildung*.

Inverted Classroom im Teilmodul Buchhaltung - aus Fehlern lernen -

Laufendes Lehrprojekt

Dipl.-Kfm. (FH) Andreas Stach

Hochschule Bochum, andreas.stach@hs-bochum.de

Prof. Dr. Carsten Theile

Hochschule Bochum, carsten.theile@hs-bochum.de

Abstract

Im Fach Buchhaltung ist an der Hochschule Bochum erstmals im Wintersemester 2023/24 das Konzept des Inverted Classrooms umgesetzt worden. Zielsetzung war die Absenkung der Abbrecher- und Durchfallquote sowie die Anhebung des Notendurchschnitts im Fach Buchhaltung. Um den Grad der Zielerreichung evaluieren zu können, wurden während der Implementierungsphase statistische Daten erhoben und ausgewertet. Im vorliegenden Beitrag stellen wir, auf Grundlage der erhobenen quantitativen Daten, die Durchführung sowie die Ergebnisse der Implementierungsphase vor. Obwohl sich unsere o.g. Erwartungen nicht erfüllt haben, lassen sich zahlreiche positive Wirkungen des Konzepts belegen. Daher gelangten wir zu der Hypothese, dass die mangelnde Zielerreichung im Wesentlichen auf eine unzureichende Umsetzung des Konzepts zurückzuführen ist: Wir konnten Fehler- und Verbesserungspotenzial identifizieren. Im kommenden WiSe sollen einige Verbesserungsmaßnahmen umgesetzt werden, die wir am Ende des Beitrags skizzieren und zur Diskussion stellen.

1 Ausgangslage

1.1 Motivation zur Änderung des didaktischen Konzepts

An der Hochschule Bochum – Hochschule für angewandte Wissenschaften – war die traditionelle Form der Lehre in der Vor-Corona-Zeit der seminaristische Unterrichtsstil: In Gruppen von bis zu 60 Studierenden wurde in Seminarräumen ein Wechsel von Vortrag und Erklärung mit Beispielen und Übungen vollzogen.

Die Einstellung des Präsenzunterrichts infolge der Pandemie führte ad hoc zum Einsatz verschiedener digitaler Formate, die überwiegend über die hochschulweit eingesetzte Lernplattform Moodle angeboten wurden. Das Spektrum reichte im Wesentlichen von der Bereitstellung bisher eingesetzter Foliensätze und Skripte, angereichert mit Aufgaben und Lösungen, bis hin zu Videos und zur Live-online-Lehre. Wie im Einzelnen die Distanzlehre aufgebaut war, war zunächst weder hochschulweit noch fachbereichsweit koordiniert. Erst im Verlauf der Pandemie wurde die Live-online-Lehre obligatorisch.

Seit Wiedereinführung der Präsenzlehre im Sommersemester (SoSe) 2022 lag hochschulweit ein Konvolut digitaler Lehrschnipsel vor, die bestenfalls – betrachtet man jede Lehrveranstaltung einzeln – in sich schlüssig aufeinander aufbauten. Die Lehrenden hatten flächendeckend erstmals umfangreiche Erfahrung in digitalen Lehrmethoden sammeln können und müssen. Angesichts dessen haben wir uns die Frage gestellt, ob eine einfache Rückkehr zum traditionellen seminaristischen Unterrichtsstil angezeigt ist oder die Verknüpfung der Präsenzlehre mit digitalen Elementen, mit denen positive Erfahrungen gemacht worden sind, möglicherweise zu besseren Studienergebnissen führen könnte. Eine rein digitale (synchrone oder asynchrone) Lehre kam für uns aus zwei Gründen nicht in Betracht: Sie scheint erstens im Hinblick auf Zufriedenheit und Präferenzen der Teilnehmenden der Präsenzlehre unterlegen (vgl. Kanning & Ohlms, 2021), und zweitens begreift sich die Hochschule Bochum traditionell als Präsenzhochschule, sodass reine Online-Lehre nicht der dortigen Lehrkultur entspricht.

Demgegenüber hat eine vor dem Start unseres Projekts veröffentlichte Review von Metaanalysen zur lernbezogenen Effektivität der Lehrformate Onlinelehre und Blended Learning (inklusive Inverted Classroom) im Vergleich zur traditionellen Präsenzlehre überwiegend kleine positive Effekte der neuen Lehrformate – insbesondere Inverted Classroom – in Bezug auf kognitive Lernergebnisse, praktische Fertigkeiten und Studierendeneinschätzungen gezeigt (Mayweg u.a., 2023). Daher haben wir uns bei der Verknüpfung der Präsenzlehre mit digitalen Elementen für das Konzept des Inverted Classrooms entschieden: Im Kern wird die Wissensvermittlung vorgelagert zum digital unterstützten Lernen des Einzelnen, während dann in der Präsenzveranstaltung mehr Raum bleibt für kollaboratives Anwenden und Üben (siehe z.B. Federau, 2023; Wipper & Schulz, 2021, S. 82-99; Handke, 2017). Die Wissenserarbeitung ist in der Anwendung des Wissens mit einem Leistungsziel verbunden (Kley, Kull & Bitterli, 2021, S. 352 m.w.N.), und der Aktivierung von Lernenden, als Voraussetzung für erfolgreiches Lernen, (Kerres, 2018, S. 172) wird mehr Zeit eingeräumt als im seminaristischen Unterrichtsstil.

1.2 Modul Buchhaltung

Das Konzept des Inverted Classrooms wurde erstmals im Wintersemester (WiSe) 2023/24 im Fach „Buchhaltung“ umgesetzt. Die Wahl fiel auf das Fach, weil die Änderungsdynamik der Fachinhalte hier recht gering ist, sodass die Halbwertszeit einmal produzierter digitaler Inhalte sehr hoch ist.

In den Bachelor-Studiengängen „Betriebswirtschaftslehre“ und „International Business and Management“ der Hochschule Bochum ist das Modul „Buchhaltung und Kostenrechnung“ im 3. Semester vorgesehen. Im WiSe wird das Modul 3-zügig und im SoSe 2-zügig angeboten. Das Modul besteht aus den beiden Veranstaltungen (Teilmodulen) „Buchhaltung“ und „Kostenrechnung“, die sowohl von unterschiedlichem Lehrpersonal als auch konzeptionell und inhaltlich unabhängig voneinander angeboten bzw. durchgeführt werden. Die Kontaktzeit in der Lehrveranstaltung „Buchhaltung“ beträgt zwei Semesterwochenstunden (= 30 h á 45 Minuten); im Selbststudium kommen noch einmal 45 h á 45 Minuten dazu. Die jeweiligen Zeiten der Lehrveranstaltung „Kostenrechnung“ haben einen doppelten Umfang. Am Ende des Semesters wird als Prüfungsform eine Klausur über 135 Minuten angeboten,

wovon 45 Minuten auf Buchhaltung entfallen. Die Klausur muss insgesamt bestanden sein, um die 7,5 ECTS zu erwerben; das Bestehen beider Teilprüfungen ist nicht erforderlich.

Im WiSe 2023/24 wurden zwei Züge mit dem Konzept des Inverted Classrooms umgesetzt, während in einem weiteren Zug eines anderen Dozenten herkömmliche Lehre praktiziert wurde.

1.3 Zielgruppenanalyse

Die Zielgruppe besteht in etwa zu gleichen Anteilen aus Männern und Frauen. Es handelt sich in der Regel um junge Erwachsene, die direkt nach der Schule oder nach einer Ausbildung an die Hochschule kommen. Das Vorwissen der Studierenden kann aufgrund unterschiedlicher Schulformen (Regelschule vs. Wirtschaftsschule) und unterschiedlichen beruflichen Kontexten (Berufserfahrung sowie Berufsschulkenntnisse) stark schwanken, ist in der Summe jedoch als eher niedrig einzustufen. Der höchste Abschluss nach dem Deutschen Qualifikationsrahmen liegt in der Regel gemäß Stufe 4 vor. Der Umgang mit IT ist dem Großteil der Teilnehmenden gut vertraut.

Die Motivation für die Wahl des Studienfachs ist eher intrinsisch anzusetzen, für das Fach Buchhaltung scheint dies jedoch nicht flächendeckend der Fall zu sein. Pro Semester gibt es ca. 250 Studierende, die potenziell an dem Kurs und der Prüfung teilnehmen können.

Lücken gibt es in den Selbstlernkompetenzen der Studierenden. Ein Grund hierfür könnte in der weiterhin starken Verbreitung direkter Instruktionen im Schulunterricht liegen (vgl. z.B. Gudjons, 2021). Einige Teilnehmende zeigen sich mit den Anforderungen an die Selbstlernkompetenz im hochschulischen Kontext noch überfordert und benötigen ein niedrigschwelliges Lernangebot.

1.4 Lernziele

Sämtliche Lehr-/Lernziele wurden so formuliert, dass sie sowohl bei den Lehrenden als auch bei den Studierenden deckungsgleich, nachvollziehbar und überprüfbar sind (vgl. Kerres, 2018, S. 304 ff.). Bei der Ausgestaltung dieser wurde der Ansatz des Constructive Alignments verfolgt, damit die Studierenden am Ende der Lehrveranstaltung in der Lage sind, die tatsächlichen Anforderungen der Prüfungsleistung zu bewerkstelligen (vgl. Kerres, 2021, S. 241 ff.; Kergel & Heidkamp-Kergel, 2020, S. 26-27).

Gemäß der Bloomschen Taxonomie wurden folgende kognitive und affektive Lernziele definiert (vgl. Bloom, 1956). Die Definition der Lernziele erfolgt als Richtziel und Grobziele, um sowohl den Zielhorizont der gesamten Lehrveranstaltung als auch die Ziele der einzelnen Lerneinheiten zu definieren (vgl. Kergel & Heidkamp-Kergel, 2020, S. 23-24; Modlinger, 2020, S. 65-66).

Das Richtziel beinhaltet sowohl kognitive als auch affektive Inhalte und lautet:

„Am Ende der Lehrveranstaltung kennen die Studierenden die Grundsätze des Rechnungswesens und sind in der Lage, diese in unternehmerischen Geschäftsvorfällen oder Fallstudien zielgerichtet und selbstständig anzuwenden.

Sie sind in der Lage die üblicherweise in Unternehmen anfallenden Geschäftsvorfälle durch die Bildung der richtigen Buchungssätze in die Sprache des Rechnungswesens zu überführen.

Ferner erkennen die Studierenden die Notwendigkeit von buchhalterischen Themen in wirtschaftlichem Kontext und setzen sich zielgerichtet mit diesen Inhalten auseinander.

Nach ANDERSON & KRATHWOHL (2001) soll jeweils folgende Kompetenzstufen erreicht werden (vgl. auch Kerres, 2021, S. 105 – 106):

Kognitiv: Kompetenzstufe 4 – Analysieren

Affektiv: Kompetenzstufe 3 – Werten

Die einzelnen Grobziele aus den Lerneinheiten sollen das Richtziel unterstützen und werden hier nur exemplarisch angeführt:

„Am Ende dieses Kapitels kennen Sie das Prozedere für das Inverted Classroom-Konzept im Fach Buchhaltung. Sie wissen, dass Sie sich für jede Präsenzveranstaltung im Vorfeld vorbereiten müssen, um aktiv an diesen teilnehmen zu können und den Lernerfolg zu gewährleisten.“

„Am Ende des Kapitels 4 kennen Sie erste Bestandskonten und deren Aufbau. Sie können erste Buchungssätze bilden und in T-Konten übertragen. Ferner kennen Sie die "+--+"-Regel.“

„Nach Bearbeitung des Kapitels kennen Sie Erfolgskonten und können diese von Bestandskonten abgrenzen. Sie wissen, wie Sie auf Erfolgskonten buchen und können die Konten zum Ende einer Periode abschließen. Ferner sind Sie in der Lage, die Auswirkungen erfolgswirksamer Geschäftsvorfälle auf das Eigenkapital einzuschätzen.“

2 Durchführung

2.1 Erwartungen

Stützt man sich auf vorliegende Metaanalysen, sind mit dem Konzept des Inverted Classroom möglicherweise nur kleine, aber immerhin messbare Verbesserungen gegenüber dem seminaristischen Unterrichtsstil zu erzielen (siehe Mayweg u.a., 2023). Wir sind daher von vornherein nicht davon ausgegangen, große Effekte, etwa bei den Lernergebnissen, in den Daten zu finden. Gleichwohl sind wir mit drei Erwartungen in das Projekt gestartet, die sich auf eine Überlegenheit ggü. dem traditionellen seminaristischen Unterrichtsstil beziehen:

- (1) Verringerung der Abbrecherquote während des Semesters: Jene, die den Kurs begonnen haben, sollten ihn auch mit der Klausur beenden.
- (2) Signifikante Reduzierung der Durchfallquote in der Klausur, bezogen auf das Teilgebiet Buchhaltung.
- (3) Signifikante Verbesserung des Notendurchschnitts, bezogen auf das Teilgebiet Buchhaltung.

Gerade zu (2) und (3) können belastbare Daten erhoben werden. Auch ein Vergleich zu den Vorsemestern ist hier möglich. Bei (1) lässt sich ein Vergleich zu Vorsemestern nicht darstellen, weil aus diesen Semestern keine validen Daten vorliegen.

2.2 Setting und Ablauf

Der überwiegende Fachinhalt der Veranstaltung Buchführung wurde auf zehn Lerneinheiten heruntergebrochen. Auf acht dieser Einheiten wurde das Konzept des Inverted Classrooms angewandt, bei dessen Umsetzung wir uns an HANDKE orientiert haben (Handke, 2020, S. 87 ff.). Der Aufbau der Lehr-/Lerninhalte folgt einem expositorischen Ansatz, um zum einen den hierarchischen Aufbau des Themengebiets und zum anderen das relativ niedrige Niveau in der Selbstlernkompetenz der Zielgruppe zu berücksichtigen. Die einzelnen Lerneinheiten sind relativ klein und werden mit zunehmender Kursdauer komplexer sowie umfangreicher (vgl. Kerres, 2021, S. 140). Konkret wurden den Studierenden wöchentlich folgende Materialien über Moodle für die Selbstlernphase zur Verfügung gestellt:

- (1) Video (ca. 30 Minuten) mit H5P-Elementen¹, produziert vor Greenscreen im Medienkompetenzzentrum der Hochschule, frei eingesprochen vom (immer gleichen) Dozenten.
- (2) Der dem Video zugrundeliegende Foliensatz.

¹ Als H5P-Elemente wurden Verständnisfragen in die Videos eingefügt, die den Studierenden ein direktes Feedback darüber ermöglichen sollen, ob Inhalte korrekt verstanden und verinnerlicht wurden. Ergänzt wurden diese Fragen von Verlinkungen zu externen Quellen, wie z.B. Gesetzestexten, sowie zu bereitgestellten Übungsaufgaben.

- (3) Übungsaufgaben (Moodle-Tests), direkt zu bearbeiten unter Moodle, mit automatischer Ergebnismitteilung und teilweise mit Lösungshinweisen bzw. -erklärungen.
- (4) Komplexere Aufgaben, die in Gruppen bearbeitet werden sollten, um deren Lösung bzw. Lösungsansätze in der Präsenzveranstaltung von Studierenden vorstellen und in der Gruppe diskutieren zu lassen.
- (5) Gelegentlich: vertiefende Literatur (kleinere, didaktisch aufbereitete Texte).

In den ersten beiden Veranstaltungen war die Lehre noch herkömmlich. Erläutert wurde, warum es Buchhaltung gibt und welchen Einfluss die Digitalisierung sowie der Einsatz von KI auf Buchhaltung haben. Außerdem wurde erläutert, wie man vom Sachverhalt über den Beleg und den Buchungssatz zum Hauptbuch kommt.

In der dritten Veranstaltung wurde das Konzept des Inverted Classrooms ausführlich erläutert und die Erwartungshaltung zur Mitarbeit kommuniziert. Es wurde insbesondere darauf hingewiesen, dass eine Zusammenfassung des Fachinhalts der Videos auf der nächsten Präsenzveranstaltung nicht erfolgen würde. Die Bearbeitung zu (4) setze aber eine vorherige intensive Beschäftigung mit den Materialien (1) bis (3) und ggf. (5) voraus. Die Lösungen zu (4) würden auf in der Präsenzveranstaltung zusammengefasst bzw. erarbeitet.

In der Präsenzphase der acht Veranstaltungen, in denen das Konzept des Inverted Classrooms umgesetzt worden ist, konnten die Studierenden zu Beginn jeder Veranstaltung Fragen stellen zu den in den Videos kommunizierten Fachinhalten. Sodann bestanden die Präsenzveranstaltungen aus der Lösung bzw. der Lösungsbesprechung zu (4). Eine Wiederholung des Stoffes, der über die Videos vermittelt worden ist, wurde – wie zuvor angekündigt - nicht durchgeführt. Auch schriftliche Lösungen zu (4) wurden nicht verteilt oder bereitgestellt, um die Entwicklung eigenständiger Lösungen der Studierenden sowie die Partizipation am Kurs positiv zu beeinflussen.

Die letzten beiden Veranstaltungen dienten gezielt der Klausurvorbereitung. Hierzu konnte die spezielle Klausurbearbeitungstechnik anhand alter Klausuraufgaben eingeübt werden.

Das Konzept wurde auf zwei Parallelkurse angewendet: Ein Montagskurs (beginnend am 9. Oktober 2023), der, bedingt durch den Stundenplan, 14-tägig mit 4 Stunden stattfand, und ein Dienstagskurs (beginnend am 10. Oktober 2023), der wöchentlich mit 2 Stunden angeboten wurde. Beide Gruppen haben denselben Moodle-Kurs genutzt.

3 Ergebnisse

3.1 Zwischen-Feedback

3.1.1 Die Fragen

Nach den ersten beiden Inverted Classroom Veranstaltungen wurden die Studierenden um Feedback gebeten. Bei der Konstruktion des Fragebogens haben wir uns von aktuellen, hierzu vorliegenden Forschungsergebnissen inspirieren lassen (Kärchner, Gehle & Schwinger, 2022), unseren Fragebogen jedoch bewusst einfach und niedrigschwellig gehalten in der Hoffnung, so eine hohe Rücklaufquote zu erreichen. Insgesamt 33 Studierende (von ca. 250 eingeschriebenen Personen in Moodle bzw. von zusammengefasst ca. 50 Studierenden in den beiden Präsenzveranstaltungen) haben Angaben zu folgenden - in Moodle bereitgestellten - zehn Aussagen gemacht:

- 1) Ich war an folgenden Veranstaltungen anwesend:
- 2) Ich habe das Lehrkonzept „Inverted Classroom“ verstanden.
- 3) Die technische Qualität der Videos zu Kapitel 4 und 5 benote ich mit: (Schulnoten)
- 4) Die fachinhaltliche Qualität der Videos zu Kapitel 4 und 5 benote ich mit: (Schulnoten)
- 5) Die Übungsaufgaben zu Kapitel 4 und 5 zur Vorbereitung der Präsenz waren für mich klar und verständlich.
- 6) Die Übungsaufgaben zu Kapitel 4 und 5 zur Vorbereitung der Präsenz konnte ich gut allein erarbeiten.

- 7) Die Übungsaufgaben zu Kapitel 4 und 5 zur Vorbereitung der Präsenz habe ich vor der Präsenz intensiv genug bearbeitet.
- 8) Ich habe zur Bearbeitung der Übungsaufgaben zu Kapitel 4 und 5 zur Vorbereitung der Präsenz andere Informationsmaterialien (z.B. Bücher, Youtube-Videos, Chat mit Kommilitonen) genutzt.
- 9) Die Präsenzveranstaltungen zu Kapitel 4 und 5 haben mir fachlich geholfen.
- 10) Bitte benoten Sie das hier umgesetzte Konzept des „Inverted Classrooms“ insgesamt nach dem Schulnotenprinzip.

Darüber hinaus wurden die Studierenden in einem Freitextfeld um „Lob, Kritik und Anregungen“ gebeten. Davon machten 22 Studierende Gebrauch.

3.1.2 Videos

Konzept: Da die Wissensvermittlung konzeptionell und schwerpunktmäßig über Videos erfolgte, waren für uns die Antworten zu den Fragen 3 und 4 zentral. Unabhängig davon, dass wir von der technischen und fachinhaltlichen Qualität der Videos überzeugt waren, muss diese Überzeugung auch bei den Studierenden vorliegen. Anders gewendet: Das Konzept des Inverted Classrooms ist u.E. von vornherein zum Scheitern verurteilt, wenn die Art der Wissensvermittlung in der Selbstlernphase nicht zu überzeugen vermag.

