



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Einsatzpotenziale von künstlicher Intelligenz in der Lehrkräftebildung – Wirkung und Nutzung einer KI-basierten Assistenz für die Unterrichtsplanung im Kontext von Digitalisierung und Nachhaltigkeit

Kumulative Dissertation

zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil)

Eingereicht an der Fakultät für Erziehungswissenschaft

am Fachbereich 3: *Berufliche Bildung und Lebenslanges Lernen*

der Universität Hamburg

Vorgelegt von Julia Pargmann

Hamburg, Mai 2025

Erstgutachter: Prof. Dr. Florian Berding

Zweitgutachter: Prof. Dr. Werner Kuhlmeier

Drittgutachter: Prof. Dr. Josef Guggemos

Mündliche Gutachterin: Prof. Dr. Vera Braun

Datum der Disputation: 17.10.2025

„... hier war ich, obwohl ich fast überhaupt nichts sehen konnte, mitten im hier und jetzt statt wie sonst im Wenn und Aber.“

Mariana Leky, Was man von hier aus sehen kann

Das Vorhaben, zu promovieren, mag der einen wahnsinnig vorkommen, der anderen sterbenslangweilig. Was es aber immer ist, ist ein andauerndes Auf und Ab, eine Welle, die dich mitreißt, dir den Boden unter den Füßen wegzieht, dir Salz in die Augen treibt und dich erst wieder loslässt, wenn du nicht mehr weißt, wo oben und unten ist. Und man findet Ausflüchte, so viele, wie noch nie zuvor, Ausflüchte, sich anderen Dingen zu widmen, oder es eben nicht zu tun. Ja, es ist wahr, manchmal sieht man neben diesem ganzen Vorhaben nichts anderes mehr. Manchmal spürt man sich nur ganz entfernt und der Kopf ist unglaublich voll. Aber das, was immer hilft, ist, einen Schritt vor den anderen zu machen. Im hier und jetzt zu sein. Eine Herausforderung, eine Seite nach der anderen. Und dann hat man es am Ende geschafft – ohne Wenn und Aber.

Ich danke...

...meinen beiden Betreuern Florian Berding und Werner Kuhlmeier, die mich immer wieder inspiriert, beraten oder auf den Boden der Tatsachen zurückgeholt haben. Eure Expertise ergänzt sich so gut, dass ich mir keine besseren Betreuer hätte wünschen können. Danke für eure Geduld, eure offenen Ohren, eure Unterstützung.

...allen anderen Weggefährtinnen und Weggefährten, die mich bei den verschiedenen Teilprojekten dieser Arbeit unterstützt haben: Ob am Institut für Berufs- und Wirtschaftspädagogik (Danke Anna, Vera, Lucas!), an der Universität Oldenburg (Danke Karin!), der Universität Graz (Danke Elisabeth!), überall durfte ich meine Ideen teilen, sie diskutieren und wachsen lassen.

...meinen Freundinnen und Freunden, dass sie mir zugehört haben, auch wenn sie meistens nicht den Hauch einer Ahnung hatten, wovon ich eigentlich rede. Aber man kennt es: Manchmal reicht es schon aus, das ganze einfach auszusprechen. Danke Juliane, Sophia, Talina, Caro, Valerie, Simon, Sarah und Manuel, auch dafür, dass ihr mich manchmal dazu gezwungen habt, einfach mal an was anderes zu denken.

...meiner Mama, dass sie mir von klein auf gezeigt hat, dass Frauen* alles sein können. Deinen Mut bewahre ich mir tief in meinem Herzen und erinnert mich immer wieder an all die Frauen*, die vor uns waren und nach uns kommen.

...meinem Papa, dass er mich stets in meinen verrückten Unterfangen unterstützt hat, auch wenn er sie nicht immer gut fand. Du hast mir gezeigt (und tust es auch noch immer), dass Wachstum keine Altersgrenzen kennt.

...meinem Bonuspapa für seine Geduld. Ich bewundere dich für deine Konsequenz, nach deinen Werten zu handeln.

...meinen Geschwistern, dass ihr mich immer abgelenkt habt, wenn ich es brauchte und mir immer wieder die Leichtigkeit des Seins zeigt, wenn ich Dinge mal wieder zu ernst nehme.

...meinem Mann Lukas. Du hast immer an mich geglaubt, egal wie aussichtslos ich auch war. Du hast mich motiviert, mich herausgefordert, dir unzählige Präsentationen angehört, Texte gelesen, Argumentationen nachvollzogen, mit mir gelacht und geweint. Danke, dass du immer für mich da bist und meine innere Stimme für mich hörbar machst.

Zu guter Letzt – und vielleicht ungewöhnlich – möchte ich mir selbst danken. Trotz aller Unterstützung habe ich es letztendlich vor allem mir selbst und meiner harten Arbeit zu verdanken, dieses Kapitel angefangen und auch zu Ende gebracht zu haben. Frauen* tendieren dazu, ihre Erfolge anderen Personen zuzuschreiben. Ich denke, da gibt es Handlungsbedarf. Also: Danke an mich selbst. Vergangenheits-Julia könnte es nicht fassen.

Inhalt

Abkürzungsverzeichnis	II
Abbildungsverzeichnis	III
Tabellenverzeichnis	III
1 Einleitung	1
1.1 Zur Problemstellung	1
1.2 Zum Aufbau der Arbeit	9
2 Theoretische und empirische Rahmungen	16
2.1 Unterrichtsplanung	16
2.2 Künstliche Intelligenz als Gegenstandsbereich	23
2.3 (Digitales und KI-gestütztes) Feedback	37
2.4 Aktuelle Herausforderungen in der beruflichen Bildung: Digitalisierung und Nachhaltigkeit im Rechnungswesen- und Controllingunterricht	44
2.5 Vorstellung der KI-basierten didaktischen Assistenz EDDA	50
3 Beiträge der kumulativen Dissertation	56
3.1 Beitrag 1	56
3.2 Beitrag 2	57
3.3 Beitrag 3	58
3.4 Beitrag 4	59
4 Diskussion	60
4.1 Gesamtdiskussion der Ergebnisse	60
4.2 Limitationen und Forschungsdesiderata	73
5 Fazit	74
Literaturverzeichnis	77
Anhang	102
A Liste von Publikationen zum Dissertationsthema	
B Beitrag 1	
C Beitrag 2	
D Beitrag 3	
E Beitrag 4	
F Zusammenfassung	
G Summary	

Abkürzungsverzeichnis

IKT	Informations- und Kommunikationstechnologien
BIBB	Bundesinstitut für Berufsbildung
KMK	Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland
UPK	Unterrichtsplanungskompetenz
KI	Künstliche Intelligenz
EDDA	Elektronisch-didaktische Assistenz für die Unterrichtsplanung
AI	Artificial Intelligence
NLP	Natural Language Processing
BERT	Bidirectional Encoder Representations
GPT	Generative Pre-Trained Transformers
DSGVO	Datenschutz-Grundverordnung
LLM	Large Language Model
BBiG	Berufsbildungsgesetz
ERP	Enterprise Resource Planning
KMU	Klein- und mittelständische Unternehmen
BBNE	Berufliche Bildung für eine nachhaltige Entwicklung
AOI	Area of Interest
NKPI	Nachhaltigkeits-Key Performance Indikatoren
SDG	Sustainable Development Goal

Abbildungsverzeichnis

Abbildung 1. Argumentationsbogen	1
Abbildung 2. Forschungsdesign der Arbeit	9
Abbildung 3. Nutzung professionellen Wissens nach Oevermann (1996).....	17
Abbildung 4. Arten des maschinellen Lernens im Überblick (basierend auf Kleesiek et al., 2020, Russel & Norvig, 2021 und Welsch et al., 2018)	25
Abbildung 5. Standardberufsbildpositionen für eine digitalisierte Arbeitswelt (Bundesinstitut für Berufsbildung, 2021).....	47
Abbildung 6. Standardberufsbildpositionen für Umweltschutz und Nachhaltigkeit (Bundesinstitut für Berufsbildung, 2021).....	49
Abbildung 7. Ergebnisübersicht.....	53
Abbildung 8. Übersicht über die grafischen Analyseergebnisse von EDDA	54
Abbildung 9. Beispiel für die Klappboxen mit ausgeklappter Rückmeldung	55

Tabellenverzeichnis

Tabelle 1. Überblick über die Beiträge und ihre verschiedenen theoretischen und methodischen Zugänge sowie zentralen Ergebnisse (Stand: 06.05.2025)	12
Tabelle 2. Synopse ausgewählter Studien zur Unterrichtsplanung	21
Tabelle 3. Synopse ausgewählter Studien zu KI und Lernförderlichkeit	33
Tabelle 4. Synopse ausgewählter Studien zu digitalem und KI-gestütztem Feedback	41
Tabelle 5. Überblick über die verschiedenen Interaktionstypen	68

1 Einleitung

1.1 Zur Problemstellung

Der Unterricht ist das zentrale Mittel des Lernorts Schule, um das normative Ziel der beruflichen Bildung zu erreichen: die Beförderung der beruflichen Handlungskompetenz (Abbildung 1, Feld 1). Diese wird definiert als die „Bereitschaft und Befähigung des Einzelnen, sich in beruflichen, gesellschaftlichen und privaten Situationen sachgerecht durchdacht sowie individuell und sozial verantwortlich zu verhalten“ (KMK, 2021, S. 15). Dabei erhöht hochwertiger Unterricht die Chance auf eine umfangreiche berufliche Handlungskompetenz (Abbildung 1, Feld; Riebenbauer et al., 2024). Um die gewünschte Handlungsfähigkeit herbeizuführen, müssen Auszubildende auch auf die Bewältigung zeitgemäßer Herausforderungen vorbereitet werden (KMK, 2021). Dies sind aktuell vor allem die Digitalisierung und die Nachhaltigkeit (Abbildung 1, Feld 3).

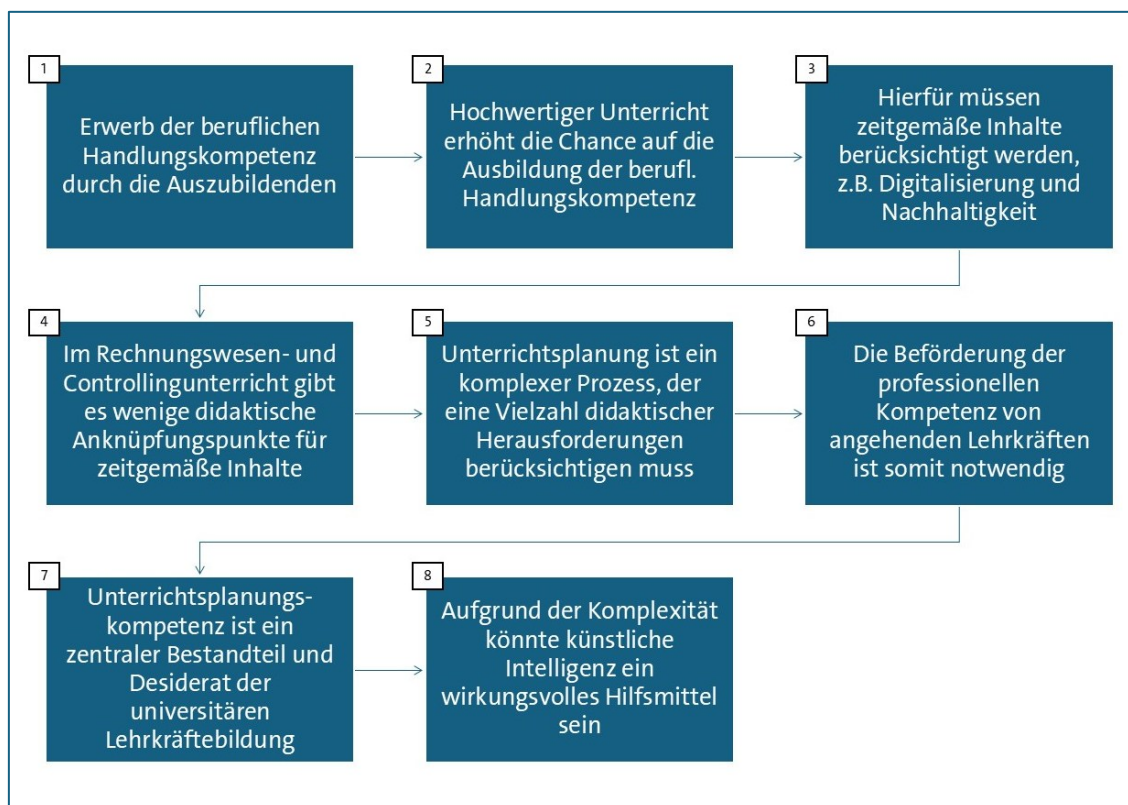


Abbildung 1. Argumentationsbogen

Der Digitalisierungsbegriff bezieht sich heute vor allem auf das Handeln in einer Welt, in der Informations- und Kommunikationstechnologien (IKT) für immer mehr Prozesse genutzt werden (Traum et al., 2017). Ursprünglich meinte der Begriff jedoch nur den Wandel vom Analogen ins Digitale. Heute hat die Digitalisierung weitreichende Implikationen für

die Berufswelt und verändert immer mehr Tätigkeiten (Dengler & Matthes, 2021). Ähnliche Auswirkungen hat auch die Nachhaltigkeit, welche vor allem die Verschränkung der Dimensionen Ökologie, Ökonomie und Soziales meint (Pufé, 2017). Im Sinne eines nachhaltigen Wirtschaftens kann dies z.B. bedeuten, andere Zielgrößen als nur die ökonomische bei der Entscheidungsfindung zu berücksichtigen (Berding et al., 2018). Im Kontext der Nachhaltigkeit verändern sich ebenso wie in der Digitalisierung ganze Tätigkeitsbereiche und unternehmerische Prozesse (Pargmann & Berding, 2025; Schaltegger, 2016). Während vormals die Berücksichtigung von Nachhaltigkeit z.B. im Controlling keine Rolle spielte, ist sie nun zunehmend ein konkretes Erfordernis bei der Datenzusammenstellung, sodass auch Controller:innen vermehrt nachhaltigkeitsrelevante Daten aggregieren und auswerten müssen (Sailer, 2020).

Auch das Bundesinstitut für berufliche Bildung (BIBB) hat die Relevanz von Digitalisierung und Nachhaltigkeit erkannt und die Veränderungen in die Standardberufsbildpositionen aufgenommen. Curricular sind die beiden Handlungsfelder seit 2021 für neue und neu geordnete Berufe über die Standardberufsbildpositionen *Umweltschutz und Nachhaltigkeit* und *Digitalisierte Arbeitswelt* vertreten. Die Standardberufsbildpositionen stellen berufsübergreifende Ausbildungsinhalte dar, die für „jede qualifizierte Fachkraft ein unverzichtbares Fundament kompetenten Handelns“ sind (Bundesinstitut für Berufsbildung, 2021, S. 5). Die Novelle dieser Inhalte erfordert sowohl im Betrieb als auch in der Schule inhaltliche und methodische Anpassungen. In den Rahmenlehrplänen der beruflichen Bildung gibt es trotzdem nur geringfügige Anhaltspunkte (Bundesinstitut für Berufsbildung, 2021; Holst, 2022). Damit liegt es an einzelnen Schulen oder Lehrkräften, sich mit den Inhalten auseinanderzusetzen und abzuleiten, wie diese in den Unterricht integriert werden können und eigene Professionalisierungsbedarfe abzuleiten (Holst, 2022).

Dieses allgemeine Desiderat für alle berufsbildenden Bildungsgänge kann ergänzt werden durch eines für die wirtschaftsberufliche Domäne. Das Rechnungswesen und Controlling nimmt einen wesentlichen Teil wirtschaftsberuflicher Curricula ein (Brötz et al., 2015), es ist ein zentrales Instrument zur Förderung des wirtschaftlichen Denkens (Chiapello, 2009; Preiß & Tramm, 1996). Dies unterstreicht die Bedeutung von Rechnungswesen und Controlling für die Beförderung der beruflichen Handlungskompetenz. Exemplarische Forschungsarbeiten umfassen Erkenntnisse über die didaktische Gestaltung des Unterrichts (Seifried, 2004), Interesse und Lernmotivation (Greimel-Fuhrmann, 2008), typische Fehler (Türling et al., 2011), Langeweile (Kögler & Wuttke, 2012), Effekte des Verhaltens von

Lehrkräften, Motivation und Leistungen (Helm, 2015), Grundvorstellungen (Berding, 2019), die Gestaltung von Lernaufgaben (Stütz, 2024), das Erleben des Unterrichts (Stütz, 2021) oder allgemeine Entwicklungen in der Unterrichtsplanung (Riebenbauer, 2021). Erkenntnisse zur Unterrichtsplanung im Kontext von Digitalisierung und Nachhaltigkeit fehlen (Abbildung 1, Feld 4; Pargmann et al., 2022; Pargmann et al., 2023; Pargmann et al., im Druck).

Betriebswirtschaftlich sind die beiden Geschäftsbereiche gleichzeitig optimal zur Realisierung von Transformationsprozessen hin zu einer nachhaltigen oder digitalisierten Unternehmung geeignet, da sie die Entscheidungsfindung vorbereiten, informieren und unterstützen (Schaltegger, 2016). Digitale Technologien bieten hierfür neue Möglichkeiten zur Unterstützung des Rechnungswesens und Controllings, z.B. durch die Datenanalyse (Klein & Küst, 2020). Digitalisierung wirkt disruptiv auf ganze Berufsfelder und Branchen und verändert Tätigkeiten (Dengler & Matthes, 2021). So reduziert sich zunehmend der Routinegrad der Tätigkeiten und einfachere Arbeiten werden eher automatisiert. Vor allem in sogenannten Helferberufen liegt der Anteil der Aufgaben, der potenziell substituierbar ist, bei 57% (Grienberger et al., 2024). Hier gingen die Zahlen zuletzt jedoch zurück, während in Expertenberufen nun 35,8% der Tätigkeiten substituierbar sind (26,1% im Jahr 2019, siehe Dengler & Matthes, 2021). Digitalisierung betrifft somit de facto alle Berufe und Branchen.

Gleichzeitig spielt Nachhaltigkeit eine entscheidende Rolle, da zukünftige Fachkräfte mit der Herausforderung konfrontiert werden, wirtschaftliche Entscheidungen unter Berücksichtigung ökologischer und sozialer Aspekte zu treffen (Hantke & Pranger, 2019). So können z.B. Daten über die Nachhaltigkeitsleistung eines Unternehmens aufbereitet und zur Bewertung der Umsetzung der Nachhaltigkeitsstrategie genutzt werden. In der Fachwissenschaft und aus der Fachpraxis werden explizite Anforderungen an die zukünftigen Mitarbeiter:innen deutlich (Pargmann & Berding, 2025; Pargmann et al., 2023). Jedoch ist nicht klar, wie eine Umsetzung von Digitalisierung und Nachhaltigkeit im Rechnungswesen- und Controllingunterricht *inhaltlich* ausgestaltet werden kann, welche dezidierten Kompetenzen bei den Lernenden befördert werden sollen und welche Implikationen dies für die Unterrichtsgestaltung hat (Pargmann & Berding, 2025; Pargmann et al., 2023). Die ohnehin schon hohe Komplexität des Unterrichts steigt somit, genauso wie die der Unterrichtsplanung.

Zur Unterrichtsplanung gehören alle Tätigkeiten, die den Unterricht an sich vorbereiten (Sandfuchs, 2004). Damit ist sie ein essenzieller Bestandteil der Professionalisierung von Lehrkräften (Abbildung 1, Feld 6; KMK, 2019). Für einen hochwertigen Unterricht muss insgesamt eine Vielfalt didaktischer Prinzipien berücksichtigt werden und eine gute Planung sollte wiederum zu einem guten Unterricht führen (Abbildung 1, Feld 5; Riebenbauer et al., 2024; Wilbers, 2023). Die Unterrichtsplanung wurde somit einerseits deshalb als Gegenstandsbereich ausgewählt, weil durch einen hochwertigen Unterricht die Chance auf eine langfristig tragfähige berufliche Handlungskompetenz realisiert werden kann. Andererseits ist sie als Bestandteil der Professionalisierung von Lehrkräften aber auch zentrale Aufgabe der universitären Lehrkräftebildung, indem sie auf die spätere Kerntätigkeit von Lehrkräften vorbereitet; das Unterrichten (Abbildung 1, Feld 7; Baumert & Kunter, 2006). Auch die KMK (2022, S. 3) formuliert im Kompetenzbereich *Unterrichten* der Standards für die Lehrkräftebildung eindeutig, dass Lehrkräfte als „Fachleute für das Lehren und Lernen“ gelten. In der darauffolgend erläuterten Kompetenz wird ein direkter Bezug zur Unterrichtsplanung und -durchführung nach aktuellen didaktischen und wissenschaftlichen Erkenntnissen hergestellt.

Nach Sandfuchs (2004) ist eine sorgfältige Unterrichtsvorbereitung notwendig, um die komplexen Prozesse des Lehrens und Lernens bestmöglich zu optimieren. Hierfür wird im Rahmen des Lehramtsstudiums die Unterrichtsplanungskompetenz (UPK) befördert. Ein bewährtes Mittel zum *Planen lernen* sind Unterrichtsentwürfe. Diese sind ein zentrales Element des professionellen Handelns von Lehrkräften (König & Rothland, 2022). Das professionelle Handeln umfasst neben der Planung an sich auch die Auswertung und Überarbeitung von Unterricht. Trotz der allgemeinen Anerkennung der Relevanz der UPK fehlt in der Literatur eine einheitliche Definition. Einige Autor:innen sehen sie als eigenständige Kompetenz an, andere als Bestandteil der professionellen Kompetenz von Lehrkräften (Baumert & Kunter, 2006; Klusmeyer, 2021; König & Rothland, 2022; Shulman, 1987). Übereinstimmende Facetten sind die fachliche, fachdidaktische und allgemeine pädagogischen Kompetenzen. Die Forschung zu diesem wichtigen Bereich der Lehrkräftebildung ist in der beruflichen Bildung bisher jedoch unterentwickelt. Ein Großteil der Studien stammt aus den 2010er Jahren, was auf eine bislang nachrangige Relevanz der Thematik hindeutet (Reinisch, 2021).

Auch wenn in den letzten Jahren neue Publikationen hinzugekommen sind, die aktuelle Dimensionen und Modelle der Unterrichtsplanung beleuchten, bestehen gerade im

Kontext der Digitalisierung deutliche Desiderata (Klusmeyer, 2021; Riebenbauer, 2021; Söll, 2021). Denn die Studien betonen die Notwendigkeit, moderne Entwicklungen in die Unterrichtsplanung zu integrieren. Es zeigt sich, dass die Auseinandersetzung mit der Unterrichtsplanung und deren Prinzipien für Lehramtsstudierende von entscheidender Bedeutung ist. Sie bilden die Grundlage für die Entwicklung einer handlungsvorbereitenden und reflektierenden Wissensbasis, die für die erfolgreiche Ausführung von Lehrtätigkeiten unerlässlich ist (Söll & Klusmeyer, 2018). Eine aktuelle Studie von Riebenbauer et al. (2025) belegt für die Wirtschaftspädagogik, dass zentrale didaktische Planungsdimensionen häufig nur unzureichend umgesetzt werden. Es besteht somit eine eindeutige Notwendigkeit, der Beförderung der UPK mehr Beachtung zu schenken. Da gezielte Unterrichtsplanungsaktivitäten zur Verbreitung von Innovationen beitragen können (Aprea, 2020; Klusmeyer & Söll, 2021), ist es in Bezug auf die Integration aktueller Herausforderungen vielversprechend, gezielte Planungen zu Rechnungswesen- und Controllingthemen zu unternehmen, die Digitalisierung und Nachhaltigkeit berücksichtigen und somit bewirken, dass Lehramtsstudierende diese Themen von Anfang an mitdenken. Hierfür muss die UPK weiterentwickelt werden, indem sich explizit auf die spezifischen Anforderungen des Rechnungswesen- und Controllingunterrichts konzentriert wird.

Bei der Professionalisierung sind regelmäßige qualitative Rückmeldungen über die Lehr-Lernprozesse lernförderlich, z.B. in Form von Feedback (Wisniewski et al., 2019). Feedback wird nach Hattie und Timperley (2007) als Bereitstellung von Informationen über bestimmte Aspekte einer Leistung verstanden. Wisniewski et al. (2019) zufolge ist mit einer Effektstärke von $d=0,99$ besonders solches Feedback lernförderlich, welches möglichst viele Informationen über die Ergebnisse, deren Zustandekommen und konkrete Handlungsbedarfe beinhaltet. Im Vergleich hierzu hat korrigierendes Feedback nur einen mittleren Effekt von $d=0,46$. Um lernförderlich zu sein, sollte Feedback möglichst kriterienorientiert, spezifisch, detailliert, individuell und auf Augenhöhe formuliert werden (Wisniewski et al., 2019). Außerdem sollte es Informationen über die Leistungen an sich liefern und formativ entlang des Lernprozesses erfolgen (Narciss, 2013; Schartel, 2012; Shute, 2008).

Aufgrund vielfältiger struktureller und systemischer Aspekte, wie zeitlichen Restriktionen der Dozierenden, hohen Lehrdeputaten oder der Veranstaltungsgröße sind die Möglichkeiten für individualisiertes Feedback in der jedoch stark Hochschullehre begrenzt (Henderson et al., 2019). Speziell in der hochschuldidaktischen Lehrkräftebildung kann demnach

die Nutzung von KI-Feedback positive Effekte auf die Kompetenzentwicklung der Lehramtsstudierenden haben (z.B. Er et al., 2024 oder Dai et al., 2024). Denn Metaanalysen legen nahe, dass auch digitales Feedback mindestens genauso, wenn nicht sogar lernförderlicher ist als traditionelles Feedback (Mertens et al., 2022; van der Kleij et al., 2015).

Es zeichnen sich somit vielfältige Herausforderungen für den Professionalisierungsprozess von Lehrkräften ab. Um diesen Herausforderungen gerecht zu werden, sollten neue Technologien als Unterstützung für die Unterrichtsplanung herangezogen werden, um auch die oben angesprochenen Innovationen in die Lehrkräftebildung zu integrieren. Dazu zählen aktuell vor allem Werkzeuge auf Basis künstlicher Intelligenz (KI) (Abbildung 1, Feld 8). Rich (1983, S. 1) definiert KI wie folgt: „Artificial intelligence is the study of how to make computers do things at which, at the moment, people are better“. Forschungen zu KI existieren bereits seit den 1940er Jahren, aufgrund der technologischen Möglichkeiten können viele Ansätze jedoch erst im aktuellen Jahrzehnt für den breiten Markt verfügbar gemacht werden (Russel & Norvig, 2021). KI bietet vielfältige Möglichkeiten, Lehr-Lernprozesse zu verändern, zu unterstützen und zu optimieren. Für die professionelle Kompetenz von Lehrkräften bedeutet dies eine Erweiterung, z.B. im Rahmen der Beförderung von KI-Kompetenzen (Bundesregierung, 2020; Pargmann & Berding, 2022; Zhai et al., 2021).

Jedoch bergen verschiedene Technologien verschiedene Chancen und Grenzen, sowohl auf der allgemeinen Ebene als auch speziell für den Bildungsbereich. Aktuell bündeln sich die Forschungsbemühungen im Bereich der generativen KI, vor allem mittels ChatGPT (Beheshti, 2023; Jacobsen & Weber, 2024; Jeon & Lee, 2023; Kaplan-Rakowski et al., 2023; Khowaja et al., 2024; Lee & Moore, 2024; Yiheng Liu et al., 2023; Yue Liu et al., 2023; Strzelecki, 2023, 2024; Strzelecki & ElArabawy, 2024; van den Berg & Du Plessis, 2023; Wu et al., 2023). Generative KI-Systeme erzeugen z.B. neue Texte, Bilder oder Videos auf Basis von Algorithmen der Wahrscheinlichkeitsrechnung (Beheshti, 2023). In Abhängigkeit von der Eingangsfrage, dem sogenannten Prompt, werden die Ergebnisse generiert, die mit höchster Wahrscheinlichkeit die Anforderungen des Nutzenden erfüllen. Solche Anwendungen bergen die Herausforderung, dass sie schlecht reproduzierbare Inhalte generieren, die zudem ein hohes Risiko für Halluzinationen, also verzerrte oder falsche Aussagen, aufweisen (Beutel et al., 2023; Salvagno et al., 2023).

Eine Alternative zu generativen KI-Systemen stellen analytische Systeme dar, die mittels maschinellen Lernens Muster in Daten erkennen und diese daraufhin in vorgefertigte

Kategorien einordnen, auf deren Basis Rückmeldungen gegeben werden (Beheshti, 2023). Spezialisierte KI-Systeme haben darüber hinaus den Vorteil, dass aufgrund ihrer Funktionsweise weniger Energie verbrauchen und damit potenziell nachhaltiger sind. Aktuell gibt es zu analytischen KI-Systemen im Bildungsbereich nur wenige Erkenntnisse, speziell im Bereich der beruflichen Bildung sind keine Arbeiten bekannt. Analytische KI-Plattformen haben gegenüber generativen Systemen jedoch den Vorteil, dass sie jederzeit reproduzierbares Feedback auf Expertenniveau geben können, ohne, dass Nutzerprofile erstellt werden müssen.

Losgelöst von der Art der KI-Anwendung besteht das Risiko, dass die KI einen Bias entwickelt, also vorurteilsbehaftete und im Extremfall diskriminierende Ergebnisse produziert (Bostrom & Yudkowsky, 2014; Girasa, 2020; Luan et al., 2020). Dies macht einen unreflektierten und unkontrollierten Einsatz im Bildungskontext herausfordernd. Ein weiterer Aspekt ist der des Datenschutzes; so nutzen einige beliebte KI-Plattformen die Nutzungsdaten und -ergebnisse, um die KI-Ergebnisse zu verbessern (Kloos & Schmidt-Bens, 2020; Lauterjung, 2020; Vincent-Lancrin & van der Vlies, 2020). Daher sollten keine sensiblen Daten eingespeist werden, was gerade im Bildungsbereich nicht einfach ist.

Erste Forschungsarbeiten zu KI-basiertem Feedback zeigen auf, dass es durchaus vergleichbare Ergebnisse zu traditionellem Feedback erzeugen kann (Hooda et al., 2022; Jacobsen & Weber, 2024; Lee & Moore, 2024; Sembey et al., 2024). Zu KI-basiertem Feedback und dessen Effekte auf die Unterrichtsplanung bzw. den Einfluss auf die Professionalisierung von Lehrkräften liegen aktuell noch keine Erkenntnisse vor.

Eine analytische KI-Plattform, die auf Basis der Erkenntnisse zur Feedbackforschung und der Mediendidaktik (siehe hierzu z.B. R. E. Mayer & Fiorella, 2021a) von Forschenden an der Universität Hamburg in Kooperation mit den Universitäten Oldenburg und Graz entwickelt wurde, ist die elektronisch-didaktische Assistenz zur Unterrichtsplanung, kurz EDDA (Pargmann et al., 2024). In EDDA können (angehende) Lehrkräfte (aktuell nur der Wirtschaftspädagogik) ihre Unterrichtsplanungen und -materialien hochladen und hinsichtlich der Umsetzung verschiedener allgemein- und fachdidaktischer Planungsdimensionen überprüfen lassen. Die Nutzer:innen erhalten dann ein passendes Feedback mit einer Definition der Kategorie, einer Einordnung des Ergebnisses, einer Begründung für diese Einordnung und konkrete Überarbeitungshinweise. Die Überarbeitungshinweise basieren auf der Analyse häufiger Fehler durch das Projektteam, auf deren Basis Expertenfeedback formuliert und

im System hinterlegt wurde. Außerdem erhalten die Nutzer:innen Hinweise zu weiterführender Literatur, um individuelle Vertiefungen zu ermöglichen. EDDA wurde mit dem Gedanken konzipiert, folgende Zielgruppen zu unterstützen (Pargmann et al., 2024):

- Im Lehramtsstudium: zur individuellen Auseinandersetzung mit dem eigenen Unterrichtsentwurf und zur beratenden Auseinandersetzung mit den Entwürfen in einer größeren Lerngruppe
- Im Vorbereitungsdienst: Entwicklung von Unterrichtsentwürfen unter aktuellen fachdidaktischen und lehr-lernpsychologischen Erkenntnissen durch die Lehrkräfte im Vorbereitungsdienst, zielorientiertere Beratung zu Unterrichtsplanungen durch die Ausbildungslehrkraft
- In der Berufstätigkeit: kontinuierliche Fortbildung und Inspiration zur Umsetzung neuer fachdidaktischer Forschungserkenntnisse und Herausforderungen wie Digitalisierung und Nachhaltigkeit, neue didaktische und methodische Impulse

Diese Kooperation zwischen KI und Mensch, in der die jeweiligen Stärken genutzt werden, entspricht dem von Seufert et al. (2021) formulierten Partnerschaftsprinzip, welches als Ziellinie der KI-Nutzung beschrieben wird.