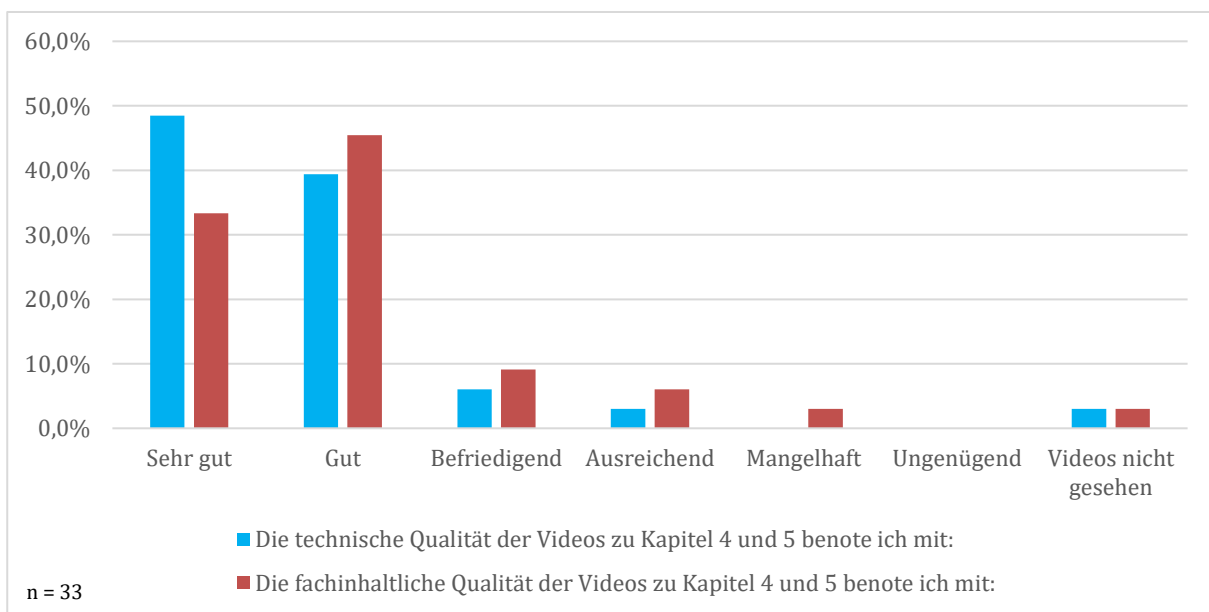


Abb. 1: Technische und fachinhaltliche Qualität der Videos

Auswertung: Von den 33 Studierenden, die an der Feedbackrunde teilgenommen haben, hat nur ein Studierender die Videos nicht gesehen. Fast 90 % der Studierenden beurteilten die technische Qualität der Videos als sehr gut oder gut, und fast 80 % beurteilten die fachinhaltliche Qualität als sehr gut oder gut.

Würdigung: Weil Studierende kaum in der Lage sein dürften, die fachinhaltliche Qualität im Sinne einer objektiven Güte beurteilen zu können, vermuten wir, dass in der Beantwortung der Frage vor allem zum Ausdruck kommt, ob Studierende glauben, den Fachinhalt verstanden zu haben. Obwohl wir uns bei der Frage nach der fachinhaltlichen Qualität noch bessere Ergebnisse gewünscht hätten, haben wir dies insgesamt als Bestätigung unserer Arbeit aufgefasst. Weitere Videos wurden auch erst im Anschluss an das Zwischen-Feedback produziert, um Rückmeldungen berücksichtigen zu können.

Maßnahmen: Eher Keine. Wir denken noch darüber nach, die Videos – insbesondere bei den ersten Lerneinheiten – in noch kleinere Einheiten aufzuteilen. Der Vorteil kann darin gesehen werden, eine Konzentrationspause zu erzwingen und Studierende leichter an diese Art des Lernens zu gewöhnen. Auf der anderen Seite steht es auch bei längeren Videos jedem Studierenden frei, eine Pause einzulegen. Eine Neuproduktion der Videos ist nicht erforderlich.

3.1.3 Übungsaufgaben zur Vorbereitung der Präsenz

Konzept: Nachdem die Studierenden die Videos gesehen und die zugehörigen (niedrigschwelligen) Moodle-Tests (3) inklusive Feedback bearbeitet haben, sollten sie nach unserer Erwartung – jedenfalls grundsätzlich - in der Lage sein, die (anspruchsvolleren) Übungsaufgaben (4) zur Vorbereitung der Präsenzveranstaltung zu verstehen und zu lösen oder zumindest einer Lösungsidee zuführen zu können. Ob dies erreicht werden konnte, sollte mit den Fragen 5 bis 7 überprüft werden.

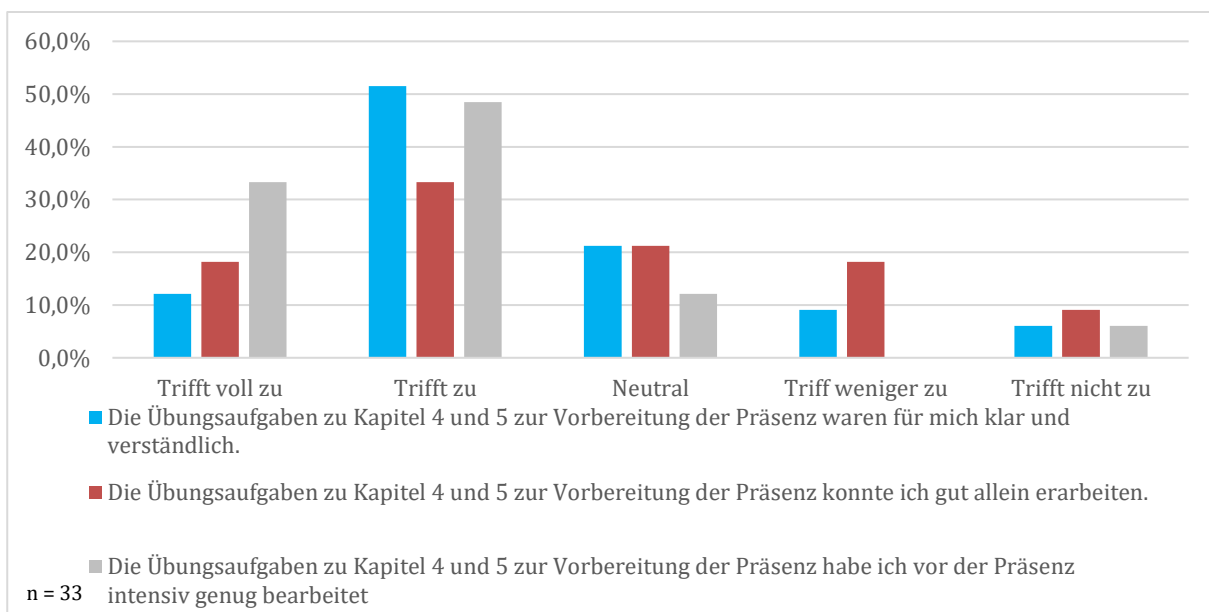


Abb. 2: Rezeption der Übungsaufgaben zur Vorbereitung der Präsenzveranstaltungen

Auswertung: Obwohl die Übungsaufgaben immerhin mit über 60 % als klar und verständlich beurteilt worden sind (trifft voll zu oder trifft zu, Frage 5), konnten sie nur zu knapp 50 % gut allein erarbeitet werden (Frage 6). Besonders auffällig ist dieses Ergebnis auch, da die Übungsaufgaben von über 80 % der Teilnehmenden nach eigener Einschätzung intensiv genug bearbeitet worden sind (Frage 7).

Vorbemerkung: Der Aufbau von nicht fachspezifischen Kompetenzen zur selbstständigen Erarbeitung von Inhalten kann durch den didaktischen Aufbau des Kurses Buchhaltung, dem das Prinzip „vom Einfachen zum Komplexen“ zugrunde liegt, lediglich unterstützt werden. Eine explizite Ersts Schulung dieser Kompetenzen kann nicht geleistet werden. Vielmehr sollten diese Grundlagen mit Abschluss der Schule oder nach den ersten zwei Semestern so weit aufgebaut sein, dass diese Vorgehensweise nicht nur bekannt, sondern von den Studierenden auch verinnerlicht worden ist.

Würdigung: Die Bereitschaft zur intensiven Beschäftigung mit den Übungsaufgaben lag offensichtlich vor (Frage 7), aber der selbst eingeschätzte Erfolg der Studierenden (Frage 6) ist aus unserer Sicht deutlich zu niedrig. Ein Grund kann darin gesehen werden, dass die Übungsaufgaben zu wenig klar und verständlich waren (Frage 5). Möglicherweise gab es auch eine zu hohe Diskrepanz im Anforderungsprofil zwischen den (niedrigschwelligen) Moodle-Tests und den (anspruchsvolleren) Übungsaufgaben. Drittens ist nicht auszuschließen, dass Studierende

nach dem Videokonsum zwar glauben, ein Thema verstanden zu haben, sie aber gleichwohl noch nicht in der Lage sind, diese Fachinhalte auf die Anwendungspraxis beziehen zu können. Ferner könnten Studierende - auch im dritten Semester - noch mit der eigenständigen Auseinandersetzung und Reflektion von Lehrinhalten überfordert sein.

Maßnahmen: Im Erstellungsprozess der Moodle-Tests und der Übungsaufgaben haben wir zwar eine fachinhaltliche Kontrolle durchgeführt, aber keine didaktische. Wir werden dies für den nächsten Durchlauf ändern, um bessere Ergebnisse im Hinblick auf die Fragen 5 und 6 erzielen zu können: In einer kleinen Gruppe von drei oder vier Studierenden werden im Vorfeld sowohl die Moodle-Tests als auch die Übungsaufgaben auf Verständlichkeit, Aufbau und Niveausteigerung getestet. Getrennt davon ist es bei diesen Fragen zukünftig geplant, die Expertise von zwei weiteren Dozierenden einzuholen.

3.1.4 Präsenzveranstaltung

Konzept: Die Studierenden hatten eine Woche vor der jeweiligen Veranstaltung Zeit, die Videos, Moodle-Tests und die Übungsaufgaben zu bearbeiten. In der dann folgenden Präsenzveranstaltung der jeweiligen Kurse wurden nur die Übungsaufgaben besprochen. Eine Zusammenfassung des Lehrstoffs wurde nicht geboten. Diese Vorgehensweise ist, wie oben schon ausgeführt, in der dritten Veranstaltung auch so kommuniziert worden.

In der Präsenzveranstaltung haben sich die Studierenden in Gruppen zusammengesetzt, um Ergebnisse auszutauschen und gegenseitig zu prüfen. Der Dozent ging von Gruppe zu Gruppe und gab, wo nötig, Hilfestellung. Sodann sollten Studierende ihre Lösungen vor der gesamten Gruppe vorstellen. Im Ergebnis ist der Erfolg dieser Vorgehensweise als mäßig einzustufen. Die Bereitschaft zur Vorstellung eigener Ergebnisse war nur rudimentär vorhanden, sodass überwiegend der Dozent die Ergebnisdarstellung moderierte. Auf diese Weise wurden die Lösungen zu den Übungsaufgaben erarbeitet und vermittelt. Eine schriftliche Muster-Lösung ist, aus oben genannten Gründen, nicht angeboten worden.

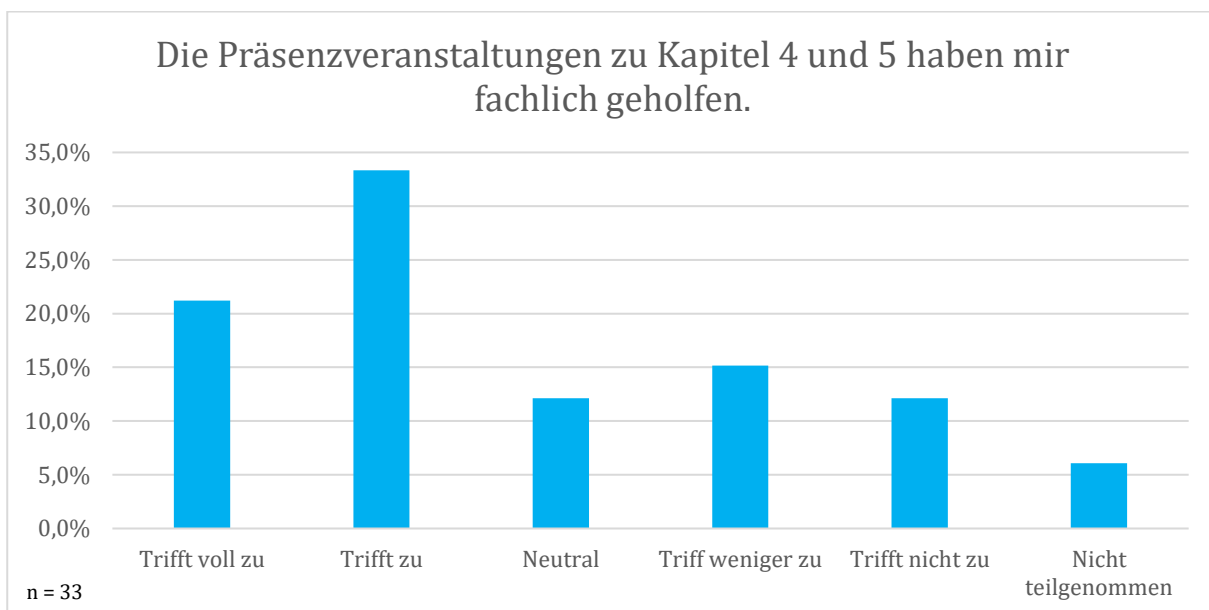


Abb. 3: Beurteilung der Präsenzveranstaltungen

Auswertung: Knapp 55 % der Studierenden geben an, dass Ihnen die Präsenzveranstaltungen fachlich geholfen haben (trifft voll zu oder trifft zu).

Würdigung: Uns ist diese Zustimmungsquote viel zu gering. Gewünscht hätten wir uns mindestens 80 %. Wir führen die geringe Zustimmungsquote (1) auf den Ablauf der Veranstaltung und (2) auf unsere Weigerung, (Muster-)Lösungen anzubieten, zurück.

Zu (1): Der 14-tägig stattfindende Montagskurs (mit 4 Stunden) war mit rund 40 Studierenden relativ gut ausgelastet. Das Gruppenkonzept (10 Gruppen) ist jedoch von einem einzigen Dozenten qualitativ nicht auf dem benötigten Niveau zu bewerkstelligen, da eine individuelle Gruppenbetreuung nicht ausreichend möglich ist. Ferner war es häufig eine Herausforderung für die Studierenden, dass innerhalb der vierstündigen Veranstaltung gleich zwei (aufeinander aufbauende) Kapitel bearbeitet werden mussten. Sofern der vorgelagerte Inhalt nicht vollständig durchdrungen wurde, führte dies auch zu Problemen beim Verständnis des darauf aufbauenden Inhalts.

In der wöchentlichen Dienstagsgruppe (2 Stunden) waren nur 11 Studierende anwesend. Hier wurden drei Gruppen gebildet, die Betreuung durch den Dozenten war hier deutlich individueller. Zudem hatte jede Präsenzveranstaltung nur ein Thema zum Gegenstand.

Zu (2): Die Weigerung, schriftliche (Muster-)Lösungen in Moodle bereitzustellen, ist auch im Freitextfeld des Zwischen-Feedbacks sehr oft moniert worden. Studierende hätten sich offensichtlich wohler gefühlt, wenn sie im Nachgang der Präsenzveranstaltung noch eine Lösung hätten herunterladen können.

Maßnahmen: Wir haben – als Sofortmaßnahme - ab Kapitel 6 sowohl das Kleingruppenkonzept als auch die Lösungsvorträge durch Studierende in der Präsenzveranstaltung aufgegeben. Dies führte zu einem entspannteren Lernklima sowie weniger Irritation innerhalb der Studierendengruppe.

Die Nichtherausgabe der Lösungen haben wir beibehalten, um die Studierenden zur Teilnahme an den Präsenzveranstaltungen zu motivieren. Tatsächlich ist die Teilnahmefrequenz nicht weiter eingebrochen. Die Nichtherausgabe der Lösungen ist aber bis zum Schluss von Studierenden kritisiert worden.

Wir werden darüber hinaus im nächsten Durchlauf Folgendes ändern: Die im Vorfeld von den Studierenden zur Vorbereitung auf die Präsenzveranstaltungen zu bearbeitenden Aufgaben werden erstens weniger komplex sein, und die Lösungen werden zweitens - **nach** den Präsenzveranstaltungen - bereitgestellt. Dies gilt jedoch nicht für die Angabe des Lösungswegs, wo dieser erforderlich ist. Wenn beispielsweise die Angabe von Buchungssätzen gefragt ist, aus denen ein einfacher Jahresabschluss erstellt werden soll, werden wir als Lösung nur den Jahresabschluss bereitstellen. Wir werden das künftig auch im Vorfeld so kommunizieren. Darüber hinaus wird es – drittens – weitere (komplexere) Übungsaufgaben geben, die erst in der Präsenzstunde bekanntgegeben werden. Noch nicht entschieden ist, ob auch für diese Aufgaben Lösungen bereitgestellt werden. Schließlich werden wir in Abstimmung mit der Stundenplanung keine Veranstaltungen mehr im 14tägigen Rhythmus durchführen.

3.1.5 Gesamtbeurteilung und Stellungnahmen:

Wir wollten wissen, wie die Studierenden insgesamt das umgesetzte Konzept des Inverted Classrooms beurteilen:

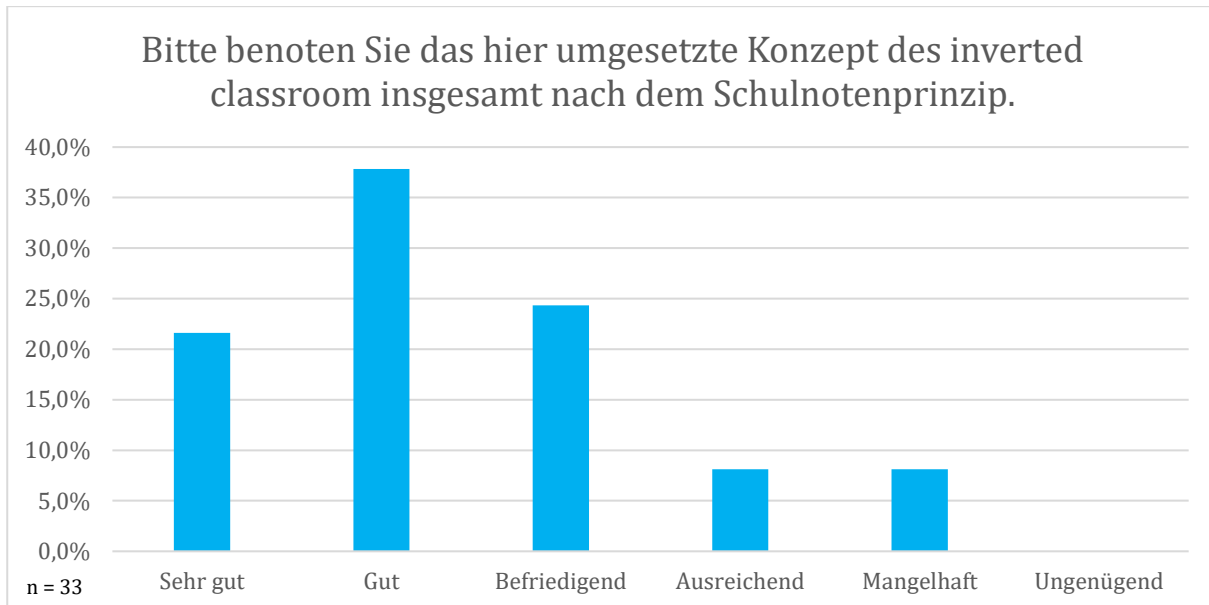


Abb. 4: Gesamtbeurteilung des Konzepts des Inverted Classrooms

Knapp 60 % der Studierenden benoteten das umgesetzte Konzept des Inverted Classrooms als sehr gut oder gut. Die hohe Zustimmungsrates bekräftigte uns den Weg weiterzugehen.

Die Studierenden konnten ihr Feedback auch in Freitextfeldern äußern. Uns ist positiv aufgefallen, dass 22 von 33 Studierenden, also 2/3, von dieser Möglichkeit Gebrauch gemacht haben. Noch mehr hat uns überrascht, dass die Stellungnahmen sehr ausführlich waren. 14 der Stellungnahmen begrüßten explizit das Konzept, nur eine lehnt es ab: „normale Vorlesungen mit dann anschließenden Übungsaufgaben zuhause, die dann kurz in der Präsenz besprochen werden, würden mir besser gefallen.“ Demgegenüber hier einige beispielhafte Stellungnahmen der Befürwortenden: „Ich finde es cool, weil es einfach mal ein bisschen Abwechslung bringt.“ – „Generell finde ich das Inverted Classroom-Konzept super und würde mir wünschen, dass auch andere Dozenten dies in Zukunft anbieten.“ – „Das Konzept des Inverted Classroom(es) ist super; es regt einen mehr an Vorbereitungen für die Vorlesung zu erledigen, sowie Nachbearbeitungen.“ – „Grundsätzlich mag ich das System des "Inverted Classroom", weil es mich motiviert mich intensiv auf die Vorlesungen vorzubereiten.“ – „Tolles Konzept. So lernt und bleibt der Stoff gleich viel besser hängen, da man sich es quasi zuerst selbst beibringen muss.“ – „Ich finde das Konzept des Inverted Classrooms sehr gut geplant, ausgearbeitet und durchgesetzt. Als Studentin zwingt es mich förmlich dazu, das wöchentlich zu bearbeitende Material auch wirklich zu bearbeiten. Des Weiteren hilft es mir sogar dabei, das Lernen für dieses Fach nicht aufzuschieben.“

Sowohl die expliziten Befürwortenden als auch jene, bei denen eine Zuordnung nicht möglich war, gaben insbesondere Hinweise zur Durchführung der Präsenzveranstaltungen, etwa zur zeitlichen Verteilung der Gruppenphase im Vergleich zur gemeinsamen Besprechung der Übungsaufgaben. Vor allem aber regten beinahe alle Rückmeldungen an, Lösungen zu den Übungsaufgaben bereitzustellen. Warum wir das noch nicht umgesetzt haben, gleichwohl aber künftig umsetzen werden, haben wir oben schon erläutert.