Aus dem hier dargelegten Argumentationsbogen (Abbildung 1) und den abgeleiteten *Dconsiderata* ergibt sich für die vorliegende Arbeit nachfolgende Fragestellung:

Wie kann eine KI-Plattform in der wirtschaftsberuflichen Lehrkräftebildung gestaltet und genutzt werden, um Studierende bei ihrer Unterrichtsplanung zu aktuellen Themen zu unterstützen?

Gleichwohl ergeben sich folgende Teilfragestellungen:

1. Wie ist der Stand der Digitalisierung im Rechnungswesen und Controlling? Welche Inhalte sollten in den Unterricht integriert, welche Kompetenzen befördert werden?
2. Wie ist der Stand der Nachhaltigkeit im Rechnungswesen und Controlling? Welche Inhalte sollten in den Unterricht integriert, welche Kompetenzen befördert werden?
3. Wie wirksam ist die Nutzung der KI-Plattform als Feedbackgeberin im Vergleich zu rein menschlichem Feedback?

4. Wie interagieren Studierende mit der Plattform und welche Nutzungstypen gibt es?

Im Laufe der vorliegenden Arbeit werden diese Fragestellungen mittels vier ausgewählter veröffentlichter und unveröffentlichter Zeitschriftenbeiträge bearbeitet. Daher wird im nächsten Abschnitt zunächst der weitere innere Aufbau der Dissertation entwickelt. Das zweite Kapitel setzt den inhaltlichen Rahmen, indem theoretische und empirische Bezüge der Zeitschriftenbeiträge vorbereitet und vertieft werden. Es folgen die einzelnen Zeitschriftenbeiträge mit einer kurzen Zusammenfassung, Einordnung und Herausstellung des Eigenanteils der Autorin in Kapitel 3. In der anschließenden Diskussion in Kapitel 4 werden zentrale Erkenntnisse hervorgehoben, Konsequenzen abgeleitet, Limitationen herausgestellt und Desiderata für Forschung und Praxis formuliert. Die Arbeit schließt mit einem Fazit in Kapitel 5.

1.2 Zum Aufbau der Arbeit

Die innere Struktur dieser Arbeit orientiert sich an dem zugrundeliegenden Forschungsdesign in Abbildung 2:

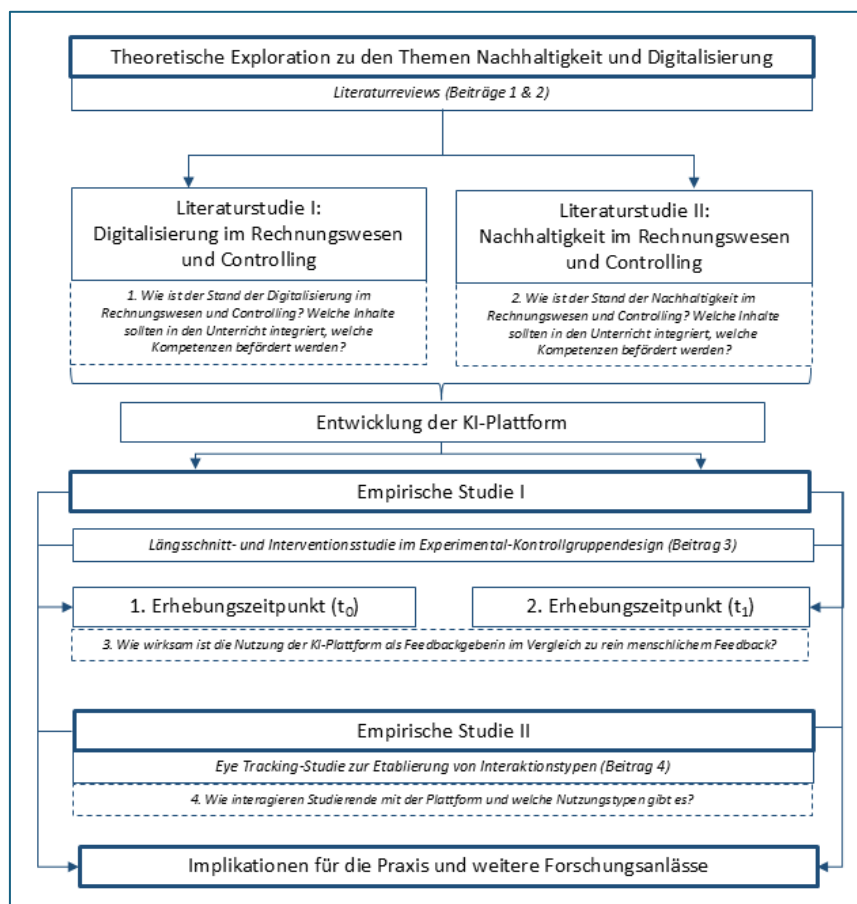


Abbildung 2. Forschungsdesign der Arbeit

Die initiale theoretische Exploration mündet in zwei Literaturreviews (Beitrag 1 und 2), die funktional den inhaltlichen Bezugsbereich dieser Arbeit abstecken (Abbildung 2, oberer Teil). Das Ziel der beiden Literaturreviews ist es, geeignete Inhalte für die Rückmeldungen der Software EDDA zu erarbeiten und so einen Beitrag zur Grundlegung der Software zu leisten (Abbildung 2, mittlerer Teil). Indirekt werden dadurch auch die beiden empirischen Studien zur Wirksamkeit des KI-Einsatzes vorbereitet. Beide Studien basieren auf dem gleichen Erhebungszeitraum. Außerdem werden die Fragestellungen 1 und 2 bearbeitet.

In der darauffolgenden empirischen Studie I im Experimental-Kontrollgruppendesign planen Studierende der Wirtschaftspädagogik einen nachhaltigkeitsorientierten Rechnungswesen- oder Controllingunterricht und überarbeiten ihn über ein Semester hinweg immer wieder in Bezug auf variierende didaktische Schwerpunkte (Abbildung 2, unterer Teil). Auch Digitalisierung ist integriert, um vertiefte Bezüge zu den Standardberufsbildpositionen herzustellen (siehe dazu Abschnitt 2.4). Beitrag 3 schildert die auf EDDA basierende längsschnittliche Interventionsstudie im Experimental-Kontrollgruppendesign. Zunächst wird die KI-Plattform vorgestellt und die didaktischen Kriterien, zu denen sie Feedback gibt, werden erläutert. Der Beitrag trägt zur Beantwortung der dritten Fragestellung bei. Es werden Unterrichtsentwürfe der Studierenden gesammelt und auf ihre Umsetzung von bestimmten Planungskriterien hin analysiert. Parallel zur Interventionsstudie wurden im Rahmen der empirischen Studie II in der Experimentalgruppe freiwillige Teilnehmende mit mobilen Eye Trackern begleitet (Beitrag 4), um Fragestellung 4 zu bearbeiten (Abbildung 2, unterer Teil). Aus den Erkenntnissen der Beiträge 1-4 werden abschließend Implikationen für Praxis und Forschung abgeleitet (Abbildung 2, unten). Das Forschungsvorhaben stellt ferner Bezüge zu verschiedenen theoretischen und methodischen Zugängen her. Diese Bezüge werden im zweiten Kapitel, im Rahmen der theoretischen und empirischen Rahmungen, entwickelt.

Einen Überblick aller Beiträge, dem Veröffentlichungsstatus, den Beitragenden, der Fragestellung, der Methodik sowie einer Kurzfassung der Ergebnisse gibt Tabelle 1. Ziel ist es, durch die Gestaltung der Tabelle einen Überblick über den inneren Zusammenhang der Arbeiten aufzuzeigen. Weitere Publikationen mit Bezug zum Dissertationsthema sind Anhang A zu entnehmen. Die Beiträge 1 und 2 entwickeln somit das theoretische Fundament des exemplarischen Bezugsbereichs dieser Arbeit, der Integration von Digitalisierung und Nachhaltigkeit in den Rechnungswesenunterricht. Die Ergebnisse der Literaturreviews werden in das Feedback integriert, das die KI-Plattform EDDA gibt (siehe Abschnitt 2.5). Um

die Plattform zu evaluieren, wird in Beitrag 3 eine Interventionsstudie im Experimental-Kontrollgruppen-Design vorgestellt und ausgewertet, die die Unterrichtsplanungen von Masterstudierenden der Wirtschaftspädagogik über ein Semester hinweg beinhaltet. Zuletzt sollen in Beitrag 4 verschiedene Interaktionstypen mit der KI-Plattform identifiziert werden, um Ansatzpunkte für eine personalisierte Nutzungserfahrung zu entwickeln.

Tabelle 1. Überblick über die Beiträge und ihre verschiedenen theoretischen und methodischen Zugänge sowie zentralen Ergebnisse (Stand: 06.05.2025)

Stichprobe	Fragestellung(en) des Beitrags (Gekürzt)	Bezug zur Fragestellung der Dissertation	Methodik	Theoriebezüge	Ergebnisse
1. Pargmann, J., Riebenbauer, E., Flick-Holtsch, D. & Berding, F. (2023). Digitalisation in accounting: a systematic literature review of activities and implications for competences. <i>Empirical Research in Vocational Education and Training</i> , 15(1). https://doi.org/10.1186/s40461-023-00141-1					
N=70 Beiträge	Welche Bereiche des Rechnungswesens und Controllings sind von Digitalisierung betroffen und welche Digitalisierungsgrade sind bei den Tätigkeiten vorzufinden?	Zentrale inhaltliche Digitalisierungsaspekte für den Wirtschaftsunterricht werden abgeleitet	Systematisches Literaturreview mit inhaltlich-strukturierender qualitativer Inhaltsanalyse	Stadien der Digitalisierung im Rechnungswesen und Accountings von Bleiber (2019), Stufen der Automatisierung nach Aepli et al. (2017).	Vor allem Nicht-Routinetätigkeiten werden sich verändern. Beliebte Technologien sind Automationen, KI und Big Data. Viele Betriebe sind noch in frühen Digitalisierungsstadien. Lernende benötigen Datenkompetenzen. Außerdem müssen sie unternehmerische (Wertschöpfungs-)Prozesse verstehen und ihre Tätigkeiten darin einordnen können, um Zusammenhänge zu durchdringen.

Stichprobe	Fragestellung(en) des Beitrags (Gekürzt)	Bezug zur Fragestellung der Dissertation	Methodik	Theoriebezüge	Ergebnisse
2. Pargmann, J. & Berding, F. (angenommen in <i>Frontiers in Education</i>). Sustainability content knowledge for teachers in accounting and controlling: A systematic literature review on industry approaches, practices and desired competences.					
N=79 Beiträge	Wie wird Nachhaltigkeit in Rechnungswesen- und Controllingliteratur diskutiert? Welche Werkzeuge, Prozesse und Praktiken sind relevant? Welche nachhaltigkeitsbezogenen Kompetenzen müssen befördert werden?	Zentrale inhaltliche Nachhaltigkeitsaspekte für den Wirtschaftsunterricht werden abgeleitet	Systematisches Literaturreview mit inhaltlich-strukturierender qualitativer Inhaltsanalyse nach PRISMA-Richtlinien	Kompetenzdimensionen nach Roth (1971), Domänenbezüge nach Winther und Achtenhagen (2008), Taxonomie der Geschäftsprozesse nach Gadatsch (2017) und Unternehmensbereiche nach Wöhe et al. (2020), Nachhaltige Geschäftsmodelle nach Perridon et al. (2017)	Nachhaltigkeit im Rechnungswesen und Controlling wird eher unstrukturiert integriert. Die benötigten Kompetenzen variieren stark. Unternehmen entwickeln ihre eigenen Reporting-Strukturen und Kennzahlen.

Stichprobe	Fragestellung(en) des Beitrags (Gekürzt)	Bezug zur Fragestellung der Dissertation	Methodik	Theoriebezüge	Ergebnisse
3. Pargmann, J., Berding, F., Rebmann, K. & Riebenbauer, E. (im Review bei <i>Empirical Research in Vocational Education and Training</i>). How AI feedback supports lesson planning in vocational teacher education: A longitudinal intervention study using an analytical AI platform.					
N=103 Studierende der Wirtschaftspädagogik (Master)	Wie unterscheiden sich Experimental- und Kontrollgruppe in der Umsetzung der didaktischen Kriterien? Wie beeinflusst die Nutzung der KI-Plattform die Motivation der Studierenden, ihre KI-Literacy, ihre Einstellungen zu KI, ihre digitalen Kompetenzen und ihre Selbstwirksamkeit mit digitalen Technologien (Metadimensionen)?	Empirische Überprüfung der Wirksamkeit des KI-Feedbacks der Plattform bei Einsatz in einem Didaktikseminar	Studie im Experimental-Kontrollgruppen-Design Auswertung der Unterrichtsentwürfe: Zweiebenen-Probit-Regression. Auswertung der latenten Konstrukte (Metadimensionen) Motivation, KI-Literacy, Selbstwirksamkeit mit Technologien, KI-Einstellungen und digitalen Kompetenzen mittels latenten Wachstumsmodells (Zwei Ebenen)	Unterrichtsplanungsmodelle wie Klafki (1963), Heimann et al. (1965), Schulz (1980) oder überblicksartig und domänenspezifisch Wilbers (2023). Erkenntnisse aus der Feedbackforschung, wie z.B. Narciss (2013), Prilop und Weber (2023) und digitales Feedback, wie Cai et al. (2023) oder Steiss et al. (2024)	Die Ergebnisse der Experimentalgruppe sind in einigen Dimensionen signifikant besser als die der Kontrollgruppe. Dies deutet darauf hin, dass das Feedback von der KI-Plattform mindestens genauso wirksam, wenn nicht wirksamer ist als menschliches Feedback. Die für die Metadimensionen verwendeten Messinstrumente haben sich als ungeeignet für den Längsschnitt herausgestellt. Daher sind Aussagen nur für 7 der 24 Konstrukte möglich. Die Intervention beeinflusst lediglich die Fähigkeiten zum Erkennen von KI und der Emotionsregulation positiv.

Stichprobe	Fragestellung(en) des Beitrags (Gekürzt)	Bezug zur Fragestellung der Dissertation	Methodik	Theoriebezüge	Ergebnisse
4. Pargmann, J., Berding, F., Riebenbauer, E. & Rebmann, K. (eingereicht im <i>Journal of New Approaches in Educational Research</i>). Engaging with AI: How student teachers interact with an analytical AI platform for feedback and the effect on their lesson plans – An eye tracking study.					
N=32 Studierende der Wirtschaftspädagogik (Master)	Welche Interaktionstypen mit der KI-Plattform können identifiziert werden? Wie verändern sich diese Interaktionstypen während des Semesters? Welcher Interaktionstyp produziert bessere Unterrichtsentwürfe?	Empirische Entwicklung der Nutzungstypen und deren Effekt auf die Integration planerischer Merkmale	Auswertung von Eye Tracking Daten mittels Sequenzanalyse (Gabadinho et al., 2011) und Probit Regression	Forschungsstand zur Expertiseentwicklung (Brückner & Zlatkin-Troitschanskaia, 2024; Gegenfurtner et al., 2023) und bisherigen Einsätzen von Eye Tracking Studien in der Lehrkräftebildung und der beruflichen Bildung (Eklund et al., 2020; Jarodzka et al., 2021)	Es können insgesamt drei Arten von Typen identifiziert werden: jene, die feedbackorientiert interagieren und vor allem die Empfehlungen zur Überarbeitung lesen, jene, die eher ergebnisorientiert interagieren und sich primär die Bewertung des aktuellen Stands anschauen sowie ein gemischter Typ, der anhand der Ergebnisse auswählt, welche Rückmeldungen angeschaut werden sollen. Der feedbackorientierte und der gemischte Typ wirken sich positiv auf die Umsetzung einiger planerischer Merkmale aus.

Im nachfolgenden Kapitel wird nun der theoretische und empirische Rahmen für den weiteren Verlauf dieser Arbeit gesetzt. Hierzu wird zunächst in Abschnitt 2.1 eine Einführung in die Thematik der Unterrichtsplanung gegeben.

2 Theoretische und empirische Rahmungen

2.1 Unterrichtsplanung

Eines der Kerngeschäfte von Lehrkräften ist das Unterrichten (Baumert & Kunter, 2006; Bromme, 1997; König & Rothland, 2022). Unabhängig vom lerntheoretischen Zugang und der Perspektive der Lehrkraft auf dem Spektrum der Wissensvermittlung bis hin zur Lernprozessbegleitung gilt es, Unterricht zu planen, durchzuführen, auszuwerten und zu überarbeiten (KMK, 2022). Sandfuchs (2004, S. 348) formuliert hierzu:

„Als Unterrichtsvorbereitung oder Unterrichtsplanung bezeichnen wir alle dem Unterricht vorausgehenden Maßnahmen, die Lehren und Lernen im Unterricht selber optimieren sollen. (...) Unterricht als zielgerechtes Lehren und Lernen ist ein hochkomplexer Prozess, zu dessen Gelingen Planung eine notwendige Voraussetzung ist.“

Da das Lehren als professionelle Handlung von Lehrkräften verstanden wird, ist das Erlernen der Gestaltung von Lehr-Lehrprozessen somit ebenfalls wichtiger Bestandteil ihres Professionalisierungsprozesses. Dies schlägt sich nieder in der UPK (König & Rothland, 2022). Obwohl der Begriff in der Literatur Verwendung findet, existiert weder eine einheitliche Konzeptualisierung noch eine einheitliche Begriffsverwendung (König & Rothland, 2022). Während Baumert und Kunter (2006) die UPK als Teilbereich der professionellen Kompetenz von Lehrkräften betrachten, sieht Rothland (2022) sie als eigenständige Kompetenz. Gemein haben die verschiedenen Modellierungsversuche eine Reihe an Kompetenzfacetten: fachliche, fachdidaktische sowie allgemein pädagogische Kompetenzen werden benötigt (Baumert & Kunter, 2006; Klusmeyer, 2021; Shulman, 1987; Söll, 2021; Werneke & Zierer, 2017). König und Rothland (2022, 780f.) schlagen als Ergebnis der Darstellung des aktuellen Forschungsstands zur UPK folgende Definition vor, die sich aus den verschiedenen Operationalisierungen der Datenbasis ergibt:

„Insbesondere mit Bezug auf das Expertise-Paradigma und den kompetenzorientierten Ansatz bezeichnet UPK erlernbare kognitive Fähig- und Fertigkeiten (*cognitive abilities and skills*) (Hervorhebung im Original) sowie damit verbundene affektiv-motivationale Dispositionen von Lehrkräften, die sich inhaltlich und funktional auf Anforderungen der Planung von Unterricht beziehen und die erfolgreiche Bewältigung solcher Anforderungen wahrscheinlich machen.“

Aus diesem Vorschlag ergibt sich eine Wechselwirkung zwischen den Voraussetzungen, unter denen ein Unterricht zu planen ist, und der Ableitung von Konsequenzen aus diesen. Eingearbeitet in das Modell der professionellen Kompetenz von Lehrkräften nach Stender et al. (2015) bedeutet dies, dass im Planungsprozess das Wissen der Lehrkraft in eine Sequenz aus Handlungen überführt wird. Dies geschieht, indem durch wiederholte Übung im Sinne des *Planen Lernens* Handlungswege verinnerlicht werden. Die dadurch beförderte professionelle Kompetenz bietet Sicherheit in zukünftigen Handlungen. Bewährte Konzepte werden in zukünftigen Handlungen berücksichtigt, nicht erfolgreiche Konzepte hingegen reflektiert und angepasst. Allerdings gibt das Modell keine Auskunft darüber, wie die Struktur dieser Handlungen modelliert werden kann (Pargmann et al., 2024). Hier bietet sich der strukturanalytische Ansatz von Oevermann (1996) an. Dieser benötigt wissenschaftliches Wissen, welches die inhaltliche Domäne der professionellen Kompetenz von Lehrkräften darstellt und ermöglicht, Situationen entsprechend in die Wissensbasis einzuordnen und zu deuten (Baumert & Kunter, 2006; Koehler & Mishra, 2009). Zur Erklärung dieses Ablaufs schlägt Oevermann (1996) den in Abbildung 3 dargestellten Mechanismus vor, welcher Aufschluss darüber gibt, wie professionelles Wissen in konkreten Situationen genutzt wird.

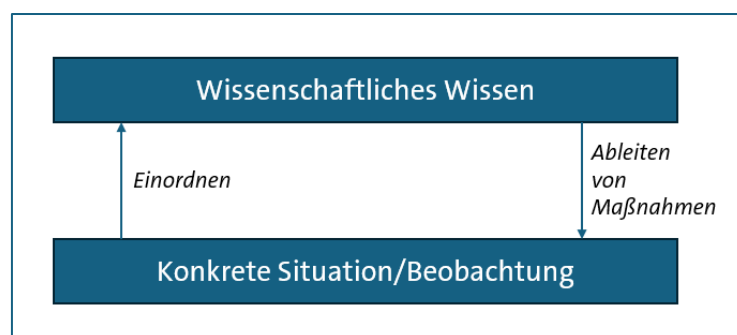


Abbildung 3. Nutzung professionellen Wissens nach Oevermann (1996)

Der Ablauf ist folgender: Konkrete Situationen oder Beobachtungen werden in wissenschaftliches Wissen eingeordnet, um Maßnahmen abzuleiten, die in konkreten Situationen genutzt werden können (Oevermann, 1996). So kann eine Lehrkraft eine Schülerin dabei beobachten, wie sie Zwischengespräche führt und dies in ihr professionelles Wissen zur Individualisierung einordnen. Sie zieht vielleicht den Schluss, dass die Schülerin bereits fertig mit der Aufgabe ist, oder aber sich ablenken möchte. Hierzu werden weitere Aspekte der Situation gedeutet und in Wissensbestände eingeordnet. Mit dieser Diagnostik kann eine Lehrkraft nun eine Maßnahme ableiten, wie Vertiefungsaufgaben bereitzustellen oder die Schülerin zu ermahnen. Bezogen auf die Unterrichtsplanung bedeutet dies, dass

bereits im Vorhinein die möglichen Handlungen im Unterricht antizipiert werden, Rückschlüsse gezogen werden und die Planung entsprechend angepasst wird. Eine solcher Kreislauf entwickelt sich jedoch nur durch wiederholte Auseinandersetzung mit verschiedenen Planungsdimensionen. Das wissenschaftliche Wissen wird somit aus zwei Perspektiven genutzt: Einerseits interventionspraktisch durch Rekonstruktion des Einzelfalls, sodass dieser subsumiert werden kann, andererseits ingenieural, d.h. zur Deduktion von Konsequenzen (Oevermann, 2008).

Ohne Zweifel ist hierfür auch eine praktische Komponente erforderlich, doch die gedankliche Durchdringung der Planungen soll die Verhaltenssicherheit der Lehrkraft erhöhen und somit im späteren Unterricht eine zielgerichtete Transformation in konkrete Handlungen und auch eine gewisse Flexibilität durch Handlungssicherheit herbeiführen (Riebenbauer et al., 2025). Die Unterrichtsplanung im Sinne des Übens im Professionalisierungsprozess soll eben diese Sicherheit in didaktischen Entscheidungen unterstützen. Die fachdidaktische universitäre Ausbildung ermöglicht hierfür eine intensive Auseinandersetzung mit wissenschaftlichen (in diesem Fall didaktischen und pädagogisch-psychologischen) Konzepten und die Erprobung des Treffens entsprechender Entscheidungen – entsprechend der Subsumtion nach Oevermann (1996). Eine solche Perspektive auf den Unterrichtsplanungsprozess verdeutlicht die Rolle wissenschaftlicher Konzepte für die Professionalisierung von (angehenden) Lehrkräften. Entlang des Professionalisierungsprozesses geschieht eine Transformation vom *Planen lernen* hin zum *Planen*, welche sich auch in der Tiefe der Elaboration unterscheiden (Riebenbauer et al., 2025) und eine Reflexion der Planungsentscheidungen im Sinne einer Bewusstmachung erfordern (Bromme & Seeger, 1979). Umgesetzt wird dieser Prozess zumeist mittels Unterrichtsentwürfen, welche das zentrale Medium zur Förderung der UPK in allen Phasen der Lehrkräftebildung sind (Reinisch, 2021; Söll & Klusmeyer, 2018; Stender et al., 2015).

Zur Ableitung und Begründung von Planungsentscheidungen steht eine Vielzahl (fach-)didaktischer Modelle und Konzepte zur Verfügung, die auf verschiedenen theoretischen Strömungen basieren. Von allgemeinen Modellen wie dem Hamburger Modell (Schulz, 1980), dem Berliner Modell (Heimann et al., 1965) oder der didaktischen Analyse nach Klafki (1963) hin zu domänenspezifischen Planungsmodellen wie dem Didaktikmodell von Wilbers (2023), stehen unterschiedliche Perspektiven auf den Unterricht und seine Planung zur Verfügung. Auch wenn die Spezifikationen der Bezugsdimensionen der Modelle variieren, so sind doch Überschneidungen identifizierbar (Schmaltz, 2019). Dies sind

Überlegungen zur Lerngruppe, zu den Lern- bzw. Kompetenzzielen, dem Inhalt, dem Verlauf des Unterrichts sowie zum Medien- und Methodeneinsatz. So wird überwiegend von inhaltlichen und lerngruppenspezifischen Voraussetzungen ausgehend ein zielorientierter Prozess geplant und mit adäquatem Medien- und Materialeinsatz unterstützt. Um also die Ableitung von Maßnahmen zu ermöglichen, gilt es für die planende Person, sich geeignete Wissensbestände für die einzelnen Planungsdimensionen anzueignen und – im Sinne einer andauernden Professionalisierung – diese regelmäßig zu reflektieren und zu aktualisieren. Planungsmodelle zu nutzen, bietet somit „erhebliche Orientierungspotenziale“ (Nickolaus, 2009, S. 55). Ein Ansatzpunkt zur Lösung ist, verschiedene Dimensionen zu kombinieren, denen empirisch eine hohe Wahrscheinlichkeit zugesprochen wird, zu einem hochwertigen Unterricht zu führen und damit auch die Chance auf einen Lernerfolg bei den Lernenden bieten (Riebenbauer et al., 2025).

In einem kürzlich erschienenen Artikel haben Riebenbauer et al. (2025) den aktuellen Stand der Forschung dargelegt und diskutiert, welche Planungsmerkmale berufsbildender Unterricht generell berücksichtigen kann, um zu hochwertigem Unterricht zu führen. Diese Merkmale gliedern sich in allgemeine Anforderungen an die Lernergebnisse, um aktuelle gesellschaftliche Transformationsprozesse bewältigen zu können, allgemeine Merkmale der Unterrichtsplanung sowie Merkmale der Gestaltung von Lehr-Lernprozessen¹:

- Merkmale der Lernergebnisse: Nachhaltigkeit (17 Nachhaltigkeitsziele und Integration von Nachhaltigkeit in die berufliche Handlungskompetenz) und Digitalisierung (inhaltlich und methodisch)
- Merkmale der Lehr-Lernprozessgestaltung: Handlungsorientierung, Problemorientierung, Prozessorientierung, Motivationsförderung
- Merkmale der Unterrichtsplanung: Strukturierung der Unterrichtseinheit, Aktivierung der Lernenden, inhaltliche Vernetzung

Die Dimensionen wurden mit dem Ziel ausgewählt, möglichst viele Facetten der etablierten Didaktikmodelle mit dem normativen Ziel der Beförderung der beruflichen Handlungskompetenz zu vereinen (Riebenbauer et al., 2025). Um in diesem Sinne eine langfristige Handlungsfähigkeit bei den Lernenden zu befördern, gilt es, auch aktuelle Entwicklungen und Transformationsprozesse zu integrieren. Daraus lässt sich die Notwendigkeit zur

¹ Für die detaillierte Herleitung der einzelnen Kriterien und Dimensionen sei an dieser Stelle auf den Beitrag von Riebenbauer et al. (2025) verwiesen.

Integration von Innovationen ableiten, welches dem Handlungsfeld *Innovieren* der KMK (2022) zugeordnet werden kann. Die Auseinandersetzung mit diesen Planungsprinzipien, didaktischen Grundsätzen und unterrichtlichen Anforderungen ist für (angehende) Lehrkräfte also auch notwendig, um eine „handlungsvorbereitende und reflektierende Wissensbasis zu schaffen“ (Söll & Klusmeyer, 2018, S. 75).

Die im Rahmen dieses Abschnitts beschriebenen Modelle und Dimensionen bilden insgesamt einen wesentlichen Rahmen für die Planung von Unterricht, insbesondere in der Lehrkräftebildung. Dennoch weist die empirische Berufsbildungsforschung im Bereich der Unterrichtsplanung und -gestaltung ein wesentliches Defizit auf. So ließen die Veröffentlichungszahlen zur Thematik über die letzten Jahrzehnte nach (Aprea, 2014, 2018; Haas, 2005; Koch-Priewe, 2000; Seifried, 2009; Söll & Klusmeyer, 2018; Stender et al., 2015; Wernke & Zierer, 2017; Zierer et al., 2015). Dies ist ein Indiz dafür, dass die UPK von nachrangiger Relevanz war (Reinisch, 2021). Erst in den letzten Jahren haben sich neue Publikationen zum Thema ergeben (Klusmeyer, 2021; Riebenbauer, 2021; Söll, 2021). Um für den weiteren Verlauf dieser Arbeit relevante Anknüpfungspunkte und Desiderata zu entwickeln, soll der aktuelle Forschungsstand im Folgenden aufgearbeitet werden. Einen Überblick über die ausgewählten Studien, die Methodik sowie zentrale Befunde bietet darüber hinaus Tabelle 2. Dabei wurden nur Studien aufgenommen, die relevant für die Fragestellungen dieser Arbeit sind. Für eine detaillierte Übersicht über die Beförderung der verschiedenen Teildimensionen der UPK sei auf Riebenbauer (2021) und König und Rothland (2022) verwiesen.

Tabelle 2. Synopse ausgewählter Studien zur Unterrichtsplanung

Studie	Stichprobe	Methodik	Ausgewählte Ergebnisse
Haas (2005)	N=39 Personen	Erhebung der alltäglichen Unterrichtsplanung mittels Beobachtung und der Methode des lauten Denkens	• Keine längeren Planungsentwürfe auf Makro- und Meseobene, abseits vom sog. „Stoffverteilungsplan“ • Klare Inhalts- und Zielbereiche, wobei die Ziele implizit berücksichtigt werden (Wissensebene), Schulbücher als Hauptmedium, Vorkenntnisse und Bedingungen der Lernenden wenig berücksichtigt • Wenig bis kein Bezug zu gezielten Planungsdimensionen
Seifried (2009)	N=225 Personen	Schriftliche Befragung, Konstruktinterviews mit n=21 Personen aus der Stichprobe	• Intensive Planungsaktivitäten vor allem bei Lehrkräften im Vorbereitungsdienst, etablierte Lehrkräfte nehmen eher Modifikationen vor • Lernziele gewinnen mit zunehmender Berufserfahrung an Bedeutung • Didaktische Modelle werden nur in Teildimensionen angewendet, Lehrkräfte verlassen sich eher auf ihre Planungsrouitinen
Stender (2014)	N=51 Studierende, N=48 Lehrkräfte im Vorbereitungsdienst, N=fertige Lehrkräfte	Erarbeitung des Transformationsmodells der Unterrichtsplanung und empirische Überprüfung dessen im Rahmen des Unterrichtsfachs Physik mittels Vignettentests (Querschnittuntersuchung)	• Mit steigender Erfahrung steigt auch die Qualität der Handlungsskripte im Unterricht, ergo die Überführung prozeduralen Wissens in konkrete Handlungen • Es existieren Zusammenhänge zwischen dem Professionswissen, der Motivation und den Handlungsskripten • Überführung des Professionswissens in konkrete unterrichtliche Handlung benötigt Motivation und die Fähigkeit zur Selbstregulation
Backfisch et al. (2020)	N=94 Personen	Effekte des professionellen Wissens und motivationaler Einstellungen auf die Fähigkeit, digitale Technologien in Unterrichts-entwürfe zu integrieren	• Die Einstellungen zur Nützlichkeit digitaler Technologien beeinflussen die Integration in den Unterrichtsentwurf • Das professionelle Wissen beeinflusst dies nicht. • Motivationale Aspekte sind relevanter als das professionelle Wissen.

Riebenbauer (2021)	N=131 Personen	Umsetzungen der unterrichtsplanerischen Anforderungen mit Verlaufsplan werden inhaltsanalytisch kodiert, diese beziehen sich auf die Phasen und die Elemente der Planungen; Längsschnittstudie	• Im Laufe des Studiums steigt der Umfang der Verlaufspläne • Inhalte nehmen den größten Teil der Pläne ein, gefolgt von den Methode, dann den Medien, den Sozialformen, den Zielen und zuletzt den Rahmenbedingungen des Unterrichts
Riebenbauer et al. (2025)	N=702 Unterrichtsentwürfe	Umsetzung unterrichtsplanerischer Anforderungen, didaktischer Kriterien und Transformationserfordernissen (Digitalisierung, Nachhaltigkeit) in Unterrichtsentwürfen und dazugehörigen Materialien von Wirtschaftspädagogikstudierenden mittels Inhaltsanalyse; Typenbildung mittels latenter Klassenanalyse	• Die meisten Unterrichtsentwürfe berücksichtigen die analysierten Planungsprinzipien nur geringfügig; einige Kategorien zu Transformationsaspekten werden gar nicht behandelt • Nur eine kleine Gruppe an Entwürfen erfüllt die Prinzipien in mittlerer – manchmal aber auch höherer Ausprägung • Die Umsetzung der Planungen in den Unterrichtsmaterialien gelingt nicht immer oder die Umsetzung wird in den Plänen nicht entsprechend ausformuliert • Insgesamt sprechen Ergebnisse für eher gering ausgeprägte Professionalität bei der Planung von Unterricht; Anwendung im Studium scheint nicht zu gelingen; (Selbst-)Reflexionsanlässe sind notwendig

Aus diesen Studien ergibt sich, dass der Inhalt der Unterrichtsstunde für (angehende) Lehrkräfte ein größeres Gewicht einnimmt als die Eruerung des konkreten Unterrichtsziels. Erst nach Abklärung der Inhalte werden dann Überlegungen zu anderen Dimensionen getätigt, sie werden somit priorisiert. Um dennoch die Chance auf hochwertigen Unterricht zu realisieren, kann es sinnvoll sein, eine gewisse Planungsroutine (i. S. v. Handlungsskripten) zu entwickeln, in der geübt wird, didaktische Entscheidungen kontinuierlich zu hinterfragen und zu überarbeiten.

Insgesamt scheinen Unterrichtsentwürfe ein adäquates Mittel zu sein, um die Dimensionen und Interdependenzen von Unterrichtsplanung zu verdeutlichen. Während die Nutzung eines spezifischen Modells weniger relevant zu sein scheint, ähneln sich die Dimensionen, die konkret bei der Planung berücksichtigt werden, sowohl über verschiedene Herangehensweisen als auch über die empirischen Ergebnisse hinweg.

Wie in der Studie von Riebenbauer et al. (2025) deutlich wurde, existieren Herausforderungen im Planungsprozess. An dieser Stelle können Technologien auf Basis von KI helfen, zur Professionalisierung angehender Lehrkräfte beizutragen. Daher erfolgt nun eine Einführung in den Gegenstand der KI.

2.2 Künstliche Intelligenz als Gegenstandsbereich

Um die Diskussion um den Begriff der *künstlichen* Intelligenz nachzuvollziehen, bedarf es zunächst einer Annäherung an den Begriff der *Intelligenz* an sich, auch wenn beide eng miteinander verwoben sind (Buxmann & Schmidt, 2021). Görz et al. (2021, S. 6) beschreiben Intelligenz „als Erkenntnisvermögen, als Urteilsfähigkeit, als das Erfassen von Möglichkeiten, aber auch als das Vermögen, Zusammenhänge zu begreifen und Einsichten zu gewinnen“. Aufbauend auf dieser konzentrierten Sichtweise beschreibt McCarthy (2004, S. 2) KI als „science and engineering of making intelligent machines, especially intelligent computer programs“, wobei Methoden verwendet werden, die nicht unbedingt menschlichen Verhaltensweisen entsprechen müssen. Kleesiek et al. (2020, S. 24) fügen hinzu, dass KI vor allem die „Simulation intelligenten Verhaltens in Computern“ umfasst. Diese Nachahmung wird effektiv durch maschinelles Lernen umgesetzt, das Lernprozesse basierend auf Datensätzen beschreibt, welche es – stark verkürzt – wiederum ermöglichen, Prognosen zu erstellen (Buxmann & Schmidt, 2021). Forschungen zu KI sind nicht neu, das Forschungsfeld existiert bereits seit den 1940er Jahren (Haenlein & Kaplan, 2019). Aufgrund limitierter technischer Kapazitäten bis in die 1980er Jahre eher theoretischer Natur, entwickelte sich

mit dem technologischen Fortschritt auch die KI-Forschung ab den 1990er Jahren mit wachsender Schnelligkeit.

Arten von KI-Anwendungen

Kategorisch existieren drei Ebenen von KI: die schwache (narrow AI), starke (general AI) und die Superintelligenz (Massmann & Hofstetter, 2020). Schwache KI (narrow AI) ist für einen spezifischen Anwendungsbereich konzipiert und kann allein in diesem Bereich menschliche Intelligenz erreichen oder übertreffen. In der beruflichen Bildung kann dies z.B. ein intelligentes Tutoringsystem sein. Solch ein System kann Aufgaben auf individuelle Schwächen der Lernenden zuschneiden und gezielte Übungen oder Erklärungen anbieten.

Auf der nächsten Ebene befindet sich die starke KI (general AI), die mindestens ebenso intelligent ist wie ein Mensch und in der Lage ist, eine Vielzahl von Aufgaben genauso gut zu erfüllen (Massmann & Hofstetter, 2020). Dies wäre z.B. ein persönlicher, virtueller Lernassistent, der ähnlich wie ein menschlicher Lerncoach fungiert. Dieser Assistent kann Feedback geben oder Fragen zu einer Vielzahl von Themen beantworten, den Lernfortschritt der Lernenden überwachen und auf individuelle Probleme eingehen. Aktuell ist die Forschung noch nicht auf diesem Niveau angekommen, auch bestehende Feedbackroboter greifen noch auf eine begrenzte Datenbasis zurück und sind nur für enge Anwendungskontexte einsetzbar (siehe hierzu z.B. Donnermann et al., 2022).

Die höchste Ebene wird von der künstlichen Superintelligenz eingenommen, die der menschlichen Intelligenz bei weitem überlegen ist (Massmann & Hofstetter, 2020). Diese Ebene ist aktuell rein theoretisch. Im Anwendungskontext wären z.B. Konzepte denkbar, in denen die pädagogische Diagnostik, Lernstandserhebungen, didaktische Gestaltung von Unterrichtseinheiten verschiedenster Fächer und Lernfelder von Systemen übernommen werden, die Lernende ihr Leben lang begleiten und die Lernerfahrungen vollständig personalisieren. Während schwache KI bereits aktiv in der Bildung eingesetzt wird, bleibt die starke KI aktuell noch eine Zielvision und die Superintelligenz eine theoretische Möglichkeit der Zukunft.

Funktionsweise von KI-Anwendungen

KI-Systeme basieren auf maschinellem Lernen, um Vorhersagen machen zu können. Maschinelles Lernen bezeichnet Algorithmen, die in unstrukturierten Datensätzen (z.B. Texte wie Unterrichtsentwürfe oder Handlungsprodukte von Lernenden) Muster erkennen und

anhand dieser Muster dann wiederum Kategorien zuordnen und Voraussagen ableiten (Aggarwal, 2018; Ertel, 2016; Moring, 2023). Es gibt drei primäre Arten des maschinellen Lernens (Abbildung 4), die im Kontext dieser Arbeit relevant sind: das überwachte Lernen, das unüberwachte Lernen und das verstärkende Lernen. Jede dieser Arten soll im Folgenden kurz definiert und für den Bereich der beruflichen Bildung eingeordnet werden.

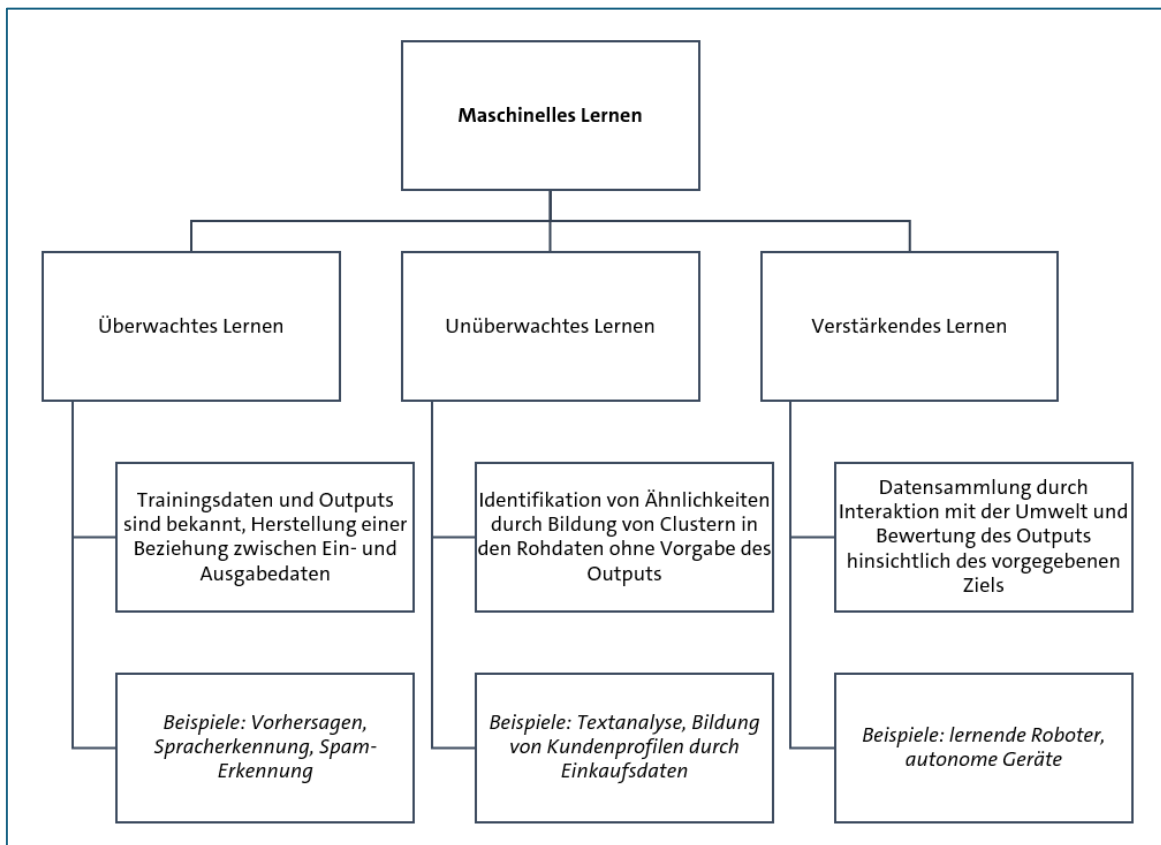


Abbildung 4. Arten des maschinellen Lernens im Überblick (basierend auf Kleesiek et al., 2020, Russel & Norvig, 2021 und Welsch et al., 2018)

Das überwachte Lernen charakterisiert sich dadurch, dass sowohl die Trainingsdaten als auch der gewünschte Output vorgegeben werden (Russel & Norvig, 2021; Welsch et al., 2018). Die KI stellt also Beziehungen zwischen den Ein- und Ausgabedaten her und erkennt somit Muster. Diese Logik wird z.B. bei der Erstellung von Marktvorhersagen auf Basis vergangener Ereignisse getätigt, denn die Algorithmen erkennen Wiederholungen, respektive Muster, und erkennen den Zeitpunkt, wann ein Szenario eintritt, also das gewünschte Ausgabeergebnis erreicht wird. In der beruflichen Bildung kann überwachtes maschinelles Lernen z.B. dazu genutzt werden, um den Lernprozess von Schüler:innen vorherzusagen. So können historische Daten über Prüfungsergebnisse oder die Mitarbeit genutzt werden, um ein Modell zu trainieren, welches wahrscheinliche Schwächen identifiziert, indem Muster in bisherigen Leistungen erkannt werden.

Beim unüberwachten Lernen werden Trainingsdaten eingegeben, aber keine gewünschten Outputs festgelegt (Welsch et al., 2018). So werden Muster erkannt, die dann in Outputs synthetisiert werden, die nicht vorgegeben sind. Es werden sogenannte Cluster gebildet, wie bei der Erstellung von Kundenprofilen und der Erstellung personalisierter Rabattcoupons auf Basis des vorherigen Kaufverhaltens. In der beruflichen Bildung können Clusteranalysen dazu genutzt werden, um ähnliche Lernverhaltensweisen zu erkennen und daraufhin personalisierte Lernempfehlungen oder Lernpläne zu erstellen.

KI-Systeme auf Basis des verstärkenden Lernens werden nicht mittels eines klassischen Trainingsdatensatzes trainiert, sondern ein Agent, z.B. ein Roboter, sammelt Umweltdaten und lernt aus der Interaktion mit ihr, während ein bestimmtes Ziel verfolgt wird, z.B. bei selbstfahrenden Kraftfahrzeugen. Der Agent vergleicht, ob das gewünschte Ergebnis (z.B. ohne Unfall von A nach B zu kommen) erreicht wurde und bewertet dann die Vorgehensweise (Russel & Norvig, 2021). Im Bildungsbereich kann ein solcher Agent z.B. ebenfalls ein Roboter sein, der verschiedene Lernstrategien bei den Schüler:innen befördert und die am besten funktionierenden Strategien dann für weitere Instruktionsprozesse beibehält.

Die Rolle der Verarbeitung natürlicher Sprache

Auch wenn die Forschungen zur künstlichen Intelligenz bereits seit Mitte des letzten Jahrhunderts andauern und in verschiedensten anderen Bereichen KI-Tools erfolgreich implementiert werden (Ertel, 2025), kann die Technologie erst seit kurzem für den Lehrberuf nutzbar gemacht werden. Da in diesem Bereich häufig mit Texten gearbeitet wird, muss die hierfür genutzte Sprache erst für den Algorithmus verständlich werden. Dafür wird die natürliche Sprachverarbeitung (engl. Natural Language Processing oder NLP) genutzt (Chowdhary, 2020; Lane et al., 2019). Hierfür müssen Texte kontextualisiert werden, wofür technisch die sogenannte Self-Attention notwendig ist (Vaswani et al., 2017). Dies ist ein Mechanismus, der es einem KI-Modell ermöglicht zu gewichten, wie wichtig die einzelnen Elemente der Eingabesequenz für die gesamte Eingabe sind. Dadurch können Kontexte erkannt und extrahiert werden. Im NLP-Bereich dominieren aktuell sogenannte Transformermodelle als präferierte Architektur (Tunstall et al., 2022). Im Gegensatz zu früheren Modellen verarbeiten Transformermodelle alle Eingabedaten gleichzeitig, wodurch die Verarbeitungsgeschwindigkeit verbessert wird (Vaswani et al., 2017). So können sie Kontexte und Bedeutungen besser erfassen als vorherige Architekturen. Ein mögliches Endprodukt ist ein vortrainiertes Sprachmodell, wie z.B. BERT, gBERTbase oder RoBERTA (Devlin et al., 2018;

Tunstall et al., 2022). Auch die zuletzt populär gewordenen GPT-Modelle von OpenAI (2023) gehören zu dieser Kategorie. Aufgrund der Novität der Technologie ergeben sich auch besondere Chancen und Grenzen, die es zu betrachten gilt, wenn der Einsatz im Bildungsbereich in Erwägung gezogen wird.

Chancen und Grenzen von KI-Anwendungen im Allgemeinen

Eine der größten Chancen liegt in der Fähigkeit von KI-Systemen, große Datenmengen effizient zu analysieren und nutzbare Informationen daraus zu extrahieren (Ifenthaler & Yau, 2021). Die darüber eröffneten Möglichkeiten zur Verbesserung von Entscheidungsprozessen und zur Entdeckung neuer Muster und Zusammenhänge sind ansonsten nur mit großem Ressourcenaufwand realisierbar. Hierfür wird z.B. Learning Analytics genutzt, eine Technologie, mit der lernprozessbezogene Daten gesammelt, analysiert und genutzt werden, um Lehr-Lernprozesse zu individualisieren, auszuwerten und zu optimieren (Gašević et al., 2017). Lehr-Lernprozesse werden somit mittels digitaler Technologien strukturiert und der Lehrperson stehen aggregierte Nutzungsdaten zur Verfügung, um Rückschlüsse für die weitere Prozessgestaltung ziehen zu können (Ifenthaler & Yau, 2021; Vincent-Lancrin & van der Vlies, 2020). Durch KI-Systeme können außerdem repetitive, zeitintensive Aufgaben automatisiert werden, wodurch Routinetätigkeiten minimiert und der Freiraum für komplexere und kreativere Tätigkeiten vergrößert wird (Owan et al., 2023). So können Lehrkräfte z.B. administrative Aufgaben zumindest teilweise auslagern und so Zeit für die Unterrichtsentwicklung freisetzen. Darüber hinaus ermöglichen KI-Systeme personalisierte Lehr-Lernprozesse, indem sie sich an die individuellen Gegebenheiten einer Person anpassen (Wirzberger & Schwarz, 2021). Im Bildungsbereich etwa können Lerninhalte auf den Kenntnisstand und Lernstil einzelner Lernender zugeschnitten werden.

In der Diskussion um Grenzen von KI-Anwendungen wird häufig das sogenannte Bias thematisiert, das Girasa (2020) als vorgefasste Meinungen oder Einstellungen gegenüber einzelnen Personen oder Gruppen beschreibt. Diese Verzerrungen können sowohl positiv als auch negativ die Wahrnehmung der betroffenen Personen beeinflussen und bergen daher Diskriminierungspotenziale. Auch die Verknüpfung von Namen mit bestimmten Emotionen oder Verhaltensformen stellen eine Grenze dar (Kiritchenko & Mohammad, 2018). Die Algorithmen können problematische Muster und Verknüpfungen erlernen, wenn die zugrunde liegenden Daten solche Verzerrungen enthalten (Ngige et al., 2021). Ein weiteres Problem, das sich auch bei Transformermodellen zeigt, ist die Blackbox-Natur dieser

Technologien (Klugman, 2021; Simbeck et al., 2019). Obwohl der Self-Attention Mechanismus es durch die Gewichtung der Eingabeelemente transparenter ermöglicht, die Ergebnisse nachzuvollziehen (Vaswani et al., 2017), erfordert dies ein umfassendes Verständnis für die Funktionsweise der Modelle. Die größte Herausforderung liegt jedoch in der benötigten Datenmenge für das Training der KI. Speziell in der Erziehungswissenschaft liegen oft nur kleine, unausgewogene Datensätze vor (Stütz et al., 2022). Seltener Muster können damit schneller übersehen werden (Haixiang et al., 2017).

Die Nutzung von KI-Systemen ist somit geprägt von verschiedenen Chancen und Grenzen. In Ergänzung zu diesen allgemeinen Ausführungen gibt es auch Aspekte, die charakteristisch für die verschiedenen Arten von KI-Systemen sind, in diesem Fall differenziert zwischen generativer und analytischer KI.

Chancen und Grenzen generativer KI-Anwendungen

Die Chancen generativer KI liegen primär in der Individualisierbarkeit der Ergebnisse (Adiguzel et al., 2023; Pargmann & Berding, 2022). Lehrende können z.B. Inspiration für die Unterrichtsentwicklung erhalten oder zielgruppengerechtes, differenziertes Material erstellen lassen. Lernende können durch generative KI-Systeme individualisiertes Feedback erhalten und sich bei der Erstellung von z.B. Grafiken, Präsentationen oder Bildern kreativ verwirklichen. Aufgrund der Funktionsweise generativer KI-Systeme ist jedoch ein kompetenter Umgang mit der Technologie zwingend notwendig (Long & Magerko, 2020). Denn sogenannte Halluzinationen erzeugen Falschaussagen oder Fehler, die kritisch reflektiert werden müssen (Sun et al., 2024). Außerdem benötigen generative KI-Systeme immense Datenmengen, um trainiert zu werden und qualitativ zufriedenstellenden Output zu generieren (Feuerriegel et al., 2024). Hierfür werden mitunter auch Urheberrechte missachtet, sodass die Nutzung von KI-generierten Ergebnissen, egal ob schriftlich oder grafisch, rechtliche Grauzonen tangiert (Dornis, 2025; Konertz & Schönhof, 2024). Darüber hinaus bergen generative KI-Systeme auch ökologische Herausforderungen (Khowaja et al., 2024). So haben z.B. Luccioni et al. (2024) gezeigt, dass Mehrzweckanwendungen wie ChatGPT mehr Energie verbrauchen als spezialisierte Modelle, selbst wenn die Anzahl der zu analysierenden Parameter begrenzt ist.

Chancen und Grenzen analytischer KI-Anwendungen

Obwohl aktuell generative KI den öffentlichen und Forschungsdiskurs dominiert, existieren darüber hinaus noch analytische KI-Systeme, welche z.B. in Form von Expertensystem

schon länger verfügbar sind (Beheshti, 2023). Diese erkennen Muster in Datensätzen und erstellen dann Bewertungen, überprüfen die Erfüllung von Kriterien oder prognostizieren Ereignisse. Diese Systeme basieren ebenfalls auf maschinellem Lernen, dabei greifen sie auf statistische Analyseverfahren zurück. Der Vorteil analytischer KI-Systeme liegt vor allem darin, dass die Analyseergebnisse reproduzierbar sind, da kein Zufall eingestreut werden muss (Yang & Zhu, 2024). Dadurch steigt die Verlässlichkeit der Ergebnisse. Auch das Risiko für Halluzinationen sinkt, da keine neuen Daten generiert werden. Um jedoch aussagekräftige Ergebnisse zu erzeugen, benötigen sie große Datenmengen, um Muster zu erkennen (Šíma, 1995; Yoon et al., 1994). Durch die Erkennung von Mustern, wie bestimmten Präferenzen im Material, sind außerdem personalisierte Empfehlungen möglich (Hui & Xiao, 2023). Allerdings sind sie oft nur für einen engen Anwendungsbereich brauchbar (Yang & Zhu, 2024). Im Vergleich zu generativen KI-Systemen verbrauchen analytische Systeme weniger Energie, da sie mit reduziertem Rechen- und Datenaufwand betrieben werden können.

Es zeichnet sich ab, dass die Nutzung von KI im Bildungsbereich, egal ob als Lerngegenstand oder als Methode, einer durchdachten Strategie bedarf und dass auch auf den Bildungsbereich spezialisierte Anwendungen ihre Berechtigung haben. Daraus ergeben sich spezifische Anforderungen für KI-Systeme im Bildungsbereich.

Anforderungen an KI-Systeme im Bildungsbereich

Auf den Bildungsbereich angewendet, ergeben sich für Luan et al. (2020) sechs kritische Stellschrauben von KI, die vor allem aufgrund der heterogenen Lerngruppen und der personenbezogenen Daten zu durchdenken sind:

1. Die Bildung entwickelt sich hin zu personalisierten Lehr-Lernansätzen, die die individuellen Voraussetzungen sowie Stärken und Schwächen der Lernenden berücksichtigen sollen.
2. Der Einsatz von KI sollte aus verschiedenen Perspektiven betrachtet und interdisziplinär geplant werden.
3. Die Auswahl der für KI erhobenen Daten muss sorgfältig erfolgen, was ein durchdachtes Design der Systeme erfordert. Zudem sollten KI-Systeme Kontextfaktoren wie den kulturellen Hintergrund der Lernenden einbeziehen.
4. Ethische Fragen müssen behandelt werden, um Probleme wie Verzerrungen und Diskriminierung zu vermeiden.

5. Lehr-Lerntheorien müssen erweitert werden, um die digitale Dimension zu integrieren, damit Lernende aktive Gestalter:innen ihrer Bildungserfahrung werden und nicht nur passive Konsumenten von KI-Systemen sind.
6. Der Schutz sensibler personenbezogener Daten, die bei der Nutzung von KI-Systemen gesammelt, verarbeitet und gespeichert werden, muss unbedingt gewährleistet sein.

Darüber hinaus gilt es in der beruflichen Bildung immer auch, die Berufswelt zu betrachten. Allerdings ist die Domänenspezifität in den gängigsten KI-Tools nicht vorhanden (Yue Liu et al., 2023). Für die Hochschulbildung ergänzen Zawacki-Richter et al. (2019) vier Nutzungsdimensionen von KI: a) die Profilerstellung und Vorhersage, b) die Beurteilung und Bewertung, c) adaptive Systeme und Personalisierung und d) intelligente tutorielle Systeme. Je nach Dimension ergeben sich daraus unterschiedliche Implikationen, wenn sie gemeinsam mit den Stellschrauben von KI im Bildungsbereich (Luan et al., 2020) gedacht werden. So ist in allen Dimensionen auf die ethischen Restriktionen zu achten, da es gerade bei der Profilerstellung in Dimension a) oder bei der Beurteilung in Dimension b) zu diskriminierenden Situationen kommen kann. Doch die Nutzung von KI-Systemen für die Personalisierung von Lernmöglichkeiten bietet eine weitreichende Chance, die sehr standardisierte Hochschulbildung für verschiedene Personengruppen zugänglicher zu machen, indem z.B. Sachverhalte in einfacherer Sprache wiedergegeben werden, Alternativtexte und Audiodateien für Menschen mit Behinderungen zur Verfügung gestellt werden oder auch asynchrone Lerngelegenheiten für Menschen nutzbar gemacht werden, die aus verschiedenen Gründen die Präsenzangebote nicht wahrnehmen können.

Daraus ergeben sich im Umkehrschluss einige zentrale Anforderungen an KI-Systeme im Bildungsbereich. Diese sollten auf die Bedürfnisse der Lernenden zugeschnitten sein und ein gewisses Personalisierungspotenzial besitzen (Luckin et al., 2016). Darüber hinaus sind KI-Systeme mit Bedacht so zu entwickeln, dass nur wirklich notwendige Daten erhoben und gespeichert werden. Lernenden sollte immer auch die Möglichkeit eingeräumt werden, einer Nutzung zu widersprechen, idealerweise in Form einer informierten Zustimmung (Mah et al., 2023; Rodríguez-Barroso et al., 2020). Hier gilt es auch die aktuelle Legislative zu berücksichtigen; der kürzlich erschienene EU AI Act bildet hier eine weitere Regulationsebene neben der Datenschutz-Grundverordnung (DSGVO).

KI-Kompetenzen als Teil der professionellen Kompetenz von Lehrkräften

Die plötzliche Verfügbarkeit von KI-Systemen macht die Förderung digitaler Kompetenzen und speziell einer KI-Literacy bei Lehrenden und Lernenden notwendig, damit diese den Erfordernissen einer digitalen Transformation gewachsen sind (Ng et al., 2021; Zhang & Zhang, 2024). Long und Magerko (2020, S. 2) definieren KI-Literacy als „a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace“. Dem von ihnen vorgeschlagenen Framework fehlt es jedoch an Domänenspezifität (Knoth et al., 2024). Ähnliche Bedarfe wurden schon im Zuge der digitalen Transformation deutlich (Gerholz et al., 2022). Somit bedarf es einer methodisch-didaktischen und einer inhaltlich-curricularen Betrachtung dieser Perspektiven (Freiling & Porath, 2020; Happ et al., 2024). So ist speziell für die berufliche Bildung die Integration berufsweltlicher Veränderungen in die Ausbildung zu nennen, ergo die Veränderungen, mit denen Auszubildende und Fachkräfte in der Arbeitswelt konfrontiert sind (Meiners et al., 2022). Daraus ergibt sich auch für die Lehrkräfte in der beruflichen Bildung ein konkreter Weiterbildungsbedarf: Es gilt festzustellen, welche Möglichkeiten und Grenzen der Einsatz von KI in der Lehrkräftebildung einerseits, im Unterricht andererseits bietet.

Lernförderlichkeit von KI-Anwendungen

Ein zentrales Argument zur Nutzung von KI in Bildungsprozessen ist die gesteigerte Lernförderlichkeit. Lernen kann nach der kognitiven Entwicklungstheorie von Piaget (1977) als Prozess verstanden werden, bei dem kognitive Schemata entweder durch Assimilation – die Integration neuer Informationen in vorhandene Strukturen – oder durch Akkommodation – die Anpassung bestehender Schemata zur Aufnahme neuer Informationen – verändert werden. Konkret wird dieser Prozess durch kognitive Konflikte oder Perturbationen ausgelöst, sodass neues Wissen konstruiert wird, indem Informationen entsprechend verarbeitet werden (Dörner, 1979). KI-Anwendungen, die in diesem Sinne lernförderlich sind, lösen Perturbationen aus, die Lernende dazu anregen, mentale Modelle weiterzuentwickeln.

Übersichtsarbeiten zur Lernförderlichkeit von digitalen Technologien im Allgemeinen legen nahe, dass diese einen durchaus signifikanten Effekt haben können (Bernard et al., 2009; Manaff & Halabi Azahari, 2024; Schmid et al., 2009; Sormunen et al., 2021; Tamim et al., 2011; van der Kleij et al., 2015). Hierbei gilt es jedoch zu beachten, dass der Einsatz didaktisch und curricular begründet stattfinden sollte, auch in der Hochschulbildung

(Kelland & Kanuka, 2008). Außerdem müssen sowohl Lehrende als auch Lernende über die notwendigen Kompetenzen verfügen oder diese müssen dediziert befördert werden (Richter et al., 2024). Um einen Eindruck davon zu erhalten, ob der Einsatz von KI ähnliche Effekte haben kann, wie die Nutzung digitaler Technologien im Allgemeinen, wird in Tabelle 3 der aktuelle Forschungsstand zur Lernförderlichkeit von KI dargestellt und anschließend kurz eingeordnet.