Die Präsenzveranstaltungen zu den Kapiteln 4 und 5 waren weniger stark frequentiert als die Veranstaltungen 1 bis 3. Es ist auch in herkömmlich durchgeführten Veranstaltungen nicht ungewöhnlich, dass der Besuch im Laufe des Semesters nachlässt. Allerdings war der Einbruch im Dienstagskurs deutlich: von etwa 35 auf 11 Teilnehmende. Montags war der Einbruch (von 60 auf 40) in etwa normal. Eine Stellungnahme machte uns aufmerksam auf einen

Aspekt, den wir so nicht erwartet hatten: „Außerdem kann ich mir vorstellen das einige nicht zur Vorlesung kommen, gerade weil man da ja nur noch in den Gruppen arbeitet.“ Diese Stellungnahme ist zwar eine Einzelaussage, aber es mag sein, dass Studierende im dritten Semester nach der Corona-Zeit nicht an das kollaborative Arbeiten in Gruppen gewöhnt sind und diese Art der Zusammenarbeit vermeiden wollen oder keinen Mehrwert darin sehen. Generell gibt es zahlreiche Auswirkungen der Pandemie (z.B. Kontaktprobleme, erschöpfte Studierfähigkeit, Ängste usw.), die die Studierenden belasten und zu einem deutlichen Anstieg bei Lern- und Arbeitsstörungen sowie Motivationsproblemen führen (vgl. (K. Schmieade, J. Küster, S. Stiehler, 2023, S. 32–50). Dies kann bspw. darin erkannt werden, dass Studierende Angst davor haben, eigene Wortbeiträge zu leisten oder auch in Gruppenarbeiten überfordert sind. Im schlimmsten Fall kann dies zum Fernbleiben in den Veranstaltungen führen.

3.2 Statistische Auswertung am Ende des Kurses

3.2.1 Vorbemerkung

Moodle zählt die einzelnen Klicks als Zugriffe auf den Kurs: Wenn ein Studierender den Kurs aufruft und bei einer Moodle-Test-Aufgabe 15 Klicks hervorruft, dann zählt Moodle jeden dieser Klicks. Aus diesem Grund wurden insgesamt knapp 79.000 Klicks erfasst. Da das Aufrufen von Downloads in der Regel nur 1-3 Klicks verursacht, die Teilnahme an einem Test hingegen zu sehr vielen Aufrufen führt, bestehen Grenzen in der Würdigung der Erfassung von Aktivitäten. Andererseits kommen die Inhalte (Videos, Tests, Downloads etc.) insgesamt relativ gleichmäßig in den einzelnen Kapiteln vor, sodass dennoch durchaus Rückschlüsse auf das Verhalten der Teilnehmenden gezogen werden können.

Ebenso gibt es Grenzen bei der Berücksichtigung der Ergebnisse und den erreichten Prozents in den Aktivitäten, beispielsweise bei den Moodle-Tests. Die Ermittlung der Durchschnittswerte beruht auf den vorliegenden abgeschlossenen Aktivitäten. Sofern eine Aktivität abgeschlossen wurde, fließt dieses Ergebnis, unabhängig davon, ob alle oder z.B. nur eine Testaufgabe bearbeitet wurden, in die Ermittlung der Durchschnittsbewertungen ein. Daraus resultieren Schwierigkeiten in der Einordnung und Vergleichbarkeit der verschiedenen Durchschnittsnoten, da sie auf unterschiedlicher Anzahl an Ergebnissen basieren können.

3.2.2 Zugriffe auf den Kurs (Aktivitätserfassung)

Der Moodle-Kurs wurde ab dem 18.9.2023 (Woche 1) bereitgestellt. Ab Woche 3 – eine Woche vor Vorlesungsbeginn – haben wir die Zugriffe erfasst:

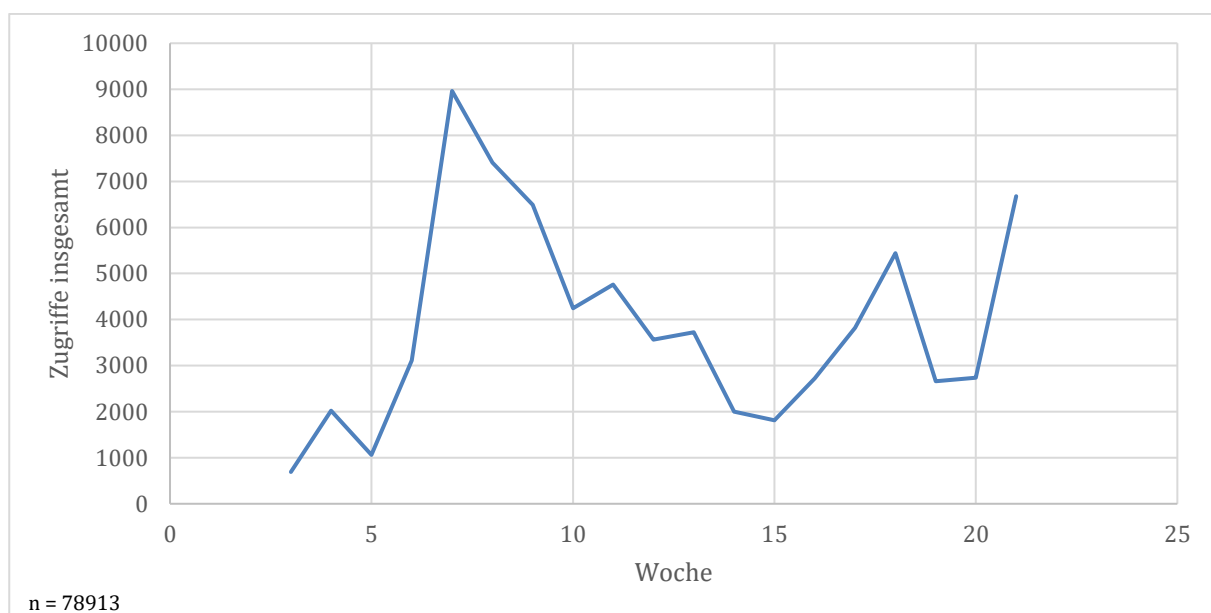


Abb. 5: Zugriffe auf den Moodle-Kurs (Aktivitätserfassung)

In der in der Grafik angegebenen Woche 6 haben wir – in Präsenz – das Konzept des Inverted Classrooms erläutert. In Woche 7 startete das Konzept mit Kapitel 4. Das erklärt den hohen Peak von 9.000 Zugriffen. Zu diesem Zeitpunkt waren 221 Studierende im Moodle-Kurs eingeschrieben, was einen durchschnittlichen Wert von rund 40,7 Klicks pro Studierenden ergibt. Da es jedoch auch Studierende gibt, die sich nur in den Kurs eingeschrieben haben und dann nicht mehr aktiv waren, ist die Anzahl der Klicks pro tatsächliche Nutzer höher einzuschätzen.

Sodann sind die Zugriffe jedoch zunächst gesunken, um kurz vor der Klausur wieder – erwartungsgemäß - deutlich anzusteigen. Hier gilt es herauszufinden, welche Gründe für den zwischenzeitlichen Rückgang ursächlich sind. Bei der Adaption der Gründe für einen Studienabbruch (vgl. Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH [DZHW], 2017, 11ff.) auf das vorliegende Konzept kommen aus unserer Sicht folgende Gründe infrage:

- Generell nachlassendes Interesse.
- Eine Fehleinschätzung der Komplexität der Inhalte (am Anfang erscheint es eher trivial) könnte zu der Annahme führen, dass die Bearbeitung der Inhalte nicht notwendig oder kurz vor der Klausur ausreichend sei. Dies führt zu einer nachlassenden mentalen Anstrengung sowie nachlassender Auseinandersetzung mit den Inhalten (vgl. auch Kerres, 2018, S. 179–180).
- Eine Überforderung der Studierenden mit der regelmäßigen Auseinandersetzung mit den Inhalten oder
- Eine sich im Zeitablauf entwickelnde generelle Ablehnung des Lehrkonzeptes.

Wir haben zwar darauf verzichtet festzuhalten, wie viele Klicks pro Lerneinheit mindestens erforderlich sind, um diese einmal abzuschließen. Dennoch hätten wir erwartet, dass die Klicks in etwa in einer Range von 5.000 bis 9.000 auch in künftigen Wochen liegen. Ungeachtet der Ausreißer in den Wochen 14 und 15 – das war die Weihnachtszeit vom 18.12. bis 31.12. – hat uns der kontinuierliche Rückgang schon überrascht, zumal die Anzahl der Studierenden, die im Moodle-Kurs eingeschrieben waren, kontinuierlich (trotz einiger Austragungen aus dem Kurs) bis Ende Januar auf max. 256² gestiegen ist. So lagen beispielsweise in Woche 12 nur noch 3.566 Klicks von 245 Studierenden vor, was für diesen Zeitraum einen durchschnittlichen Wert von nur noch rund 14,5 Klicks pro Studierenden ergibt. Das ist auch deshalb beachtlich, weil bei neu eingeschriebenen Studierenden davon ausgegangen werden kann, dass sie – zumindest teilweise – die bisher verpassten Lerneinheiten nachholen würden.

Auf der anderen Seite ist die Nutzung des Moodle-Kurses insgesamt als hoch einzuschätzen. In der Phase der herkömmlichen Präsenzlehre lagen die Zugriffe nur zwischen 1.000 und 2.000. Zu diesem Zeitpunkt gab es noch keine Videos oder Moodle-Test, d.h. die Studierenden haben sich nur die Gliederung, Foliensätze und Lehrtexte herunterladen können. Nach Beginn des Inverted Classrooms-Konzepts mit Videos und Moodle-Tests lagen die Zugriffe immer über dem vorherigen Niveau konventioneller Lehre. Daraus lässt sich vorsichtig eine stete und im Vergleich zur herkömmlichen Präsenzlehre höhere Auseinandersetzung mit den Inhalten ableiten.

3.2.3 Teilnahme an Klausur in Abhängigkeit von Moodlefortschritt und Moodlebewertungen

Wir wollten wissen: Haben (1) die mengenmäßige Nutzung *aller* Moodle-Materialien („Moodlefortschritt“) und (2) die Ergebnisse der Moodle-Tests („Moodlebewertungen“) Einfluss auf die Bereitschaft, am Ende des Semesters die Präsenzklausur mitzuschreiben?

² Dies Wert repräsentiert diejenigen Studierenden, die max. zu einem **Zeitpunkt** im Moodlekurs eingeschrieben waren. Bei der Betrachtung von **Zeiträumen** kann dieser Wert größer sein, da in dieser Spanne auch etwaige Abmeldungen aus dem Kurs noch berücksichtigt werden.

Zu (1) Moodlefortschritt

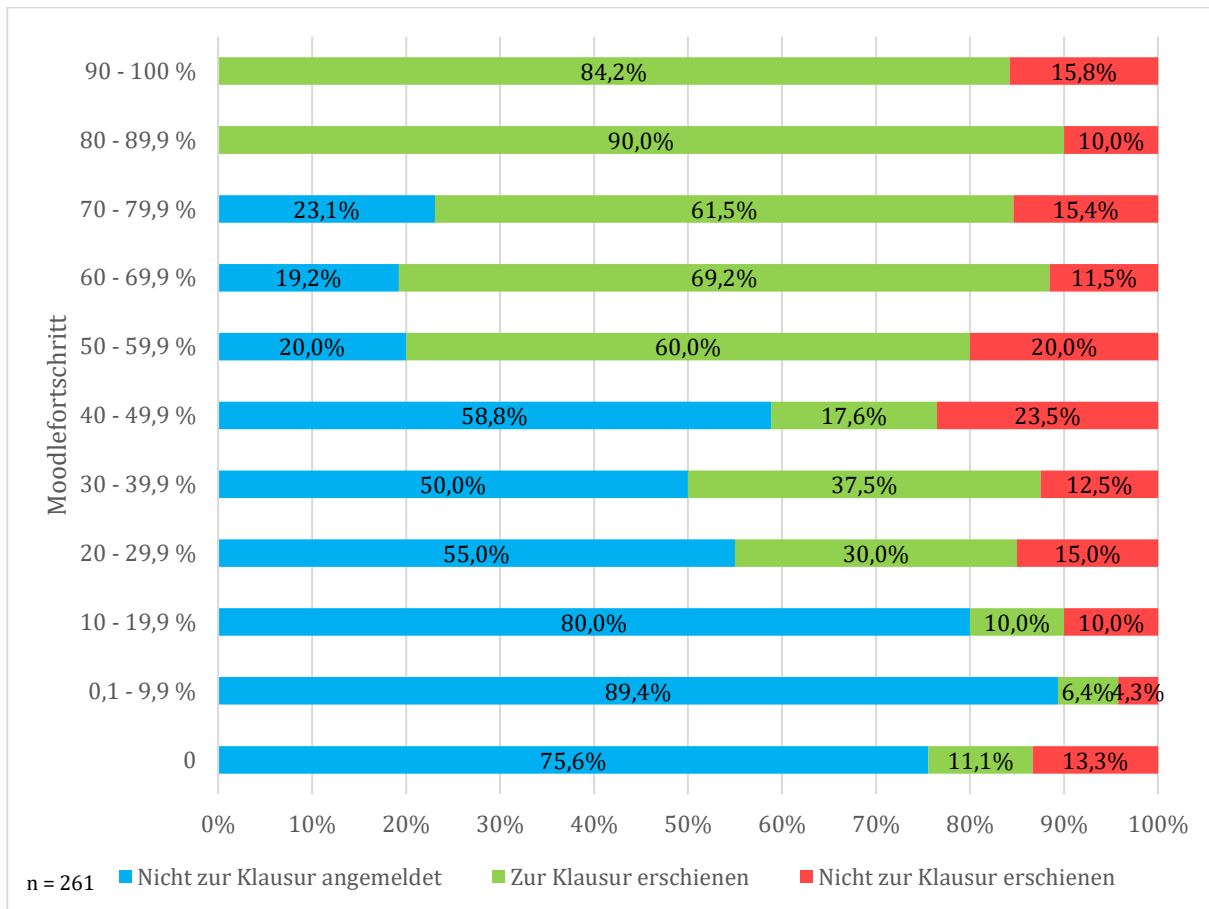


Abb. 6: Teilnahme an Klausur nach Moodlefortschritt

Von jenen Studierenden, die mindestens 80 % der bereitgestellten Moodle-Materialien bearbeitet haben, haben sich alle zur Klausur angemeldet. Umgekehrt: Von jenen Studierenden, die zwar mehr als 0, aber weniger als 10 % der bereitgestellten Materialien bearbeitet haben, haben sich knapp 90 % nicht zur Klausur angemeldet. In der Gruppe von 10 % bis 20 % der Materialbearbeitung waren es 80 %, die sich nicht angemeldet haben.

Auffällig ist, dass von der Gruppe der im Moodle-Kurs eingeschriebenen Studierenden, die nichts bearbeitet haben, sich immerhin knapp 25 % zur Klausur angemeldet haben (wobei mehr als die Hälfte davon nicht zum Klausurtermin erschienen sind). Wir können davon ausgehen, dass es sich hier um Studierende handelt, die die Veranstaltung zum Teil im herkömmlichen Parallelkurs des anderen Dozenten besucht haben oder die infolge einer Ausbildung der Auffassung waren, auch ohne Moodle-Unterstützung die Klausur angehen zu können.

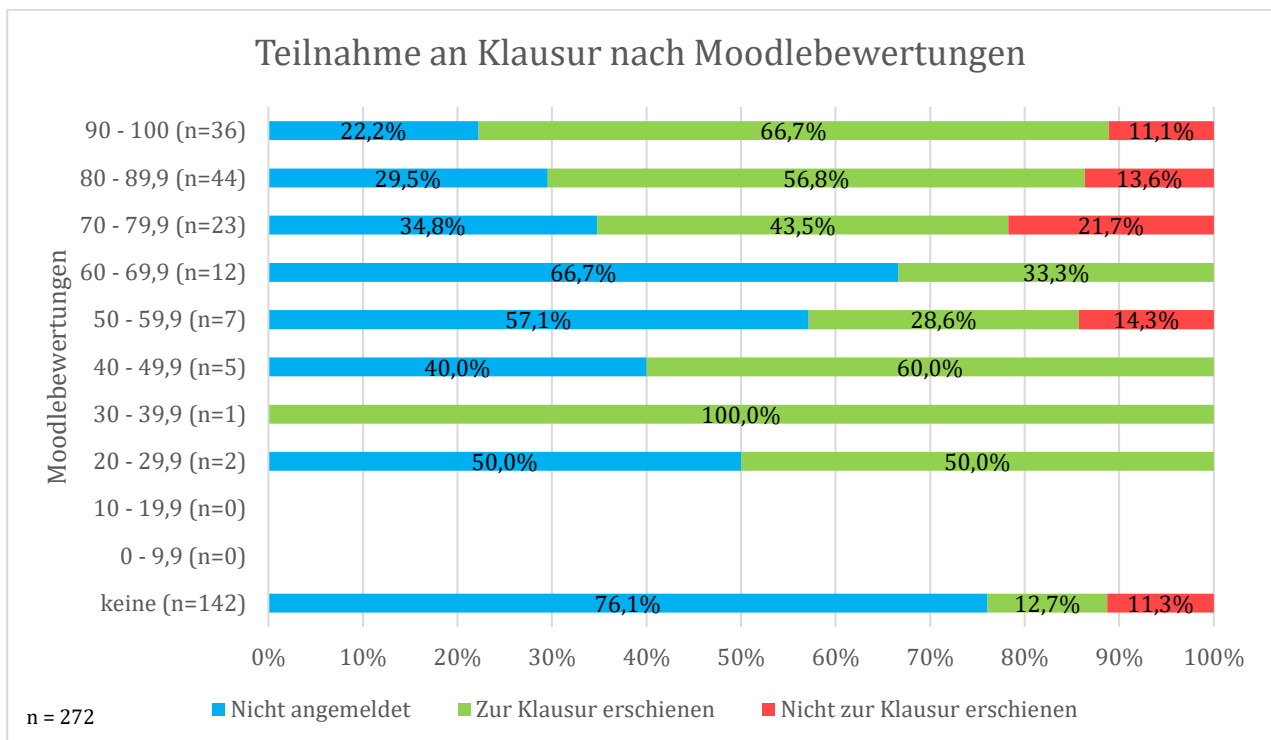
Zu (2) Moodlebewertungen (Moodle-Test-Ergebnisse)³

Abb. 7: Teilnahme an Klausur nach Moodlebewertungen

Insgesamt 142 Studierende haben die Moodle-Tests nicht bearbeitet. Von den verbleibenden 130 Studierenden sind Ergebnisse von besser als 20 % erzielt worden, wobei die Grundgesamtheit jener, die Ergebnisse bis 60 % erzielt haben, recht gering ist. Hinsichtlich der darüberliegenden Ergebnisse lässt sich sagen: Mit der Güte positiver Moodlebewertungen steigt auch die Bereitschaft, sich zur Klausur anzumelden. Bei der großen Gruppe derjenigen, die mehr als 70 % erzielt haben, sinkt auch kontinuierlich das Fernbleiben von der Klausur.

Etwas vereinfacht bestätigt die Auswertung die These, dass Studierende, die sich sicher fühlen, eher bereit sind, sich zur Klausur anzumelden und auch zu erscheinen.

³ Die Abschlussverfolgung und somit der Kursfortschritt müssen im Gegensatz zu den Bewertungen in Moodle manuell markiert werden. Es kann also sein, dass jemand eine Bewertung erhalten, aber den Inhalt in Moodle nicht als abgeschlossen markiert hat. Daher gibt es Unterschiede in der Grundgesamtheit der betrachteten Daten von (1) im Vergleich zu (2).

3.3 Klausurergebnisse

3.3.1 Klausurergebnisse in Korrelation zum Kursfortschritt und zu den Moodle-Test-Ergebnissen

Nachdem im vorherigen Abschnitt gezeigt wurde, dass die Bereitschaft, sich zur Klausur anzumelden, sowohl mit dem Kursfortschritt als auch mit den Moodlebewertung ansteigt, wollten wir nun wissen, ob diese Elemente auch positiv mit den Klausurergebnissen korrelieren. Dazu haben wir Kreuztabellen gebildet, die darstellen, in welchem Zusammenhang (1) das Klausurergebnis zum allgemeinen Kursfortschritt sowie (2) zu den durchschnittlich erzielten Bewertungen in den angebotenen Moodle-Inhalten (Moodle-Tests und H5P-Elemente in Form von Kontrollfragen in den Videos) steht. Da die Klausurergebnisse aufgrund der prozentualen Betrachtungsweise sehr heterogen und die Anzahl der Studierenden mit gleichen Ergebnissen sehr gering sind, haben wir die Daten in 25%-Clustern gebündelt, um belastbarere Aussagen zu generieren.

Zu (1): Klausurergebnis und Moodle-Fortschritt

Aus Abb. 8 wird deutlich, dass das Bestehen der Klausur bzw. ein gutes Klausurergebnis stark von der Bearbeitung der Moodle-Inhalte abhängen:

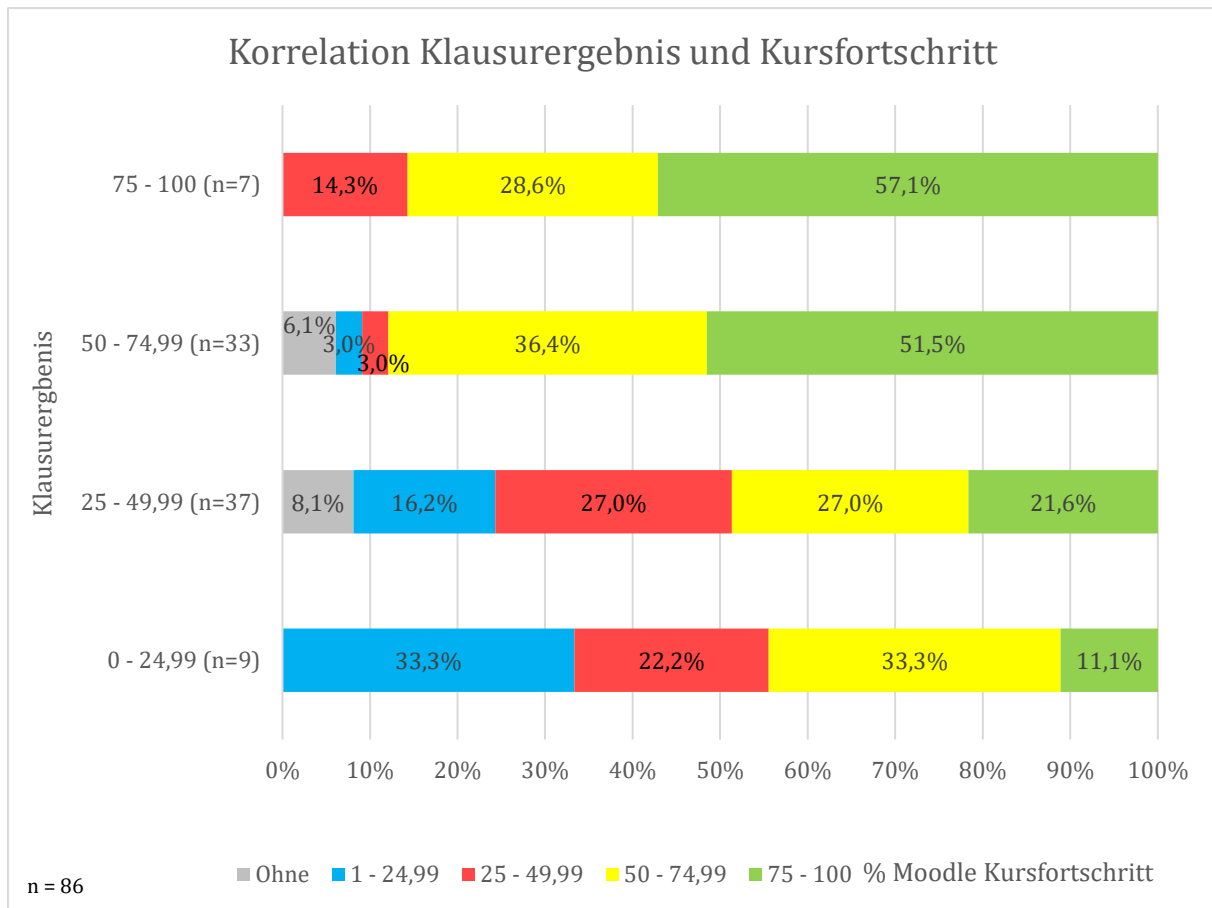


Abb. 8: Korrelation Kursfortschritt und Klausurergebnis

In der – allerdings kleinen - Gruppe derjenigen, die 75 – 100 Prozent der möglichen Punkte in der Klausur erreicht haben, hat der Großteil der Studierenden mit rund 85 % mindestens die Hälfte der bereitgestellten Inhalte bearbeitet. Diejenigen, die den Kurs nicht oder nur wenig (bis 25 %) bearbeitet haben, sind hingegen bei diesen guten Klausurergebnissen nicht vertreten.

Ein ähnliches Bild zeigt sich auch bei denjenigen, die 50 – 75 % in der Klausur erreicht haben. Auch hier hat der überwiegende Teil mit knapp 88 % den Kurs im großen Umfang absolviert. Mit schlechter werdenden Klausurergebnissen verschieben sich die Verhältnisse, und der Anteil derjenigen, die sich nicht ausgiebig mit den Inhalten beschäftigt haben, steigt auf ca. 50 %. Unter 3.2.3 haben wir ausgeführt, dass die Anmeldungen zur Klausur eher bei denen erfolgen, die sich aktiv vorbereitet haben. Dann kann im Umkehrschluss folgende These aufgestellt werden: Wenn alle Teilnehmenden des Moodlekurses auch die Klausur geschrieben hätten, wären im unteren Bereich der Klausurergebnisse noch deutlicher jene zu finden gewesen, die sich nicht oder unzulänglich über Moodle vorbereitet haben.

Zu (2): Klausurergebnis und Moodle-Test-Ergebnis

Auch bei der Betrachtung der in Moodle erhaltenen Bewertungen der Tests sowie der H5P-Elemente und der in der Klausur erzielten Prozentpunkte (Klausurergebnis, siehe Abb. 9) lassen sich starke Korrelationen feststellen.

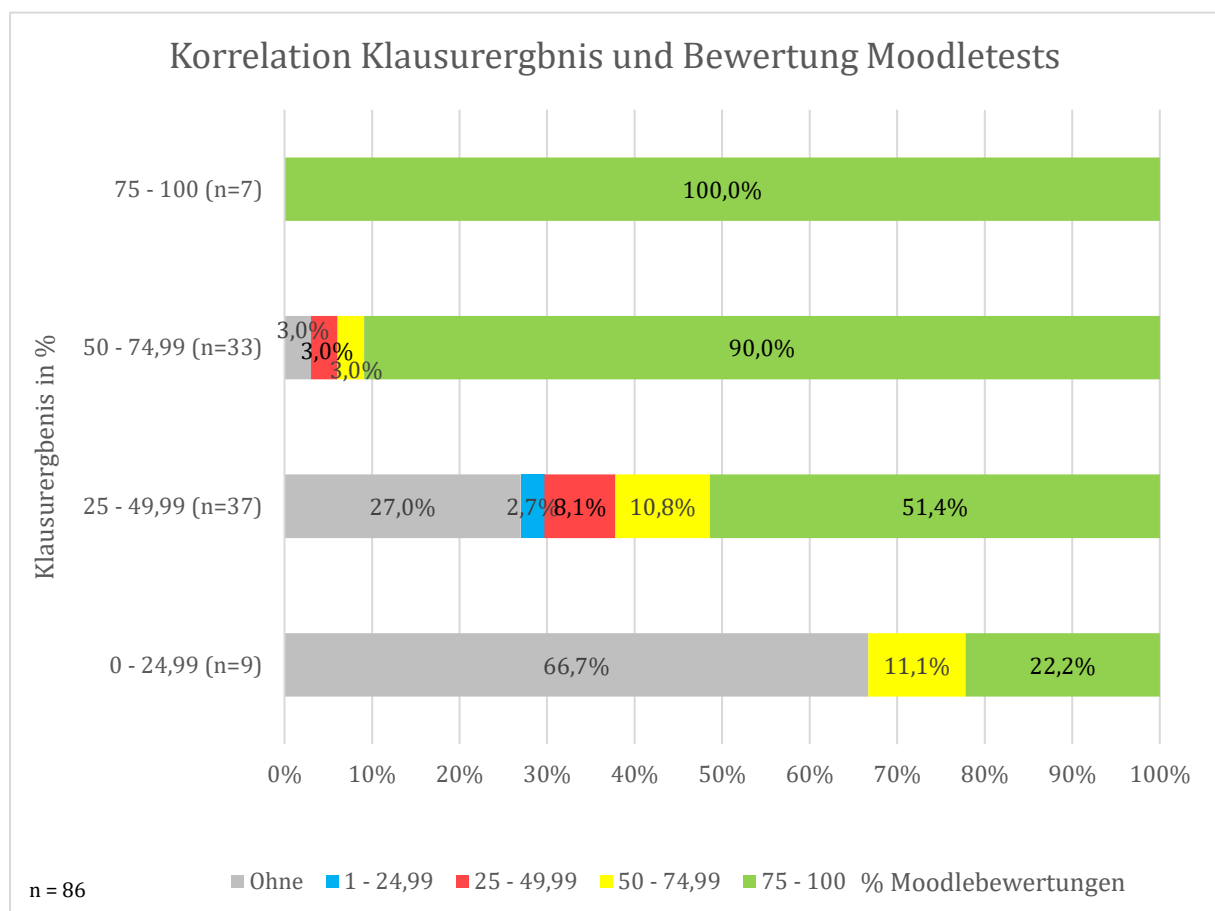


Abb. 9: Korrelation Klausurergebnis und Bewertung der Moodle-Tests

In der Gruppe derjenigen mit einem guten bis sehr gutem Klausurergebnis sind ausschließlich Studierende vertreten, die auch in Moodle mindestens 75 % der erreichbaren Punkte erreicht haben. Auch im nächsten Cluster zeigt sich der starke Zusammenhang zwischen den Bewertungen in Moodle und dem Bestehen der Klausur.

Im Bereich derjenigen, die nur 25 – 50 % in der Klausur erreicht haben, stellt sich jedoch die Frage, warum ca. 62 % derjenigen, die in Moodle gute Bewertungen erhalten haben, die Klausur dennoch nicht bestehen konnten. Dazu gibt es verschiedene Erklärungsansätze:

- Insgesamt zehn Studierende, von denen der Großteil gute Moodle-Bewertungen und einen Kursfortschritt größer als 75 % nachweisen, sind nur denkbar knapp gescheitert, da lediglich 2,5 Punkte zum Bestehen fehlten.
- Diejenigen, die gute Bewertungen in Moodle erhielten, haben die Inhalte evtl. mehrfach bearbeitet, sodass ein Wiederholungseffekt bei genau diesen Tests eingetreten sein könnte, der zu guten Ergebnissen geführt hat. Es ist aber fraglich, ob die Fachinhalte wirklich durchdrungen worden sind.
- Auch könnte das Niveau der Inhalte in Moodle über den tatsächlichen Lernstand getäuscht haben, wenn keine, wie im Konzept verankert, weitere Auseinandersetzung mit den bereitgestellten Lernangeboten in der Präsenz stattgefunden hat. Dies führt dazu, dass die Studierenden die Komplexität in der Klausur, die sich an den Präsenzaufgaben orientiert hat, nicht bewältigen konnten.
- Gute Bewertungen können auch daraus resultieren, dass die korrekten Ergebnisse der Moodle-Inhalte unter den Studierenden ausgetauscht wurden, also tatsächlich keinen eigenen Lernerfolg abbilden.
- Da in das Modulergebnis nur ein Anteil von 1/3 Buchhaltung einfließt, hingegen 2/3 Kostenrechnung und das Bestehen jedes Teilmodulfachs nicht erforderlich war, ist es für Studierende rational, sich in der Klausursituation zunächst auf die Bearbeitung der Fragen zur Kostenrechnung zu konzentrieren. Möglicherweise war dann am Ende der Klausurbearbeitung zu wenig Zeit, um den Teil Buchhaltung zu bearbeiten.

Das insgesamt schlechte Abschneiden derjenigen, deren Ergebnisse auch in Moodle nicht befriedigend waren, ist hingegen erwartbar gewesen.

Maßnahmen: Im Folgesemester werden wir die Studierenden über die Ergebnisse dieser Auswertung in Kenntnis setzen, insbesondere über die positiven Klausurwirkungen der Mitarbeit. Wir werden sie auffordern, sich selbstständig und regelmäßig am gesamten Umfang des Lehrkonzepts zu beteiligen. Um die Frequenz der Beteiligung der Studierenden zu erhöhen (und die Abbrecherquote zu senken), werden wir folgende wesentliche Einzelmaßnahmen ergreifen:

- Vergabe von personalisierten Aufgaben mit der Verpflichtung, die eigenen Ergebnisse im bereitgestellten Forum zu veröffentlichen und die Ergebnisse von anderen Studierenden zu kommentieren. In diesem Zusammenhang wird es im Vorfeld eine Einweisung über ein kollaboratives, wertschätzendes, aber kritisches Miteinander geben. Um die Einhaltung dieser Umgangsformen sicherzustellen, wird das Forum durch ausgewählte Moderator:innen begleitet.
- Inhalte werden nur noch für einen Zeitraum von 3 Wochen veröffentlicht und dann, wenn sie nicht erfolgreich bearbeitet worden sind, deaktiviert. Dies soll eine regelmäßige Bearbeitung der Inhalte forcieren und somit die Gefahr von Prokrastination verringern. Zu einem späteren Zeitpunkt werden die Inhalte dann für die Vorbereitungen auf die Klausur wieder öffentlich zugänglich gemacht.

Ein möglicher Nachteil dieser Maßnahmen ist darin zu sehen, dass sich einige Studierende evtl. unterfordert fühlen, wenn diese bereits über umfangreiches Vorwissen verfügen und sich dennoch mit den Inhalten auseinandersetzen müssen. Sie könnten gelangweilt sein und die Inhalte insoweit weniger konzentriert bearbeiten. Wir wollen proaktiv auf diese Möglichkeit hinweisen und das Argument umkehren: Wer fachinhaltlich bereits gut vorbereitet ist, sollte auch in der Lage sein, bei vermeintlich oder tatsächlich einfachen Tests gute und schnelle Ergebnisse zu erzielen.

Bei allen Bemühungen um eine hohe Anschlussfähigkeit von Lehr-Lern-Konzepten an die Bedürfnisse und Kompetenzen der Studierenden, wird sich keine Lösung etablieren lassen, die alle Studierenden gleichermaßen zum Bestehen der Klausur befähigt oder einen Abbruch des Seminars vollständig verhindert. Dennoch sehen wir deutliches Optimierungspotenzial in unserem Ansatz und halten die geplanten Maßnahmen für geeignet, um weitere Verbesserungen der Quoten in den kommenden Semestern herbeizuführen.

Klausurergebnisse im Vergleich zu früheren Ergebnissen bei herkömmlicher Lehre

Haben sich die Durchfallquoten im Vergleich zur bisherigen traditionellen Lehre verringert und die Durchschnittsnoten verbessert? Wir haben hierzu die beiden letzten Buchhaltungskurse desselben Dozenten aus dem SoSe 2022 (erstes Präsenzsemester nach Corona, aber noch mit unter Corona zwischenzeitlich eingeräumten Freiversuch) und dem SoSe 2019 (letztes Präsenzsemester vor Corona) herangezogen. Den Buchhaltungskurs in der Corona-Zeit haben wir nicht betrachtet, weil er weder dem Konzept des Inverted Classrooms noch dem der herkömmlichen Lehre entsprach.

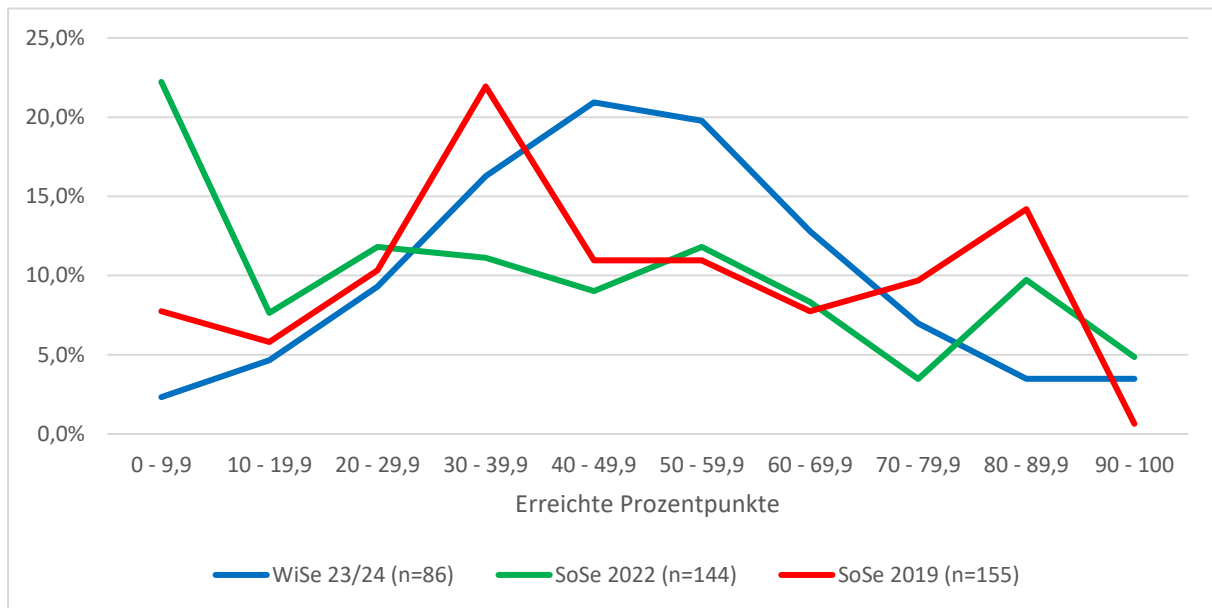


Abb. 10: Vergleich der Klausurergebnisse mit vorherigen Semestern

Bei den Klausurergebnissen aus dem WiSe 2023/24 liegt im Vergleich zu den Vorsemestern erstmals eine Normalverteilung vor. In den erreichten Prozentpunkten bis 30 % liegt der Anteil der Klausuren aus dem WiSe 2023/24 immer niedriger als in den Vergleichssemestern. Tendenziell umgekehrt verhält es sich allerdings im oberen Segment jener, die mehr als 70 % der Punkte erreicht haben. Auffällig ist auch, dass die Anzahl jener, die die Klausur im WiSe 2023/24 mitgeschrieben haben, deutlich niedriger ist, als die Zahl derjenigen, die sich im Moodle-Kurs eingeschrieben haben. Wir vermuten, dass sich die Studierenden anhand der direkten Feedbacks in Moodle im WiSe 2023/24 im Vorfeld zumindest teilweise selbst besser einschätzen konnten – sei es anhand der Moodle-Test-Ergebnisse oder anhand der schlichten Wahrnehmung des (dann nicht bearbeiteten) Materialumfangs. Oft wurde möglicherweise dann die Konsequenz gezogen, nicht an der Klausur teilzunehmen.

Gestützt wird diese Hypothese auch durch die höheren Teilnehmerzahlen der Vorsemester. In diesen gab es weder Moodle-Tests, noch – über den Foliensatz hinausgehende – Materialien. Das kann dazu geführt haben, dass Studierende, selbst wenn sie die Präsenzveranstaltungen nicht hinreichend besucht haben, gleichwohl in Unkenntnis ihres eigenen (geringen) Leistungsvermögens an der Klausur teilgenommen haben. Das würde dann auch erklären, warum die Klausurergebnisse in den Vorjahren im unteren Bereich so schlecht ausgefallen sind.

Hinzu kommt: Im SoSe 2022 gab es noch die während Corona eingeführte Regelung eines Freiversuchs. Die Wahrscheinlichkeit, auch bei unzureichender Vorbereitung von diesem Gebrauch zu machen, könnte durch diese Sonderregelung erhöht gewesen sein.

Auch bei der Betrachtung der Mittelwerte (MW) der erzielten Prozentwerte in den Klausuren können Unterschiede identifiziert werden. Im SoSe 2019 lag dieser bei 46,5, im SoSe 2022 bei 39,7 und im WiSe 2023/24 bei 48,5 Prozent. Der vergleichsweise gute Wert aus dem WiSe 2023/24 sollte nicht betrachtet werden, denn der hohe Grad schwacher Ergebnisse (zum Teil mit null Punkten) in den beiden Vergleichssemestern ziehen den dortigen Mittelwert natürlich sehr nach unten.

Ziel in den folgenden Semestern ist es, dass mehr Studierende die Klausur bestehen und auch die Anzahl der guten Klausurergebnisse ansteigt. Das würde in einer Verschiebung der Normalverteilung nach rechts zum Ausdruck kommen.

4 Zusammenfassung: Konsequenzen für ein Relaunch

Wir gehen von der These aus, dass Menschen, die sich für ein Studium entschieden haben, dabei für sich etwas mitnehmen wollen, d.h. an Kompetenzerwerb interessiert sind. Dabei unterstellen wir nicht eine flächendeckende intrinsische Motivation. Wir sehen es als unsere Aufgabe an, auch jene, die möglicherweise zunächst rein extrinsisch motiviert sind, zu erreichen und im Idealfall durch unser Lehrhandeln zu einer stärkeren intrinsischen Motivation anzuregen.

Das Konzept des Inverted Classrooms eignet sich aus unserer Sicht eher hierzu als klassische seminaristische Veranstaltungen. In diesem haben Studierende geringere Möglichkeiten, den eigenen Lernstand zu überprüfen, weil sie bei der Bearbeitung anwendungsbezogener Aufgaben allein gelassen werden. Demgegenüber ist beim Inverted Classroom die vorherige (alleinige) Erarbeitung der Fachinhalte die Voraussetzung, um Erfolgserlebnisse bzw. Selbstwirksamkeit (und damit einen Motivationsschub) bei den in der Gruppe zu bearbeitenden bzw. zu besprechenden anwendungsbezogenen Aufgaben zu erzielen (vgl. Kerres, 2018, S. 33–34).