Tabelle 3. Synopse ausgewählter Studien zu KI und Lernförderlichkeit

Studie	Stichprobe	Methodik	Ausgewählte Ergebnisse
Celik et al. (2022)	Nicht angegeben	Systematisches Literaturreview von 37 Studien zur Frage, wie KI-Tools die 21st Century Skills befördern können	- Folgende 21st Century Skills können durch die Nutzung von KI angesprochen werden: Kommunikation, Kollaboration, Kreativität, IT-Fähigkeiten, Kritisches Denken, Problemlösen - Vor allem intelligente Tutor-Systeme scheinen durch ihr individuelles Feedback zu diesen Fähigkeiten beizutragen
Gerlich (2025)	N=666 Teilstichprobe von n=50 für Interviews	Betrachtung des Zusammenhangs zwischen der Nutzung von KI-Systemen und den Fähigkeiten des kritischen Denkens in einem Mixed-Methods-Design; Auswertung mittels ANOVA und Korrelationsanalyse und Regressionsanalyse, sowie qualitativer Themenanalyse von Interviews mit Teilstichprobe	- Negative Korrelation zwischen häufiger KI-Nutzung und den Fähigkeiten des kritischen Denkens durch sogenanntes <i>Cognitive Offloading</i>, also der Abwälzung kognitiven Aufwands auf die KI, ohne das Ergebnis zu reflektieren - Jüngere Teilnehmende verlassen sich mehr und unreflektierter auf KI-Ergebnisse und zeigen niedrigere Fähigkeiten im kritischen Denken - Höhe Bildungsabschlüsse wirken sich positiv auf die Fähigkeit kritischen Denkens aus, unabhängig von der KI-Nutzung - Nutzung von KI-Systemen kann intensive kognitive Anstrengung verhindern
Hangen und Wuttke (2024)	N=83 in der Fragebogenstudie N=32 in der	Wirksamkeit eines 9-wöchigen Online-Trainings zur Förderung des KI-Wissens, der KI-Einstellungen sowie Erhebung von Veränderungen in der Motivation durch das Training; Auswertung mittels Fragebogenstudie	- Keine signifikante Wissensveränderung über Anwendungsmöglichkeiten von KI, aber das selbstberichtete Wissen über Unternehmensgründungen nimmt zu - Intrinsische Motivation sinkt leicht - Programm effektiv bei der Vermittlung von Start-Up-bezogenem Wissen, aber nicht bei KI-Wissen

	Inter- viewstu- die	und Interviewstudie mit einer Teil- stichprobe	
Jia und Tu (2024)	N=637	Fragebogenstudie zum Selbstbericht der KI-Fähigkeiten, der Lernmotivation, Auswertung via explorativer Faktoranalyse, konfirmatorischer Faktoranalyse sowie eines Strukturgleichungsmodells	• KI-Fähigkeiten an sich haben keinen Effekt auf das kritische Denken • KI-Fähigkeiten können den kognitiven Wissenserwerb verändern, indem die Fähigkeit zum kritischen Denken gezielt gesteigert wird • Dies geschieht durch eine Steigerung der Selbstwirksamkeit und der Lernmotivation
Lin et al. (2021)	N=91	18-Wochen-Kurs zu Artificial Intelligence of Things (Verknüpfung intelligenter Gegenstände) unter Nutzung von Augmented Reality, Auswertung mittels eines Least Squares Strukturgleichungsmodells	• Die Anwendung hat die Performanzebene des rechnerischen Denkens (computational thinking) positiv beeinflusst, aber nicht die Kognitionsebene • Programmierfähigkeiten werden nicht positiv beeinflusst, hierfür werden trotzdem Kurse benötigt • Die kognitive Auslastung durch die Technologie wirkte sich nicht negativ aus
Riski und Nuryanto (2024)	N=144	KI-basierte Lerninhalte werden genutzt, um den Effekt auf den Lernerfolg zu betrachten; quasi-experimentelles Design; Auswertung mittels ANOVA und T-Tests; Begleiterhebung für metakognitive Aspekte und Motivation	• Lernende der Interventionsgruppe schnitten signifikant besser ab beim Post-Test als Lernende der Kontrollgruppe • Ergebnisse der Begleitstudie ergaben, dass die Interventionsgruppe dem Lernen mit KI gegenüber positiver eingestellt sind, höhere Werte bei der Metakognition (z.B. Bewusstsein und Steuerung des Denkprozesses) erreichten und sich überwiegend motivierter fühlten

Roldán-Álvarez und Mesa (2024)		Entwicklung eines intelligenten Tutor-Systems für individuelle Unterstützung; Auswertung der Testergebnisse am Ende des Jahres und Abgleich auf statistische Unterschiede zu den vorangegangenen Jahren ohne Tutor-System	Durch die Einführung des intelligenten Tutor-Systems konnten signifikant positive Ergebnisse bei den Lernergebnissen der teilnehmenden Studierenden erreicht werden. Diese waren höher als in den Jahren zuvor, in denen es noch kein intelligentes Tutor-System gab
Schmidt (2024)	N=153	Entwicklung eines Messinstruments zum Grundlagenwissen von KI von angehenden Lehrkräften; Studie zur Förderung von kognitiven und non-kognitiven Kompetenzfacetten mittels Selbstbericht im Pre-Post-Design	- Kognitive Kompetenzfacetten zu KI sind bei den angehenden Lehrkräften schwach ausgeprägt - Non-kognitive Kompetenzfacetten sind eher positiv, aber die Stichprobe ist den Entwicklungen von KI gegenüber kritisch - Im Pre-Post-Test ist ein leichter Zuwachs des Grundlagenwissens erkennbar - Bei den non-kognitiven Facetten liegen keine signifikanten Veränderungen vor
L. Wang und Zhao (2024)	Nicht angegeben	Metaanalyse mit 50 Studien von 2010 bis 2023	- Die Nutzung von KI-Anwendungen hat einen starken signifikanten Effekt auf den Lernerfolg ($g=0,825$), der Effekt auf die Kognition ($g=0,907$) ist stärker als auf die Non-Kognition ($g=0,774$) - Kein signifikanter Unterschied in den Effekten verschiedener KI-Anwendungen auf den Lernerfolg (z.B. Agentensysteme/Roboter mit $g=0,886$ oder intelligente Lernsysteme mit $g=0,942$) - Signifikante Unterschiede in den Effektstärken verschiedener Bildungsbereiche bestätigt; Effekte in der Vorschule ($g=0,989$) und an der Universität ($g=1,021$) stärker als in der Grund- ($g=0,677$) und Oberschule ($g=0,634$)
Zheng et al. (2023)	N=2908	Metaanalyse mit 24 Studien von 2001-2020	- KI hat einen starken Effekt ($g=1,022$ in der Mittelstufe, $g=0,955$ in der höheren Bildung) auf den Lernerfolg ab der Mittelstufe, aber nur einen kleinen auf die Wahrnehmung des Lernerfolgs - Problembasiertes Lernen ($g=1,044$) und situiertes Lernen ($g=1,071$) mit KI weisen starke Effekte auf, diese sind nicht signifikant

			• KI als intelligenter Tutor ($g=0,639$) und als intelligentes Lernwerkzeug ($g=0,809$) sind nicht so effektive Rollen von KI wie die KI als Berater für Rahmenbedingungen ($g=2,875$) • KI als Evaluationstool ($g=0,836$), für personalisierte Empfehlungen ($g=1,084$) und als Tutor ($g=0,935$) sind die effektivsten KI-Anwendungen, diese sind nicht signifikant • Systeme mit virtuellen Assistenzen sind am effektivsten ($g=2,059$), neben diagnostischen Systemen ($g=1,661$) und einer Mischung ($g=0,862$), Bewertungssysteme sind am wenigsten effektiv ($g=0,443$), diese sind nicht signifikant • Expertensysteme ($g=0,642$) sind nicht so effektiv wie Systeme auf NLP-Basis ($g=1,071$) oder gemischte Systeme ($g=0,919$), diese sind nicht signifikant • Keine signifikanten Unterschiede zwischen den Effekten von experimentellen ($g=1,173$) und quasi-experimentellen ($g=0,743$) Studiendesigns • Die Stichprobengröße beeinflusst die Effektstärke signifikant; größere Stichproben weisen stärkere Effekte auf (stärkste Effekte zwischen 51-100 Teilnehmenden ($g=0,718$) und 101-300 Teilnehmenden ($g=1,380$))
--	--	--	--

Durch die Darstellung des aktuellen Forschungsstands zur Lernförderlichkeit von KI wird ersichtlich, dass die Nutzung von KI ähnlichen Prämissen unterliegt, wie der Einsatz digitaler Technologien im Allgemeinen. Es bedarf einer konkreten didaktischen Implementation und einer weiterhin erfolgenden Begleitung durch menschliche Ansprechpartner:innen. Seufert et al. (2021) schlagen hierfür das Konzept der Augmentation, ergo der Partnerschaft von Mensch und Maschine vor, in welcher jeweilige Stärken genutzt werden, um entsprechende Schwächen auszugleichen.

In Anbetracht der Studienlage aus Tabelle 3 sollten KI-basierte Anwendungen für den Bildungsbereich so aufgesetzt werden, dass sie den Kontext von Texten erfassen können (Zheng et al., 2023). Ferner sollten sie individualisierte Rückmeldungen ermöglichen, um die Vorteile von KI zu nutzen. Die Interaktion mit einem spezifischen Agenten, z.B. einem Chatbot, ist nicht zwingend notwendig (L. Wang & Zhao, 2024). Problematisch für die Nutzung von KI-Systemen in Lehr-Lernprozessen ist, dass die Nutzung generativer KI-Systeme negative Effekte auf die Fähigkeit zum kritischen Denken haben kann (Gerlich, 2025). Es ist somit sinnvoll, analytische KI-Systeme zu nutzen, deren Rückmeldungen selbstständig in die Handlungsprodukte eingearbeitet werden müssen, indem z.B. intelligente Expertensysteme genutzt werden. Zwar geben Zheng et al. (2023) Hinweise darauf, dass Expertensysteme nicht ganz so lernförderlich sind, wie generative Systeme, L. Wang und Zhao (2024) können diese Ergebnisse aber nicht reproduzieren. Unter diesen Umständen ist der Einsatz von KI-Systemen, egal welcher Art, insgesamt genauso lernförderlich oder sogar lernförderlicher, wie traditionelle Unterrichtssettings (L. Wang & Zhao, 2024). An der Universität Hamburg wurde auf Basis dieser Erkenntnisse in einem Gemeinschaftsprojekt mit den Universitäten Oldenburg und Graz, Österreich, die elektronisch-didaktische Assistenz EDDA entwickelt. Mittels analytischer KI gibt sie Studierenden Feedback zu verschiedenen allgemein- und fachdidaktischen Dimensionen der Unterrichtsplanung (für Details siehe Abschnitt 2.5).

Feedback kann, wie eingangs erläutert, unter gewissen Umständen das Lernen fördern. Im folgenden Abschnitt erfolgt nun die Darstellung des aktuellen Stands der Diskussion zur Feedbackforschung, inklusive neuerer Erkenntnisse zu digitalem und KI-Feedback.

2.3 (Digitales und KI-gestütztes) Feedback

Im Sinne von Hattie und Timperley (2007, S. 81) wird Feedback verstanden als eine Information, die jemandem über Aspekte der eigenen Leistung oder des eigenen Verständnisses

gegeben wird. Diese bloße Information bewirkt nicht automatisch eine Verbesserung, kann es aber unter bestimmten Umständen tun (Shute, 2008). Diese Umstände werden als Feedbackkriterien bezeichnet.

Konzeptuell wird in verschiedene Feedbackrichtungen differenziert: das *Feed-up*, das *Feed-back* und das *Feed-forward* (Hattie & Timperley, 2007). Das *Feed-Up* stellt Informationen über Standards, Ziele und Kriterien bereit, das *Feed-back* enthält Informationen zum Stand der Zielerreichung und das *Feed-forward* entwickelt konkrete Handlungsbedarfe, um den Lernstand an die Lernziele anzunähern. Diese drei Bausteine von Feedback sind insofern empirisch tragfähig, als dass sie in darauffolgenden Studien zur Wirksamkeit von Feedback in dieser oder ähnlicher Form bestätigt wurden. Studien von Beaumont et al. (2011), Lemmrich et al. (2024), Narciss et al. (2021) oder Wisniewski et al. (2019) ergaben ähnliche Kriterien für lernförderliches Feedback. Dies sollte a) formativ, ergo entlang des Lernprozesses gegeben werden, um eine Implementation zu ermöglichen, b) zeitnah an die Leistung gegeben werden, c) einem bestimmten Format folgen, d) kommunizierten Kriterien entsprechen und e) konkrete Ansatzpunkte für Überarbeitungen enthalten. In Anknüpfung daran ergänzt Jonsson (2013), dass der/die Empfänger:in das Feedback auch als nützlich erachten sollte.

In ihrer ersten Metaanalyse zur Lernförderlichkeit verschiedener didaktischer, methodischer und pädagogischer Instrumente untersuchten Hattie und Timperley (2007) auch, wie sich verschiedene Merkmale von Feedback auf dessen Wirksamkeit auswirken. So arbeiteten sie heraus, dass korrigierendes Feedback am lernförderlichsten ist. Auch eine positive Wirkung von technologieunterstütztem Feedback konnte nachgewiesen werden. Einige Moderatorvariablen, wie Feedbacktyp, Studiendesign, Feedbackrichtung und Feedbackkanäle wurden von Wisniewski et al. (2019) erneut untersucht, diesmal mit dem Ergebnis, dass sich alle Variablen außer des Feedbackkanals positiv auf die Lernförderlichkeit von Feedback auswirken.

Zusammenfassend sollte Feedback also spezifisch, detailliert und individuell sein, es sollte auf Augenhöhe mit den Lernenden gegeben werden und auf konkreten Kriterien basieren. Darüber hinaus sollte sich das Feedback auf die Leistung und nicht auf die Person konzentrieren (Narciss, 2013; Schartel, 2012). Das präferierte Format ist ein formatives, entlang des Lernprozesses. Außerdem sollte es *aktuell* sein. Hierfür sind Beobachtungen nötig, denn das tatsächliche Leistungsniveau sollte mit dem erwarteten Niveau abgeglichen werden

(Geyskens et al., 2012). Die daraus abgeleiteten Handlungsbedarfe liefern daraufhin Informationen zur Verbesserung. Ein lernförderliches Feedback sollte auch als ein solches wahrgenommen werden – es bietet sich somit an, Feedbackzeitpunkte zu kommunizieren, so dass das Feedback angenommen und akzeptiert wird (Underwood, 2008). Demnach sollte Feedback informativ und nach Möglichkeit *selbstreguliert* verfügbar sein (Wisniewski et al., 2019). Hierbei ist jedoch zu beachten, dass das individuelle Selbstbewusstsein die Bereitschaft beeinflusst, Feedback an- und aufzunehmen (Kluger & DeNisi, 1996). Ein hohes Maß an Ängsten² beeinflusst die Aufnahme z.B. negativ, da jene Lernende eher dazu neigen, negative Konfrontation zu vermeiden. Es gilt somit, Varianzen im Selbstkonzept bei der Planung von Feedbackgelegenheiten mitzudenken. Neben der Art des Feedbacks ist auch die feedbackgebende Person ein relevanter Faktor für die Lernförderlichkeit. Feedback von Expertinnen: führt zu größeren Verbesserungen als Feedback von sogenannten Peers, den anderen Lernenden der Lerngruppe (Prilop & Weber, 2023).

Traditionelles Feedback bietet somit vielfältige Ansatzpunkte, um Lernprozesse zu verbessern, sofern es gewisse Kriterien erfüllt. Durch die digitale Transformation stehen mittlerweile in vielen Lerngruppen vielfältige Technologien zur Verfügung, durch die ebenfalls Feedback gegeben werden kann. Tatsächlich gibt es konkrete Hinweise darauf, dass digitales Feedback ähnliche oder sogar bessere Auswirkungen auf das Lernen hat als traditionelles Feedback (Cai et al., 2023; Riedel & Kleppsch, 2021; van der Kleij et al., 2015). Ferner ist korrigierendes Feedback weniger effektiv als elaboriertes, digitales Feedback (Mertens et al., 2022; van der Kleij et al., 2015; Wagner et al., 2024). Hier besteht ein Unterschied zur Studie von Hattie und Timperley (2007), in der noch korrigierendes Feedback zu präferieren war. Primär die allgemeinen Feedbackdimensionen der *Aktualität* und der *Selbstregulierung* sind hier als Vorteile digitalen Feedbacks hervorzuheben.

Ähnlich wie im vorherigen Abschnitt, haben die Forschungs- und Entwicklungsaktivitäten im KI-Bereich auch zu Innovationen in der Feedbackforschung geführt (siehe hierzu auch Hooda et al., 2022, Hsia et al., 2024, Lee & Moore, 2024, Sembey et al., 2024 oder Shaik et al., 2022). Während Lehrkräfte zuvor verschiedenste Probleme antizipieren mussten, um elaboriertes digitales Feedback vorzuformulieren, können KI-Anwendungen genutzt

² Im Originalbeitrag wird hier von *anxiety* gesprochen, einem Begriff, dessen Übersetzung ins Deutsche herausfordernd ist. Während *anxiety* als Angststörung pathologisiert werden kann, ist der Begriff auch Teil des allgemeinen englischen Sprachgebrauchs, entspricht im Deutschen aber eher Unsicherheiten, Beklemmungen oder Nervosität.

werden, um adäquate Varianten zu generieren und immer wieder anzupassen. Williams (2024) subsumiert theoriebasiert, dass drei traditionellen Feedbackbereiche (Konsistenz, Effektivität und Aktualität) von der KI-Nutzung profitieren. Während die Konsistenz des Feedbacks durch kommunizierte Kriterien erhöht wird, kann die Effektivität durch die Anpassung an eine tragfähige Feedbackstruktur erhöht werden. KI-basiertes Feedback ist zudem überwiegend zeitnah und selbstreguliert verfügbar, sodass die Lernenden dies eigenständig abrufen oder generieren lassen können (Er et al., 2024; Lee & Moore, 2024).

Bestehende Feedbackforschung fokussiert aktuell hauptsächlich große Sprachmodelle (LLM) und generative KI (Dai et al., 2024; Jacobsen & Weber, 2024; Jeon & Lee, 2023; Steiss et al., 2024). Steiss et al. (2024) zeigen, auch KI-basiertes Feedback lernförderlich ist, da die menschliche Rückmeldung das Feedback, in der Studie konkret von ChatPT, kaum übertrifft. Jacobsen und Weber (2024) ergänzen, dass gut gepromptetes Feedback von generativer KI in einigen Dimensionen sogar von höherer Qualität ist als traditionelles Peer- oder Expertenfeedback, die Qualität ist aber vom verwendeten Sprachmodell abhängig. Ferner liegen Erkenntnisse über die Veränderungen von KI-Feedback auf Ebene der Lehrkraft vor: Lee und Moore (2024) entwickeln, dass sich die Rolle der Lehrkraft bei der Nutzung von KI-Feedback verändert. Diese entwickelt sich von der feedbackgebenden, zur feedbackerklärenden Person. Sie vertieft nach Bedarf Details, unterstützt bei der Nutzung des Tools oder beantwortet Fragen zum Feedback. Je nach Qualität des Feedbacks muss sie dieses auch weiterentwickeln, indem sie im ersten Schritt Feedbackkriterien festlegt und im zweiten Schritt im Sinne der oben entwickelten Dimensionen konkrete Handlungsbedarfe für die Lernenden identifiziert.

Die Verwendung von generativer KI für Feedback wirft jedoch, analog zu den Herausforderungen in Abschnitt 2.2, einige Probleme für den Einsatz im Unterricht auf, z. B. mangelnde Reproduzierbarkeit, intransparente Trainingsdaten, die Reproduktion von Verzerrungen oder Fragen des Datenschutzes und der Ethik (Banihashem et al., 2022). Diesen gilt es im Rahmen der Feedbackerstellung gezielt zu begegnen.

Um den aktuellen Forschungsstand der Feedbackforschung mit einem fokussierten Blick auf digitales und KI-Feedback zusammenzufassen und zentrale Desiderata für die Lehrkräftebildung zu identifizieren, gibt Tabelle 4 einen Überblick über ausgewählte Forschungsarbeiten.

Tabelle 4. Synopse ausgewählter Studien zu digitalem und KI-gestütztem Feedback

Studie	Stichprobe	Methodik	Ausgewählte Ergebnisse
Cai et al. (2023)	N=14.287	Metaanalyse mit 61 Studien von 1988 bis 2021 zur Untersuchung der Lernförderlichkeit von digitalem Feedback	- Feedback hat einen mittleren Effekt auf den Lernerfolg ($g=0,44$) - Erklärendes Feedback hat den stärksten Effekt von allen Feedbackarten - Feedback in kombinierten Lernumgebungen hat einen stärkeren Effekt als Feedback in On-linesettings
Dai et al. (2024)	N=103	Auswertung von KI-generiertem Feedback für ein Seminarprojekt, welches mit Feedback von Lehrpersonal abgeglichen wurde, Vergleich von Lehrpersonal, GPT-3.5 und GPT-4-Feedback, Kodierung des Feedbacks auf einer 5-stufigen Skala sowie Berechnung der Übereinstimmung der Feedbacks	- Beide GPT-Feedbacks werden als lesbarer bewertet als das menschliche Feedback - Beide GPT-Feedbacks geben mehr Kontext zu ihrem Feedback - Menschliches Feedback beinhaltet nicht immer Feed-up, KI-generiertes Feedback enthält dies hingegen fast immer, dies trifft aber nicht auf Feed-back zu - KI-generiertes Feedback enthält weniger Feed-forward als das menschliche - GPT-4 formuliert bessere Feedbacks als GPT-3.5 und das menschliche Feedback - Schwächen des menschlichen Feedbacks vor allem in der Zieldimension
Er et al. (2024)	N=54	Vergleich von menschlichem und KI-generiertem Feedback zu Studierendenlösungen und deren Auswirkungen auf Abschlussprüfungen und Laborperformance mittels ANOVA; Messung der Feedbackwahrnehmung der Studierenden	- Studierende nehmen sowohl menschliches als auch KI-generiertes Feedback positiv wahr - KI-Feedback wurde als weniger fair, hilfreich, entwicklungsorientiert und ermutigend wahrgenommen, nur die Wahrnehmung der Dimension <i>hilfreich</i> ist statistisch signifikant - Studierende, die menschliches Feedback erhalten haben, schnitten signifikant besser in den summativen Leistungsüberprüfungen ab - Ergebnisse werden auf domänenspezifisches Wissen der menschlichen Feedbackgebenden zurückgeführt
Lee und Moore (2024)	Nicht angegeben	Systematisches Literaturreview mit 10 Studien von 2019 bis 2023; Ziel: Schaffung eines Überblicks der	Generative KI reduziert den Workload der Feedbackgebenden, welche sich daraufhin mit anderen Aspekten ihrer Tätigkeit beschäftigen können (z.B. Feedback vertiefen, Lehre verbessern, etc.)

		Wirkung von generativer KI zur automatischen Feedbackgenerierung in der Hochschulbildung	- Generative KI-Systeme verbessern Transparenz der Kommunikation, bieten kognitive und emotionale Unterstützung und fördern eine kreative Lernumgebung - Generative KI-Systeme schaffen produktivere Lernumgebungen, sollten aber noch besser in die Lernprozesse integriert werden
Lim et al. (2020)	N=86	Vergleichsstudie mit verschiedenen didaktischen Settings und Feedbackoptionen; Ziel: u.a., welche Rückschlüsse Lernende aus selbstreguliertem, personalisiertem Feedback ziehen; Auswertung mittels epistemischer Netzwerkanalyse und qualitativer Inhaltsanalyse	- Personalisiertes Feedback steigert die Motivation der Lernenden und hilft ihnen, die Aufgabe zu bewältigen - Feedback sollte klar sein, Handlungsbedarfe identifizieren und positiv formuliert sein - Curriculare Einbettung von Feedback macht dessen Nutzung zur Gewohnheit - Wichtig: Rückmeldungen müssen korrekt und passend sein
Maier und Klotz (2022)	Nicht angegeben	Systematisches Literaturreview mit 39 Studien aus 2012 bis 2022; Ziel: Entwicklung von Bestandteilen personalisierten Feedbacks	- Personalisiertes Feedback wird auf Makro-, Meso- und Mikroebene der digitalen Lernumgebungen gegeben, also für verschiedene Schritte im Lernprozess - Die Personalisierung des Feedbacks bezieht sich auf den aktuellen Lernstand und Daten zum Lernverhalten, aber nicht auf die Emotionen, Lernziele oder die Persönlichkeit des Lernenden - Die meisten Studien berichten positive, gemischte oder neutrale Lernförderlichkeit von personalisiertem Feedback - Nur wenige Studien bieten eine theoretische oder empirische Begründung für die Formulierung ihres Feedbacks
Mertens et al. (2022)	Nicht angegeben	Metaanalyse von 77 experimentellen Studien aus 1968 und 2021; Ziel: Wirksamkeit verschiedener computerbasierter Feedbackarten; Auswertung mittels Frequentist network Metaanalyse	- Elaboriertes Feedback ist am effektivsten bei einfacheren (z.B. Erinnerung, $g=0,714$) und komplexeren (z.B. Transfer, $g=0,462$) Lernzielen - Auch das Kennen der richtigen Lösung ($g=0,489$) und Answer-until-correct-Feedback ($g=0,554$) weisen kleine bis große Effektstärken auf

			• Das Kennen des Ergebnisses ist weniger effektiv darin, das Erreichen von Lernzielen zu fördern ($g=0,237$ bei einfachen, $g=0,136$ bei komplexeren Lernzielen) • Ergebnisse für komplexere Lernziele nur für elaboriertes und Answer-until-correct-Feedback signifikant • Bei hohem Vorwissen ist elaboriertes Feedback sowohl bei einfachen als auch komplexen Lernzielen am effektivsten ($g=0,74$ respektive $g=0,32$) • Bei niedrigem Vorwissen und einfachen Lernzielen ist Answer-until-correct ($g=0,77$) leicht effektiver (elaboriertes Feedback: $g=0,72$), bei komplexen Lernzielen ist elaboriertes Feedback am effektivsten ($g=0,66$)
Sembey et al. (2024)	Nicht angegeben	Systematisches Literaturreview von 38 Beiträgen aus 2016 bis 2021; Ziel: Erkenntnisse über die Nutzung von digitalen Technologien in Feedbackprozessen der Hochschulbildung	• Positiver Effekt auf den Lernprozess bei der Nutzung von KI • Hohe Übereinstimmung zwischen menschlichen und KI-generierten Urteilen • KI-Nutzung aber nicht näher spezifiziert • Keine Effektstärken berechnet
van der Kleij et al. (2015)	Nicht angegeben	Metaanalyse von 40 Studien aus 1968 bis 2012; Ziel: Effekte von Feedback in computerbasierten Lernumgebungen und deren Auswirkungen auf Lernergebnisse aggregieren; Inhaltsanalytische Auswertung und Berechnung der Effektstärken	• Elaboriertes Feedback am effektivsten ($g=0,49$) • Sofortiges Feedback am effektivsten ($g=0,59$) • Aber: Art des elaborierten Feedbacks variiert stark, hier ist keine einheitliche Zuordnung möglich (Beispiel: Bezieht sich das elaborierte Feedback auf die Lösung einer Aufgabe oder einen Lernprozess im Gesamten, variieren die Effektstärken)
Z. Wang et al. (2024)	N=111	Vergleichsstudie mit verschiedenen Feedbackvariationen	• Feedback, welches sowohl die richtige Lösung als auch die Fehler des Lernenden beinhaltet, hat einen positiveren Effekt als die anderen drei Feedbackarten (Kennen der richtigen Antwort, Hinweise zur Problemlösung und vollständige Beispiellösung) • Ein solches Feedback ist außerdem weniger kognitiv auslastend und eher kognitiv stimulierend, außerdem konnten die Lernenden es besser auf ihre Ergebnisse anwenden

Aus Tabelle 4 ergibt sich, dass die Nutzung von KI eine vielversprechende Feedbackgelegenheit darstellt. Aus den vorliegenden Arbeiten zeichnet sich jedoch ein deutliches Forschungsdesiderat in Bezug auf intelligente Tutoresysteme ab. Während in den vorliegenden Arbeiten vor allem generative KI-Systeme beforscht wurden, gibt es wenig Erkenntnisse zum Einsatz KI-basierter Tutoresysteme. Dies kann auch damit zusammenhängen, dass generative KI-Systeme flächendeckend verfügbar sind und schnell für Feedbackzwecke nutzbar gemacht werden können. Für die Entwicklung einer KI-basierten Feedbackassistentenz sollten folgende Aspekte handlungsleitend sein:

- Wahrung des Partnerschaftsprinzips (Abschnitt 2.2) durch die Nutzung einer KI-Assistentenz
- KI-basiertes Feedback sollte den gleichen Prinzipien folgen, wie traditionelles Feedback, aber die Vorteile der Technik nutzen, wie z.B. die Personalisierung des Feedbacks

Nachdem der Forschungsstand zur Lernförderlichkeit von Feedback mittels verschiedener Kanäle umrissen wurde, wird in Abschnitt 2.4 in den Beispielbereich dieser Arbeit eingeführt: Digitalisierung und Nachhaltigkeit im Rechnungswesen und Controlling.

2.4 Aktuelle Herausforderungen in der beruflichen Bildung: Digitalisierung und Nachhaltigkeit im Rechnungswesen- und Controllingunterricht

Die inhaltliche Ausgestaltung der beruflichen Bildung wird in Deutschland über die Curricula geregelt. Für den Lernort Schule sind dies auf der Makroebene die Rahmenlehrpläne, auf der Mesoebene die schulinternen Curricula sowie die Unterrichtssequenzplanungen der Lehrkräfte(-teams) und auf der Mikroebene die einzelne Unterrichtsstunde. Diese Curricula, die die Berufsbildpositionen konkretisieren, werden durch weitere Ordnungsmittel ergänzt. Eines dieser Ordnungsmittel sind die Standardberufsbildpositionen, welche definiert werden können als domänenübergreifende Kompetenzen, die in allen Ausbildungsberufen am Lernort Betrieb befördert werden sollen (Bundesinstitut für Berufsbildung, 2021). Für den Lernort Schule bestehen aus zwei Gründen Handlungsbedarfe: Einerseits aufgrund der Vorgaben zur Lernortkooperation (§2 BBiG), andererseits aufgrund der Beförderung der beruflichen Handlungskompetenz als normatives Ziel der Ausbildung, welche dazu beiträgt, dass Lernende zukünftigen Herausforderungen und Veränderungen adäquat begegnen können (Euler, 2019). Als im Jahr 2021 die neuen Standardberufsbildpositionen veröffentlicht wurden, enthielten diese als Reaktion auf globale Transformations-

prozesse erstmals dedizierte Ausführungen zu Digitalisierung und Nachhaltigkeit (Bundesinstitut für Berufsbildung, 2021). Für einen zeitgemäßen Unterricht müssen auch inhaltliche Neuerungen in die Planungen integriert werden.

Aktuelle Herausforderungen im Rechnungswesen und Controlling

Der Wandel zur Informationsgesellschaft wirkt sich zunehmend auch auf die erfordernten Kenntnisse der Auszubildenden aus. Unternehmen verlangen umfassendere Controllingkenntnisse, die über die bloße Buchhaltung hinausgehen (Schöning et al., 2020). Ein solches Desiderat ist nicht neu, es wurde bereits vor über 20 Jahren formuliert (Seifried, 2004). Diese Anforderungen spiegeln sich auch curricular wider: Kaufmännische Steuerung und Kontrolle nehmen einen wesentlichen Anteil wirtschaftsberuflicher Curricula ein (Brötz et al., 2015). Dabei sind sowohl das Rechnungswesen als auch das Controlling aus unternehmerischer Sicht entscheidend für den betrieblichen Erfolg, aus Lernendensicht aber auch zur Förderung des wirtschaftlichen Denkens (Chiapello, 2009; Preiß & Tramm, 1996).

Das Rechnungswesen nach Baetge et al. (2024, S. 1) „erfasst, speichert und verarbeitet betriebswirtschaftlich relevante quantitative Informationen über realisierte oder geplante Geschäftsvorgänge und -ergebnisse“. Zwei zentrale Bereiche können unterschieden werden: das interne und das externe Rechnungswesen. Während im internen Rechnungswesen vor allem Buchhaltung, Kosten- und Erlösrechnungen oder Planrechnungen wie die Finanzplanung über Aufwendungen und Erträge sowie Kosten und Leistungen stattfinden, ist das externe Rechnungswesen als Instrument der Finanzbuchhaltung darauf ausgerichtet, den Jahresabschluss vorzubereiten, die Gewinn- und Verlustrechnung aufzustellen und Rechenschaft darüber abzulegen, welche Einnahmen und Ausgaben das Unternehmen hat (Corsten, 2020; Wöhe et al., 2020).

Während im realen Unternehmen das Rechnungswesen den Unterstützungsprozessen zuzuordnen ist, ist das Controlling ein Steuerungsprozess (Gadatsch, 2017). Wöhe et al. (2020, S. 178) definieren das Controlling im weiteren Sinne als „Summe aller Maßnahmen, die dazu dienen, die Führungsbereiche Planung, Kontrolle, Organisation, Personalführung und Information so zu koordinieren, dass die Unternehmensziele optimal erreicht werden“. Im engeren Sinne kann die Definition von Horváth (2019) herangezogen werden, nach welcher Controlling die Unternehmensführung unterstützt, indem Ziele geplant, kontrolliert und informiert werden. Controlling ist somit ein Informationssystem, das der Koordination und der Vorbereitung unternehmerischer Entscheidungsfindung dient. Dabei gilt es,

strategisches und operatives Controlling zu differenzieren (Horváth, 2019). Strategisches Controlling wird charakterisiert durch einen langfristigen Planungshorizont, welcher der Existenzsicherung und Erfolgsentwicklung dient. Im operativen Controlling spielt die Sicherung der Wirtschaftlichkeit betrieblicher Prozesse eine existenzielle Rolle; der Planungshorizont ist kurz- bis mittelfristig und zentrale Stellschrauben sind Gewinn, Liquidität und Rentabilität (Horváth, 2019).

Nachdem die allgemeine Relevanz des Bezugsbereichs hergeleitet und dieser abgegrenzt wurde, wird im Folgenden detaillierter auf die beiden Begriffe *Digitalisierung* und *Nachhaltigkeit* eingegangen, indem sie definiert und für die berufliche Bildung entwickelt werden.

Aktuelle Herausforderung 1: Digitalisierung

Für die Digitalisierung existieren vielfältige Definitionsansätze (Traum et al., 2017). Ursprünglich primär technisch als einfache Umwandlung des Analogen ins Digitale verstanden, reicherte sich der Digitalisierungsbegriff mit zunehmenden Implikationen und Perspektiven zu einem weitreichenden, gesamtgesellschaftlichen Transformationsprozess mit vielfältigem Innovationspotenzial an (Wiesböck & Hess, 2020) Mit Blick auf das arbeitende Individuum definieren Traum et al. (2017, S. 4) Digitalisierung ergänzend als

„die Einführung bzw. verstärkte Nutzung von Informations- und Kommunikationstechnologien (IKT) durch (arbeitende) Individuen, Organisationen, Wirtschaftszweige und Gesellschaften mit den charakteristischen Folgen der Beschleunigung, zunehmenden Abstraktheit, Flexibilisierung und Individualisierung von Prozessen und Ergebnissen“.