Unsere Analyse hat gezeigt: Wer die Inhalte aktiv bearbeitet, erzielt bessere Ergebnisse als jene, die dies nicht tun. Diese Erkenntnis erscheint trivial und gilt vermutlich nicht nur für den Inverted Classroom, sondern auch für die herkömmliche Lehre. Allerdings ist der Moodle-unterstützte Inverted Classroom nicht nur geeignet, diesen Zusammenhang belastbar zu belegen, sondern auch für Studierende bereits im Semesterverlauf den positiven Nutzen ihrer Mitarbeit gespiegelt zu bekommen. Hierin steckt ein Moment der Selbstwirksamkeit, das sich positiv auf die Motivation auswirken kann.

Unter 2.1 haben wir unsere Erwartungen und Ziele formuliert. Wir wollten (1) die Verringerung der Abbrecherquote während des Semesters, (2) die Reduzierung der Durchfallquote und (3) die Verbesserung des Notendurchschnitts erreichen. Obwohl sich der Notendurchschnitt (3) leicht verbessert hat, haben wir keines der avisierten Ziele mit signifikanten Werten erreicht. Dennoch ist das Konzept des Inverted Classrooms nicht per se gescheitert, denn wir konnten belegen, dass jene, die sich auf das Konzept eingelassen haben, durchweg gute Ergebnisse erzielt haben. Die mangelnde Zielerreichung führen wir daher auf unsere unzureichende Umsetzung des Konzepts zurück.

Folglich werden wir am Konzept festhalten, müssen es aber verbessern. Unser Ziel im kommenden WiSe ist, (1) die Mitmacherquote zu erhöhen und (2) Studierende dazu zu befähigen, bessere Klausurergebnisse zu erzielen. Wir haben oben bereits einige Maßnahmen genannt und fassen diese jetzt in einer Gesamtbetrachtung zusammen.

Für die erfolgreiche Durchführung des Inverted Classrooms ist es zwingend erforderlich, dass die Studierenden sich eigenständig und zielorientiert mit **allen** Modulinhalten auseinandersetzen. Wir werden im kommenden Semester die Besonderheiten des Lehrkonzepts und die Notwendigkeit des Selbststudiums als Vorbereitung für die Präsenz noch deutlicher in unserer Konzept-Einführungsveranstaltung herausstellen. Diese Veranstaltung wird erneut aufgezeichnet, um Studierende, die nicht teilgenommen haben, mit den notwendigen Informationen zu versorgen.

Wir erwägen die Einführung eines Pre-Tests, der den Leistungsstand der Studierenden vor Modulbeginn dokumentiert. Auf diese Weise können ggf. jene, die bereits über gute Vorkenntnisse verfügen (etwa aufgrund einer kaufmännischen Ausbildung), zielgerichtet auf für sie noch notwendige Lerneinheiten gelenkt werden. Auch könnte der Pre-Test geeignet sein, jedem Studierenden die zu erwerbenden Kompetenzen vor Augen zu führen. Schließlich wäre es auch eine gute Möglichkeit, um bereits früh geeignete Kandidat:innen für etwaige Gruppenarbeiten oder die Rolle als Moderator:in zu identifizieren.

Generell – mit Ausnahme ggf. jener Bereiche, die aufgrund des Pre-Tests für Studierende mit Vorbildung identifiziert worden sind – werden die jeweiligen Kursabschnitte nur noch für einen Zeitraum von drei Wochen zur Verfügung gestellt und nachfolgende Inhalte nur nach Abschluss der vorgehenden Einheit freigeschaltet. Dies soll für eine kontinuierliche Bearbeitung der Inhalte sorgen und dient ebenso als Element der Qualitätssicherung, da wir mit einer größeren Datenbasis besser in der Lage sind, Schwachstellen in den Inhalten zu identifizieren.

Die bereitgestellten Videos bilden eine gute fachinhaltliche Basis. Obwohl wir selbst nicht proaktiv ins Auge fassen, die Videos neu zu produzieren, wollen wir diese Frage ergebnisoffen lassen. Im September 2024 wird eine kleine Arbeitsgruppe den Kurs evaluieren. Wir wollen insbesondere den Zeitaufwand erfassen, die inhaltliche Güte überprüfen und hinterfragen, ob der Niveauanstieg (von Moodle-Tests zu den in der Präsenz zu bearbeitenden Aufgaben) zu abrupt erfolgt. Die Arbeitsgruppe wird sich aus zwei oder drei BA-Studierenden aus dem 5. Semester und aus einem Master-Studierenden zusammensetzen. Sollte sich hieraus ergeben, dass schon an den Videos angesetzt werden muss – etwa durch prägnantere H5P-Inhalte oder durch kürzere Zuschnitte – werden wir dies umsetzen.

Die Moodle-Tests werden weiterhin als elementarer Bestandteil des Systems erhalten bleiben und bilden eine der ersten Feedbackmöglichkeiten für die Studierenden. In Abhängigkeit der Ergebnisse der o.g. Arbeitsgruppe werden die Fragen überarbeitet und ggf. schon in den Tests verschiedene Niveaustufen etabliert.

Dies gilt auch für die weiterführenden Aufgaben als Vorbereitung für die Präsenz. Ferner werden wir hier zukünftig Lösungshinweise bereitstellen, um den Studierenden mehr Orientierung und Sicherheit zu geben. Tendenziell wird es in diesem Zusammenhang auch zur Aufgabe gehören, diese im implementierten Forum zu veröffentlichen und die Lösungen von anderen Teilnehmenden kommentieren zu lassen.

Für die Arbeit in der Präsenzlehre könnten interessierte Studierende eine Rolle als Moderator:in übernehmen, um die Gruppenarbeiten in insgesamt großen Kursen effektiver zu gestalten. Diese Rolle kann auch rotieren. Wir halten die Maßnahme für geeignet, um das Wir-Gefühl und die Motivation im Kurs zu steigern. Es kann verdeutlicht werden: Wenn sich jeder einbringt, werden insgesamt bessere Ergebnisse erzielt.

Durch eine schon beschlossene Änderung der Studiengangprüfungsordnung müssen erstmals im WiSe 2025/26 beide Teilprüfungen – Kostenrechnung und Buchhaltung – für das Bestehen des Moduls bestanden werden. Diese von uns nicht zu verantwortende Änderung wird sicherlich dazu beitragen, die Beschäftigung mit Buchhaltung zu steigern.

Viele unserer jetzt angedachten Maßnahmen führen zu Mehraufwand. Dieser Mehraufwand soll aber einmalig sein. Einmal aufgesetzt, soll das Konzept des Inverted Classrooms nicht zu einer zusätzlichen Inanspruchnahme von Ressourcen führen. Wir gehen davon aus, dass der Kurs Buchhaltung von den Dozierenden mithilfe von etwa 4 h/Woche Zuarbeit von studentischen Hilfskräften bewerkstelligt werden kann.

Literaturverzeichnis

- Abeysekera, L., Dawson, P. (2014): Motivation and cognitive load in the Inverted Classroom: definition, rationale and a call for research, *Higher Education Research & Development*, Vol. 34: 1-14
- Anderson, L. W., Krathwohl, D. R. (2001). *A Taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Bloom, B. (1956). *Taxonomy of educational objectives. Vol. 1: Cognitive domain*. McKay, New York.
- Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH. (2017). *Motive und Ursachen des Studienabbruchs an baden-württembergischen Hochschulen und beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher*. Verfügbar unter: https://mwk.baden-wuerttemberg.de/fileadmin/redaktion/m-mwk/intern/dateien/pdf/Studium_und_Lehre/Studie_Motive_und_Ursachen_des_Studienabbruchs_an_baden-w%C3%BCrttembergischen_Hochschulen_Langversion.pdf
- Federau, M. (2023): *Grundlagen der Wirtschaft und finanziellen Unternehmensführung: Ein feedbackbasierter Inverted Classroom-Ansatz*. In: Behringer, S. (Hrsg), *Konferenzband CARF Luzern*.
- Gudjons, Herbert (2021): *Frontalunterricht - neu entdeckt. Integration in offene Unterrichtsformen*. 4., aktualisierte Auflage. Bad Heilbrunn: UTB; Verlag Julius Klinkhardt.
- Handke, J. (2017): *Gelingensbedingungen für den Inverted Classroom*. In: Zeaiter, S./Handke, J. (Hrsg.), *Inverted Classroom – The Next Stage. Konferenzband zur 6. ICM Fachtagung in Marburg 2017*. Tectum, Baden-Baden.
- Handke, J. (2020): *Handbuch Hochschullehre Digital*, 3. Aufl., Tectum, Baden-Baden.
- Kärchner, H., Gehle, M., Schwinger, M. (2022): *Entwicklung und Validierung des Modulare Fragebogens zur Evaluation digitaler Lehr-Lern-Szenarien (MOFEDILLS)*, *ZeHf*: 62-84.
- Kanning, U., Ohlms, M (2021): *Hochschullehre in Zeiten von Corona*, *Wirtschaftspsychologie*: 44-55.
- Kerres, M. (2018). *Mediendidaktik. Konzeption und Entwicklung digitaler Lernangebote* (De Gruyter Studium, 5. Auflage). Berlin: De Gruyter Oldenbourg.
- Kerres, M. (2021). *Didaktik. Lernangebote gestalten* (UTB. Erziehungswissenschaft). Münster, New York, Münster: UTB; Waxmann.
- Kley, C./Kull, S./Bitterli, C. (2021): *Erfahrungsbericht Onlineunterricht im Wahlpflichtmodul "Kunst der Finanzanlage"*. In: Behringer, S. (Hrsg). *Konferenzband CARF Luzern*.
- Kley, C., Kull, S. & Bitterli, C. (2021). *CARF Luzern 2021 Konferenzband. Controlling. Accounting&Audit. Risk&Compliance. Finanzen*. Verfügbar unter: <https://www.hslu.ch/-/media/campus/common/files/dokumente/w/ifz/seminare-konferenzen/carf/carf-2021-konferenzband-klein.pdf>
- Mayweg, E., Enders, N., Bohndick, C., Rückmann, J. (2023): *Online, blended oder Präsenz? Ein systematisches Literaturreview von Metaanalysen zur Effektivität hochschulischer Lehrformate*, *ZeHf*: 96-122.
- Modlinger, D. (2020). *eLearning und Mobile Learning – Konzept und Drehbuch*. *Handbuch für Medienautoren und Projektleiter* (X.media.press, 3., vollst. überarb. und akt. Auflage 2020). Wiesbaden: Springer Fachmedien Wiesbaden GmbH; Springer Vieweg.
- Schmieade, K., Küster, J., Stiehler, S. (2023). *Auswirkungen der Corona-Pandemie auf die psychische Gesundheit Studierender*. K. Schmieade, J. Küster, S. Stiehler *Eine Qualitative Inhaltsanalyse von Anfragen in der Psychosozialen Beratungsstelle (PSB)*.
- Wipper, A., Schulz, A. (2021): *Digitale Lehre an der Hochschule. Vom Einsatz digitaler Tools bis zum Blended-Learning-Konzept*, 1. Aufl., Verlag Barbara Budrich, Opladen.

Fit für die Nachhaltigkeitstransformation? – Zur Integration von Nachhaltigkeitskompetenzen in der Controllinglehre

Research Paper

Prof. Dr. Ute Vanini

Fachhochschule Kiel, Institut für Controlling, Sokratesplatz 2, D-24149 Kiel, E-Mail: ute.vanini@fh-kiel.de

Annika Jacobsen

Fachhochschule Kiel

Abstract

Die regulatorischen und kapitalmarktorientierten Anforderungen an die Nachhaltigkeitsberichterstattung und -steuerung steigen z.B. durch die EU-Taxonomie, die Corporate Social Responsibility Directive (CSRD) oder die European Sustainability Reporting Standards (ESRS) beständig an. Zudem integrieren Unternehmen immer mehr Nachhaltigkeitsziele wie z.B. den CO₂-Fußabdruck in ihre Unternehmenssteuerung. Daher wird vielfach eine (stärkere) Beteiligung des Controllings an der Nachhaltigkeitstransformation von Unternehmen gefordert, z.B. bei der Erhebung und Qualitätssicherung von Nachhaltigkeitskennzahlen. Es stellt sich die Frage, inwieweit zukünftige Controller über die dafür notwendigen Kompetenzen verfügen. Der Beitrag untersucht daher 459 Modulbeschreibungen von Controlling- und Nachhaltigkeitsmodulen betriebswirtschaftlicher Studiengänge an deutschsprachigen Hochschulen und wertet diese inhaltsanalytisch aus. Die Ergebnisse zeigen eine geringe und uneinheitliche Abdeckung von Konzepten und Instrumenten des Nachhaltigkeitscontrollings in der Controllinglehre. Dabei werden überwiegend grundlegende Begriffe und Konzepte des Nachhaltigkeitscontrollings und -reportings aufgegriffen, während regulatorische Anforderungen oder spezifische Instrumente wie die Wesentlichkeitsanalyse vereinzelt oder gar nicht als Modulinhalte benannt werden.

1 Problemstellung

Die Anforderungen von Regulatoren, Investoren, Banken und Kunden an die Nachhaltigkeitstransparenz von Unternehmen haben seit der Verabschiedung des Green Deals der Europäischen Union (EU) 2019, der eine Senkung der CO₂-Emissionen bis 2030 um mindestens 55 % im Vergleich zu 1990 sowie die Klimaneutralität der Mitgliedsstaaten bis 2050 zum Ziel hat, beständig zugenommen. Beispiele sind die EU Sustainable Finance Initiative, die EU-Taxonomie zur Klassifikation von nachhaltigen und nicht-nachhaltigen Wirtschaftsaktivitäten sowie die Corporate Sustainable Reporting Directive (CSRD) und ihre Umsetzung durch die European Sustainability Reporting Standards (ESRS) (für einen Überblick über die rechtlichen Entwicklungen (Sopp et al., 2023, S. 1-59)). Die Implikationen für die Berichterstattung von Unternehmen und damit für den Finanzbereich im Allgemeinen und das Controlling im Speziellen sind vielfältig. So müssen zukünftig Finanzunternehmen und Kapitalgesellschaften, die bestimmte Größenkriterien erfüllen, einen Nachhaltigkeitsbericht in ihren Lagebericht integrieren, in dem sie u.a. über wesentliche Informationen zu ihren Nachhaltigkeitszielen, wesentliche nachhaltigkeitsbezogene Risiken sowie Key Performance Indicators (KPI) zu diesen Größen berichten müssen (Sopp et al., 2023, S. 69 ff.). Aufgrund der o.g. Anforderungen wird zunehmend die Forderung erhoben, dass insbesondere das Controlling eine zentrale Informations- und Unterstützungsfunktion bei der Nachhaltigkeitstransformation von Unternehmen übernehmen soll (Littan et al., 2022). Dies betrifft neben der Bewertung der finanziellen Materialität von Informationen (Outside-In-Perspektive) auch die Bewertung der Impact Materialität für die Stakeholder (Inside-Out-Perspektive), die Strukturierung des Datenmanagements und die Sicherstellung eines prüffähigen Daten- und Kennzahlensets zur Nachhaltigkeitsleistung sowie die Integration von Nachhaltigkeitszielen in die Unternehmenssteuerung einschließlich der Anreiz- und Vergütungssysteme (AK Finanzierung der Schmalenbach-Gesellschaft, 2022, S. 10 f.; Kämmler-Burrak & Bauer, 2022).

Trotz der steigenden Anforderungen von Regulatoren, Investoren und Kunden an die Nachhaltigkeitsberichterstattung und das Nachhaltigkeitsmanagement zeigen Studien eine bislang eher geringe Beteiligung des Controllings an diesen Themen (Apitz et al., 2023; Ascani et al., 2021; Kämmler-Burrak et al., 2022; Petersen et al., 2021; Reimer et al., 2021). Als ein möglicher Grund werden fehlende Kompetenzen der Controller in der Nachhaltigkeitsbewertung und -steuerung vermutet. Erste internationale Studien deuten in der Tat auf einer eher geringe Integration von Konzepten und Methoden des Nachhaltigkeitscontrollings in die Rechnungswesen- und Controllinglehre an Hochschulen hin (Botes et al., 2014; Boulianne & Keddie, 2018; Boulianne et al., 2018). Für deutschsprachige Hochschulen liegen derartige Untersuchungen bislang nicht vor. Dies stellt insofern ein Problem dar, da die grundlegende Ausbildung von Controllern vielfach im Rahmen eines Studiums erfolgt und daher unklar ist, inwieweit die Controller gut auf die an sie in der Unternehmenspraxis gestellten Anforderungen der Nachhaltigkeitstransformation vorbereitet sind.

Dieser Beitrag untersucht daher die Frage, ob und wie Aspekte des Nachhaltigkeitscontrollings in die Controllinglehre von Hochschulen in der DACH-Region integriert sind. Zur Beantwortung dieser Forschungsfrage wurden 459 Modulbeschreibungen von Controlling- und Nachhaltigkeitsmodulen in betriebswirtschaftlichen Studiengängen analysiert. Die Ergebnisse deuten auf eine geringe quantitative und qualitative Abdeckung des Nachhaltigkeitscontrollings in der Controllinglehre hin. Dabei werden überwiegend grundlegende Begriffe und Aspekte des Nachhaltigkeitscontrollings und -reportings aufgegriffen, während regulatorische Anforderungen oder spezifische Instrumente wie die Wesentlichkeitsanalyse vereinzelt oder gar nicht als Modulinhalt benannt werden. Die Ergebnisse leisten daher einen wichtigen Beitrag zur Diskussion der Zukunftsfähigkeit der Controlling-Curricula und ihrer Weiterentwicklung, um die Employability zukünftiger Controller zu verbessern, damit diese ihren Beitrag zu einer erfolgreichen Nachhaltigkeitstransformation der Unternehmen leisten können.

Der Rest des Beitrags ist wie folgt aufgebaut: Zunächst werden ein Überblick über den Stand der Forschung zur Beteiligung des Controllings an der Nachhaltigkeitssteuerung von Unternehmen und zur Integration von Konzepten und Instrumenten des Nachhaltigkeitscontrollings in die Controllinglehre an Hochschulen gegeben und eine Forschungslücke herausgearbeitet. Im 3. Kapitel werden die Stichprobe der untersuchten Modulbeschreibungen hergeleitet und der methodische Analyseansatz – insbesondere das Kategoriensystem für die durchgeführte Inhaltsanalyse

– beschrieben. Anschließend werden in Kapitel 4 die Ergebnisse der quantitativen und qualitativen Inhaltsanalyse vorgestellt und in Kapitel 5 diskutiert. Zudem werden Implikationen für die Controllinglehre und -Forschung abgeleitet und wesentliche Limitationen der eigenen Studie benannt.

2 Stand der Forschung

Die Auswertung von Daten sowie die Generierung und Bereitstellung entscheidungsrelevanter Informationen für das Management ist grundsätzlich eine der wesentlichen Aufgaben des Controllings. Das Ziel ist dabei, die Qualität und Rationalität von Managemententscheidungen zu verbessern und dadurch die Unternehmensziele besser zu erreichen (Weber & Schäffer, 2022, S. 43 ff.). Aufgrund des Management Approach müssen vielfach auch entscheidungsrelevante Informationen in der Geschäftsberichterstattung von Unternehmen berichtet werden, so dass das Controlling auch zur externen Berichterstattung beiträgt.

Durch die Nachhaltigkeitstransformation sowie die Umsetzung der regulatorischen Anforderungen der CSRD und der ESRS an die Nachhaltigkeitsberichterstattung wird der Aufgabenbereich des Controllings wesentlich erweitert. So wird der Umfang der berichtspflichtigen Nachhaltigkeitsinformationen u.a. von einer Wesentlichkeitsanalyse bestimmt. Zur Bewertung der finanziellen Materialität für das Unternehmen kommt jetzt die Bewertung der sogenannten Impact-Materialität aus der Perspektive der Stakeholder hinzu (doppelte Wesentlichkeit). Desweiteren sollen die Nachhaltigkeitsinformationen sowohl kurz- als auch mittel- und langfristige Zeithorizonte und auch die gesamte Lieferkette des Unternehmens abdecken, was ebenfalls eine wesentliche Perspektiverweiterung für die betroffenen Unternehmen und ihr Controlling darstellt. Zudem geben die ESRS konkrete Vorgaben, wie diese Informationen zu erheben und zu berichten sind. Zudem wird zukünftig der Nachhaltigkeitsbericht verpflichtend einer (eingeschränkten) Prüfung durch den Wirtschaftsprüfer unterzogen (Sopp et al., 2023, S. 72-85).

Neben den Anforderungen aus der externen Nachhaltigkeitsberichterstattung erweitert sich der Aufgabenbereich des Controllings auch aus einer stärkeren Integration von Nachhaltigkeitsaspekten in die Geschäftsmodelle und Strategien des Unternehmens. Ausgehend von den Erwartungen der Stakeholder werden im normativen Nachhaltigkeitsmanagement eine grundsätzliche Positionierung des Unternehmens zur Nachhaltigkeit entwickelt und in die Mission, Vision und ggf. das Leitbild integriert. Das strategische Nachhaltigkeitscontrolling unterstützt dann das Management bei der Ableitung einer Nachhaltigkeitsstrategie und strategischen Nachhaltigkeitszielen. Häufig müssen zusätzlichen zu den finanziellen Zielen soziale und ökologische Ziele festgelegt und in das Ziel- und Leistungsmessungssystem einschließlich der Anreiz- und Vergütungssysteme integriert werden. Zudem müssen die Nachhaltigkeitsziele durch geeignete nichtfinanzielle Kennzahlen gemessen und nach festgelegten Qualitätsstandards erhoben und berichtet werden (Baumüller et al., 2023; Vanini, 2022, S. 177 f.). Wesentliche Instrumente sind z.B. Stakeholder- und Wesentlichkeitsanalysen sowie die Integration der Nachhaltigkeit in eine Balanced Scorecard (BSC). Zudem müssen dann für die Nachhaltigkeits-KPI in der Planung Vorgabewerte abgeleitet werden, deren Erreichung kontrolliert und die Zielerreichung mit entsprechenden Anreizen belohnt werden (Nachhaltigkeitssteuerung). Dabei muss eine prüfungsfähige Qualität der Daten und Kennzahlen sichergestellt werden (AK Finanzierung der Schmalenbach-Gesellschaft, 2022, S. 10 f.; Littan et al., 2022). Zudem müssen das Management und die Stakeholder durch die interne und externe Nachhaltigkeitsberichterstattung über die Nachhaltigkeitsleistung informiert und Nachhaltigkeitsziele systematisch in das Investitionscontrolling integriert werden (Sailer, 2022, S. 165 ff.).

Trotz der o.g. Herausforderungen zeigen Studien eine bislang unterdurchschnittliche Beteiligung des Controllings an der Nachhaltigkeitsberichterstattung und -steuerung (Apitz et al., 2023; Ascani et al., 2021; Kämmler-Burrak et al., 2022; Petersen et al., 2021; Reimer et al., 2021). So ermitteln Reimer et al. (2021) auf Grundlage ihrer Online-Befragung von 291 Unternehmen aus der DACH-Region, dass nur 50% über ein externes Nachhaltigkeitsreporting verfügen. 55% der großen, aber nur 11% der kleinen Unternehmen haben sich quantifizierte Nachhaltigkeitsziele gesetzt. Nur 12% der Unternehmen kaskadieren Nachhaltigkeitsziele über die Unternehmenshierarchie, nur 9% incentivieren die Zielerreichung im Management. Zudem treibt nur in 10% der Unternehmen das Controlling Nachhaltigkeitsthemen (Reimer et al., 2021, S. 16 ff.). Die Ergebnisse werden durch eine Studie des Internationalen

Controller Vereins bestätigt. Hier geben nur bei 11% der Befragten an, dass die Verantwortung für das Nachhaltigkeitscontrolling ausschließlich beim Unternehmenscontrolling liegt. Wenn das Controlling an der Nachhaltigkeitstransformation beteiligt ist, übernimmt es vor allem die Integration von ökologischen Aspekten in Kennzahlen und Kennzahlensysteme (27%), die Investitionsbewertung (25%) sowie in strategische Steuerungsinstrumente (23%) (Kämmler-Burrak et al., 2022, S. 20 f.). Auch im Investitionscontrolling werden Nachhaltigkeitsaspekte trotz ihrer hohen strategischen Bedeutung oftmals als unwichtig eingestuft und daher in Investitionsrechenverfahren nicht berücksichtigt (Apitz et al., 2023, S. 23 f.)

Als wesentliche Gründe für die bisher geringe Beteiligung des Controllings an der Nachhaltigkeitstransformation gelten u.a. eine hohe Auslastung der Controller mit Routineaufgaben des finanziellen Reportings, fehlende Kompetenzen insbesondere bei der Bewertung der ökologischen und der sozialen Nachhaltigkeitsleistung und ein unzureichendes Nachhaltigkeitsbewusstsein (Ascani et al., 2021; Petersen et al., 2021; Williams, 2015, S. 267 ff.). So ist der Controller-Alltag durch einen straffen Terminkalender mit regelmäßigen finanziellen Analysen und Berichtszyklen stark vorstrukturiert, so dass der zeitliche Spielraum für eine aktivere Rolle im Nachhaltigkeitscontrolling fehlt. Ein weiterer Hinderungsgrund ist die starke Ausrichtung vieler Controller auf das operative Geschäft (Petersen et al., 2021, S. 11 f.).

Die herausragende Bedeutung von Nachhaltigkeitskompetenzen im Controlling wird auch von Kämmler-Burrak et al. (2022, S. 35 ff.) betont. So schätzen 78% der von ihnen befragten Controller das allgemeine Verständnis von Nachhaltigkeitsanforderungen (Sustainability Literacy) für die Controllertätigkeit als sehr oder eher relevant ein. Der Fähigkeit zum Systemdenken wird von 67% eine sehr hohe bzw. hohe Relevanz beigemessen. Zentral für das Nachhaltigkeitsreporting und -controlling werden zudem die Kenntnis bzw. das Verständnis von Nachhaltigkeitsanforderungen von Regulatoren, Kunden, Banken und des Kapitalmarkts als wesentliche Stakeholdergruppen erachtet. Zudem wird auch auf das Verständnis technischer, wie beispielsweise der Messung von CO₂-Emissionen wie auch sozialer Aspekte wie z.B. Verletzungen der Menschenrechte hingewiesen. Insgesamt wird von Controllern gefordert, fachliche Kenntnisse im Bereich der Nachhaltigkeitsberichterstattung aber auch ein grundlegendes Verständnis von Nachhaltigkeitsanforderungen entlang der Wertschöpfungskette aufzubauen. Letztendlich muss der Controller in der Lage sein, die Qualität der Nachhaltigkeitsdaten und -kennzahlen in derselben Qualität wie die der finanziellen Kennzahlen sicherzustellen, um die Anforderungen der Wirtschaftsprüfer zu bestehen (Vanini et al., 2023, S. 104 ff.). Die Nachhaltigkeitstransformation erfordert darüber hinaus eine intensive Kommunikation und Abstimmung des Controllings mit anderen Funktionsbereichen im Unternehmen, so dass auch eine spezifische Kommunikationsfähigkeit des Finanzbereichs im Allgemeinen und des Controllings im Speziellen erforderlich ist (AK Finanzierung der Schmalenbach-Gesellschaft, 2022, S. 13 ff.).

Aus den um Nachhaltigkeitsaspekte erweiterten Aufgaben- und damit Kompetenzprofilen von Controllern ergibt sich die Notwendigkeit, die Controllinglehre an Hochschulen um fachliche aber auch methodische und kommunikative Nachhaltigkeitsthemen und -ansätze zu erweitern. Insbesondere grundlegende Konzepte und Instrumente des Nachhaltigkeitscontrollings wie z.B. die Wesentlichkeitsanalyse oder die Messung von ESG-Kennzahlen und regulatorische Anforderungen an die Nachhaltigkeitsberichterstattung sollten systematisch in die Controllinglehre integriert werden (Vanini et al., 2023). In diesem Zusammenhang wird von einer Sustainability Literacy (Hahn & Reimsbach, 2014, S. 55 ff.) der Controller gesprochen.

Internationale Studien zeigen allerdings eine eher geringe Abdeckung von Konzepten und Instrumenten eines Nachhaltigkeitscontrollings in betriebswirtschaftlichen Studiengängen. So führen Mburayi und Wall (2018) eine Literaturübersicht über Studien aus den Jahren 2010 bis 2017 durch, die sich mit der Integration von Nachhaltigkeit in Business- und Management-Studiengänge mit einem besonderen mit besonderem Fokus auf Rechnungswesen- und Controlling-Curricula befassen. Ihre Ergebnisse zeigen, dass das Rechnungswesen bei der Integration von Nachhaltigkeit hinter anderen Managementdisziplinen zurückbleibt. Andere Studien unterstützen das Ergebnis, dass Nachhaltigkeitsaspekte, wenn überhaupt, nur rudimentär in Rechnungswesenmodulen in betriebswirtschaftlichen Studiengängen integriert sind (Botes et al., 2014; Boulianne & Keddie, 2018; Boulianne et al., 2018; Cho et al., 2020; Gomes et al., 2021; Marx & van der Watt, 2013). Wenn Nachhaltigkeitsaspekte in die Rechnungswesen- bzw.

Controllinglehre enthalten sind, zeigt die Forschung kein klares Bild, auf welchem Studienniveau (Bachelor, Master, Postgraduate) bzw. mit welchem Verpflichtungsgrad (Pflicht- oder Wahlmodul) dies erfolgt. So ermitteln einige Studien, dass an Universitäten, an denen prominente Wissenschaftler auf dem Gebiet des Nachhaltigkeitscontrollings tätig sind, Nachhaltigkeitsthemen oft nicht auf Bachelor- sondern eher auf Masterniveau gelehrt werden (Khan, 2013). Im Gegensatz dazu haben Cho et al. (2020) in ihrer Umfrage unter akademischen Lehrkräften und Forschern an nordamerikanischen Universitäten festgestellt, dass Nachhaltigkeitsthemen vor allem auf der Undergraduate-Ebene in einem eigenem Wahlmodul gelehrt werden und nicht Bestandteil grundlegender Pflichtmodule des Rechnungswesens bzw. Controllings sind (Cho et al., 2020).

Auf Modulebene zeigen Studien, dass die Integration von Nachhaltigkeitsaspekten bisher hauptsächlich von den Interessen einzelner Akademiker getrieben wurde (Matten & Moon, 2004; Stubbs & Schapper, 2011). Nachhaltigkeitsthemen werden vielfach eher oberflächlich und als Zusatz behandelt, und ihre Integration in die vorhandenen Modulthemen ist häufig nicht tief genug, um den Studierenden ein ganzheitliches Nachhaltigkeitsverständnis zu vermitteln. Darüber hinaus wählen und integrieren die Lehrenden Nachhaltigkeitsthemen meist nach ihren individuellen Präferenzen und auf Ad-hoc-Basis (Botes et al., 2014). Dies lässt den Schluss zu, dass zu mindestens international die notwendigen Kenntnisse und Fähigkeiten für ein erfolgreiches Nachhaltigkeitscontrolling nicht systematisch in die Controllinglehre integriert sind. So kommen Matten und Moon (2004) zu dem Schluss, dass die Nachhaltigkeitscontrolling-Ausbildung nicht umfassend institutionalisiert ist und sich eher in einem Pionierstadium befindet.

Der geringe Grad der Integration von Nachhaltigkeitsaspekten in die Rechnungswesen- und Controllinglehre wird durch verschiedene Barrieren erklärt, z.B. eine starke Betonung von Techniken und Instrumenten des Rechnungswesens für die Finanzberichterstattung sowie eine Überfrachtung des Lehrplans (Botes et al., 2014; Marx & van der Watt, 2013). Ein weiterer Grund könnte in der traditionell wertorientierten Perspektive des Management Accountings bzw. Controllings liegen. Daher ist die Messung der Auswirkungen von Managemententscheidungen auf den finanziellen Unternehmenswert (Wertrelevanz) das entscheidende Kriterium für die Integration von Themen, Methoden und Instrumenten in die Curricula von betriebswirtschaftlichen Studiengängen (Hahn & Reimsbach, 2014). Zudem wird eine fehlende Aufnahme von Themen, Methoden und Instrumenten des Nachhaltigkeitscontrollings in internationale Lehrbüchern bemängelt, was deren systematische Integration in die Controllinglehre erschwert (Vanini & Bochert, 2024).

Als Fazit der durchgeführten Literaturanalyse soll an dieser Stelle festgehalten werden, dass zwar einerseits eine stärkere Beteiligung von Controllern an der Nachhaltigkeitssteuerung und – berichterstattung gefordert wird, es aber unklar ist, inwieweit diese über die dafür notwendigen Fach-, Methoden- und Sozialkompetenzen verfügen. Auch wenn erste internationale Studien auf eine eher geringe Integration von Nachhaltigkeitsaspekten in die Rechnungswesen- und Controllinglehre schließen lassen, liegt eine umfassende Untersuchung für den deutschen Sprachraum bislang nicht vor. Dies ist insofern problematisch, da Controller ihre grundlegende Ausbildung mehrheitlich im Rahmen ihres Studiums erwerben und daher aus den Ergebnissen wesentliche Implikationen in Bezug auf die Zukunftsfähigkeit und die Weiterentwicklung der Controlling-Curricula sowie -Module betriebswirtschaftlicher Studiengänge abgeleitet werden können. Zudem basieren bisherige Untersuchungen vor allem auf Interviews, großzahligen Befragungen und der Analyse von Lehrbüchern (Botes et al., 2014; Boulianne et al., 2018; Cho et al., 2020; Vanini & Bochert, 2024). Eine systematische Analyse von Modulbeschreibungen, die einen detaillierteren Einblick in die Konzeption und Inhalte von Lehrveranstaltungen geben, erfolgte bisher nicht.

3 Forschungsmethode und Stichprobe

Zur Analyse des Umsetzungsstandes von Anforderungen an sowie Konzepten und Instrumenten des Nachhaltigkeitscontrollings in die Controllinglehre von betriebswirtschaftlichen Studiengängen an Hochschulen in der DACH-Region wurden geeignete Modulbeschreibungen erhoben und ausgewertet. Während andere Studien Befragungen

oder Interviews durchführen oder Lehrbücher analysieren (Vanini & Bochert, 2024), verwenden wir Modulbeschreibungen, da diese zum einen standardisiert und daher für eine quantitative Inhaltsanalyse gut geeignet sind. Zum anderen geben sie nicht nur detailliert über Lerninhalte sondern auch über weitere Modulmerkmale wie z.B. Lehr- und Prüfungsmethoden oder vergebene Leistungspunkte Auskunft, die ebenfalls Rückschlüsse über die quantitative und qualitative Abdeckung von Nachhaltigkeitscontrolling in den analysierten Modulen zulassen. Daher werden Modulbeschreibungen auch von anderen Studien als Datengrundlage genutzt (Olalla & Merino, 2019).

Für die Auswahl der Stichprobe wurden zunächst grundlegende Kriterien in Bezug auf die zu analysierenden Studiengänge festgelegt. In Anlehnung an (Gomes et al., 2021) wurden betriebswirtschaftliche Studiengänge, Management-Studiengänge sowie Studiengänge des Bereichs Finance, Accounting, Controlling und Taxation (FACT) ausgewählt. Zusätzlich wurden wirtschaftswissenschaftliche Studiengänge mit betriebswirtschaftlichem Schwerpunkt einbezogen. Als Hilfsmittel für das Zusammenstellen der Hochschul- beziehungsweise Studiengangliste wurden drei verschiedene Internetportale genutzt, die über eine Suchfunktion für Hochschulen und Studiengänge verfügen und als Orientierungshilfe für Studieninteressierte dienen. Für Deutschland wurde die Internetseite Hochschulkompass.de verwendet, die von der Hochschulrektorenkonferenz organisiert wird und Informationen über staatliche und staatlich anerkannte deutsche Hochschulen bereitstellt. Für die Schweiz wurde auf die Internetseite swissuniversities.ch zugegriffen, die von der Rektorenkonferenz der schweizerischen Hochschulen organisiert wird. Für die Studiengänge aus Österreich steht leider keine offizielle Seite der Rektorenkonferenz als Hilfsmittel zur Verfügung. Die Seite studieren.at der TarGroup Media GmbH bietet jedoch auch einen Überblick über in Österreich vorhandenen Hochschulen und wurde dementsprechend verwendet.

Zunächst war geplant, dass über die jeweiligen Suchfunktionen der Portale unter Verwendung festgelegter Suchkriterien die entsprechenden Studiengänge herausgefiltert werden sollten. Da die Seiten allerdings über zum Teil stark voneinander abweichende Suchkriterien auf der Ebene der Studiengänge verfügten, bestand die Gefahr, dass die finale Stichprobe in sich nicht valide wäre. Es wurde daher entschieden, die Suche auf Hochschulebene durchzuführen. Hierfür wurden für jedes der drei Länder alle Hochschulen mit wirtschaftswissenschaftlichen Studiengängen herausgefiltert. Anhand dieser Liste wurden die Internetseiten der Hochschulen im Hinblick auf ihr Studienangebot untersucht. Alle Studiengänge, die den vorher festgelegten Kategorien - betriebswirtschaftlich, wirtschaftswissenschaftlich mit Schwerpunkt BWL, Management sowie FACT-Studiengänge - zugeordnet werden konnten, wurden in einer ersten Studiengangliste festgehalten. Diese Liste umfasste 587 verschiedene Studiengänge aus Deutschland, der Schweiz und Österreich. Im nächsten Schritt wurde geprüft, ob für diese Studiengänge die zugehörigen Modulhandbücher als PDF-Dateien verfügbar waren. Nach einer umfangreichen Recherche konnten für insgesamt 330 der 587 Studiengänge vor Beginn des Sommersemester 2023 die Modulhandbücher heruntergeladen werden. Für die anderen Studiengänge waren die Modulhandbücher nicht frei sondern nur mit einem Hochschulzugang zugänglich oder sie lagen nur in Form von Texten auf den Internetseiten der Hochschulen vor, ohne dass die Möglichkeit bestand, sie in Dateiformat offline verfügbar zu haben.

Ein besonderes Problem stellte der unterschiedliche Umfang der Modulbeschreibungen dar. Speziell die Modulhandbücher der schweizerischen Studiengänge stellten sich als sehr knapp dar. Da die Stichprobe zu gering ausgefallen wäre, wenn lediglich solche Modulhandbücher ausgewählt worden wären, die im Hinblick auf die ihre Informationen vollständig sind, wurde ein Richtwert von 80 Prozent der Informationen festgelegt. Hierbei wurde für die Auswahl der Studiengänge und ihrer Modulhandbücher nur geprüft, ob die Informationen grundsätzlich vorhanden waren, nicht jedoch in welchem Umfang. Schließlich enthielt die der Analyse zugrundeliegende Studiengangliste 175 Bachelor- und Masterstudiengänge, die sich auf insgesamt 96 Hochschulen verteilten. Insgesamt wurden 459 Modulbeschreibungen ausgewertet.

Für die Erhebung und Auswertung der Informationen aus den Modulbeschreibungen wurde ein Mixed Methods-Ansatz aus qualitativer und quantitativer Inhaltsanalyse verwendet (Rädiker & Kuckartz, 2019, S. 181 ff.). Vor der Inhaltsanalyse mussten ein Codesystem sowie ein Kodierleitfaden entwickelt und anschließend die einzelnen Modulbeschreibungen codiert werden (Rädiker & Kuckartz, 2019, S. 67 ff.). Das Ergebnis ist der folgenden Tabelle zu entnehmen.

Tabelle 1: Codesystem und deskriptive Ergebnisse der Inhaltsanalyse

Code	Subcode 1. Stufe	Subcode 2. Stufe	Anzahl Codes
1. Hochschul- typ	1.1 Universität		30
	1.2. Hochschule für angewandte Wissenschaft, Fachhochschule		66
2. Studiengang	2.1. Bezeichnung	2.1.1 BWL 2.1.2 WiWi, 2.1.3 FACT, 2.1. 4 Man	111 / 14 / 30 / 20
	2.2. Abschluss	2.2.1 Bachelor 2.2.2 Master	102 73
3. Modul	3.1. Modulbezeichnung	3.1.1 Controllingmodul 3.1.2 Nachhaltigkeitsmodul	403 56
	3.2. Modulart	3.2.1 Pflichtmodul 3.2.2 Wahlmodul	207 252
	3.3 Lehrform	3.3.1 Vorlesung	232
		3.3.2 Übung	189
		3.3.3 Seminar	195
		3.3.4 andere	41
	3.4 Nachhaltigkeitsspezifische Lerninhalte	3.4.1 Allgem. Nachhaltigkeitscon.	44
		3.4.2 Anforderungen	14
		3.4.3 Normative und strategische Nachhaltigkeitssteuerung	18
		3.4.4 Operatives Nachhaltigkeits- controlling	17
		3.4.5 Nachhaltiges Investitions- controlling	7
		3.4.6. Nachhaltiges Risikocon.	2
		3.4.7. Nachhaltigkeitsreporting	31
		3.4.8 Sustainable Finance	10
		3.4.9 Nachhaltigkeitsmanagement	24
		2.4.10 sonstiges	27
	3.5 Literatur	3.5.1 Nachhaltigkeits-spezifisch	44
		3.5.2 Rechnungswesen-spezifisch	271
		3.5.3 andere	50
	3.6 Art der Prüfung	3.6.1 Klausur	306
		3.6.2 Präsentation	84
		3.6.3 Hausarbeit	62
		3.6.4 Portfolio	37
		3.6.5 Mündliche Prüfung	33
		3.6.6 Sonstige	83
	3.7. Umfang der Leistungspunkte	3.7.1. bis 5 LP	241
		3.7.2. 6-10 LP	190
		3.7.3. 11-15 LP	10
		3.7.4. > 15 LP	4

Zunächst wurden wesentliche Merkmale der Hochschule (Universität/ Hochschule für angewandte Wissenschaften oder Fachhochschule) sowie des Studiengangs (Bezeichnung sowie Abschluss) codiert, da angenommen wurde, da diese möglicherweise einen Einfluss auf den Umfang der Integration von Nachhaltigkeitsthemen in die Controllinglehre aufweisen. Der Tabelle lässt sich entnehmen, dass die Stichprobe hauptsächlich aus allgemeinen betriebswirtschaftlichen Studiengängen (63%) gefolgt von FACT-Studiengängen (17%) besteht, 58% sind Bachelor- und 42% Masterstudiengänge.