Es geht somit um das Individuum als in der Digitalisierung agierende Personen, deren Handlungsmöglichkeiten sich verändern und erweitern. Digitalisierung bringt auch die Veränderung von beruflichen Tätigkeiten mit sich (Dengler & Matthes, 2021; Klein & Küst, 2020; Pargmann et al., 2023). Durch Veränderungen der Tätigkeitsstrukturen können drei Digitalisierungsstadien unterschieden werden: Substitution (Ersatz analoger Daten durch digitale), Prozessänderungen (veränderte Abläufe durch digitale Technologien) und Innovation (neue Prozesse und Ergebnisse) (Bleiber, 2019; Pargmann et al., 2023). Daraus ergeben sich auch Veränderungen *in* der Tätigkeit an sich (Dengler & Matthes, 2021). So leiten Aepli et al. (2017) verschiedene Tätigkeitstypen ab, die von variierenden Digitalisierungsgraden tangiert werden und damit unterschiedlich stark von Digitalisierungsprozessen beeinflusst werden.

Für die berufliche Bildung werden über die oben beschriebenen Standardberufsbildpositionen die digitalisierungsbedingten Veränderungen mittels Kompetenzen für eine *digitalisierte Arbeitswelt* operationalisiert (Abbildung 5). Diese sollen die Lernenden darauf vorbereiten, in den verändernden Tätigkeitsstrukturen und -typen handlungsfähig zu bleiben.

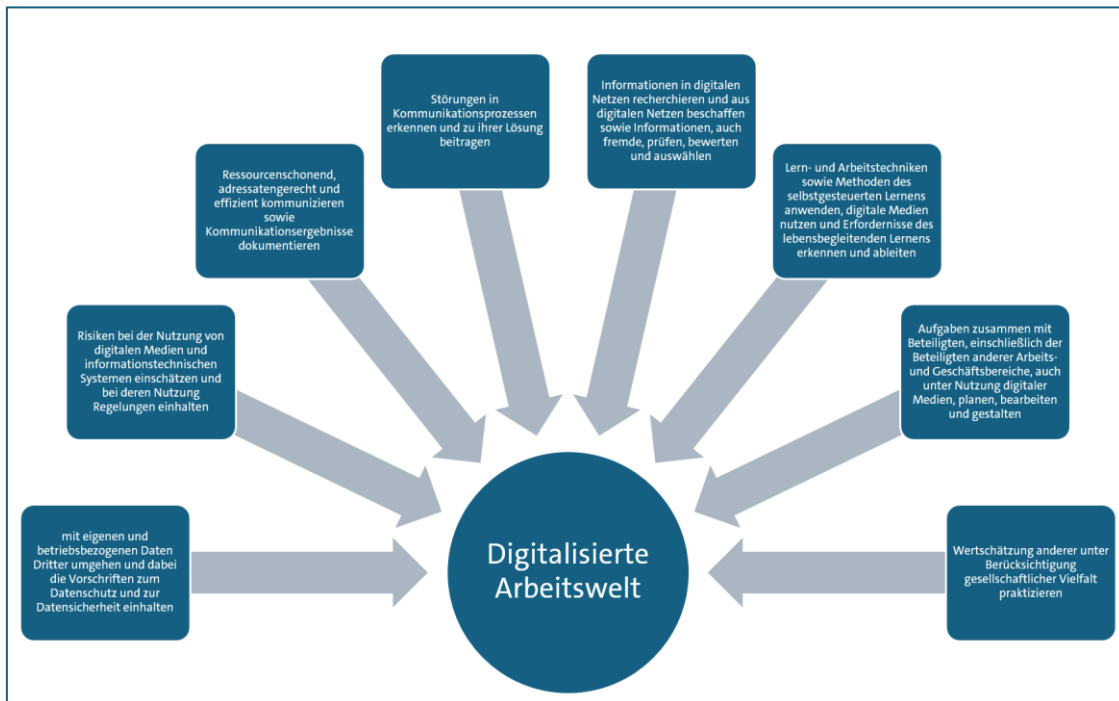


Abbildung 5. Standardberufsbildpositionen für eine digitalisierte Arbeitswelt (Bundesinstitut für Berufsbildung, 2021)

Die Standardberufsbildpositionen spiegeln den Konsens wider, in inhaltliche und methodische Digitalisierungsbezüge zu differenzieren (Euler, 2018; Euler & Wilbers, 2019; Freiling & Porath, 2020). Dies meint die Nutzung von digitalen Technologien als didaktisches Werkzeug im Unterricht einerseits und die inhaltliche Digitalisierung im Sinne des Arbeitens in der digitalisierten Welt andererseits. Inhaltliche Digitalisierungsbezüge sind z.B. die Abwägung von Risiken der Technologienutzung oder die Nutzung spezieller beruflicher Technologien, wie Tabellenkalkulation, ERP-Systeme oder berufs- und branchenspezifische Software wie Lagermanagementsysteme. Zur methodischen Gestaltung von Lehr-Lernprozessen gehören z.B. der Erwerb von Lern- und Arbeitstechniken oder die Nutzung von allgemeinen Technologien wie Präsentationen, digitalen Arbeitsblättern, Online-Quizze, Serious Games, Lernvideos oder E-Books.

Aktuelle Herausforderung 2: Nachhaltigkeit

Gemeinhin wird der Begriff der Nachhaltigkeit als Vereinigung der dreifachen Ziellinie, der Triple Bottom Line, aus Ökonomie, Ökologie und Sozialem verstanden (Pufé, 2017). Um

dorthin zu gelangen, bedarf es einer *nachhaltigen Entwicklung*, die die World Commission on Environment and Development (1987, S. 43) definiert als eine Entwicklung, „die den Bedürfnissen der heutigen Generation entspricht, ohne die Möglichkeiten künftiger Generationen zu gefährden, ihre eigenen Bedürfnisse zu befriedigen und ihren Lebensstil zu wählen“. Dieser Prozess muss von Wandel und Innovation geprägt und durch ein nachhaltiges Wirtschaften unterstützt werden. Hierfür müssen Unternehmen Verantwortung für die von ihnen produzierten Umwelt- und Sozialkosten übernehmen, die sie direkt oder indirekt berühren (Beckmann & Schaltegger, 2014). Eine Möglichkeit, dies zu realisieren, ist das Nachhaltigkeitscontrolling, welches systematisch nachhaltigkeitsrelevante Daten aufbereitet (Sailer, 2020). Die Europäische Union fördert das Erfordernis des nachhaltigen Wirtschaftens weiter durch die Einführung der Pflicht zur Nachhaltigkeitsberichterstattung, welche einen direkten Bezug zum Nachhaltigkeitscontrolling herstellt. Die benötigten Daten müssen gesammelt, analysiert und aufbereitet werden, um im Nachhaltigkeitsbericht extern kommuniziert werden zu können. Diese Pflicht gilt jedoch nur für Großunternehmen. Da jedoch 68,1% der Auszubildenden im Jahr 2022 in kleinen und mittelständischen Unternehmen (KMU) arbeiten, ist es dennoch notwendig, auch die konkreten Handlungsmöglichkeiten dieser Auszubildenden zu berücksichtigen (Bundesinstitut für Berufsbildung, 2024, S. 191). Denn diese sind bisher nicht zur Nachhaltigkeitsberichterstattung verpflichtet (Braun & Senger, 2022). Aufgrund des Bildungsziels der beruflichen Handlungskompetenz sollte der Unterricht diese Aspekte dennoch integrieren, um die Auszubildenden auch für zukünftige Situationen in anderen Betrieben handlungsfähig zu machen.

In der beruflichen Bildung existieren seit einigen Jahren Forschungsbemühungen zu einer beruflichen Bildung für nachhaltige Entwicklung (BBNE), welche gestaltungsorientiert ausgerichtet ist und zumeist auf betrieblicher Ebene geschieht (Kuhlmeier & Vollmer, 2018). Rahmengebend hinsichtlich des Ausgestaltungspotenzials für die berufliche Bildung sind erneut die Curricula. Eine Analyse der Rahmenlehrpläne ergab jedoch nur implizite Bezugspunkte von Nachhaltigkeit, die über den reinen Umweltschutz hinausgehen (Holst, 2022; Otte & Singer-Brodowski, 2017). Damit bleibt die Thematisierung von Nachhaltigkeit als mehrdimensionales Konzept in den fachlichen Inhalten weitgehend ein Desiderat (Hantke, 2020). Herangezogen werden kann auch hier die neue Standardberufsbildposition *Umweltschutz und Nachhaltigkeit* (Abbildung 6).

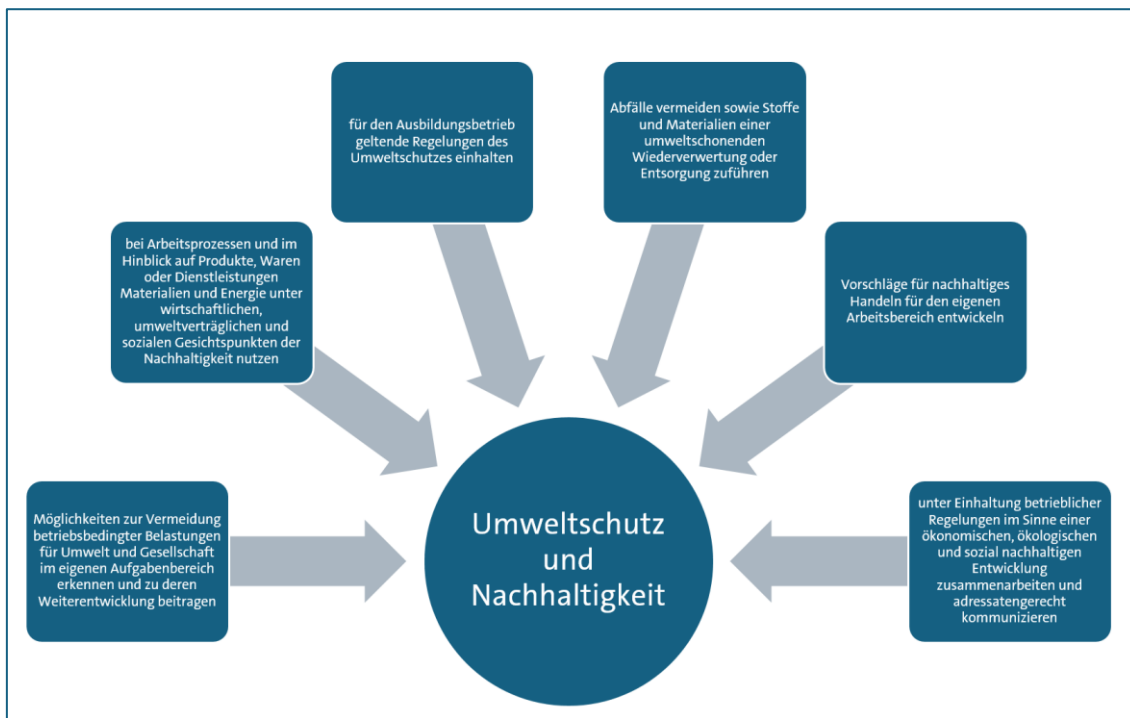


Abbildung 6. Standardberufsbildpositionen für Umweltschutz und Nachhaltigkeit (Bundesinstitut für Berufsbildung, 2021)

Ähnlich wie bei der Standardberufsbildposition zur Digitalisierung steht auch in diesen Ausführungen die Kompetenzentwicklung der Lernenden im Zentrum; so sollen diese befähigt werden, Möglichkeiten und Implikationen zu entwickeln und Veränderungsprozesse zu motivieren. Die Abwägung der eigenen und der betrieblichen Interessen wird durch die Entwicklung von Vorschlägen zum nachhaltigen Handeln im eigenen Arbeitsbereich berücksichtigt. Somit wird gezielt darauf hingearbeitet, Widersprüche und Spannungsfelder zu bearbeiten. Für nachhaltig handlungsfähige Mitarbeitende wird das Konzept der *Sustainable Change Agents* vorgeschlagen, dies sind Mitarbeitende, die die notwendigen Kompetenzen besitzen, um Widerstände zu bewältigen, Wandel voranzutreiben und als gute Beispiele voranzugehen (Berding, Slopinski & Frerichs, 2020; Heiskanen et al., 2016).

Implikationen der beiden aktuellen Herausforderungen für den Rechnungswesen- und Controllingunterricht

Digitalisierung und Nachhaltigkeit eröffnen jeweils umfangreiche Diskussionsbereiche für die berufliche Bildung im Allgemeinen und für den Rechnungswesen- und Controllingunterricht im Speziellen. Trotz dessen gestaltet sich die Studienlage zu ihrer unterrichtlichen Integration diffus (Pargmann et al., 2022). Im Rechnungswesen- und Controllingunterricht werden grundlegende Prinzipien und Techniken der internen und externen Rechnungslegung sowie der kaufmännischen Steuerung in Unternehmen behandelt (Seifried, 2004).

Ziel des Unterrichts ist es, betriebliche Prozesse zu analysieren, zu planen, zu steuern und zu kontrollieren, indem unternehmerische Informationen genutzt werden.

Explizite Bezüge zur *Digitalisierung* im Rechnungswesenunterricht existieren aktuell primär im Kontext digitaler Arbeitsblätter (Berding, Jahncke & Albersmann, 2020) oder Enterprise-Resource-Planning-Systeme (ERP-Systeme) mit losem Bezug zum Rechnungswesen (Frötschl, 2015; Hommel, 2019; Spener & Schumann, 2021). Für *Nachhaltigkeit* im Rechnungswesenunterricht sind zum aktuellen Zeitpunkt keine Studien bekannt, lediglich zu Nachhaltigkeit in der wirtschaftsberuflichen Ausbildung an sich (Börner & Brötz, 2015; Minnameier, 2022). Übertragungen aus der internationalen Forschung sind nicht ohne weiteres möglich, da diese überwiegend aus der Hochschuldidaktik stammen (Cho & Costa, 2024; Khosa et al., 2024). Aufgrund der Lernfeldorientierung und der Dualität der Berufsausbildung in Deutschland sind hochschuldidaktische Erkenntnisse lediglich als Ansatzpunkte für weitere Forschung nutzbar. Konkret für das deutsche Berufsbildungssystem entwickeln Riebenbauer et al. (2022) auf Basis des Literaturreviews aus Beitrag 1 dieser Arbeit fachwissenschaftliche Implikationen der Digitalisierung für die Lehrkräftebildung im Bereich des Rechnungswesens, für die Nachhaltigkeit leiten Pargmann et al. (im Druck) Perspektiven für eine nachhaltigkeitsorientierte Rechnungswesendidaktik ab.

In den vergangenen Abschnitten wurden theoretische und empirische Rahmungen entwickelt, in deren Kontext KI für die Professionalisierung von Lehrkräften nutzbar gemacht werden kann. Die Erkenntnisse aus diesen Ausführungen wurden zur Entwicklung der elektronisch-didaktischen Assistenz für die Unterrichtsplanung, kurz EDDA, genutzt. Im folgenden Abschnitt soll EDDA nun kurz in Konzeption und Gestaltung vorgestellt werden, da sie für die empirischen Erprobungen im weiteren Verlauf der Arbeit genutzt wurde.

2.5 Vorstellung der KI-basierten didaktischen Assistenz EDDA

Entstehung, Funktionsweise und Gestaltung

EDDA wurde in einem gemeinschaftlichen Vorhaben der Universitäten Hamburg, Oldenburg und Graz (Österreich) auf Basis des maschinellen Lernens entwickelt (Pargmann et al., 2024). Die Plattform soll angehende und praktizierende Lehrkräfte darin unterstützen, besser im Planen von Unterricht, im Treffen von didaktischen Entscheidungen und der Integration von aktuellen Herausforderungen zu werden. Die Kernidee der Plattform basiert auf der in Abschnitt 2.2 entwickelten Perspektive auf das Modell zur Rekonstruktion didaktischer Entscheidungen von Oevermann (1996). Auf dieser Basis können Nutzer:innen

Unterrichtsskizzen, Entwürfe sowie zugehörige Materialien hochladen, didaktische Kriterien als Analyseschwerpunkte auswählen und dann personalisiertes Feedback erhalten. EDDA wurde entwickelt, um entlang aller Phasen der Lehrkräftebildung zur Professionalisierung beizutragen. Damit bietet EDDA Nutzungsmöglichkeiten für Lehramtsstudierende, Lehrkräfte im Vorbereitungsdienst und fertig ausgebildete Lehrkräfte, die sich z.B. bezüglich Innovationen wie Digitalisierung und Nachhaltigkeit fortbilden wollen oder neue didaktische Impulse erhalten möchten. Denn durch die Orientierung an Oevermanns (1996) Ansatz berücksichtigt EDDAs Feedback aktuelle fachdidaktische und fachwissenschaftliche Forschungserkenntnisse, ermöglicht also eine zeitgemäße Einordnung der Unterrichtsplanungen und -materialien.

Die Plattform wurde mit 5.983 Datensätzen trainiert. Diese umfassen 884 Entwürfe, 420 Unterrichtsmaterialien und 4.679 einschlägige Schulbuchaufgaben. Hierfür wurden die Datensätze hinsichtlich 10 Haupt- und 57 Unterkategorien manuell kodiert und in die Software eingebaut, welche darauf hin Muster in den Datensätzen erkennt. Auf Basis dieses Trainings kann EDDA dann die Kodierungen in neuen hochgeladenen Dokumenten nachbilden (für Details zur technischen Umsetzung siehe Pargmann et al., 2024). Die Auswahl der Kategorien orientiert sich an etablierten allgemeinen lernpsychologischen und fachdidaktischen Prinzipien in der (aktuell nur wirtschaftsberuflichen) Unterrichtsplanung. Die detaillierte Herleitung der Kriterien erfolgte durch Riebenbauer et al. (2024).

Die allgemeinen lernpsychologischen Prinzipien sind:

- Basisdimensionen des Unterrichts: Strukturierung, Aktivierung und Vernetzung (Riebenbauer, 2021)
- Problemorientierung (Dörner, 1979; Piaget, 1977)
- Motivationsförderung (Bloemen, 2011; Deci & Ryan, 1993)

Folgende fachdidaktische Prinzipien wurden ausgewählt:

- Prozessorientierung (Bloemen, 2011; Stütz et al., 2022)
- Handlungsorientierung und Authentizität (Aebli, 2011; Brown et al., 1989; Hacker & Sachse, 2014)
- Integration von Nachhaltigkeit (Berding, Slopinski & Frerichs, 2020; United Nations, 2023)
- Inhaltliche und methodische Digitalisierung (Bernard et al., 2009; Krämer et al., 2017; Schmid et al., 2009; Tamim et al., 2011)

Anhand dieser Kriterien wurde der jeweilige aktuelle Stand der Forschung aufgearbeitet, um qualifizierte Rückmeldungen zu geben. Das Feedback, das EDDA den Nutzer:innen dann gibt, wurde anhand der Kriterien für gutes Feedback (Abschnitt 2.3) formuliert und grafisch sowie schriftlich aufbereitet. Es folgt dem sogenannten *Embodiment Principle*, welches besagt, dass die Aussagen aus Perspektive der Technologie (Ich-Form) formuliert werden sollten (Fiorella & Mayer, 2021a). Darüber hinaus wurden bei der Gestaltung der Benutzeroberfläche folgende Prinzipien der kognitiven Theorie multimedialen Lernens nach R. E. Mayer und Fiorella (2021a), die für einen möglichst hohen Lernerfolg bei der Nutzung digitaler Lehr-Lernplattformen sorgen sollen:

- *Multimedia-Prinzip*: Durch die Kombination von Worten und Bildern wird oft besser gelernt als durch Worte allein (R. E. Mayer, 2021)
- *Prinzip der Expertiseumkehr*: Es gibt individuelle Unterschiede in der Effektivität von multimedialen Lerngelegenheiten, die abhängig von der Expertise der Nutzer:innen sind, es muss also die Möglichkeit geben, in gewisser Weise die Inhalte selbst auswählen zu können (Kalyuga, 2021)
- *Redundanzprinzip*: Menschen lernen besser, wenn dieselbe Information nicht in verschiedenen Formaten präsentiert wird, also nicht grafisch *und* schriftlich (Kalyuga & Sweller, 2021)
- *Prinzip der räumlichen und zeitlichen Nähe*: Zusammengehörige Inhalte werden auch zusammen präsentiert (Fiorella & Mayer, 2021b)
- *Segmentierungsprinzip*: Lernende sollten im eigenen Tempo den Fortschritt in der Interaktion mit der Lernplattform bestimmen können (R. E. Mayer & Fiorella, 2021b)

Die Arbeit mit EDDA

Nachdem der Unterrichtsentwurf und ggf. zugehörige Materialien auf der Startseite hochgeladen wurden, kann die Nutzerin bzw. der Nutzer die Schwerpunkte auswählen, die analysiert werden sollen. Dies realisiert das *Prinzip der Expertiseumkehr*, da unterschiedliche Schwerpunkte unterschiedlich viel Überarbeitungsbedarf bedeuten. Nach erfolgter Auswahl startet der Analyseprozess. Dabei werden die hochgeladenen Texte vom Algorithmus in Zahlensequenzen transformiert, die den Inhalt für die KI lesbar machen. Dann wird die Umsetzung der didaktischen Merkmale durch den Algorithmus eingeschätzt. Für diese Einschätzungen werden dann passende Empfehlungen ausgegeben. Diese basieren auf Expertenfeedback, um eine möglichst hohe, konsistente Qualität und Verlässlichkeit zu

gewährleisten (Pargmann et al., 2024). Wenn der Upload beendet ist, erfolgt eine automatische Weiterleitung zur Ergebnisübersicht (Abbildung 7), in welcher ein Gesamtergebnis (obere, längliche Kachel) sowie die Einzelergebnisse (untere, einzelne Kacheln) für alle ausgewählten Schwerpunkte angezeigt wird. Die Kacheln sind je nach Ergebnis verschiedenfarbig eingefärbt, um den Umsetzungsgrad der Kriterien auch farblich widerzuspiegeln. Hervorzuheben ist, dass selbstverständlich nicht immer alle Kriterien vollständig in einem einzigen Entwurf umgesetzt werden können und sollten. Es geht jedoch vorrangig darum, zu reflektieren, ob das Ergebnis der Intention entspricht. Außerdem sollten die Kriterien über mehrere Unterrichtssequenzen oder Ausbildungsjahre ausgewogen verteilt werden, sodass durch die Nutzung von EDDA auch überprüft werden kann, ob alle Aspekte gleichermaßen berücksichtigt wurden.



Abbildung 7. Ergebnisübersicht

Durch einen Klick auf „Details“ erfolgt dann eine Weiterleitung auf die jeweilige Kategorienseite, die grafische und schriftliche Rückmeldungen beinhaltet (*Multimedia-Prinzip*). Jeder didaktische Schwerpunkt hat eine eigene Seite, um die *zeitliche und räumliche Nähe* zu gewährleisten. Zunächst erfolgt die grafische Aufbereitung der Ergebnisse, der Kohärenz von Entwurf und ggf. zusätzlichem Material und der Sicherheit der Analyseergebnisse (Abbildung 8). Die Ergebnisanalyse (Abbildung 8a) soll eine einfache Möglichkeit sein, um einen Überblick über die Umsetzung der einzelnen Teilkriterien zu gewinnen. Im Beispiel erreicht der Entwurf auf keiner Teilkategorie eine hohe Bewertung. Im Sinne des *Redundanzprinzips* werden die grafisch dargestellten Inhalte nicht nochmal schriftlich dargestellt. Welche konkreten Handlungsbedarfe bestehen, wird in der schriftlichen Rückmeldung aufgezeigt.

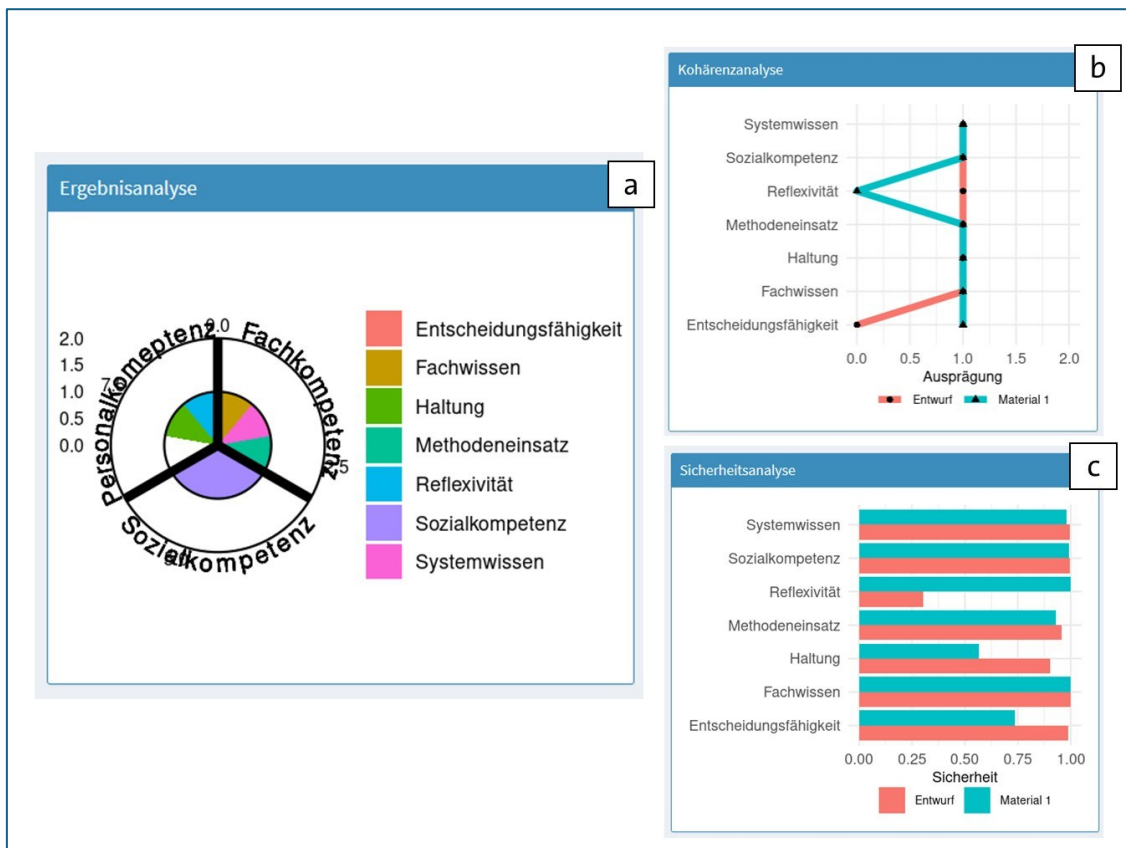


Abbildung 8. Übersicht über die grafischen Analyseergebnisse von EDDA

Daneben befindet sich die Kohärenzanalyse (Abbildung 8b). Diese zeigt auf, wie stark die Bewertungen für Entwurf und Material übereinstimmen bzw. divergieren. Diese Abbildung kann dafür genutzt werden, zu reflektieren, ob die (fehlende) Umsetzung der Kategorie im Material gewünscht war oder nicht. Im Beispiel wäre dies etwa bei der Reflexivität der Fall, diese ist im Entwurf besser geplant als im Material umgesetzt.

Die letzte grafische Darstellung ist die Sicherheitsanalyse (Abbildung 8c), welche zeigt, wie sicher sich EDDA bei ihrer Bewertung für die jeweilige Teilkategorie für alle hochgeladenen Dokumente ist. Hier ist sich EDDA z.B. bei den Kategorien Systemwissen und Sozialkompetenz sehr sicher, bei der Reflexivität jedoch nur im Material.

Neben den grafischen Repräsentationen gibt es ebenfalls ausklappbare Boxen mit einer schriftlichen Rückmeldung (Abbildung 9). Diese sind standardmäßig eingeklappt, um eine eigenständige Auseinandersetzung mit dem Feedback zu ermöglichen (*Segmentierungsprinzip*). Die Rückmeldung umfasst eine allgemeine Beschreibung der Kategorie, eine kurze Begründung für die Bewertung sowie konkrete Überarbeitungsempfehlungen. Im vorliegenden Beispiel wurde die kritische Haltung nicht mit einem Nachhaltigkeitsbezug befördert, d.h. dieser sollte, wenn erwünscht, hergestellt werden.

The image shows a user interface for EDDA. On the left, there is a vertical list of seven blue buttons, each with a white plus sign: 'Fachwissen', 'Systemwissen', 'Methodeneinsatz', 'Sozialkompetenz', 'Entscheidungsfähigkeit', 'Kritische Haltung', and 'Reflexivität'. An arrow points from the 'Kritische Haltung' button to a larger panel on the right titled 'Kritische Haltung'.

Kritische Haltung

Allgemeine Beschreibung
Lernende benötigen eine offene und kritische persönliche Haltung, um die Komplexität der nachhaltigen Transformation bewältigen zu können. Hierzu gehört der Umgang mit bestehenden Lösungen, Routinen und Wissensbeständen und die eigene Einstellung gegenüber Veränderungen und Unsicherheiten. Sie müssen die Fähigkeit entwickeln, Wandlungsbedarfe und Transformationspunkte im eigenen Handlungsbereich und auch darüber hinaus zu erkennen und bestehende Routinen hinterfragen zu können. Dies ist notwendig, weil häufig konträre Sichtweisen und Gestaltungsideen aufeinandertreffen.

Meine Empfehlungen für Sie
In Ihrem Unterrichtsentwurf wird eine offene und kritische Haltung bereits berücksichtigt. Zu diesem Schluss komme ich, weil Sie Aufgabenstellungen oder Gestaltungsansätze gewählt haben, in denen die Lernenden Bestehendes kritisch hinterfragen. Hierfür nutzen Sie bspw. Operatoren wie Stellung nehmen, Diskutieren, Beurteilen, Rückmeldung geben, Bewerten, Entwickeln, Umwandeln oder Gestalten. Das ist schonmal gut!

Versuchen Sie doch nun, einen impliziten oder expliziten Nachhaltigkeitsbezug herzustellen:

- Implizit: Dies gelingt leicht, indem Sie Ihre Handlungssituation in ein nachhaltig wirtschaftendes Modellunternehmen einbetten und die Produkte und Dienstleistungen nachhaltiger gestalten. So müssen die Lernenden Nachhaltigkeit automatisch in ihren Entscheidungen mitdenken, um im Sinne des Unternehmensleitbildes zu handeln.
- Explizit: Regen Sie Lernende an, bestehende Wirtschaftsweisen und Produktgestaltungen des Unternehmens hinsichtlich Fragestellungen wie bspw. den Arbeitsbedingungen, dem Energiekonsum oder der Recyclingfähigkeit zu hinterfragen, zu überarbeiten und neu zu gestalten. Auch als Standard hingenommene Prämissen wie die Gewinnmaximierung können kritisch hinterfragt und diskutiert werden.

Sie müssen einen nachhaltigkeitsorientierten Unterricht im Rechnungswesen oder Controlling planen? Auch da habe ich einige Anregungen für Sie:

► Nachhaltigkeitsorientierte offene und kritische Haltung im Rechnungswesen und Controlling

Weiterführende Quellen

Abbildung 9. Beispiel für die Klappboxen mit ausgeklappter Rückmeldung

In diesem Abschnitt wurde die KI-Assistenz für die Unterrichtsplanung EDDA hinsichtlich ihrer Gestaltung und des Arbeitsprozesses präsentiert. In EDDA manifestieren sich insgesamt verschiedenste Erkenntnisse der Abschnitte 2.1 bis 2.4 zu unterrichtsplanerischen Merkmalen (Pargmann et al., 2024; Riebenbauer et al., 2024), der Forschung zu analytischer und generativer KI (Beheshti, 2023), zur Gestaltung digitaler Lehr-Lernplattformen (R. E. Mayer & Fiorella, 2021a), der Formulierung von (KI-gestütztem) Feedback (Narciss, 2013; Sembey et al., 2024; Wisniewski et al., 2019) und zur Integration aktueller Herausforderungen in den wirtschaftlichen Berufsschulunterricht (Pargmann & Berding, 2025; Pargmann et al., 2023). Im nachfolgenden Kapitel werden nun darauf aufbauend die Beiträge vorgestellt und abgedruckt, die gemeinsam mit den Ausführungen dieses Kapitels zur späteren Diskussion und Beantwortung der zentralen Fragestellungen genutzt werden.

3 Beiträge der kumulativen Dissertation

3.1 Beitrag 1

Der erste Beitrag dieser Arbeit mit dem Titel „Digitalisation in accounting: a systematic literature review of activities and implications for competences“ erschien 2023 in der Zeitschrift *Empirical Research in Vocational Education and Training* in Koautorenschaft mit Elisabeth Riebenbauer, Doreen Flick-Holtsch und Florian Berding. Der Abdruck des Beitrags befindet sich in Anhang B.

Der Eigenanteil der Autorin an dem Beitrag stellt sich wie folgt dar: Sie kodierte, analysierte und interpretierte die Daten und verfasste federführend das Manuskript, verantwortete die Überarbeitung sowie die Einreichung. Elisabeth Riebenbauer, Doreen Flick-Holtsch und Florian Berding haben das Projekt initiiert, den theoretischen Hintergrund und die Methodologie entwickelt und einen Beitrag zum Anfertigen und Überarbeiten des Manuskripts geleistet.

Der Beitrag nimmt Bezug zur Digitalisierung als aktuelles Thema im Rechnungswesen und Controlling. Neue Technologien verändern die Tätigkeiten von Mitarbeitenden. Das Konzept der beruflichen Handlungskompetenz soll Mitarbeitende dazu befähigen, langfristig handlungsfähig im Beruf zu bleiben. Um Lernende aber bedarfsgerecht auszubilden, müssen (angehende) Lehrkräfte wirtschaftlicher Fächer und Lernfelder Kenntnisse über aktuelle Entwicklungen und Herausforderungen der Tätigkeitsbereiche haben. Aus diesem Grund widmet sich der Beitrag der Analyse unternehmerischer Tätigkeiten im Rechnungswesen und Controlling, klassifiziert diese u.a. in Routine- und Nicht-Routinetätigkeiten und ordnet Unternehmen einem Stadium der Digitalisierung zu. Darüber hinaus werden Chancen und Risiken der Digitalisierung betrachtet, um eine kritisch-diskursive Perspektive auf klassische Inhalte und Tätigkeiten im Rechnungswesen und Controlling entwickeln zu können. Diese Ziele werden mittels eines systematischen Literaturreviews bearbeitet, welches N=70 fachwissenschaftliche Publikationen inhaltsanalytisch ausgewertet. Am Ende werden die Ergebnisse kritisch eingeordnet und hinsichtlich ihrer Implikationen für die Tätigkeitsbereiche von Mitarbeitenden im Rechnungswesen und Controlling sowie die Ausbildung dieser bewertet.

3.2 Beitrag 2

Der zweite Beitrag dieser Arbeit trägt den Titel „Sustainability content knowledge for teachers in accounting and controlling: A systematic literature review on industry approaches, practices and desired competences“. Dieser wurde von der Zeitschrift *Frontiers in Education* zur Veröffentlichung angenommen. Verfasst wurde er gemeinsam mit Florian Berding. Der Abdruck des Beitrags befindet sich in Anhang C.

Der Eigenanteil der Autorin stellt sich folgendermaßen dar: Sie hat das Vorhaben initiiert, den theoretischen Hintergrund und die Methodik entworfen, die Daten recherchiert, kodiert, analysiert und interpretiert und federführend das Manuskript erstellt. Florian Berding hat in der Entwicklung der Idee als Sparringpartner fungiert und die Methodik gemeinsam mit der Autorin weiterentwickelt. Er hat außerdem Überarbeitungen am Manuskript vorgenommen.

Der Beitrag eruiert den Stand der Nachhaltigkeit in der fachwissenschaftlichen Diskussion des Themenbereichs Rechnungswesen und Controlling. Nicht nur aufgrund der Agenda 2030 und der daraus entstehenden gesellschaftlichen, politischen sowie unternehmerischen Relevanz einer Begrenzung des Klimawandels ergibt sich für Unternehmen die Herausforderung, Nachhaltigkeit in ihre Geschäftspraktiken zu integrieren. Während es hierfür vielfältige Ansatzpunkte über alle Geschäftsbereiche gibt, werden primär das Rechnungswesen und Controlling als wirkungsvolle Optionen identifiziert. Aufgrund des diffusen Status Quo wird ein Literaturreview mit N=79 fachwissenschaftlichen Publikationen durchgeführt, welches darauf abzielt, Entwicklungslinien, Verortungsmöglichkeiten und Kompetenzbedarfe zu identifizieren. Zuletzt werden zentrale Erkenntnisse abgeleitet und mit Bezug auf die wirtschaftliche Berufsbildung diskutiert.

3.3 Beitrag 3

Der vierte Beitrag mit dem Titel „How AI feedback supports lesson planning in vocational teacher education: A longitudinal intervention study using an analytical AI platform“ wurde bei der Zeitschrift *Empirical Research in Vocational Education and Training* gemeinsam mit Florian Berding, Karin Rebmann und Elisabeth Riebenbauer eingereicht und befindet sich aktuell im Reviewverfahren. Der Abdruck des Beitrags befindet sich in Anhang D.

Die Autorin dieser Arbeit entwickelte das Design der Interventionsstudie gemeinsam mit Florian Berding, sie wählte außerdem die latenten Konstrukte/die verwendeten Erhebungsinstrumente aus. Darüber hinaus bereitete sie die Daten auf, kodierte sie, wertete sie aus und interpretierte sie. Außerdem verfasste sie das Manuskript des Beitrags. Florian Berding unterstützte sie bei statistischen Fragen und der Umsetzung in R und fungierte als theoretischer und methodischer Sparringpartner. Karin Rebmann und Elisabeth Riebenbauer unterstützten die Manuskripterstellung durch ihr kritisches Feedback. Außerdem stellten sie zeitliche Ressourcen zur Durchführung der Interventionsstudie sowie der Entwicklung von EDDA zur Verfügung.

Dieser dritte Beitrag untersucht, welche Effekte das Feedback der KI-Plattform auf die Umsetzung von (fach-)didaktischen und allgemeinen planerischen Dimensionen hat. Die Nutzung von KI für Feedbackzwecke birgt zahlreiche Potenziale, aber auch Grenzen. Diese werden im Beitrag zusammengefasst (siehe dazu auch Abschnitt 2.3). Es lässt sich ableiten, dass erste Studien mit generativen KI-Systemen Hinweise darauf liefern, dass KI-Feedback genauso gute oder sogar bessere Effekte bei Lernenden hat als rein menschliches Feedback (Cai et al., 2023). Da generative KI-Systeme aber weitreichende ethische, nachhaltigkeits- sowie literalitätsbezogene Auswirkungen hat, ist es eine sinnvolle Alternative, eine analytische KI zu nutzen, die hochwertiges Feedback gibt. Ob diese Lösung empirisch tragfähig ist und wie sie sich auf KI-Literacy, Motivation, Einstellungen zu Technologien und digitalen Kompetenzen auswirkt, wird in diesem Beitrag betrachtet. Außerdem werden Implikationen für Design und Inhalt von analytischen KI-Plattformen abgeleitet und im Kontext der Studienlage diskutiert.

3.4 Beitrag 4

Der letzte Beitrag dieser Arbeit „Engaging with AI: How student teachers interact with an analytical AI platform for feedback and the effect on their lesson plans – An eye tracking study“ wurde im *Journal of New Approaches in Educational Research* gemeinsam mit Florian Berding, Karin Rebmann und Elisabeth Riebenbauer eingereicht. Der Abdruck des Beitrags befindet sich in Anhang E.

Die Autorin dieser Arbeit hat die Studie konzipiert, die Daten erhoben, kodiert und ausgewertet sowie den Beitrag verfasst. Florian Berding hat durch sein Feedback zum Studienkonzept und zum Manuskript sowie Überarbeitungshinweise zum Artikel beigetragen und als Sparringpartner für statistikbezogene Fragen fungiert. Ebenfalls unterstützte er bei der syntaktischen Umsetzung in R. Karin Rebmann und Elisabeth Riebenbauer unterstützten die Manuskripterstellung durch ihr Feedback. Außerdem stellten sie zeitliche Ressourcen zur Durchführung der Interventionsstudie sowie der Entwicklung von EDDA zur Verfügung.

Eye Tracking-Studien erlauben es den Forscher:innen, Nutzungsdaten von Plattformen und Reizen (einem sogenannten Stimulus) zu erheben, ohne, dass diese von der nutzenden Person reflektiert oder verbalisiert werden müssen. Dies hat den Vorteil, dass Rückschlüsse über die Nutzungsintensität gezogen werden können, ohne, dass Teilnehmende aufgrund äußerer Einflüsse oder sozialer Erwünschtheit verzerrte Perspektiven berichten. Da personalisiertem Lernen große Vorteile zugeschrieben werden (Maier & Klotz, 2022), sollten auch digitale Anwendungen, soweit möglich, personalisierbare Interaktionen ermöglichen oder diese im Sinne der Bildung für die Nutzer:innen erlernbar machen. Die Interventionsstudie im dritten Beitrag dieser Arbeit hat zwar die Wirksamkeit von EDDA untersucht, die Daten erlauben jedoch keine Rückschlüsse auf die Art der Interaktion und ob unterschiedliche Interaktionstypen unterschiedlich wirksam sind. Daher wurde im Rahmen dieses Beitrags eine Eye Tracking-Studie durchgeführt, bei der die Lernenden mit der Plattform interagieren sollten. Verschiedene Interaktionstypen betrachten Abbildungen unterschiedlich lang und intensiv, lesen Texte auf unterschiedliche Art und interagieren insgesamt unterschiedlich mit dem System. Zur vergleichenden Analyse wurden verschiedene sogenannte Areas of Interest (AOI) festgelegt und kodiert. AOIs sind vom Forschenden festgelegte Bereiche im Blickfeld. Der Beitrag diskutiert die Charakteristika der verschiedenen Interaktionstypen, ihre Entwicklung im Laufe des Semesters und die Effekte auf die Planungsentwürfe.

4 Diskussion

4.1 Gesamtdiskussion der Ergebnisse

Die vorliegende Arbeit will herausarbeiten, wie eine KI-Plattform in der wirtschaftsberuflichen Lehrkräftebildung gestaltet und genutzt werden kann, um Studierende bei ihrer Unterrichtsplanung zu aktuellen Themen zu unterstützen. Um sich dieser Fragestellung zu nähern, wurden vier Teilfragestellungen formuliert, die im Folgenden mithilfe der eingebrachten Beiträge beantwortet werden sollen. Außerdem werden Implikationen für die Entwicklung und Nutzung von KI-Systemen und KI-Feedback für die Unterrichtsplanung abgeleitet.

Zusammenfassung und Diskussion der Ergebnisse von Beitrag 1

Die ersten beiden Beiträge beschäftigen sich mit den fachwissenschaftlichen Erfordernissen zur Integration von Digitalisierung und Nachhaltigkeit in den Rechnungswesen- und Controllingunterricht. Dieser Bereich ist deshalb ein so interessanter Beispielbereich, weil er einerseits einen relevanten curricularen Umfang hat (Brötz et al., 2015), andererseits aber auch das wirtschaftliche Denken befördert (Preiß & Tramm, 1996). Gleichzeitig liegen nur veraltete respektive wenig zu aktuellen didaktischen Fragestellungen passende Konzepte vor, die die Integration aktueller Herausforderungen erschweren (Pargmann et al., 2022). Der erste Beitrag erarbeitet in Form eines systematischen Literaturreviews den aktuellen Stand der fachwissenschaftlichen Diskussion zur Digitalisierung im Rechnungswesen und Controlling.

Zentrales Ergebnis des Beitrags sind zwei ineinandergreifende Transformationsebenen: Die betrieblichen Digitalisierungsstadien einerseits, sich daraus entwickelnde Veränderungen in den konkreten beruflichen Tätigkeiten andererseits. So sind zwar die meisten Unternehmen noch auf der Ebene der Substitution (Ersatz analoger durch digitale Prozesse), daraus ergibt sich aber für die Tätigkeiten eine durchaus kritische Veränderung: denn durch die Substitution, ergo den Austausch analoger gegen digitaler Prozesse, und auch einer Automatisierung von Abläufen, ergeben sich Zeitersparnisse, die wiederum Möglichkeiten für eine inhaltliche Veränderung der Arbeit eröffnen. Diese Einsparpotenziale führen zu weitreichenden Veränderungen und ergeben auch Qualifikationsbedarfe bei den Mitarbeitenden (Dengler & Matthes, 2021). Somit verändern sich auch die benötigten Kompetenzen, die im Sinne der Beförderung der beruflichen Handlungskompetenz auch in der Ausbildung befördert werden müssen. Dies sind vor allem Datenkompetenzen und ein gutes

Verständnis der Wertschöpfungsprozesse des Unternehmens (Riebenbauer et al., 2022). Dabei werden in der wirtschaftsberuflichen Bildung zwar Inhalte über Lernfelder prozessorientiert gedacht, gerade im Rechnungswesen- und Controllingunterricht aber nicht unbedingt auch so umgesetzt (Stütz, 2024). Hier gehen die Anforderungen der Berufswelt und die Unterrichtsinhalte auseinander.

Neben diesen Auswirkungen auf die *Arbeitsorganisation* wurden auch Auswirkungen auf die *Aufgabentypen* identifiziert. Im Controlling und im internen Rechnungswesen dominieren schwer digitalisier- und automatisierbare Nicht-Routineaktivitäten, die wiederum aber erhöhte Erfordernisse an die Kompetenzen der Mitarbeitenden stellen. Währenddessen bleiben im externen Rechnungswesen überwiegend automatisierte oder manuelle Routinetätigkeiten, wie die Rechnungserstellung und -buchung übrig. Hieraus ergeben sich ebenfalls Einsparpotenziale, die die Qualifikationsebene betreffen. Inhaltlich werden immer mehr Technologien relevant, um die Automatisierung der Tätigkeiten umsetzen, überwachen und ihre Ergebnisse bewerten zu können. Die dafür benötigte Digitalkompetenz erstreckt sich auf allgemeine wie spezifische berufliche Technologien. Da die gesamte Wertschöpfungskette digitalisiert wird, müssen vermehrt digitale Werkzeuge für das Prozessmanagement eingesetzt werden. Dies sind im allgemeinen Workflowmanagementsysteme, spezieller können es aber auch ERP-Systeme oder Buchhaltungssoftware sein.

Beantwortung von Teilfrage 1

Im Kontext der ersten Teilfrage *„Wie ist der Stand der Digitalisierung im Rechnungswesen und Controlling? Welche Inhalte sollten in den Unterricht integriert, welche Kompetenzen befördert werden?“* ergeben sich folgende Antworten:

1. Digitalisierung erfolgt über den gesamten Wertschöpfungsprozess eines Unternehmens. Umfangreiche Prozesskenntnisse müssen daher befördert werden. Gleichzeitig liegen für die Prozessorientierung im Rechnungswesen deutliche Optimierungspotenziale vor (Stütz, 2024). Es bedarf somit einer dedizierten Integration.
2. Besonders im Controlling und im externen Rechnungswesen ergeben sich Desiderata für die Adaption von Nicht-Routinefällen in den Kompetenzerwerb. Die Lernenden müssen somit noch mehr dazu befähigt werden, auf Unsicherheiten und Veränderungen zu reagieren und ihre Fähigkeiten anzupassen. Hierfür bedarf es eines umfangreichen Verständnisses wirtschaftlicher Zusammenhänge, im Sinne der

Beförderung wirtschaftlichen Denkens nach Preiß und Tramm (1996). Außerdem müssen auf inhaltlicher Ebene auch Datenkompetenzen befördert werden.

3. Immer mehr Unternehmen digitalisieren ihre Geschäftsprozesse und verwenden Technologien wie KI, Robotics oder Big Data. Die steigende Nutzung digitaler Technologien im Rechnungswesen und Controlling erfordert eine Integration relevanter Produkte im Unterricht. Die Integration sollte jedoch Teil des Lehr-Lernprozesses sein, um eine Handlungsfähigkeit bei den Lernenden zu erzeugen und keine ‚Klickschulungen‘ zu geben (Pargmann et al., 2022).

Zusammenfassung und Diskussion der Ergebnisse von Beitrag 2

Der zweite Beitrag nähert sich derselben Domäne aus der Perspektive der Nachhaltigkeit und fragt nach spezialisiertem Inhaltswissen für Lehrkräfte im Kontext von Nachhaltigkeit im Rechnungswesen und Controlling. Dies geschieht mittels eines systematischen Literaturreviews. Zentrales Ergebnis des Beitrags sind Inhalte des Rechnungswesens und Controllings, die einen unterrichtlichen Nachhaltigkeitsbezug ermöglichen. So wird Nachhaltigkeit betrieblich auf verschiedensten Ebenen und in verschiedenen Unternehmensbereichen eingesetzt. Überwiegend sind dies Führungs- und Steuerungsprozesse oder sie betreffen die gesamte Wertschöpfungskette, z.B. im Rahmen eines Beschaffungskodexes, der auch für das Controlling relevant ist. Zwar ist die vorherrschende Motivation zur Integration von Nachhaltigkeit noch immer monetärer oder rechtlicher Natur, doch gibt es aus der Perspektive der sozialen Verantwortung heraus einige Unternehmen, die auf verschiedenen Ebenen Nachhaltigkeitsbemühungen anstreben.

Bezüglich der Tätigkeiten im Controlling zeichnet sich ein Fokus auf allgemeine Kennzahlen und spezifische Nachhaltigkeits-Key Performance Indikatoren (NKPI) sowie auf das Kostenmanagement ab. Dies widerspricht in gewisser Weise der holistischen Integration von Nachhaltigkeit, die die unternehmerische Persistenz zwar priorisiert, Kosten aber individuell bewerten lässt. Dies erfordert auch Tools, die Nachhaltigkeit integrativ berücksichtigen, statt als separates Nachhaltigkeitscontrolling zu verorten. Im Rechnungswesen sind die Tätigkeiten stärker auf die Herausforderung der Erfassung nicht-monetärer Informationen ausgerichtet; speziell im Kontext buchhalterischer Erfordernisse existieren Fragen nach der Internalisierung externer, mitunter globaler Kosten. Fraglich ist, inwiefern diesbezüglich im externen Rechnungswesen überhaupt Handlungsmöglichkeiten bestehen, da dieses an geltendes Recht gebunden ist. Das Review identifiziert hierfür einige Ansatzpunkte,

die Diskussionspotenziale für den Unterricht bieten, wie z.B. die Amortisation externer Kosten über Fonds oder die Supplementierung vorgeschriebener Rechnungslegungen durch Umwelt- und Sozialbilanzen. Für den Bezugsbereich des Controllings bestehen mehr Handlungsmöglichkeiten, insbesondere Werkzeuge, deren Integration im Sinne der Handlungsorientierung lohnend erscheint, wie eigene Kennzahlensysteme und Berichtswege oder NKPI.

Die Kompetenzdimensionen, die überwiegend angesprochen werden, sind die Fach- und Methodenkompetenz. Diese äußern sich durch den Aufbau eines relevanten Nachhaltigkeitswissens, sowie durch die Fähigkeit, relevante Tools und Kennzahlen(-systeme) zu identifizieren und zu adaptieren. Dies bezieht sich auch auf notwendige Software. Speziell im Kontext des Nachhaltigkeitswissens ist es notwendig, die Bedeutung von Steuerungsinformationen einschätzen zu können und Verbindungen zu identifizieren. Im Bereich der Sozial- und Selbstkompetenzen werden vor allem Kommunikation, Kollaboration und Kooperation als zentrale Fähigkeiten herausgearbeitet, ebenso wie eine nachhaltigkeitsbezogene Lernbereitschaft. Komplexitäten und Zielkonflikte sollen verstanden und Lösungsvorschläge entwickelt werden. Auch die Abwägung der eigenen und der unternehmerischen Prinzipien und Werthaltungen sind relevante Aspekte, wenngleich nicht kritikfrei diskutiert: So wird die Frage nach dem Stellenwert der Identitätsbildung von Auszubildenden (z.B. Thole, 2021) immer wieder in Frage gestellt hinsichtlich der Koexistenz individueller und unternehmerischer Handlungsmotivation, sofern der/die Lernende als Mitarbeitender des Unternehmens agiert (Minnameier, 2022).

Beantwortung von Teilfrage 2

Auf Basis dieser Erkenntnisse ergeben sich im Kontext von Teilfrage 2 „*Wie ist der Stand der Nachhaltigkeit im Rechnungswesen und Controlling? Welche Inhalte sollten in den Unterricht integriert, welche Kompetenzen befördert werden?*“ folgende Aspekte:

4. Analog zur Digitalisierung erfordert auch die Integration von Nachhaltigkeit eine vertiefte Anknüpfung an reale Geschäfts- und Handlungsprozesse, in denen Steuerung und Kontrolle in den Fokus gerückt werden (Stütz, 2024). Denn die Einbettung von Nachhaltigkeit in verschiedene Ebenen des Controllings und auch das Rechnungswesen an sich bedarf eines Stroms an Informationen, der mithilfe des Bezugsbereichs aufgearbeitet werden muss.

5. Da Nachhaltigkeit über alle Unternehmensbereiche hinweg angesiedelt werden kann, sollte im Sinne einer Handlungsorientierung eine authentische Repräsentation von unternehmerischer Nachhaltigkeit angestrebt werden. Dies bezieht sich auch darauf, Zielkonflikte zu thematisieren und Komplexitäten herauszuarbeiten. Dabei gilt es, die individuelle Handlungsmacht zu fokussieren, um keine Überforderung zu bewirken (Schütt-Sayed et al., 2021). Dies kann z. B. vom Ausgangspunkt geschehen, ökonomische Nachhaltigkeit als primäre Zieldimension zu hinterfragen.
6. Entsprechende Kalküle, Konzepte und NKPI für die Bereiche Rechnungswesen und Controlling können genutzt werden, um eine authentischere Situierung des Lernens zu erzeugen. Das Rechnungswesen sollte sich somit aus einer nachhaltigkeitsorientierten Perspektive dem Kostenbegriff nähern und didaktische Konzepte nutzen, die alternative Sichtweisen erlauben, wie etwa das wirtschaftsinstrumentelle Rechnungswesen (Berding, 2019; Preiß & Tramm, 1996). Im klassischen Rechnungswesen- und Controllingunterricht etablierte Konzepte, wie der Return on Investment, sollten durch nachhaltigkeitsorientierte Alternativen wie den Sustainable Value ersetzt oder erweitert werden (Pargmann et al., im Druck).
7. Die im Review herausgearbeiteten Kompetenzen sind weitestgehend anschlussfähig an Kompetenzen für Sustainable Change Agents, welche jene Mitarbeitenden bezeichnen, die über Kompetenzen verfügen, eine nachhaltige Transformation voranzutreiben (Berding, Slopinski & Frerichs, 2020). Eine gezielte Beförderung ebendieser Kompetenzen scheint somit sinnvoll, um die berufliche Handlungskompetenz im Sinne einer nachhaltigen Entwicklung zu befördern.

Zusammenfassung und Diskussion der Ergebnisse von Beitrag 3

Um den Nutzen von EDDA für den konkreten Prozess der Unterrichtsplanung zu identifizieren, präsentiert der dritte eingebrachte Beitrag dieser Arbeit die Ergebnisse einer semesterbegleitenden Interventionsstudie im Experimental-Kontrollgruppendesign. Zentrale Erkenntnis des Beitrags ist, dass das Feedback von EDDA bei 14 von 57 Planungsmerkmalen wirksamer ist als das menschliche Feedback. Bei allen anderen Merkmalen ist das KI-Feedback mindestens genauso lernförderlich, wie menschliches Feedback, außer beim Merkmal *Einsatzbedingungen der Technologie*. Dort wirkt das menschliche Feedback besser. Bei der Nachhaltigkeit wird das Sustainable Development Goal (SDG) 15 besser umgesetzt als in der Kontrollgruppe, hinsichtlich der Kompetenzen für Sustainable Change Agents sind die Sozialkompetenz und die Reflexivität als Teildimension der Selbstkompetenz

hervorzuheben. Auch die Dimensionen des Informierens, Planens, Entscheidens und Bewertens bei der Handlungsorientierung sowie die Integration von Informationsströmen zur Steigerung der Situierung wurden von der Interventionsgruppe besser umgesetzt. Diese erreichte außerdem auch höhere Stufen bei der Motivationsförderung im Unterrichtsentswurf auf den Dimensionen des Kompetenzerlebens und der sozialen Eingebundenheit. Bei der Digitalisierung zeigen sich signifikante positive Unterschiede beim Technologieeinsatz. Bei der Thematisierung der Einsatzbedingungen schneidet jedoch die Kontrollgruppe besser ab. Hingegen setzt die Experimentalgruppe alle Basisdimensionen besser um.

Insgesamt zeigen sich also deutliche positive Effekte von EDDA auf die Umsetzung einiger Planungsdimensionen. Kritisch hervorzuheben ist jedoch, dass Studierende, die schon zu Beginn des Semesters die Dimensionen gut umgesetzt hatten, nur noch wenig Wachstum zeigten. Hier sollte darauf geachtet werden, dass EDDA auch für leistungsstärkere Studierende noch genügend Feedbackpotenzial bietet, indem z.B. vertiefende Reflexionsfragen gestellt werden. Alternativ sind auch vertiefte Diskussionen mit Peers oder der Seminarleitung denkbar, oder aber die Nutzung generativer KI-Systeme als Sparringpartnerin. In einer Dimension wirkt das menschliche Feedback scheinbar sogar besser als das von EDDA. Ursache hierfür kann eine individuellere Passung des Feedbacks auf den konkreten Planungskontext der Studierenden sein. Da EDDA Expertenfeedback nutzt und anhand der Umsetzung der Kategorien das passende Feedback ausspielt, fehlen inhaltliche Bezüge zur Planung der Studierenden. Einige Vorschläge sind generisch und in komplexen Anwendungszusammenhängen durchaus herausfordernd zu implementieren. Hier bedarf es einer engmaschigen Betreuung durch die Seminarleitung, die Diskussionsbedarfe erkennt und auf das Feedback von EDDA aufbaut.

In Ergänzung zur Umsetzung der Planungsdimensionen wurden Metadimensionen für den Technologieeinsatz erhoben, wie die Entwicklung der KI-Literacy, der Einstellungen zu digitalen Technologien, der eigenen Digitalkompetenzen sowie der Motivation. Die Prüfung des Messmodells ergab, dass nicht alle ausgewählten Instrumente für die längsschnittliche Erhebung über ein ganzes Semester geeignet sind. Das resultierende Messmodell konnte daher nur für ausgewählte Skalen angewendet werden. Für den längsschnittlichen Einsatz weisen die verwendeten Instrumente Grenzen auf, die einen Bedarf nach verlässlichen Alternativen signalisieren. Bei den Metadimensionen zeichnen sich nur bei zwei Subskalen der KI-Literacy signifikante Effekte ab. Dieses Ergebnis bestätigt aktuelle

Erkenntnisse aus der Technologieforschung, nach der der bloße Einsatz einer Technologie nicht notwendigerweise auch zu einer Veränderung von Kompetenzen oder Einstellungen führt (Hangen & Wuttke, 2024; Jia & Tu, 2024; Riski & Nuryanto, 2024).

Bezüglich der Integration von Nachhaltigkeit und Digitalisierung im Rechnungswesen und Controlling als Unterrichtsschwerpunkt können aufgrund der fehlenden qualitativen Auswertung der Unterrichte keine konkreten Aussagen zu Umfang und Art der Bezüge getätigt werden. Wohl aber kann festgehalten werden, dass es Unterschiede zwischen den Gruppen gibt. Vor allem im Vergleich zu der Querschnittstudie von Riebenbauer et al. (2025), die zur Umsetzung derselben Planungsdimensionen und unter Nutzung des selben Kodiermanuals wie in der vorliegenden Arbeit durchgeführt wurde, zeichnen sich in Beitrag 3 klare Verbesserungen ab. So inkludieren die Entwürfe aus der Interventionsstudie insgesamt mehr SDGs und erreichen auch insgesamt häufiger höhere Kodierstufen bei den Nachhaltigkeitskompetenzen, als die Entwürfe aus Riebenbauer et al. (2025). So integriert die Experimentalgruppe insgesamt zwölf SDGs und damit doppelt so viele wie in der Vergleichsstudie. Darüber hinaus integriert maximal 1% der Entwürfe aus der Studie von Riebenbauer et al. (2025) Nachhaltigkeitskompetenzen. In der vorliegenden Interventionsstudie sind es in der Experimentalgruppe je nach Merkmal zwischen 72% und 100% der Entwürfe, in der Kontrollgruppe immer noch zwischen 39% und 100%. Ähnliche Befunde treffen auch auf die beiden Subkategorien der Digitalisierung zu. Bei der methodischen Digitalisierung plant die Experimentalgruppe häufiger Möglichkeiten zur eigenständigen Auseinandersetzung mit der Technologie sowie zum kollaborativen Arbeiten als die Kontrollgruppe und auch als die Entwürfe aus der Vergleichsstudie. Auf der Ebene der inhaltlichen Digitalisierung werden insgesamt bei der Experimentalgruppe häufiger allgemeine und auch berufsspezifische Technologien integriert und es wird häufiger über die Ergebnisse der Technologien kommuniziert (z.B. über Auszüge aus dem ERP-System). Allerdings wird die Entwicklung von technologischen Lösungen genauso selten diskutiert, wie in der Vergleichsstudie und auch die Einsatzbedingungen der Technologie werden nur in der Kontrollgruppe in einem Entwurf thematisiert. Insgesamt haben die Lernenden beider Gruppen also erfolgreich Nachhaltigkeit und Digitalisierung in den Planungsprozess ihres Unterrichts integriert. Die Dimensionen wurden damit zum Bestandteil des *Planen Lernens*.

Beantwortung von Teilfrage 3

Verknüpft man nun die Ergebnisse des dritten Beitrags mit diesen Erkenntnissen, dann ergeben sich folgende Antworten für Teilfrage 3 „*Wie wirksam ist die Nutzung der KI-Plattform als Feedbackgeberin im Vergleich zu rein menschlichem Feedback?*“:

8. Die Nutzung von KI-basiertem Feedback ist bei 14 von 57 Merkmalen wirksamer als rein menschliches Feedback. Bei allen anderen Merkmalen ist das KI-Feedback mindestens genauso lernförderlich, wie menschliches Feedback, außer beim Merkmal *Einsatzbedingungen der Technologie*. Dort wirkt das menschliche Feedback besser. Zur lernförderlichen Verarbeitung von KI-gestütztem Feedback sollten Erfahrungsräume eröffnet werden, die den Dialog zwischen Menschen ermöglichen. Daraus ergeben sich Diskussionsanlässe über KI-generiertes Feedback und weitere Lernmöglichkeiten.
9. Bei den Metadimensionen zeigen sich nur bei der KI-Literacy positive Effekte der Intervention beim *Erkennen von KI* und bei der *Emotionsregulation bei der KI-Nutzung*. Die anderen Metadimensionen zeigen keine Gruppenunterschiede. Die Nutzung digitaler Technologien bedarf somit eines durchdachten didaktischen Konzepts, da durch die Nutzung der Technologie allein noch kein Zuwachs digitaler Kompetenzen oder KI-Literacy auszumachen ist. Auch ist der Einsatz von KI nicht automatisch förderlich für die Motivation.
10. Studierende, die bereits zu Beginn des Semesters starke Anfangswerte bei den Entwürfen hatten, zeigen über das Semester hinweg nur noch geringes Wachstum. KI-Plattformen als Feedbackmedium zu nutzen, kann zeitliche Ressourcen bei den Lehrenden freisetzen. Im Sinne der Zusammenarbeit von KI und Mensch sollten menschliches und KI-Feedback produktiv kombiniert werden: So können speziell leistungsstarke Lernende ein erstes Feedback durch EDDA erhalten, welches durch ein tiefergreifendes Feedback der Seminarleitung ergänzt werden kann, damit auch diese Lernenden weiter professionalisiert werden können.

Zusammenfassung und Diskussion der Ergebnisse von Beitrag 4

Der dritte Beitrag half zwar, die Frage nach den Effekten von EDDA zu beantworten, gibt aber noch keinen Aufschluss darüber, wie genau die Studierenden mit ihr interagieren und ob es Unterschiede zwischen den Interaktionen gibt. Dies ist deshalb interessant, weil die Interaktion mit Technologien sich im Verlauf der Expertiseentwicklung verändert und unterschiedliche Lernerfolge und Problemlöseprozesse entstehen (C. W. Mayer et al., 2023).

Der letzte, vierte Beitrag dieser Arbeit schließt diese Lücke, indem er sich mit der Interaktion von Studierenden mit EDDA und den Effekten der Interaktionstypen auf die Unterrichtsplanungen beschäftigt. Hierfür wurden 32 Studierende bei der Interaktion mit EDDA mittels eines mobilen Eye Tracking-Geräts zu zwei Erhebungszeitpunkten begleitet und hinsichtlich ihrer Interaktion mit AOIs (in t0 zum Thema Nachhaltigkeit, in t1 zur Digitalisierung) ausgewertet.