Die folgenden Codes bilden wesentliche Merkmale der einzelnen analysierten Module ab. Da sich die Forschungsfrage auf die Umsetzung von Nachhaltigkeitsaspekten in der Controllinglehre bezieht, wurde unter dem Merkmal

„Modulbezeichnung“ nach Controllingmodulen und Modulen des internen Rechnungswesens in den Studiengängen gesucht und diese mit dem Begriff „Controllingmodul“ codiert. Da zudem Aspekte des Nachhaltigkeitscontrollings auch in spezifischen Nachhaltigkeitsmodulen behandelt werden, wurde ebenfalls nach Nachhaltigkeitsmodulen gefiltert und diese als „Nachhaltigkeitsmodul“ codiert. Eine derartige Zuordnung kann darüber Aufschluss geben, ob ein Modul einen Nachhaltigkeitsschwerpunkt aufweist oder Nachhaltigkeit als eines von mehreren Themen in ein Modul integriert ist (Mburayi & Wall, 2018). Die nächste Moduleigenschaft bezieht sich auf die Art des Moduls. Pflichtmodule müssen im Laufe des Studiums belegt und erfolgreich abgeschlossen werden, während Wahlmodule nach Neigung des Studierenden belegt werden. Daher ist es möglich, dass in einem Studiengang Module zu Nachhaltigkeitsaspekten angeboten werden. Handelt es sich dabei jedoch ausschließlich um Wahlmodule, ist Nachhaltigkeit kein verpflichtender Bestandteil des Curriculums. Anschließend wird die Lehrform des Moduls untersucht. Hierbei muss erwähnt werden, dass einem Modul zum Teil mehrere Lehrformen zugeordnet sind, beispielsweise wenn eine Vorlesung durch eine Übung ergänzt wird. Neben den beiden genannten Lehrformen kommt zudem das Seminar (auch als seminaristischer Unterricht bezeichnet) in Frage. Zum Subcode „Weitere Lehrformen“ zählen z.B. Softwareschulungen oder Einzelcoachings. Diese Unterteilung ist wichtig, da sich insbesondere die interdisziplinären und kommunikativen Nachhaltigkeitskompetenzen nicht unbedingt im Rahmen von Vorlesungen vermitteln lassen.

Zentraler Analyseaspekt ist Modulinhalt. Die nachhaltigkeitspezifischen Lerninhalte wurden in Anlehnung an Sailer (2022) weiter untergliedert. Während der Inhaltsanalyse wurden die Subcodes zudem um die Kategorien „Nachhaltigkeitsmanagement“ sowie „Andere nachhaltigkeitspezifische Inhalte“ ergänzt, um für entsprechende Inhalte, wie z.B. nachhaltige Unternehmensführung oder aber grundlegendere Themen wie Nachhaltigkeitskonzepte passende Kategorien zu haben. Insgesamt unterteilt sich die Kategorie „nachhaltigkeitspezifische Inhalte“ in zehn Subcodes. Der Subcode „Allgemeines Nachhaltigkeitscontrolling“ ist für Inhaltsangaben gedacht, in denen das Nachhaltigkeitscontrolling als einzelner Begriff verwendet wird und keine weiteren Informationen über die in diesem Themengebiet behandelten Inhalte gegeben werden, so dass eine genauere Zuteilung nicht möglich ist. Die zweite Unterkategorie umfasst die Nachhaltigkeitsanforderungen an Unternehmen. Hierzu zählen sowohl regulatorische Anforderungen wie auch Anforderungen aus Nachhaltigkeitsstandards. Für Analysezwecke wurde für jede der Unterkategorien eine Liste mit Schlüsselbegriffen zusammengestellt. Im Zusammenhang mit den Unternehmensanforderungen beinhaltet die Liste die Begriffe „Nachhaltigkeitsstandard“, „EU-Taxonomie“, „Global Reporting Initiative“ beziehungsweise „GRI“ und „Nachhaltigkeitskodex“. Als dritte Unterkategorie folgt die normative und strategische Nachhaltigkeitssteuerung. Die Schlüsselbegriffe hierzu sind „strategisches Nachhaltigkeitscontrolling“, „Wesentlichkeitsanalyse“ und „Nachhaltigkeitsstrategie“. Die in den Modulhandbüchern genannten Inhalte wurden grundsätzlich auch dann einer der Unterkategorien zugeordnet, wenn in der Inhaltsangabe eine zu den Schlüsselbegriffen passende Wortkombination zu finden war, oder Synonyme verwendet wurden. Das operative Nachhaltigkeitscontrolling macht die vierte Unterkategorie aus. Es umfasst die Begriffe „Sustainability Balanced Scorecard“, „Nachhaltigkeitskennzahl“ beziehungsweise „Nachhaltigkeits KPI“, „Umweltkostenrechnung“, „Nachhaltigkeitssoftware“ sowie „operatives Nachhaltigkeitscontrolling“.

Die fünfte Unterkategorie beschreibt das nachhaltige Investitionscontrolling. Für diese Unterkategorie kamen deutlich weniger klar zuordbare Begriffe infrage, weshalb sich die Schlüsselbegriffe auf „nachhaltiges Investieren“ oder „nachhaltige Investition“ und auf den Begriff „nachhaltigkeitsorientierte Nutzwertanalyse“ beschränken. Selbiges gilt für die sechste Unterkategorie, das nachhaltige Risikocontrolling. Auch hier standen nur wenige Begrifflichkeiten zur Auswahl, so dass sich die Liste der Schlüsselbegriffe auf „nachhaltige Risiken“ und „Nachhaltigkeitsrisiko“ beschränkt. Der siebte Unterpunkt im Rahmen der nachhaltigkeitspezifischen Inhalte ist das Nachhaltigkeitsreporting. Hierbei kamen sowohl die Begriffe „Nachhaltigkeitsreporting“ wie auch „Nachhaltigkeitsberichterstattung“ und „Nachhaltigkeitskommunikation“ zum Tragen. Außerdem gehören „Nachhaltigkeitsbericht“, „GRI“ beziehungsweise „Global Reporting Initiative“ und „Nachhaltigkeitskodex“ zur Liste der Schlüsselbegriffe. In diesen Punkten überschneiden sich die beiden Unterkategorien Unternehmensanforderungen sowie Nachhaltigkeitsreporting, da es speziell im Rahmen der Berichterstattung zahlreiche Anforderungen an Unternehmen gibt. Die achte Unterkategorie gilt dem Bereich Sustainable Finance. Zu den Schlüsselbegriffen in diesem Bereich gehören sowohl „Sustainable Finance“ wie auch „Green Finance“ und der Begriff „ESG“. Der Bereich Sustainable Finance bildet

den Abschluss des Teils der nachhaltigkeitspezifischen Inhalte, die an die Gliederung von Sailers Nachhaltigkeitscontrolling angelehnt sind. Von den zusätzlichen Unterkategorien wurde die Kategorie der anderen nachhaltigkeitspezifischen Inhalte bereits näher erläutert. Die Unterkategorie Nachhaltigkeitsmanagement bezieht sich auf Inhalte, die in den Bereich des Managements fallen, jedoch nicht in den Bereich Controlling. Dazu gehören zum Beispiel nachhaltiges Personalmanagement oder nachhaltiges Produktionsmanagement.

Zudem wurde die in den Modulbeschreibungen empfohlene Literatur codiert. Hier erfolgt eine Unterscheidung nach nachhaltigkeitspezifischer Literatur, Literatur zum Rechnungswesen sowie Controlling und anderer Literatur. Die Klassifizierung der einzelnen Literaturangaben erfolgt anhand des in der Modulbeschreibung genannten Titels. Abschließend wurden die Kategorien „Art der Prüfung“ sowie „Umfang der Leistungspunkte codiert. Analog zur Lehrform lassen sich bestimmte Kompetenzen wie z.B. kritisches Denken oder kommunikative Fähigkeiten nicht durch alle Prüfungsformen prüfen, so eignet sich beispielsweise eine Klausur vor allem zur Überprüfung von erworbenen Fachkompetenzen. Die letzte Kategorie sollte vor allem einen Rückschluss über die Relevanz der nachhaltigkeitspezifischen Inhalte im Kontext des gesamten Studiums ermöglichen.

Das folgende Kapitel gibt einen Überblick über die zentralen deskriptiven und qualitativen Ergebnisse der Inhaltsanalyse sowie der durchgeführten Zusammenhangsanalyse.

4 Ergebnisse

4.1 Deskriptive Ergebnisse

Insgesamt wurden 459 Modulbeschreibungen von 96 Hochschulen (30 Universitäten und 66 Fachhochschulen bzw. Hochschulen für angewandte Wissenschaften) und aus 175 Studiengängen analysiert. Dabei handelt es sich 280 Bachelor- und 179 Mastermodule sowie 207 Pflicht- und 252 Wahlmodule. Neben 403 Controllingmodulen wurden auch 56 Nachhaltigkeitsmodule untersucht. Dominante Lehrform in der Controllinglehre ist nach wie vor die Vorlesung (232) gefolgt vom Seminar (195) und der Übung (189), dementsprechend ist auch die Klausur die mit Abstand häufigste Prüfungsform (306) gefolgt von der Präsentation (84) und sonstigen Prüfungsformen (83). Die meisten Module haben einen quantitativen Umfang von 5 LP (241) bzw. 6-10 LP (190).

In 84 der untersuchten Controlling- und Nachhaltigkeitsmodule werden Nachhaltigkeitscontrolling-spezifische Lerninhalte behandelt, darunter sind 31 Controlling- und 53 Nachhaltigkeitsmodule. Das bedeutet aber auch im Umkehrschluss, dass 375 (82%) der Module das Thema Nachhaltigkeitscontrolling nicht aufgreifen. Insbesondere in 372 Controllingmodulen wurden keine Nachhaltigkeitsthemen so intensiv behandelt, dass dies in der Modulbeschreibung vermerkt war. Von den 53 Nachhaltigkeitsmodulen wiesen nur 17 einen eindeutigen Controllingbezug im Titel auf und befassten sich schwerpunktartig mit Themen des Nachhaltigkeitscontrollings und -reportings. Bei den übrigen 36 Modulen handelt es sich eher um allgemeine Nachhaltigkeitsmodule, CSR Controlling- oder Sustainable Finance-Module. Interessanterweise verwenden nicht einmal alle Module, in den nachhaltigkeitspezifische Inhalte vermittelt werden, auch entsprechende nachhaltigkeitspezifische Literaturhinweise (44). Vor allen in den Controllingmodulen dominieren erwartungsgemäß Rechnungswesensspezifische Literatur, die nach unserer Systematik auch Kostenrechnungs- und Controllinglehrbücher einschließt. Überraschenderweise sind nicht in allen Modulbeschreibungen entsprechende Literaturhinweise hinterlegt.

4.2 Nachhaltigkeitsspezifische Lerninhalte

In 84 der untersuchten Module werden nachhaltigkeitspezifische Lerninhalte behandelt, darunter sind 31 Controlling- und 53 Nachhaltigkeitsmodule. Die Verteilung der einzelnen Lerninhalte lässt sich der folgenden Tabelle entnehmen.

Tabelle 2: Nachhaltigkeitsspezifische Lehrinhalte in Controlling- und Nachhaltigkeitsmodulen

Nachhaltigkeitsspezifische Lerninhalte	Controllingmodul	Nachhaltigkeitsmodul	Summe
1. Allgemeines Nachhaltigkeitscontrolling	24 (6%)	20 (36%)	44
2. Anforderungen	1 (0%)	13 (23%)	14
3. Normative und strategische Nachhaltigkeitssteuerung	3 (1%)	15 (27%)	18
4. Operatives Nachhaltigkeitscontrolling	3 (1%)	14 (25%)	17
5. Nachhaltiges Investitionscontrolling	0 (0%)	7 (13%)	7
6. Nachhaltiges Risikocontrolling	0 (0%)	2 (4%)	2
7. Nachhaltigkeitsreporting	5 (1%)	26 (46%)	31
8. Sustainable Finance	0 (0%)	9 (16%)	9
9. Nachhaltigkeitsmanagement	3 (1%)	21 (38%)	24
10. Andere nachhaltigkeitspezifische Inhalte	2 (0%)	25 (45%)	27
Summe	31	53	

Im Durchschnitt werden 2,3 Nachhaltigkeitsthemen pro Modul behandelt, wobei die Zahl der Themen zwischen 1 und 6 Themen schwankt. Hauptsächlich werden grundlegende Begriffe und Aspekte des Nachhaltigkeitscontrollings und des Nachhaltigkeitsreportings sowie andere nachhaltigkeitspezifische Inhalte behandelt. Auffällig ist der relative geringe Abdeckungsgrad der (regulatorischen) Anforderungen sowie der konkreten Instrumente des strategischen und operativen Nachhaltigkeitscontrollings einschließlich des Investitions- sowie des Risikocontrollings in den Modulbeschreibungen. Zudem sind Aspekte des Nachhaltigkeitscontrollings sehr wenig in Controllingmodule integriert, sondern werden eher in spezifischen Nachhaltigkeitsmodulen behandelt, wo sie auch mit anderen Themen wie dem Nachhaltigkeitsmanagement oder anderen nachhaltigkeitsbezogenen Inhalten vermischt werden.

Zudem werden nachhaltigkeitspezifische Lehrinhalte etwas häufiger in spezifischen Wahlmodulen (61%) und in Seminarform (51%) gelehrt als allgemeine Controlling-Module. 43% der Module mit nachhaltigkeitspezifischen Lehrinhalten werden aber auch als Vorlesung, teilweise in Kombination mit einer Übung abgehalten. Daher überrascht es auch wenig, dass die Klausur mit 50% die dominierende Prüfungsform ist, gefolgt von der Präsentation (25%) und der Portfolioprüfung (17%). Überraschenderweise spielen Hausarbeiten mit 8% nur eine untergeordnete Rolle als Prüfungsform.

Die geringe Bedeutung von nachhaltigkeitspezifischen Lehrinhalten spiegelt sich auch in der relativ geringen Zahl der Leistungspunkte der dazugehörigen Module wider:

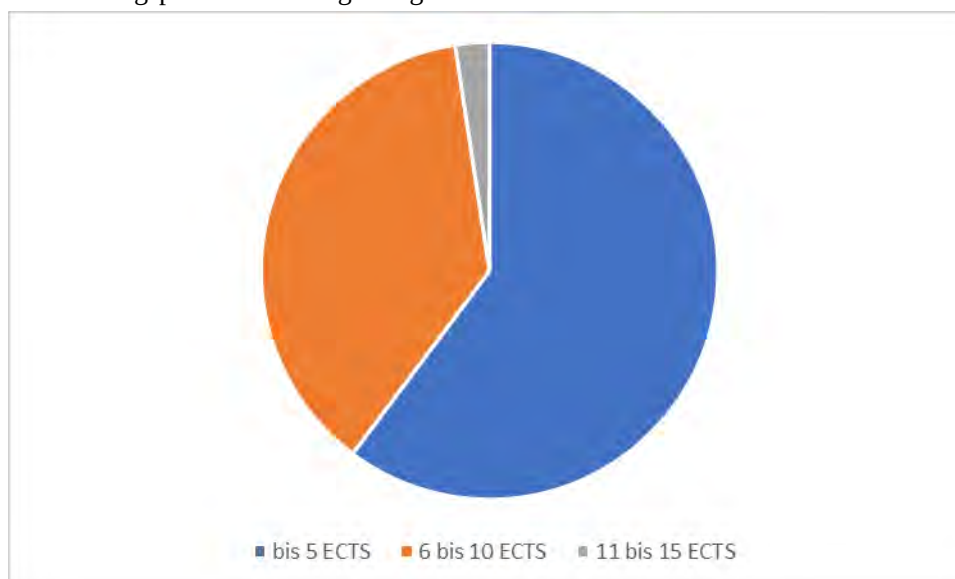


Abbildung 1: Verteilung der Leistungspunkte in den Modulen mit nachhaltigkeitspezifischen Lerninhalten

In 60% der analysierten Module mit nachhaltigkeitspezifischen Inhalten werden max. 5 LP vergeben, in 37% 6 bis 10 LP. Gemessen an den vergebenen LP haben die Module mit nur controllingspezifischen Inhalten mit 53 % bzw. 41 % eine etwas höhere quantitative Wertigkeit.

Nur 17 der Nachhaltigkeitsmodule haben die Begriffe „Nachhaltigkeitscontrolling“ oder „Nachhaltigkeitsreporting“ im Modultitel und behandelt die Themen somit ausführlich. Die anderen Nachhaltigkeitsmodule beschäftigen sich eher umfassend mit den Themen Nachhaltigkeitsmanagement und Sustainable Finance. Hier ist der Themenbereich Nachhaltigkeitscontrolling bzw. -reporting – ähnlich wie in den Controlling-Modulen, ein Thema unter vielen anderen und wird dementsprechend eher oberflächlich behandelt. Teilweise gibt es doch nicht einmal einen separaten Gliederungspunkt in den Modulbeschreibungen.

Betrachtet man die 17 Module, die sich schon im Titel auf den Themenbereich Nachhaltigkeitscontrolling oder -reporting beziehen, geben diese eine Einführung der grundsätzlichen Begrifflichkeiten, der Ziele und Aufgaben des Nachhaltigkeitscontrollings sowie einen Überblick über ausgewählte Instrumente, vereinzelt werden auch regulatorische Anforderungen, grundlegende Theorien zur Integration der Nachhaltigkeit in die Unternehmenssteuerung, Inhalte einer nachhaltigkeitsorientierten Kostenrechnung oder Nachhaltigkeitskennzahlen behandelt. Bei den regulatorischen Standards werden neben der CSRD und den ESRS auch der deutsche Nachhaltigkeitskodex (DNK), die Global Reporting Initiative (GRI), die UN Sustainable Development Goals (SDGs), das Lieferkettensorgfaltspflichtengesetz sowie Standards der Klimaberichterstattung wie z.B. der Task Force of Climate-related Financial Disclosure (TCFD) aufgegriffen (für einen Überblick über die verschiedenen Standards vgl. (Velte & Stawinoga, 2019)). Bei den Controllinginstrumenten wird – sofern diese behandelt werden - auf die Balanced Scorecard (BSC), die Kostenrechnung aber auch die Budgetierung und Investitionsrechnung verwiesen. Vereinzelt erfolgt auch eine Abgrenzung von den verwandten Begriffe CSR-Controlling und Green Controlling. Strategische Instrumente wie z.B. Wesentlichkeitsanalysen werden dagegen in den Modulbeschreibungen gar nicht erwähnt. Ein einheitlicher inhaltlicher Standard von Themen des Nachhaltigkeitscontrolling im Sinne einer gemeinsamen Sustainability Literacy von Controllern ist in den Modulinhalten nicht erkennbar.

4.3 Einflussfaktoren auf die Integration nachhaltigkeitspezifischer Lerninhalte

Abschließend wurden verschiedene Einflussfaktoren auf die Aufnahme von nachhaltigkeitspezifischen Lerninhalten in ein Modul bzw. eine Modulbeschreibung untersucht.

Tabelle 3: Einflussfaktoren auf die Aufnahme von nachhaltigkeitspezifischen Lerninhalten in die Module bzw. Modulbeschreibungen

Einflussfaktor auf Nachhaltigkeits-spezifische Lerninhalte	Statistische Maße	Statistisch signifikanter Einfluss vorhanden?
Hochschultyp (Universität vs. HAW bzw. FH)	Pearson Chi Quadrat = 1,289, df 1, p (asyp., zweiseitig) = 0,2462	Nein
Studiengangsbezeichnung (FACT, Management, BWL, WiWi)	Pearson Chi Quadrat = 2,515, df = 3, p (asyp., zweiseitig) = 0,4725	Nein
Studienabschluss (Bachelor, Master)	Pearson Chi Quadrat = 1,102, df = 1, p (asyp., zweiseitig) = 0,2938	Nein
Modulbezeichnung (Controllingmodul, Nachhaltigkeitsmodul)	Pearson Chi Quadrat = 248,663, df = 1, p (asyp., zweiseitig) = 0,000	Ja
Modulart (Pflichtmodul, Wahlmodul)	Pearson Chi Quadrat = 1,403, df = 1, p (asyp., zweiseitig) = 0,2362	Nein

Der Tabelle lässt sich entnehmen, dass lediglich der Einflussfaktor Modulbezeichnung einen statistisch signifikanten Einfluss auf die Integration von nachhaltigkeitspezifischen Inhalten in die Modulbeschreibung hat. Wenig

überraschend werden Aspekte des Nachhaltigkeitscontrollings signifikant öfter in Nachhaltigkeits- als in reine Controllingmodule integriert, wogegen sowohl der Hochschultyp als auch der Studiengang bzw. -abschluss sowie die Modulart keinen Einfluss haben. Dies ist überraschend, da Wahlmodule eine flexible Integration neuer Themen in einen Studiengang zu lassen, während eine grundsätzliche Veränderung von Pflichtmodulen eines höheren formalen Aufwands bis hin zu einer Neuakkreditierung eines Studiengangs erfordern kann. Auch wäre zu vermuten gewesen, dass sich nachhaltigkeitspezifische Lerninhalte eher in Masterstudiengängen oder in speziellen FACT-Studiengängen finden lassen, da diese zu einem höheren und spezialisiertem Kompetenzerwerb von Studierenden beitragen sollen.