Für die Ableitung der Interaktionstypen wurde die statistische Methode der Sequenzanalyse genutzt, welche es ermöglicht, charakteristische Abfolgen in den einzelnen Datensätzen der Studierenden zu identifizieren und diese entsprechend zu clustern. Die Analyse wurde zunächst für beide Zeitpunkte getrennt durchgeführt, dann wurden die AOIs aggregiert, um eine übergeordnete Typologie zu entwickeln und zu überprüfen, ob die Abgrenzungen zwischen den Typen aus der getrennten Analyse erhalten bleiben. Tatsächlich konnten im Beitrag sowohl Typen auf Ebene der einzelnen Zeitpunkte als auch auf Ebene der aggregierten Daten identifiziert werden. Für alle drei Analysen (t0, t1, aggregiert) konnte eine gemeinsame Typologie entwickelt werden, welche die Studierenden in feedbackorientierte, ergebnisorientierte und gemischte Typen unterteilt. Tabelle 5 zeigt die Unterschiede zwischen den drei Typen auf.

Tabelle 5. Überblick über die verschiedenen Interaktionstypen

Feedbackre-präsentation	Feedbackorientierter Typ (Schriftlicher Typ)	Ergebnisorientierter Typ (Visueller Typ)	Gemischter Typ
Schriftlich	Hoch	Niedrig	Mäßig
Visuell	Niedrig	Hoch	Mäßig
Kombination von Feedback und Ergebnissen	Verbringt die meiste Zeit im schriftlichen Feedback	Verbringt die meiste Zeit in den Ergebnissen	Wählt schriftliches Feedback anhand der Ergebnisse aus oder balanciert zwischen Feedback und Ergebnissen

Es können aufgrund der Struktur von EDDA zwei Erklärungsansätze für die Typenbildung identifiziert werden: einerseits eine Präferenz im Feedbackumfang, andererseits in der Repräsentation. Denn die Empfehlungen beinhalten eine Definition der Kategorie, ein Beispiel sowie Verbesserungsvorschläge, während die visuellen Ergebnisse nur die Bewertung der Umsetzung der Kategorie beinhaltet. Es kann somit einerseits ein Zusammenhang damit bestehen, ob die Person konkrete Hinweise zur Überarbeitung haben möchte, oder ob sie nur sehen möchte, was sie gut umgesetzt hat und was nicht (*Feedbackumfang*).

Andererseits kann aber auch lediglich eine Präferenz für schriftliche oder grafische Bereiche vorliegen, die von den Studierenden besser verarbeitet werden können (*Feedbackrepräsentation*). Dies müsste in einer weiteren Studie überprüft werden. Neben der Entwicklung der Typen wurde auch deren Effekt auf die Unterrichtsplanungen der Personen erhoben. Hier zeichnete sich ab, dass der feedbackorientierte und der gemischte Typ regelmäßig die Kategorien besser umsetzt als der ergebnisorientierte Typ. Dieses Ergebnis wurde so erwartet, da der ergebnisorientierte Typ anhand der visuellen Ergebnisse keine konkreten Überarbeitungshinweise erhält.

Hinsichtlich der Fragestellung „*Wie interagieren Studierende mit der Plattform und welche Nutzungstypen gibt es?*“ konnten distinkte Typen identifiziert werden können, die dahingehend variieren, wie intensiv sie sich mit den einzelnen Bestandteilen von EDDA auseinandersetzen. Während manche Studierende jede Feedbackbox intensiv lesen und sich detailliert mit dem Feedback auseinandersetzen, überfliegen manche die ausführlichen Texte nur und betrachten eher die grafische Darstellung ihrer bisherigen Ergebnisse und die Punktzahl, die sie erreicht haben. Darüber hinaus kann ein Mischtyp identifiziert werden, welcher die Ergebnisse betrachtet und anhand dieser die Empfehlungen auswählt, mit denen er oder sie daraufhin interagiert.

Die Ergebnisse des vierten Beitrags deuten darauf hin, dass es im Kontext des Einsatzes von EDDA zweierlei zu berücksichtigen gibt: Einerseits sollten die Individualisierungsoptionen von EDDA weiter ausgebaut werden, vor allem die grafische Benutzeroberfläche. Andererseits sollten die Studierenden weiter in ihrem Umgang mit Feedback professionalisiert werden, um Lehr-Lernprozesse zu verbessern. Studierende sollten zwar ihre Präferenz hinsichtlich des grafischen oder schriftlichen Feedbacks ausleben können, indem sie z.B. eine Kurzfassung der Überarbeitungshinweise erhalten, doch sollten sie auch zur Auseinandersetzung mit dem Feedback bewogen werden. So ist die Diagnostik des Feedbacktyps potenziell hilfreich, um individuelle Lernanlässe zur Integration von Feedback zu entwickeln. Auch die Entwicklung eines personalisierbaren Dashboards könnte eine hilfreiche Unterstützung in der Personalisierung der Benutzererfahrung sein. Nutzer:innen von EDDA könnten die modularisierten Feedbackseiten nach ihren eigenen Präferenzen zusammensetzen und die Anzahl an grafischen Darstellungen oder die Textlänge anpassen. Außerdem sollten die Nutzer:innen wählen können, ob sie sich eine Punktzahl anzeigen lassen können. Diese deutet zwar an, ob die Umsetzung in eine bestimmte, gewünschte Richtung geht, doch da nicht jedes didaktische Kriterium in jedem Unterricht integrierbar

ist, kann sie vom eigentlichen Fokus der Plattform ablenken. Dies wird unterstützt vom ergebnisorientierten Typ, der sich stark auf die Bewertungen konzentriert. Da dieser Interaktionstyp häufiger schlechtere Ergebnisse bei den Unterrichtsplanungen hat als der feedbackorientierte oder der gemischte Typ, wäre es hilfreich, den ergebnisorientierten Typ mittels Plattformgestaltung dazu zu bewegen, sich intensiver mit dem Feedback auseinanderzusetzen.

Beantwortung von Teilfrage 4

Unter Berücksichtigung der aus dem vierten Beitrag entwickelten Ergebnisse ergibt sich für Teilfrage 4 „*Wie interagieren Studierende mit der Plattform und welche Nutzungstypen gibt es?*“ Folgendes:

11. Es konnten drei verschiedene allgemeine Interaktionstypen identifiziert werden: der feedbackorientierte Typ, der ergebnisorientierte Typ und der gemischte Typ. Sie weisen jeweils unterschiedliche Präferenzen für Feedbackumfang und -repräsentation auf. Die Interaktionstypen können auch nach Feedbackrepräsentation unterschieden werden; hier liegt der Unterschied entweder in der Bevorzugung von Text, von Grafiken oder einer Mischung aus beidem. Für beide Erklärungsansätze bedeutet dies, dass die *kognitive Auslastung der Lernenden* bei der Interaktion berücksichtigt werden sollte (R. E. Mayer & Fiorella, 2021b). Der Umfang sollte so ausführlich wie nötig, aber so gering wie möglich gehalten werden sollte. Definitionen oder Erläuterungen sollten zwar verfügbar sein, aber der Fokus des Feedbacks sollte auf der Einordnung der Ergebnisse in die Kategorie und den Überarbeitungsempfehlungen liegen. Nutzer:innen sollten ferner die gewünschte Feedbackrepräsentation wählen können.
12. Studierende des feedbackorientierten und des gemischten Interaktionstyps schnitten bei einigen Planungsmerkmalen signifikant besser ab als Studierende des ergebnisorientierten Typs. Um allen Nutzer:innen einen lernförderlichen Umgang mit der Technologie zu ermöglichen, sollte die *Benutzeroberfläche* der KI daher so individualisierbar wie möglich sein, um verschiedenen Interaktionstypen eine lernförderliche Nutzung von EDDA zu ermöglichen und weniger lernförderliche Typen zu einer Veränderung ihrer Interaktion zu bewegen. Daraus ergibt sich die Berücksichtigung verschiedener Präferenzen der Feedbackrepräsentation.

13. Studierende vom ergebnisorientierten Typ sollten wegen der weniger wahrgenommenen Überarbeitungshinweise in den Empfehlungen der KI in der Integration des Feedbacks besonders *eng betreut* werden, um zu überprüfen, dass sie auch Überarbeitungsempfehlungen für ihre Unterrichtsplanungen wahrgenommen haben. Werden Feedbackprozesse auf die KI ausgelagert und nicht durch ein passendes Seminarkonzept begleitet, können hier Defizite entstehen.

Übergreifende Diskussion der Ergebnisse

Die fünf hier diskutierten Beiträge haben mit der Beantwortung der Teilfragestellungen zur Beantwortung der übergeordneten Fragestellung beigetragen, die danach fragt, wie eine KI-Plattform in der wirtschaftsberuflichen Lehrkräftebildung gestaltet und genutzt werden kann, um Studierende bei der Unterrichtsplanung zu aktuellen Themen zu unterstützen.

Insgesamt lässt sich hierzu festhalten, dass konkrete Ansätze zur Entwicklung der Plattform, zu relevanten Inhaltsbereichen und zur Feedbackgestaltung identifiziert werden konnten. So muss die Plattform die Unterrichtsplanungen der Studierenden in aktuelle wissenschaftliche Erkenntnisse einordnen und eine Rückmeldung produzieren, welches den Kriterien für lernförderliches Feedback gerecht wird (Narciss, 2013; Wisniewski et al., 2019). Denn diese haben auch im digitalen Raum ihre Gültigkeit. Während im Kontext der Learning Analytics-Forschung (z.B. Papa et al., 2020) auf eine Erweiterung des Verständnisses digitaler Lehr-Lernprozesse hingewiesen wird, haben diese Erkenntnisse für den Anwendungsbereich dieser Arbeit keine relevanten Implikationen. Zwar wird das Feedback digital zur Verfügung gestellt, die Einarbeitung des Feedbacks erfolgt aber hauptsächlich analog. Somit sind Erkenntnisse der traditionellen Feedbackforschung anwendbar und werden auch im Kontext dieser Arbeit als tragfähig identifiziert und durch die Erkenntnisse aus den Beiträgen 3 und 4 durch neue Ergebnisse zu analytischen KI-Plattformen ergänzt.

Die Unterrichtsplanungen der Interventionsstudie, die den Kern dieser Arbeit bildet, wurden im Kontext von Nachhaltigkeit und Digitalisierung im Rechnungswesen- und Controllingunterricht angefertigt. Bei den meisten Kategorien der beiden Bezugsbereiche wirkte das KI-Feedback von EDDA genauso gut wie das menschliche Feedback, in einigen Bereichen dominierten entweder das KI-Feedback oder das menschliche Feedback. Während auf der Ebene der Digitalisierung als Lernmethode die Experimentalgruppe bei einem Merkmal, dem *Technologieeinsatz*, besser abschnitt, performte die Kontrollgruppe bei der

Digitalisierung als Lerngegenstand beim Merkmal *Rahmenbedingungen des Einsatzes* besser. Obwohl die Studierenden ein umfassendes Nachhaltigkeitswissen vorweisen, gelingt es ihnen nicht bedeutend besser, das zielgerichtete Feedback von EDDA in ihren Unterricht zu integrieren. Zwei mögliche Erklärungsansätze wären, dass es weiterhin auf a) die individuellen Präferenzen der Studierenden und b) das Wissen der Seminarleitung ankommt, ob Nachhaltigkeit integriert wird oder nicht. Zumindest wird aber die Integration von Nachhaltigkeit und Digitalisierung sowohl durch menschliches als auch durch das KI-basierte Feedback von EDDA befördert.

Es zeigt sich insgesamt, dass die vorliegende Arbeit die bestehenden Forschungsarbeiten zu KI-Feedback im Allgemeinen und zur Unterrichtsplanung im Speziellen erweitert, ergänzt und ihnen eine neue Perspektive verleiht. So ist KI-basiertes Feedback abseits bekannter großer Sprachmodelle und generativen KI-Systemen möglich, umsetzbar und gleichzeitig auch effektiv gegenüber traditionellem Feedback. Der Forschungsstand zur Unterrichtsplanung mit KI wird erweitert um analytische KI-Plattformen, die reproduzierbares Feedback erzeugen, welches auf rückverfolgbaren Quellen basiert und domänenspezifisch ist – zwei Kriterien, denen Feedback auf Basis von generativen KI-Systemen nicht durchgängig gerecht werden kann.

Doch auch diese Arbeit weist einige Grenzen auf. In Ergänzung zu den Limitationen der einzelnen Beiträge 1 bis 4, ergeben sich auch für den übergeordneten Kontext dieser Arbeit Herausforderungen, die im Folgenden dargestellt und zur Ableitung von weiteren Forschungsdesiderata herangezogen werden.

4.2 Limitationen und Forschungsdesiderata

Limitationen der Arbeit

Die vermutlich weitreichendste Limitation dieser Arbeit liegt im geringen Stichprobenumfang der Beiträge 3 und 4. Die zentralen Erkenntnisse und somit auch die herausgearbeiteten Implikationen können damit keine Allgemeingültigkeit beanspruchen. Für eine Überprüfung der Ergebnisse sollten somit weiter passende Seminare gewonnen werden, um die Stichprobe schrittweise zu vergrößern. Dies wird auch im Rahmen der Seminare an der Universität Hamburg geschehen.

Außerdem bilden die verwendeten Erhebungsinstrumente für die Erfassung der Metadimensionen in der Interventionsstudie eine deutliche Limitation. Nur sieben der 24 genutzten Konstrukte erwiesen sich als reliabel im Längsschnitt und produzierten interpretierbare Ergebnisse. Damit gibt es ein deutliches Desiderat zur Entwicklung tragfähiger, für den Längsschnitt geeigneter Erhebungsinstrumente, die die Entwicklung von KI-Literacy, Einstellungen zu KI, digitalen Kompetenzen und Motivation adäquat abbilden können. Ferner sollten die Ergebnisse von Studien, die diese Messinstrumente ohne kritische Überprüfung des Messmodells im Längsschnitt einsetzen, mit Vorsicht interpretiert werden.

Auch die genutzte Software EDDA ist nicht ohne Limitationen. Primär sind dies die schlechte Datenverfügbarkeit von Unterrichtsentwürfen zum Training der KI, die Unausgewogenheit der Daten aufgrund des Ausbildungsabschnitts der Ersteller:innen der Entwürfe sowie die Komplexität der Daten aufgrund der Textlänge und schwer abgrenzbarer Kategorien. So ist die Kodierung der Oberkategorien herausfordernd, wodurch teilweise eine verbesserungswürdige Inter-Coder-Reliabilität bedingt ist (Riebenbauer et al., 2025). Aufgrund des kleinen Trainingsdatensatzes der KI ist die Erkennung von Mustern (siehe Abschnitt 2.2) ebenfalls herausfordernder. In ihrer Studie über den Trainingsdatensatz stellten Riebenbauer et al. (2025) eine große Unausgewogenheit der Daten fest, wodurch die KI nicht alle Ausprägungen gleich gut erkennt. Dieser Umstand ergibt sich daraus, dass die Entwürfe von Studierenden stammen, die an unterschiedlichen Punkten des Professionalisierungsprozesses stehen und entsprechend unterschiedlich weit beim *Planen lernen* sind. Eine kritische Reflexion der Rückmeldungen von EDDA durch die Nutzer:innen ist somit zwingend notwendig. Hierfür sind jedoch umfangreiche Informationen zur Leistungsfähigkeit und Sicherheit der KI-Ergebnisse abrufbar. Es bedarf daher aber auch eines strukturierten Einsatzes in der universitären Lehre.

Die Unterrichtsentwürfe wurden anhand von Kodierleitfäden in der Umsetzung der didaktischen Merkmale kodiert. Eine qualitative Auswertung, z.B. hinsichtlich des Zielbezugs, der Transformation des Ziels oder der erfolgreichen didaktischen Transformation der Analysebereiche der Unterrichtsplanungen steht noch aus. Erkenntnisse hierzu würden die Ausführungen in dieser Arbeit sicherlich ansprechend illustrieren, für die allgemeine Wirksamkeit wurden sie jedoch nicht als Zieldimension herangezogen. Dennoch sollten die Planungen auch qualitativ analysiert werden, um Rückschlüsse auf weitere Herausforderungen bei der Unterrichtsplanung zu ermöglichen. Die Integration von Aspekten wie der Individualisierung des geplanten Unterrichts oder der fachliche Bezug im Sinne der Sachanalyse wurden in der Anlage der Arbeit und den einzelnen Studien zwar inhaltlich berücksichtigt, aber konnten im Rahmen der KI nicht adäquat abgebildet werden. Hierfür sind Anpassungen zum Zeitpunkt der Erstellung dieser Arbeit in der Umsetzung und sollen in Form von EDDAplus als Plug-In zur Verfügung gestellt werden.

Darüber hinaus wurden auch keine qualitativen Analysen angestellt, welche konkreten Themen bei Digitalisierung und Nachhaltigkeit im Rechnungswesen und Controlling berücksichtigt wurden. So liegen z.B. auch nach dieser Arbeit keine Ergebnisse dazu vor, welche Nachhaltigkeitsthemen präferiert geplant wurden und ob diese dem Bereich Rechnungswesen oder dem Controlling zuzuschreiben sind.

Ferner wurden die geplanten Unterrichte nicht in ihrer tatsächlichen Realisierbarkeit überprüft. Fokus dieser Arbeit ist das *Planen lernen*. Im Sinne einer vollständigen Handlung hätte die Durchführung einen relevanten Reflexionsanlass für die Studierenden geboten, der zur weiteren Überarbeitung der Planung hätte genutzt werden können, um sie weiter zu professionalisieren. Hier könnten Forschungen angeschlossen werden, die den Planungsprozess bis hin zur Realisierung und der darauffolgenden Reflexion begleiten, evaluieren und in den Kontext des Planungsprozesses an sich setzen.

Nachdem die Limitationen dieser Arbeit dargestellt wurden, wird im nachfolgenden Kapitel ein abschließendes Fazit gezogen.

5 Fazit

Die vorliegende Arbeit liefert wesentliche Erkenntnisse für die Nutzung von KI-basiertem Feedback bei der Unterrichtsplanung bei angehenden Berufsschullehrkräften. Bisherige

Studien zur Unterrichtsplanung in der Domäne ergaben Desiderata bei der Beförderung der UPK (Riebenbauer et al., 2024, 2025; Söll, 2021). Auf Basis aktueller Erkenntnisse zur lernförderlichen Gestaltung von digitalem Feedback konnte in diesem Kontext die Nutzbarkeit einer analytischen KI-Plattform für die Unterrichtsplanung betrachtet werden. Zwar wurde die Beförderung der UPK an sich nicht beforscht, jedoch konnte in einer längsschnittlichen Interventionsstudie im Experimental-Kontrollgruppendesign nachgewiesen werden, dass sich die Nutzung der KI-Plattform positiv auf die Umsetzung einiger Planungsdimensionen auswirkt. Dabei spielt die Art der Interaktion mit der Plattform eine wesentliche Rolle in der Integration der planerischen Merkmale – feedbackorientierte Interaktionstypen und jene, die ausgeglichen mit Empfehlungen und Ergebnissen interagieren, zeigen bessere Ergebnisse als ergebnisorientierte Interaktionstypen. Somit hängt der Erfolg der Nutzung der KI einerseits von der Intensität der Auseinandersetzung mit den Überarbeitungsvorschlägen ab (*Feedbackumfang*), andererseits auch von der *Feedbackrepräsentation* (visuell, schriftlich oder gemischt). Die KI-Literacy der Teilnehmenden wird nicht umfangreich befördert, was darauf hindeutet, dass ein integrativer Ansatz für die Entwicklung entsprechender Kompetenzen auch bei Lehramtsstudierenden notwendig ist. Ebenso konnte nachgewiesen werden, dass der bloße Technologieeinsatz nicht automatisch die Motivation positiv beeinflusst. Zukünftige Forschung sollte sich hier auf die Analyse der Emotionen während der Interaktion mit der KI konzentrieren, um zu ergründen, wie die Zuwächse in der Emotionsregulation durch KI-Nutzung konkret aussehen können und ob Zusammenhänge zur Motivation bestehen. Hier können weitere Eye Tracking-Studien wertvolle Einblicke liefern.

Bei der Integration von Aspekten der Nachhaltigkeit und Digitalisierung als Bezugsbereich des Unterrichtsthemas sind sowohl das KI-basierte als auch das menschliche Feedback wirksam. Damit werden sowohl Digitalisierung als auch Nachhaltigkeit stärker in die Unterrichtsentwürfe integriert, als dies die Studie von Riebenbauer et al. (2025) vermuten ließe, die insgesamt deutliche Desiderata zur Umsetzung der Kriterien formuliert. Damit ist es gelungen, explizite didaktische Bezüge zu Nachhaltigkeit und Digitalisierung herzustellen, auch wenn diese nicht durch die Intervention, sondern durch das Design der Inhalte hervorgerufen wurden.

Für die Nutzung von KI als Feedbackmedium in der Unterrichtsplanung wurden 13 Implikationen entwickelt, die nicht nur als zentrale Erkenntnisse dieser Arbeit gelesen werden können, sondern auch als Handlungsempfehlungen für Seminarleitungen, Lehrkräfte und

andere Praktiker:innen gelesen werden können, die EDDA oder eine ähnliche Technologie für ihre Lehre einsetzen möchten. Damit fungiert diese Arbeit als wertvolle Unterstützung zur Orientierung in einer Zeit, in der die technologischen Entwicklungen und der Publikationsstand zur Thematik starken Veränderungen gegenüberstehen.

Literaturverzeichnis

- Adiguzel, T., Kaya, M. H. & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), ep429. <https://doi.org/10.30935/cedtech/13152>
- Aebli, H. (2011). *Zwölf Grundformen des Lehrens: Eine Allgemeine Didaktik auf psychologischer Grundlage. Medien und Inhalte didaktischer Kommunikation, der Lernzyklus*. Klett-Cotta.
- Aepli, M., Angst, V., Iten, R., Kaiser, H., Lüthi, I. & Schweri, J. (2017). *Die Entwicklungen der Kompetenzanforderungen auf dem Arbeitsmarkt im Zuge der Digitalisierung: Schlussbericht*. Eidgenössisches Hochschulinstitut für Berufsbildung.
- Aggarwal, C. C. (2018). Machine Learning for Text: An Introduction. In C. C. Aggarwal (Hrsg.), *Machine learning for text* (S. 1–16). Springer. https://doi.org/10.1007/978-3-319-73531-3_1
- Aprea, C. (2014). Unterrichtsplanung als Designaufgabe. *Journal für LehrerInnenbildung*, 14(4), 47–50.
- Aprea, C. (2018). Instructionsdesign und Unterrichtsplanung. In H. Niegemann & A. Weinberger (Hrsg.), *Lernen mit Bildungstechnologien: Praxisorientiertes Handbuch zum intelligenten Umgang mit digitalen Medien* (S. 1–19). Springer.
- Aprea, C. (2020). Instruktionsdesign und Unterrichtsplanung. In H. M. Niegemann & A. Weinberger (Hrsg.), *Handbuch Bildungstechnologie: Konzeption und Einsatz digitaler Lernumgebungen : mit 141 Abbildungen und 17 Tabellen* (S. 171–190). Springer.
- Backfisch, I., Lachner, A., Hische, C., Loose, F. & Scheiter, K. (2020). Professional knowledge or motivation? Investigating the role of teachers' expertise on the quality of technology-enhanced lesson plans. *Learning and Instruction*, 66, 101300. <https://doi.org/10.1016/j.learninstruc.2019.101300>
- Baetge, J., Kirsch, H.-J. & Thiele, S. (2024). *Bilanzen*. IDW Verlag GmbH.
- Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P. & Biemans, H. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review*, 37, 100489. <https://doi.org/10.1016/j.edurev.2022.100489>
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>

- Beaumont, C., O'Doherty, M. & Shannon, L. (2011). Reconceptualising assessment feedback: a key to improving student learning? *Studies in Higher Education*, 36(6), 671–687. <https://doi.org/10.1080/03075071003731135>
- Beckmann, M. & Schaltegger, S. (2014). Unternehmerische Nachhaltigkeit. In H. Heinrichs & G. Michelsen (Hrsg.), *Nachhaltigkeitswissenschaften* (S. 321–367). Springer. https://doi.org/10.1007/978-3-642-25112-2_9
- Beheshti, A. (2023). Empowering Generative AI with Knowledge Base 4.0: Towards Linking Analytical, Cognitive, and Generative Intelligence. *2023 IEEE International Conference on Web Services (ICWS)* (S. 763–771). IEEE. <https://doi.org/10.1109/ICWS60048.2023.00103>
- Berding, F. (2019). *Rechnungswesenunterricht. Grundvorstellungen und ihre Diagnose* [Dissertation]. Hampp.
- Berding, F., Jahncke, H. & Albersmann, M. (2020). Dynamische und interaktive Arbeitsblätter im Rechnungswesenunterricht.: Eine Wirkungsstudie im Kontrollgruppendesign. In F. Berding, H. Jahncke & A. Slopinski (Hrsg.), *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive* (S. 277–317). Springer.
- Berding, F., Slopinski, A. & Frerichs, R. (2020). Auszubildende als zukünftige Change Agents for Sustainable Innovations. *Betriebswirtschaftliche Forschung und Praxis*, 72(3), 313–337.
- Berding, F., Slopinski, A., Gebhardt, R., Kalmutzke, F., Schröder, T., Rebmann, K., Heubischl, S. & Schlömer, T. (2018). Innovationskompetenz für nachhaltiges Wirtschaften und Instrumente ihrer Erfassung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 114(1), 47–84. <https://doi.org/10.25162/zbw-2018-0002>
- Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A. & Bethel, E. C. (2009). A Meta-Analysis of Three Types of Interaction Treatments in Distance Education. *Review of Educational Research*, 79(3), 1243–1289. <https://doi.org/10.3102/0034654309333844>
- Beutel, G., Geerits, E. & Kielstein, J. T. (2023). Artificial hallucination: GPT on LSD? *Critical Care*, 27(1), 148. <https://doi.org/10.1186/s13054-023-04425-6>
- Bleiber, R. (2019). *Digitalisierung in der Finanzbuchhaltung: Vom Status quo in die digitale Zukunft*. Haufe.

- Bloemen, A. (2011). *Lernaufgaben in Schulbüchern der Wirtschaftslehre. Analyse, Konstruktion und Evaluation von Lernaufgaben für die Lernfelder industrieller Geschäftsprozesse. Schriften zur Berufs- und Wirtschaftspädagogik*. Hampp.
- Börner, M. & Brötz, R. (2015). Nachhaltige Entwicklungsziele in der kaufmännischen Berufsausbildung – Situationsbeschreibung und Perspektiven einer naturgemäßen Berufsbildung. In R. Brötz & F. Kaiser (Hrsg.), *Kaufmännische Berufe – Charakteristik, Vielfalt und Perspektiven* (Berichte zur beruflichen Bildung, S. 245–264).
- Bostrom, N. & Yudkowsky, E. (2014). The Ethics of Artificial Intelligence. In K. Frankish & W. Ramsey (Hrsg.), *The Cambridge Handbook of Artificial Intelligence* (S. 316–334). Cambridge University Press.
- Braun, S. & Senger, E. (2022). Nachhaltigkeitsreporting 4.0. In B. Schwager (Hrsg.), *Management-Reihe Corporate Social Responsibility. CSR und Nachhaltigkeitsstandards* (S. 119–132). Springer. https://doi.org/10.1007/978-3-662-64913-8_7
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 177–212). Hogrefe Verlag für Psychologie. <https://scholar.google.com/citations?user=ouxqbwgaaaaj&hl=de&oi=sra>
- Bromme, R. & Seeger, F. (1979). *Unterrichtsplanung als Handlungsplanung: eine psychologische Einführung in die Unterrichtsvorbereitung*. Scriptor. <https://ixtheo.de/record/02371350x>
- Brötz, R., Kock, A., Annen, S. & Schaal, T. (2015). Gemeinsamkeiten und Unterschiede der kaufmännischen Ausbildungsberufe. In R. Brötz & F. Kaiser (Hrsg.), *Kaufmännische Berufe – Charakteristik, Vielfalt und Perspektiven* (Berichte zur beruflichen Bildung, S. 91–106).
- Brown, J. S., Collins, A. & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, 18(1), 32. <https://doi.org/10.2307/1176008>
- Brückner, S. & Zlatkin-Troitschanskaia, O. (2024). Effects of teacher students' study progress on their gaze behavior while solving of an economics knowledge test. *Empirical Research in Vocational Education and Training*, 16(1). <https://doi.org/10.1186/s40461-024-00172-2>
- Bundesinstitut für Berufsbildung (Hrsg.). (2021). *VIER SIND DIE ZUKUNFT: DIGITALISIERUNG. NACHHALTIGKEIT. RECHT. SICHERHEIT. Die modernisierten*

Standardberufsbildpositionen anerkannter Ausbildungsberufe.

<https://www.bibb.de/dienst/veroeffentlichungen/de/publication/download/17281>

Bundesinstitut für Berufsbildung. (2024). *Datenreport zum Berufsbildungsbericht 2024.*

Bundesregierung. (2020). *Strategie Künstliche Intelligenz der Bundesregierung: Fortschreibung 2020.*

Buxmann, P. & Schmidt, H. (2021). Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens. In P. Buxmann & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg* (S. 3–26). Springer.

Cai, Z., Gui, Y., Mao, P., Wang, Z [Zhikeng], Hao, X., Fan, X. & Tai, R. H. (2023). The effect of feedback on academic achievement in technology-rich learning environments (TREs): A meta-analytic review. *Educational Research Review*, 39, 100521.

<https://doi.org/10.1016/j.edurev.2023.100521>

Celik, I., Gedrimiene, E., Siklander, S. & Muukkonen, H. (2022). The affordances of artificial intelligence-based tools for supporting 21st-century skills. *Australasian Journal of Educational Technology*, 40(3), 1–20. <https://doi.org/10.14742/ajet.9069>

Chiapello, E. (2009). Die Konstruktion der Wirtschaft durch das Rechnungswesen. In R. Diaz-Bone & G. Krell (Hrsg.), *Diskurs und Ökonomie: Diskursanalytische Perspektiven auf Märkte und Organisationen* (S. 125–149). VS. https://doi.org/10.1007/978-3-531-91914-0_6

Cho, C. H. & Costa, E. (2024). Sustainability accounting education: challenges and outlook. *International Journal of Sustainability in Higher Education*, 25(7), 1412–1425. <https://doi.org/10.1108/IJSHE-02-2024-0152>

Chowdhary, K. (2020). *Fundamentals of artificial intelligence*. Springer. <https://doi.org/10.1007/978-81-322-3972-7>

Corsten, M. (2020). *Betriebswirtschaftslehre*. Franz Vahlen. <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6991284>

Dai, W., Tsai, Y.-S., Lin, J., Aldino, A., Jin, F., Li, T. & Gašević, D. (2024). Assessing the Proficiency of Large Language Models in Automatic Feedback Generation: An Evaluation Study. *Computers and Education: Artificial Intelligence*, 7, 100299, 1–10. <https://doi.org/10.35542/osf.io/s7dvy>

Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238.