5 Diskussion, Limitationen und Implikationen

Die wesentlichen Ergebnisse der durchgeführten quantitativen und qualitativen Inhaltsanalyse der Modulbeschreibungen lassen sich wie folgt zusammenfassen und in den Stand der Forschung einordnen:

- Gemessen an der Anzahl der Module mit nachhaltigkeitspezifischen Inhalten sowie den zugeordneten Leistungspunkten lässt sich ein insgesamt geringer quantitativer Abdeckungsgrad von Aspekten des Nachhaltigkeitscontrollings in der Controllinglehre in betriebswirtschaftlichen Studiengängen der DACH-Region feststellen. Wenn Nachhaltigkeitsthemen adressiert werden, erfolgt dies eher in spezifischen Nachhaltigkeits- als in Controllingmodulen. Angesehen von einigen Spezialmodulen zum Nachhaltigkeitscontrolling und -reporting erfolgt jedoch vielfach eine eher oberflächliche Behandlung von grundlegenden Aspekten und Begriffen des Nachhaltigkeitscontrollings und -reportings, ohne konkret auf Konzepte, Methoden und Instrumente einzugehen oder diese zu vertiefen. Eine umfassende Integration von relevanten Fach- und Methodenkompetenzen z.B. zur Durchführung von Wesentlichkeitsanalysen, der Umsetzung von regulatorischen Anforderungen in der Nachhaltigkeitsberichterstattung oder der Ermittlung von Nachhaltigkeits-KPIs kann auch für Masterstudiengänge nicht festgestellt werden. Die Ergebnisse unterstützen somit die Befunde internationaler Studien, die ebenfalls einen geringen Integrationsgrad von Aspekten des Nachhaltigkeitscontrollings in der Rechnungswesenlehre feststellen (Botes et al., 2014; Boulianne & Keddie, 2018; Boulianne et al., 2018; Cho et al., 2020).
- Der geringe Integrationsgrad ist dabei unabhängig von Hochschultyp, Studiengang, Abschluss und Modulart.
- Eine Detailanalyse der einzelnen Modulbeschreibungen mit nachhaltigkeitspezifischen Lerninhalten ließ keine inhaltliche Standardisierung erkennen. Dies lässt die Vermutung zu, dass jeder Lehrende nach seinen spezifischen Präferenzen und Kompetenzen ausgewählte Themen und Instrumente des Nachhaltigkeitscontrollings in seine Lehre integriert. Das Ergebnis unterstützt somit die Schlussfolgerung internationaler Studien, dass wir uns bei der Integration von Nachhaltigkeitsaspekten in die Rechnungswesen- bzw. Controllinglehre eher in einem Pionierstadium befinden und es keinen akademischen Konsens über notwendige Inhalte und Methoden im Sinne einer Sustainability Literacy gibt (Barber et al., 2014; Benn & Dunphy, 2009).
- Auffällig ist auch die geringe Verwendung spezifischer Literatur zum Nachhaltigkeitscontrolling, obwohl es einschlägige Lehrbücher gibt (Sailer, 2022; Sopp et al., 2023). Die geringe Verbreitung von Themen des Nachhaltigkeitscontrollings in grundlegenden Controlling-Lehrbüchern (Vanini & Bochart, 2024) erschwert zudem für motivierte Lehrende die effiziente Integration ausgewählter Themen in ihre grundlegenden Controllingmodule.
- Die Module mit nachhaltigkeitspezifischen Lerninhalten greifen bei der Vermittlung der Nachhaltigkeitskompetenzen überwiegend auf die klassischen Lehr- und Prüfungsmethoden in Form von Vorlesungen und Klausuren zurück. Hier stellt sich die Frage, ob diese Lehr- und Prüfungsformen wirklich geeignet sind, um ein interdisziplinäres und kritisches Nachhaltigkeitsverständnis der angehenden Controller zu entwickeln.

Aus den Ergebnissen lassen sich darüber hinaus Implikationen für die Controllinglehre und -forschung ableiten:

- Grundsätzlich sollte ein gemeinsames Verständnis einer Sustainability Literacy von Controllern (Berg & Lyngstadås, 2023, S. 2; Hahn & Reimsbach, 2014) in der Wissenschaft sowie der Unternehmenspraxis entwickelt werden. Beispielsweise könnten basierend auf den Anforderungen der CSRD und der EFRS ausgewählte Aspekte und Instrumente des Nachhaltigkeitscontrollings in die Controllinglehre integriert werden, wie z.B. Wesentlichkeitsanalysen oder die Messung der Impact Materialität. Hier sollten auch die Ergebnisse bisher vorliegender Studien zu notwendigen Nachhaltigkeitskompetenzen von Controllern z.B. aus der Befragung von Nachhaltigkeitsmanagern, Controllern oder Wirtschaftsprüfern aus der Unternehmenspraxis integriert werden (Vanini et al., 2023).
- Darüber hinaus gibt es bereits Vorschläge aus der Literatur zur relevanten Aspekten des Nachhaltigkeitscontrollings z.B. die Abgrenzung von verwandten Konzepten wie dem CSR- oder dem Green Controlling, eine grundlegenden theoretischen Fundierung z.B. durch die Stakeholder-Theorie, die Ableitung von Nachhaltigkeitszielen und Messung durch prüfungssichere Kennzahlen sowie deren Integration in das Performance Measurement sowie Reporting und die Anreiz- und Vergütungssysteme (Al-Hazaima et al., 2021; Vanini & Bochert, 2024, S. 12).
- Bisherige Kompetenzrahmenwerke für Controller enthalten bislang kaum spezifische Fach-, Methoden- und Sozialkompetenzen zum Nachhaltigkeitscontrolling und -reporting oder gar ganze Rollenprofile für Nachhaltigkeitscontroller (AICPA & CIMA, 2023; Institute of Management Accountants [IMA], 2019; International Group of Controlling [IGC], 2015). Daher sollten aus den o.g. Anforderungen von Regulatoren, Investoren, Banken, Kunden, der Wirtschaftsprüfer, der Wissenschaft und sonstigen Stakeholdern an das Nachhaltigkeitscontrolling und -reporting ein entsprechendes Kompetenzprofil für unterschiedliche Professionalitätsstufen z.B. „Basis“, „Fortgeschritten“ und „Experten“ entwickelt werden. Insbesondere in den grundlegenden Professionalitätsstufen ist das Nachhaltigkeitscontrolling kein eigenes separates Aufgabengebiet. Stattdessen sollten nachhaltigkeitspezifische Kompetenzen systematisch in vorhandenen Kompetenzfelder der einzelnen Kompetenzrahmenwerke wie z.B. Strategy, Planning & Performance oder Reporting & Control (IMA, 2019, p. 4) integriert werden. Gerade kleine und mittlere Unternehmen werden sich i.d.R. keinen eigenen Nachhaltigkeitscontroller leisten. Stattdessen müssen die spezifischen Anforderungen von Banken, Kunden oder der Unternehmensführung durch die klassischen Finanzcontroller umgesetzt werden. Insbesondere das grundlegende Kompetenzniveau sollte Gegenstand der Controllinglehre an Hochschulen sein, bei einer weitergehenden Controlling-Spezialisierung könnten dann höhere Kompetenzniveaus, die eher dem Rollenbild eines Nachhaltigkeitscontrollers entsprechen, erreicht werden.
- Hilfreich für eine effektive und effiziente Integration von Nachhaltigkeitsaspekten in die Controllinglehre ist eine Ergänzung vorhandener Controllinglehrbücher z.B. auch durch geeignete Instrumente des Nachhaltigkeitscontrollings oder anwendungsorientierte Fallstudien (Botes et al., 2014). Dies würde auch die Umsetzung eines Nachhaltigkeitscontrollings in der Controllinglehre weg von einem Pionier- hin zu einem mehr institutionalisiertem Stadium maßgeblich unterstützen (Barber et al., 2014; Benn & Dunphy, 2009).

Der vorliegende Untersuchungsansatz unterliegt verschiedenen Limitationen, die im Folgenden kurz vorgestellt werden:

- Das Ziel der Untersuchung war die Analyse von Modulbeschreibungen aus der gesamten DACH-Region. Allerdings sind Modulbeschreibungen aus deutschen Hochschulen stark überrepräsentiert. Das liegt zum einen an der deutlich geringeren Zahl österreichischer und Schweizer Hochschulen und Studiengänge. Zum anderen

mussten überproportional häufig Modulbeschreibungen dieser Hochschulen aufgrund einer Nichtverfügbarkeit der Modulschreibungen oder zu vieler fehlender Informationen aus der Ausgangsstichprobe entfernt werden.

- Die ausgewerteten Modulbeschreibungen wurden im Wintersemester 2022/2023 veröffentlicht. Inwieweit diese in den folgenden Semestern aktualisiert wurden, wurde nicht untersucht.
- Die Analyse basiert auf einer quantitativen und qualitativen Inhaltsanalyse von Modulbeschreibungen. Da die Qualität einer Inhaltsanalyse wesentlich von der Reliabilität der durchgeführten Kodierung abhängt, wurden 25% der Modulbeschreibung einer erneuten Kodierung unterzogen. Aufgrund der starken Normierung von Modulbeschreibungen haben die sehr wenigen Abweichungen keinen Einfluss auf die Untersuchungsergebnisse und können daher vernachlässigt werden.

Literaturverzeichnis

- AICPA & CIMA. (2023). *CGMA Competency Framework: 2023 Update*.
- AK Finanzierung der Schmalenbach-Gesellschaft (2022). ESG-Integration im Finanzbereich: Veränderungen für Prozesse, Kommunikation und die Rolle des CFO. *Schmalenbach IMPULSE*, 2(1), 1–20.
- Al-Hazaima, H., Low, M. & Sharma, U. (2021). Perceptions of salient stakeholders on the integration of sustainability education into the accounting curriculum: a Jordanian study. *Meditari Accountancy Research*, 29(2), 371–402.
- Apitz, N., Gubini, T., Knauer, T. & Winkelmann, S. (2023). Nachhaltigkeitsorientiertes Investitionscontrolling : Ergebnisse einer Befragung zur Verankerung von Nachhaltigkeit in Strategie und Investitionscontrolling. *Controlling : Zeitschrift für erfolgsorientierte Unternehmenssteuerung*, 35(1), 20–27.
- Ascani, I., Ciccola, R. & Chiucchi, M. S. (2021). A structured literature review about the role of management accountants in sustainability accounting and reporting. *Sustainability*, 13(4), 1–25. <https://doi.org/10.3390/su13042357>
- Barber, N. A., Wilson, F., Venkatachalam, V., Cleaves, S. M. & Garnham, J. (2014). Integrating sustainability into business curricula: University of New Hampshire case study. *International Journal of Sustainability in Higher Education*, 15(4), 473–493.
- Baumüller, J., Eisl, C. & Leitner-Hanetseder, S. (2023). Neue KPI für die Unternehmenssteuerung durch CSRD und ESRS. *REthinking: Finance*, 5(4), 4–9.
- Benn, S. & Dunphy, D. (2009). Action research as an approach to integrating sustainability into MBA programs: An exploratory study. *Journal of Management Education*, 33(3), 276–295.
- Berg, T. & Lyngstadås, H. (2023). We're only in it for the money? Developing sustainable literacy through management accounting curriculum. *Accounting Education: an international journal*, 32, 1–20. <https://doi.org/10.1080/09639284.2023.2270475>
- Botes, V., Low, M. & Chapman, J. (2014). Is accounting education sufficiently sustainable? *Sustainability Accounting, Management and Policy Journal*, 5(1), 95–124.
- Boulianne, E. & Keddle, S. L. (2018). Where is sustainability within the Canadian CPA education program? In *Sustainability Accounting: Education, Regulation, Reporting and Stakeholders* (S. 71–112). Emerald Publishing Limited.
- Boulianne, E., Keddle, L. S. & Postaire, M. (2018). (Non) coverage of sustainability within the French professional accounting education program. *Sustainability Accounting, Management and Policy Journal*, 9(3), 313–335.
- Cho, C. H., Kim, A., Rodrigue, M. & Schneider, T. (2020). Towards a better understanding of sustainability accounting and management research and teaching in North America: a look at the community. *Sustainability Accounting, Management and Policy Journal*, 11(6), 985–1007. <https://doi.org/10.1108/SAMPJ-08-2019-0311>
- Gomes, S. F., Jorge, S. & Eugénio, T. (2021). Teaching sustainable development in business sciences degrees: Evidence from Portugal. *Sustainability Accounting, Management and Policy Journal*, 12(3), 611–634.
- Hahn, R. & Reimsbach, D. (2014). Are we on track with sustainability literacy? A viewpoint on the influence of sustainability and accounting education on future managers' processing of sustainability information. *Journal of Global Responsibility*, 5(1), 55–67. <https://doi.org/10.1108/JGR-12-2013-0016>
- Institute of Management Accountants. (2019). *IMA management accounting competency framework*. Institute of Management Accounting Buffalo, NY, USA.
- International Group of Controlling. (2015). *Controller-Kompetenzmodell: Ein Leitfaden für die moderne Controller-Entwicklung mit Muster-Kompetenzprofilen*. Haufe.
- Kämmler-Burrak, A. & Bauer, R. (2022). Nachhaltigkeit wird Standardaufgabe im Controlling. Vom „Why“ zum „How“. *Controller Magazin*, 47(1), 22–27.

- Kämmeler-Burrak, A., Möhrer, M., Rötzel, P., Schulze, M. & Gimpl, N. (2022). *Green Controlling Studie 2022: Stand und Herausforderungen der Integration ökologischer und sozialer Aspekte in das Controlling aus Sicht der Controllingpraxis*.
- Khan, T. (2013). Sustainability accounting courses, Talloires Declaration and academic research. *International Journal of Sustainability in Higher Education*, 14(1), 42–55.
- Littan, S. H., Brohé, A., Fertig, K., Khong, C. & Friedman, J. (2022). *Management Accountants' Role in Sustainable Business Strategy: A Guide to Reducing a Carbon Footprint*.
- Marx, B. & van der Watt, A. (2013). Sustainability in accounting education: An analysis of the teaching thereof at accredited South African universities. *South African Journal of Accounting Research*, 27(1), 59–86.
- Matten, D. & Moon, J. (2004). Corporate social responsibility. *Journal of business Ethics*, 54(4), 323–337.
- Mburayi, L. & Wall, T. (2018). Sustainability in the professional accounting and finance curriculum: an exploration. *Higher Education, Skills and Work-based Learning*, 8(3), 291–311. <https://doi.org/10.1108/HESWBL-03-2018-0036>
- Olalla, C. B. & Merino, A. (2019). Competences for sustainability in undergraduate business studies: A content analysis of value-based course syllabi in Spanish universities. *International Journal of Management Education*, 17(2), 239–253.
- Petersen, H., Lühn, M., Nuzum, A.-K., Schaltegger, S. & Wenzig, J. (2021). Controller als Partner im Nachhaltigkeits-Management. *Controlling & Management Review*, 65, 8–15.
- Rädiker, S. & Kuckartz, U. (2019). *Analyse qualitativer Daten mit MAXQDA: Text, Audio und Video*. Springer.
- Reimer, M., Schäffer, U. & Weber, J. (2021). *Nachhaltige Unternehmenssteuerung: Eine Studie des WHU Controller Panels (2021)*.
- Sailer, U. (2022). *Nachhaltigkeitscontrolling* (4. Aufl.). UKV Verlag.
- Sopp, K., Baumüller, J. & Scheid, O. (2023). *Nachhaltigkeitsberichterstattung: Nichtfinanzielle Berichterstattung nach dem CSR-RUG, Neuerung durch die CSRD und die ESRS* (3. Aufl.). nwb.
- Stubbs, W. & Schapper, J. (2011). Two approaches to curriculum development for educating for sustainability and CSR. *International Journal of Sustainability in Higher Education*, 12(3), 259–268. <https://doi.org/10.1108/14676371111148045>
- Vanini, U. (2022). Nachhaltigkeitscontrolling. *WISU Das Wirtschaftsstudium*, 51(2), 175–185.
- Vanini, U. & Bochart, S. (2024). Integration of sustainability issues into management accounting textbooks. *Journal of Accounting Education*, 66, 100886.
- Vanini, U., Heimbürg, J. von & Giese, M. (2023). Ableitung von Aufgaben-und Kompetenzanforderungen im Nachhaltigkeitscontrolling. In S. Behringer (Hrsg.), *Konferenzband CARF Luzern 2023 - Controlling. Accounting. Risiko. Finanzen* (S. 93–115). Verlag IFZ.
- Velte, P. & Stawinoga, M. (2019). Harmonisierung der Klimaberichterstattung?! Einbettung in die EU-Regulierungsinitiativen, Forschungslücken und Handlungsempfehlungen. *Der Betrieb*(37), 2025–2033.
- Weber, J. & Schäffer, U. (2022). *Einführung in das Controlling* (17. Aufl.). Schäffer Poeschel.
- Williams, B. (2015). The local government accountants' perspective on sustainability. *Sustainability Accounting, Management and Policy Journal*, 6(2), 267–287.

Institut für Finanzdienstleistungen Zug IFZ



Institut für Finanzdienstleistungen Zug IFZ



Das Institut für Finanzdienstleistungen Zug IFZ ist in der Schweiz das grösste Fachhochschulinstitut im Finanzbereich. Den Erfolg hat es einerseits der Innovation und dem Engagement seiner Mitarbeitenden und Führungskräfte zu verdanken, andererseits aber auch dem Verein IFZ, den Ehemaligen der Weiterbildungslehrgänge des IFZ und der intensiven Vernetzung mit der Finanzbranche. Das Institut der Hochschule Luzern – Wirtschaft wurde im Herbst 1997 auf Initiative des Kantons Zug und der Zuger Wirtschaft gegründet. Das Department W der Hochschule Luzern ist seit 2021 durch die AACSB akkreditiert und gehört damit zu 6 % der besten Business Schools weltweit.

IFZ – Facts and Figures

- Kompetenzzentrum der Hochschule Luzern für „Finance & Banking“, „Controlling & Accounting“, «Risk & Compliance» und „Immobilien“ (Bachelor, Master, Weiterbildung, Forschung, Beratung und weitere Dienstleistungen).
- Führendes Fachhochschulinstitut in der Schweiz im Bereich Finance.
- 3 Studienrichtungen im Bachelor-Lehrgang der Hochschule Luzern („Finance & Banking“, „Controlling & Accounting“ und „Immobilien“)
- 3 konsekutive Masterlehrgänge: MSc Banking and Finance, MSc International Financial Management, MSc Real Estate
- 7 MAS Master of Advanced Studies, 8 DAS Diploma of Advanced Studies und 25 CAS Certificate of Advanced Studies
- Aktuell rund 800 Teilnehmende in der Weiterbildung (MAS, DAS, CAS)
- Seminare und Konferenzen mit über 2'000 Teilnehmenden pro Jahr
- Forschungsprojekte in Zusammenarbeit mit Wirtschaftspartnern
- Eigene Schriftenreihe mit Buchpublikationen

- Geschäftsführung für die SECA, CFO Forum Schweiz, swissVR, Zug Commodity Association – ZCA, NPO Finanzforum und SwissERM
- Anwendungsorientierte Forschungsprojekte in Zusammenarbeit mit Wirtschaftspartnern und mit Unterstützung von Innosuisse
- Transfers von wissenschaftlichen Erkenntnissen in die Praxis und von praktischen Erfahrungen in die Wissenschaft
- Lehrmittel, Fachbücher und praxisorientierte Publikationen
- Circa 100 Mitarbeitende (Dozierende, wissenschaftliche und administrative Mitarbeitende)
- 500 Lehrbeauftragte aus Wissenschaft und Praxis

Das Institut für Finanzdienstleistungen Zug IFZ hat seine Schulungs- und Büroräumlichkeiten in unmittelbarer Nähe des Bahnhofs Rotkreuz. Abgesehen von der sichergestellten Basisfinanzierung durch die Fachhochschule Zentralschweiz, den Kanton Zug und den Verein IFZ, hat sich das Institut für Finanzdienstleistungen Zug IFZ im Bereich der Weiterbildung und der Dienstleistungen voll, in der Forschung weitgehend selbsttragend zu finanzieren. Es erwirtschaftet jährlich rund CHF 10 Mio. Franken Einnahmen.

Kontakt

Hochschule Luzern - Wirtschaft

Institut für Finanzdienstleistungen Zug IFZ

Campus Zug-Rotkreuz, Suurstoffi 1,

6343 Rotkreuz ZG

+41 41 757 67 67

ifz@hslu.ch



hslu.ch/w-ifz-newsletter



blog.hslu.ch/ifz

blog.hslu.ch/retailbanking

blog.hslu.ch/investments

blog.hslu.ch/immobilienblog

blog.hslu.ch/financialmanagement

blog.hslu.ch/digitalcontrolling



twitter.com/IFZ_Zug



facebook.com/ifz.zug



linkedin.com/company/ifz-zug