- Dengler, K. & Matthes, B. (2021). *Folgen des technologischen Wandels für den Arbeitsmarkt: Auch komplexere Tätigkeiten könnten zunehmend automatisiert werden.* (IAB-Kurzbericht Nr. 13).
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* <https://arxiv.org/pdf/1810.04805>
- Donnermann, M., Schaper, P. & Lugin, B. (2022). Social Robots in Applied Settings: A Long-Term Study on Adaptive Robotic Tutors in Higher Education. *Frontiers in robotics and AI*, 9, 831633. <https://doi.org/10.3389/frobt.2022.831633>
- Dörner, D. (1979). *Problemlösen als Informationsverarbeitung.* Kohlhammer.
- Dornis, T. W. (2025). Generative AI, Reproductions Inside the Model, and the Making Available to the Public. *International Review of Intellectual Property and Competition Law*, 1–30. <https://doi.org/10.1007/s40319-025-01582-9>
- Eklund, R., Sellberg, C. & Osvalder, A.-L. (2020). From Tacit Knowledge to Visual Expertise: Eye-Tracking Support in Maritime Education and Training. In S. Nazir, T. Ahram & W. Karwowski (Hrsg.), *Advances in Intelligent Systems and Computing: Bd. 1211, Advances in Human Factors in Training, Education, and Learning Sciences: Proceedings of the AHFE 2020 Virtual Conference on Human Factors in Training, Education, and Learning Sciences.* Springer. https://doi.org/10.1007/978-3-030-50896-8_39
- Er, E., Akçapınar, G., Bayazit, A., Noroozi, O. & Banihashem, S. K. (2024). Assessing student perceptions and use of instructor versus AI-generated feedback. *British Journal of Educational Technology*, 56(3), 1074–1091. <https://doi.org/10.1111/bjet.13558>
- Ertel, W. (2016). *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung.* Springer. <https://doi.org/10.1007/978-3-658-13549-2>
- Ertel, W. (2025). *Grundkurs Künstliche Intelligenz.* Springer. <https://doi.org/10.1007/978-3-658-44955-1>
- Euler, D. (2018). Bildung in Zeiten der Digitalisierung. *Zeitschrift für Berufs-und Wirtschaftspädagogik*, 114(2), 179–190.
- Euler, D. (2019). Kompetenzorientierung in der beruflichen Bildung. In M. Apelt, I. Bode, R. Hasse, U. Meyer, V. V. Groddeck, M. Wilkesmann & A. Windeler (Hrsg.), *Springer Reference Sozialwissenschaften. Handbuch Organisationssoziologie* (S. 1–13). Springer. https://doi.org/10.1007/978-3-658-19372-0_18-1
- Euler, D. & Wilbers, K. (2019). Berufsbildung in digitalen Lernumgebungen. In M. Apelt, I. Bode, R. Hasse, U. Meyer, V. V. Groddeck, M. Wilkesmann & A. Windeler (Hrsg.),

- Springer Reference Sozialwissenschaften. Handbuch Organisationssoziologie* (S. 1–13). Springer. https://doi.org/10.1007/978-3-658-19372-0_34-1
- Feuerriegel, S., Hartmann, J., Janiesch, C. & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Fiorella, L. & Mayer, R. E. (2021a). Principles Based on Social Cues in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 277–285). Cambridge University Press. <https://doi.org/10.1017/9781108894333.029>
- Fiorella, L. & Mayer, R. E. (2021b). Principles for Reducing Extraneous Processing in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 185–198). Cambridge University Press. <https://doi.org/10.1017/9781108894333.019>
- Freiling, T. & Porath, J. (2020). Digitalisierung des Lernens – Implikationen für die berufliche Bildung. In T. Freiling, R. Conrads, A. Müller-Osten & J. Porath (Hrsg.), *Zukünftige Arbeitswelten* (S. 205–225). Springer. https://doi.org/10.1007/978-3-658-28263-9_10
- Frötschl, C. (2015). *Enterprise Resource Planning Systeme im kaufmännischen Unterricht* [Dissertation]. Universität Bamberg, Bamberg.
- Gabadinho, A., Ritschard, G., Müller, N. S. & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37. <https://doi.org/10.18637/jss.v040.i04>
- Gadatsch, A. (2017). *Grundkurs Geschäftsprozess-Management: Analyse, Modellierung, Optimierung und Controlling von Prozessen*. Springer. <https://doi.org/10.1007/978-3-658-17179-7>
- Gašević, D., Kovanović, V. & Joksimović, S. (2017). Piecing the learning analytics puzzle: a consolidated model of a field of research and practice. *Learning: Research and Practice*, 3(1), 63–78. <https://doi.org/10.1080/23735082.2017.1286142>
- Gegenfurtner, A., Gruber, H., Holzberger, D., Keskin, Ö., Lehtinen, E., Seidel, T., Stürmer, K. & Säljö, R. (2023). Towards a cognitive theory of visual expertise. In C. Damşa, A. Rajala, G. Ritella & J. Brouwer (Hrsg.), *Re-theorising Learning and Research Methods in Learning Research* (S. 142–158). Routledge. <https://doi.org/10.4324/9781003205838-10>

- Gerholz, K.-H., Schlottmann, P., Slepcevic-Zach, P. & Stock, M. (2022). Digital Literacy in der beruflichen Lehrerinnenund Lehrerbildung – Einleitung. In K.-H. Gerholz, P. Schlottmann, P. Slepcevi-Zach & M. Stock (Hrsg.), *Digital Literacy in der beruflichen Lehrer:innenbildung: Didaktik, Empirie und Innovation* (S. 11–18). wbv.
- Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies*, 15(1), 6. <https://doi.org/10.3390/soc15010006>
- Geyskens, J., Donche, V. & Petegem, P. (2012). Towards effective feedback in higher education: bridging theory and practice. *Reflecting education*, 8(1), 132–147.
- Girasa, R. (2020). *Artificial Intelligence as a Disruptive Technology: Economic Transformation and Government Regulation*. Palgrave Macmillan.
- Görz, G., Braun, T. & Schmid, U. (2021). Einleitung. In G. Görz, U. Schmid & T. Braun (Hrsg.), *Handbuch der Künstlichen Intelligenz*. De Gruyter Oldenbourg. <https://doi.org/10.1515/9783110659948>
- Greimel-Fuhrmann, B. (2008). Interesse und Lernmotivation für Rechnungswesen – Untersuchung ihrer Struktur, Entwicklung und Förderung im kaufmännischen Unterricht. *Institut für Bildungsforschung der Wirtschaft*(1).
- Grienberger, K., Matthes, B. & Paulus, W. (2024). *Folgen des technologischen Wandels für den Arbeitsmarkt. Vor allem Hochqualifizierte bekommen die Digitalisierung verstärkt zu spüren* (IAB-Kurzbericht Nr. 5). <https://doku.iab.de/kurzber/2024/kb2024-05.pdf>
- Haas, A. (2005). *Unterrichtsplanung im Alltag von Lehrerinnen und Lehrern: Eine empirische Untersuchung zum Planungshandeln von Hauptschul-, Realschul- und Gymnasiallehrern*. Roderer.
- Hacker, W. & Sachse, P. (2014). *Allgemeine Arbeitspsychologie: Psychische Regulation von Tätigkeiten*. Hogrefe.
- Haenlein, M. & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 5–14. <https://doi.org/10.1177/0008125619864925>
- Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H. & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications*, 73, 220–239. <https://doi.org/10.1016/j.eswa.2016.12.035>

- Hangen, J. & Wuttke, E. (2024). Learning about artificial intelligence - results from a training program to promote startup skills in the AI field. *Empirische Pädagogik*, 38(1), 26–46. <https://doi.org/10.62350/TOEF2442>
- Hantke, H. (2020). Zukunftsdiskurse curricular intendiert – Plädoyer für eine ehrliche Lehrplanrezeption. *bwp@ Berufs- und Wirtschaftspädagogik – online*, Spezial 17, 1–26. [http://fox.leuphana.de/portal/en/publications/zukunftsdiskurse-curricular-intendiert--pladoyer-fur-eine-ehrliche-lehrplanrezeption\(42db48ac-3fcb-4fd4-a512-e08c89948fd0\).html](http://fox.leuphana.de/portal/en/publications/zukunftsdiskurse-curricular-intendiert--pladoyer-fur-eine-ehrliche-lehrplanrezeption(42db48ac-3fcb-4fd4-a512-e08c89948fd0).html)
- Hantke, H. & Pranger, J. (2019). Nachhaltigkeits-Werte ausbilden und kommunizieren: Lernmodule zu Corporate Social Responsibility im Bereich Transport und Logistik. *Berufsbildung in Wissenschaft und Praxis*, 48(4), 29–31.
- Happ, R., Kögler, K., Schmidt, J. & Egloffstein, M. (2024). Teaching and learning with and about artificial intelligence in vocational education and training. *Empirische Pädagogik*, 38(1), 6–8. <https://doi.org/10.62350/QVHQ3765>
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heimann, P., Otto, G. & Schulz, W. (1965). *Unterricht - Analyse und Planung*. Schroedel.
- Heiskanen, E., Thidell, Å. & Rodhe, H. (2016). Educating sustainability change agents: the importance of practical skills and experience. *Journal of Cleaner Production*, 123, 218–226. <https://doi.org/10.1016/j.jclepro.2015.11.063>
- Helm, C. (2015). Reziproke Effekte zwischen wahrgenommenem LehrerInnenverhalten, intrinsischer Motivation und der SchülerInnenleistung im Fach „Rechnungswesen“. *AMS report*(111), 1–36.
- Henderson, M., Ryan, T. & Phillips, M. (2019). The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education*, 44(8), 1237–1252. <https://doi.org/10.1080/02602938.2019.1599815>
- Holst, J. (2022). *Nachhaltigkeit & BNE in der Beruflichen Bildung: Dynamik in Ordnungsmitteln, Potentiale bei Berufen, Lernorten und in der Qualifizierung von Auszubildenden*. <https://doi.org/10.17169/refubium-35827>
- Hommel, M. (2019). *Lernhandeln mit integrierter Unternehmenssoftware* [Habilitation]. Technische Universität, Dresden.
- Hooda, M., Rana, C., Dahiya, O., Rizwan, A. & Hossain, M. S. (2022). Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education.

- Mathematical Problems in Engineering*, 2022, 1–19.
<https://doi.org/10.1155/2022/5215722>
- Horváth, P. (2019). *Controlling*. Franz Vahlen.
- Hsia, L.-H., Hwang, G.-J. & Hwang, J.-P. (2024). AI-facilitated reflective practice in physical education: an auto-assessment and feedback approach. *Interactive Learning Environments*, 32(9), 5267–5286. <https://doi.org/10.1080/10494820.2023.2212712>
- Hui, H. & Xiao, L. (2023). Assistant Teaching System Design Based on Expert System. 2023 3rd International Conference on Mobile Networks and Wireless Communications (IC-MNWC) (S. 1–6). IEEE. <https://doi.org/10.1109/ICMNWC60182.2023.10435766>
- Ifenthaler, D. & Yau, J. Y.-K. (2021). Learning Analytics zur Unterstützung von Lernerfolg: Ausgewählte Ergebnisse einer systematischen Übersichtsarbeit. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, Beiheft 31, 215–236.
- Jacobsen, L. J. & Weber, K. E. (2024). *ChatGPT & Co.: Feedback as a Use Case for Large Language Models in Higher Education [Preprint]*.
<https://doi.org/10.13140/RG.2.2.17993.35685>
- Jarodzka, H., Skuballa, I. & Gruber, H. (2021). Eye-Tracking in Educational Practice: Investigating Visual Perception Underlying Teaching and Learning in the Classroom. *Educational Psychology Review*, 33(1), 1–10. <https://doi.org/10.1007/s10648-020-09565-7>
- Jeon, J. & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28, 15873–15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Jia, X.-H. & Tu, J.-C. (2024). Towards a New Conceptual Model of AI-Enhanced Learning for College Students: The Roles of Artificial Intelligence Capabilities, General Self-Efficacy, Learning Motivation, and Critical Thinking Awareness. *Systems*, 12(3), 74. <https://doi.org/10.3390/systems12030074>
- Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, 14(1), 63–76.
<https://doi.org/10.1177/1469787412467125>
- Kalyuga, S. (2021). The Expertise Reversal Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 171–182). Cambridge University Press. <https://doi.org/10.1017/9781108894333.017>

- Kalyuga, S. & Sweller, J. (2021). The Redundancy Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 212–220). Cambridge University Press.
<https://doi.org/10.1017/9781108894333.021>
- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P. & Papin, K. (2023). Generative AI and Teachers' Perspectives on Its Implementation in Education. *Journal of Interactive Learning Research*, 34(2), 313–338.
- Kelland, J. & Kanuka, H. (2008). "We just disagree:" Using deliberative inquiry to seek consensus about the effects of e-learning on higher education. *Canadian Journal of Learning and Technology*. <https://doi.org/10.21432/T26W2K>
- Khosa, A., Pandey, R. & Wilkin, C. (2024). Accounting curricula and climate-related sustainability: evidence from Australia and New Zealand universities. *Accounting Education*, 1–23. <https://doi.org/10.1080/09639284.2024.2376739>
- Khowaja, S. A., Khuwaja, P., Dev, K., Wang, W. & Nkenyereye, L. (2024). ChatGPT Needs SPADE (Sustainability, PrivAcy, Digital divide, and Ethics) Evaluation: A Review. *Cognitive Computation*, 16(5), 2528–2550. <https://doi.org/10.1007/s12559-024-10285-1>
- Kiritchenko, S. & Mohammad, S. M. (2018). *Examining Gender and Race Bias in Two Hundred Sentiment Analysis Systems*. <https://arxiv.org/pdf/1805.04508>
- Klafki, W. (1963). *Studien zur Bildungstheorie und Didaktik*. Beltz.
- Kleesiek, J., Murray, J. M., Strack, C., Kaissis, G. & Braren, R. (2020). Wie funktioniert maschinelles Lernen? *Der Radiologe*, 60(1), 24–31. <https://doi.org/10.1007/s00117-019-00616-x>
- Klein, J. & Küst, C. (2020). Wie die Digitalisierung im Rechnungswesen die Aufgaben und Anforderungen an die Mitarbeiter/-innen verändert. In F. Berding, H. Jahncke & A. Slopinski (Hrsg.), *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive* (S. 83–97). Springer.
- Kloos, B. & Schmidt-Bens, J. (2020). Datenschutz. In M. Hartmann (Hrsg.), *KI & Recht kompakt* (S. 165–208). Springer.
- Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>

- Klugman, C. M. (2021). Black Boxes and Bias in AI Challenge Autonomy. *The American Journal of Bioethics*, 21(7), 33–35. <https://doi.org/10.1080/15265161.2021.1926587>
- Klusmeyer, J. (2021). Entwicklung eines wirtschaftsdidaktischen Unterrichtsplanungsmodells auf Grundlage der Basisdimensionen lernförderlichen Unterrichts. In J. Klusmeyer & M. Söll (Hrsg.), *Unterrichtsplanung in der Wirtschaftsdidaktik: Aktuelle theorie-, empirie- und praxisbasierte Beiträge* (S. 85–122). Springer.
- Klusmeyer, J. & Söll, M. (2021). Unterrichtsplanung im Fokus wirtschaftsdidaktischer Theorie, Empirie und Praxis. In J. Klusmeyer & M. Söll (Hrsg.), *Unterrichtsplanung in der Wirtschaftsdidaktik: Aktuelle theorie-, empirie- und praxisbasierte Beiträge* (1-13). Springer.
- KMK. (2019). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung*. https://www.kmk.org/fileadmin/Daten/veroeffentlichungen_beschluesse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf
- KMK. (2021). *Handreichung für die Erarbeitung von Rahmenlehrplänen der Kultusministerkonferenz für den berufsbezogenen Unterricht in der Berufsschule und ihre Abstimmung mit Ausbildungsordnungen des Bundes für anerkannte Ausbildungsberufe*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2021/2021_06_17-GEP-Handreichung.pdf
- KMK. (2022). *Standards für die Lehrerbildung: Bildungswissenschaften*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- Knoth, N., Decker, M., Laupichler, M. C., Pinski, M., Buchholtz, N., Bata, K. & Schultz, B. (2024). Developing a holistic AI literacy assessment matrix – Bridging generic, domain-specific, and ethical competencies. *Computers and Education Open*, 6, 100177. <https://doi.org/10.1016/j.caeo.2024.100177>
- Koch-Priewe, B. (2000). Zur Aktualität und Relevanz der Allgemeinen Didaktik in der Lehrerinnenausbildung. In M. Bayer, F. Bohnsack, B. Koch-Priewe & J. Wildt (Hrsg.), *Lehrerin und Lehrer werden ohne Kompetenz? Professionalisierung durch eine andere Lehrerbildung* (S. 148–169). Klinkhardt.
- Koehler, M. & Mishra, P. (2009). What is Technological Pedagogical Content Knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60–70. <https://doi.org/10.1177/002205741319300303>

- Kögler, K. & Wuttke, E. (2012). Unterrichtliche Monotonie als Bedingungsfaktor für Schülerlangeweile im Fach Rechnungswesen. In U. Faßhauer, B. Fürstenau & E. Wuttke (Hrsg.), *Schriftenreihe der Sektion Berufs- und Wirtschaftspädagogik. Berufs- und wirtschaftspädagogische Analysen – aktuelle Forschungen zur beruflichen Bildung* (S. 75–87). Barbara Budrich. <https://doi.org/10.25656/01:7108>
- Konertz, R. & Schönhof, R. (2024). Rechtsfolgen der Urheberrechtsverletzung bei generativer Künstlicher Intelligenz: Über die Möglichkeit des „Vergessens“ in Neuronalen Netzen. *Wettbewerb in Recht und Praxis*, 70(5), 534–541.
- König, J. & Rothland, M. (2022). Stichwort: Unterrichtsplanungskompetenz. *Zeitschrift für Erziehungswissenschaft*, 25(4), 771–813. <https://doi.org/10.1007/s11618-022-01107-x>
- Krämer, H., Jordanski, G. & Goertz, L. (2017). *Medien anwenden und produzieren. Entwicklung von Medienkompetenz in der Berufsausbildung*. (Wissenschaftliche Diskussionspapiere Nr. 181).
- Kuhlmeier, W. & Vollmer, T. (2018). Ansatz einer Didaktik der Beruflichen Bildung für nachhaltige Entwicklung. In M. Casper & T. Schlömer (Hrsg.), *Didaktik der berufl. Bildung - Selbstverständnis, Zukunftsperspektiven und Innovationsschwerpunkte*. (S. 131-151). wbv.
- Lane, H., Howard, C. & Hapke, H. M. (2019). *Natural language processing in action: Understanding, analyzing, and generating text with Python*. Manning.
- Lauterjung, S. (2020). Vom smarten Berater zur smarten Maschine. In R. Buchkremer, T. Heupel & O. Koch (Hrsg.), *Künstliche Intelligenz in Wirtschaft & Gesellschaft: Auswirkungen, Herausforderungen & Handlungsempfehlungen* (S. 249–274). Springer.
- Lee, S. S. & Moore, R. L. (2024). Harnessing Generative AI (GenAI) for Automated Feedback in Higher Education: A Systematic Review. *Online Learning*, 28(3), 82–106. <https://doi.org/10.24059/olj.v28i3.4593>
- Lemmrich, S., Ehmke, T. & Reusser, K. (2024). Adaptive Lernunterstützung durch fachliche Präzision und interaktionale Qualität. *Zeitschrift für Schul- und Professionsentwicklung*, 6(2), 6–23. <https://doi.org/10.11576/pflb-6862>
- Lim, L.-A., Dawson, S., Gašević, D., Joksimović, S., Fudge, A., Pardo, A. & Gentili, S. (2020). Students’ sense-making of personalised feedback based on learning analytics. *Australasian Journal of Educational Technology*, 36(6), 15–33. <https://doi.org/10.14742/ajet.6370>

- Lin, Y.-S., Chen, S.-Y., Tsai, C.-W. & Lai, Y.-H. (2021). Exploring Computational Thinking Skills Training Through Augmented Reality and AIoT Learning. *Frontiers in Psychology*, 12, 640115. <https://doi.org/10.3389/fpsyg.2021.640115>
- Liu, Y [Yiheng], Han, T., Ma, S [Siyuan], Zhang, J [Jiayue], Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z [Zhengliang], Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T. & Ge, B. (2023). Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2), 100017. <https://doi.org/10.1016/j.metrad.2023.100017>
- Liu, Y [Yue], Yang, Z., Yu, Z., Liu, Z [Zitu], Liu, D., Lin, H., Li, M., Ma, S [Shuchang], Avdeev, M. & Shi, S. (2023). Generative artificial intelligence and its applications in materials science: Current situation and future perspectives. *Journal of Materiomics*, 9(4), 798–816. <https://doi.org/10.1016/j.jmat.2023.05.001>
- Long, D. & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (S. 1–16). ACM. <https://doi.org/10.1145/3313831.3376727>
- Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P. & Tsai, C.-C. (2020). Challenges and Future Directions of Big Data and Artificial Intelligence in Education. *Frontiers in Psychology*, 11, Artikel 580820. <https://doi.org/10.3389/fpsyg.2020.580820>
- Luccioni, S., Jernite, Y. & Strubell, E. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (S. 85–99). Association for Computing Machinery. <https://doi.org/10.1145/3630106.3658542>
- Luckin, R., Holmes, W., Griffiths, M. & Forcier, L. B. (2016). *Intelligence Unleashed. An argument for AI in Education*. Pearson.
- Mah, D.-K., Hense, J. & Dufentester, C. (2023). Didaktische Impulse zum Lehren und Lernen mit und über Künstliche Intelligenz. In C. de Witt, C. Gloerfeld & S. E. Wrede (Hrsg.), *Künstliche Intelligenz in der Bildung* (S. 91–108). Springer. https://doi.org/10.1007/978-3-658-40079-8_5
- Maier, U. & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. *Computers and Education: Artificial Intelligence*, 3, 100080. <https://doi.org/10.1016/j.caeai.2022.100080>

- Manaff, N. A. & Halabi Azahari, M. (2024). Digital technology's effect on teaching and learning. *International Journal of Research in Education Humanities and Commerce*, 5(2), 129–135. <https://doi.org/10.37602/IJREHC.2024.5211>
- Massmann, C. & Hofstetter, A. (2020). AI-pocalypse now? Herausforderungen Künstlicher Intelligenz für Bildungssystem, Unternehmen und die Workforce der Zukunft. In R. A. Fürst (Hrsg.), *Digitale Bildung und künstliche Intelligenz in Deutschland: Nachhaltige Wettbewerbsfähigkeit und Zukunftsagenda* (S. 167–220). Springer. https://doi.org/10.1007/978-3-658-30525-3_8
- Mayer, C. W., Rausch, A. & Seifried, J. (2023). Analysing domain-specific problem-solving processes within authentic computer-based learning and training environments by using eye-tracking: a scoping review. *Empirical Research in Vocational Education and Training*, 15(1). <https://doi.org/10.1186/s40461-023-00140-2>
- Mayer, R. E. (2021). The Multimedia Principle. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 145–157). Cambridge University Press. <https://doi.org/10.1017/9781108894333.015>
- Mayer, R. E. & Fiorella, L. (2021a). Introduction to Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 3–16). Cambridge University Press. <https://doi.org/10.1017/9781108894333.003>
- Mayer, R. E. & Fiorella, L. (2021b). Principles for Managing Essential Processing in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 243–260). Cambridge University Press. <https://doi.org/10.1017/9781108894333.025>
- McCarthy, J. (2004). What is artificial intelligence?
- Meiners, H., Hartmann, P., Niegemann, H., Seeber, S., Wuttke, E. & Schumann, M. (2022). Digitale Medienkompetenz als Voraussetzung für die Erstellung von Prüfungsaufgaben. In S. Schumann, S. Seeber & S. Abele (Hrsg.), *Digitale Transformation in der Berufsbildung: Konzepte, Befunde und Herausforderungen* (S. 123–144). wbv.
- Mertens, U., Finn, B. & Lindner, M. A. (2022). Effects of computer-based feedback on lower- and higher-order learning outcomes: A network meta-analysis. *Journal of Educational Psychology*, 114(8), 1743–1772. <https://doi.org/10.1037/edu0000764>
- Minnameier, G. (2022). Wie sollen junge Kaufleute über Nachhaltigkeit denken? Normative Aspekte einer Berufsbildung für nachhaltige Entwicklung. In C. Michaelis & F.

- Berding (Hrsg.), *Berufsbildung für nachhaltige Entwicklung: Umsetzungsbarrieren und interdisziplinäre Forschungsfragen*. wbv.
- Moring, A. (2023). Funktion und Einsatz von Künstlicher Intelligenz. In A. Moring (Hrsg.), *Künstliche Intelligenz und Intuition: Robuste und nachhaltige Entscheidungen in digitalen Arbeitswelten* (S. 7–70). Springer Gabler. https://doi.org/10.1007/978-3-658-42018-5_2
- Narciss, S. (2013). Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the interactive feedback model. *Digital Education Review*, 23(1), 7–26.
- Narciss, S., Hammer, E., Damnik, G., Kisielski, K. & Körndle, H. (2021). Promoting Prospective Teacher Competencies for Designing, Implementing, Evaluating, and Adapting Interactive Formative Feedback Strategies. *Psychology Learning & Teaching*, 20(2), 261–278. <https://doi.org/10.1177/1475725720971887>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Ngige, O. C., Awodele, O. & Balogun, O. (2021). Judicial Artificial Intelligence Bias: A Survey and Recommendations. *Transactions on Machine Learning and Artificial Intelligence*, 9(2), 74–86. <https://doi.org/10.14738/tmlai.92.10118>
- Nickolaus, R. (2009). Didaktische Modelle und Konzepte für die Planung und Analyse beruflicher Lehr-Lernprozesse. In B. Bonz (Hrsg.), *Didaktik und Methodik der Berufsbildung* (S. 33–59). Schneider.
- Oevermann, U. (1996). Theoretische Skizze einer revidierten Theorie professionalisierten Handelns. In A. Combe & W. Helsper (Hrsg.), *Pädagogische Professionalität* (S. 70–82). Suhrkamp.
- Oevermann, U. (2008). Profession contra Organisation? Strukturtheoretische Perspektiven zum Verhältnis von Organisation und Profession in der Schule. In W. Helsper, S. Busse, M. Hummrich & R. T. Kramer (Hrsg.), *Pädagogische Professionalität in Organisationen* (S. 55–77). VS. https://doi.org/10.1007/978-3-531-90777-2_4
- OpenAI. (2023). *GPT-4 Technical Report*. <https://doi.org/10.48550/arXiv.2303.08774>
- Otte, I. & Singer-Brodowski, M. (2017). *Verankerung von Bildung für nachhaltige Entwicklung in der dualen beruflichen Ausbildung*. <https://doi.org/10.13140/RG.2.2.13191.70563>

- Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O. & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(8), em2307. <https://doi.org/10.29333/ejmste/13428>
- Papa, R., Jackson, K. M., Brown, R. & Jackson, D. (2020). Discourse Analysis on Learning Theories and AI. In K. Arai, S. Kapoor & R. Bhatia (Hrsg.), *Intelligent Computing: Proceedings of the 2020 Computing Conference, Volume 3* (S. 665–672). Springer International. https://doi.org/10.1007/978-3-030-52243-8_50
- Pargmann, J. & Berding, F. (2022). Künstliche Intelligenz in der beruflichen Bildung - Utopie oder Dystopie? Chancen und Herausforderungen von KI als Bildungstechnologie. In M. Eckelt, T. J. Ketschau, J. Klassen, J. Schauer, J. K. Schmees & C. Steib (Hrsg.), *Berufsbildungspolitik: Strukturen - Krise - Perspektiven* (S. 215–232). wbv.
- Pargmann, J. & Berding, F. (2025). Sustainability content knowledge for teachers in accounting and controlling: A systematic literature review on industry approaches, practices and desired competences. *Frontiers in Education*.
- Pargmann, J., Berding, F., Riebenbauer, E. & Flick-Holtsch, D. (im Druck). Perspektiven für eine nachhaltigkeitsorientierte Didaktik des Rechnungswesens und Controllings. In J. Schlicht, S. Kaiser & F. Schwehm (Hrsg.), *Innovationen für Nachhaltigkeit durch berufs- und wirtschaftspädagogische Forschung*. wbv.
- Pargmann, J., Berding, F., Riebenbauer, E. & Flick-Holtsch, D. (2022). Implikationen der Digitalisierung für die Didaktik des Rechnungswesens – konzeptionelle Überlegungen und Vorstellung eines Learning Cockpits. *bwp@*(43), 1–30. https://www.bwpat.de/ausgabe43/pargmann_etal_bwpat43.pdf
- Pargmann, J., Leube, A., Berding, F., Riebenbauer, E., Rebmann, K., Slopinski, A. & Gillert, M. (2024). Die elektronisch-didaktische Assistenz (EDDA) zur Unterrichtsplanung auf Basis künstlicher Intelligenz: Funktionsweise, Anwendungsbereiche und Forschungsperspektiven. *Empirische Pädagogik*, 38(1), 118–146. <https://doi.org/10.62350/WLQO3003>
- 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). <https://doi.org/10.1186/s40461-023-00141-1>

- Perridon, L., Rathgeber, A. W. & Steiner, M. (2017). *Finanzwirtschaft der Unternehmung*. Franz Vahlen.
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking.
- Preiß, P. & Tramm, T. (1996). Die Göttinger Unterrichtskonzeption des wirtschaftsinstrumentellen Rechnungswesens. In P. Preiß & T. Tramm (Hrsg.), *Rechnungswesenunterricht und ökonomisches Denken* (S. 222–323). Gabler.
- Prilop, C. N. & Weber, K. E. (2023). Digital video-based peer feedback training: The effect of expert feedback on pre-service teachers' peer feedback beliefs and peer feedback quality. *Teaching and Teacher Education*, 127, 104099.
<https://doi.org/10.1016/j.tate.2023.104099>
- Pufé, I. (2017). *Nachhaltigkeit*. UTB.
- Reinisch, H. (2021). Unterrichtsplanung im Kontext bildungstheoretischer Grundannahmen. In J. Klusmeyer & M. Söll (Hrsg.), *Unterrichtsplanung in der Wirtschaftsdidaktik: Aktuelle theorie-, empirie- und praxisbasierte Beiträge* (S. 17–34). Springer.
- Rich, E. (1983). *Artificial Intelligence*. McGraw-Hill.
- Richter, D., Lazarides, R. & Scheiter, K. (2024). Stichwort: Professionalisierung von Lehrkräften für die digitale Transformation. *Zeitschrift für Erziehungswissenschaft*, 27(3), 613–636. <https://doi.org/10.1007/s11618-024-01242-7>
- Riebenbauer, E. (2021). *Kompetenzentwicklung im Masterstudium Wirtschaftspädagogik. Längsschnittstudie zur Unterrichtsplanung im Rechnungswesen*. wbv.
- Riebenbauer, E., Berding, F. & Holtsch, D. (2022). Fachwissenschaftliche Implikationen für die Aus- und Fortbildung von Lehrkräften zur Digitalisierung im Rechnungswesen. In K.-H. Gerholz, P. Schlottmann, P. Slepcevic-Zach & M. Stock (Hrsg.), *Digital Literacy in der beruflichen Lehrer:innenbildung: Didaktik, Empirie und Innovation* (S. 199–212). wbv.
- Riebenbauer, E., Rebmann, K., Berding, F., Leube, A., Pargmann, J., Slopinski, A. & Gillert, M. (2024). Zum Planungshandeln von Studierenden der Wirtschaftspädagogik - Teil 1: Orientierung bei der Planung mit Unterrichtsentwürfen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 4.
- Riebenbauer, E., Rebmann, K., Berding, F., Leube, A., Pargmann, J., Slopinski, A. & Gillert, M. (2025). Zum Planungshandeln von Studierenden der Wirtschaftspädagogik – Teil 2: Empirische Befunde zur Analyse von Unterrichtsentwürfen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 1.

- Riedel, J. & Kleppsch, J. (2021). Wie bereit sind Studierende für die Nutzung von KI-Technologien? Eine Annäherung an die KI-Readiness Studierender im Kontext des Projektes „tech4comp“. In H.-W. Wollersheim, M. Karapanos & N. Pengel (Hrsg.), *Bildung in der digitalen Transformation* (S. 283–292). Waxmann.
<https://doi.org/10.25656/01:26644>
- Riski, D. & Nuryanto, A. (2024). The Effect of Artificial Intelligence (AI) Based Learning Implementation on the Quality of Learning Outcomes of Vocational High School Students in the Field of Machining. *International Journal of Social Science and Human Research*, 7(09). <https://doi.org/10.47191/ijsshr%2Fv7-i09-18>
- Rodríguez-Barroso, N., Stipcich, G., Jiménez-López, D., Ruiz-Millán, J. A., Martínez-Cámara, E., González-Seco, G., Luzón, M. V., Veganzones, M. A. & Herrera, F. (2020). Federated Learning and Differential Privacy: Software tools analysis, the Sherpa.ai FL framework and methodological guidelines for preserving data privacy. *Information Fusion*, 64, 270–292. <https://doi.org/10.1016/j.inffus.2020.07.009>
- Roldán-Álvarez, D. & Mesa, F. J. (2024). Intelligent Deep-Learning Tutoring System to Assist Instructors in Programming Courses. *IEEE Transactions on Education*, 67(1), 153–161. <https://doi.org/10.1109/TE.2023.3331055>
- Roth, H. (1971). *Pädagogische Anthropologie, Band 2. Entwicklung und Erziehung. Grundlagen einer Entwicklungspädagogik*. Schroedel.
- Rothland, M. (2022). Anmerkungen zur Modellierung und Operationalisierung (allgemein-didaktischer) Unterrichtsplanungskompetenz. *Unterrichtswissenschaft*, 50(3), 347–372.
- Russel, S. J. & Norvig, P. (2021). *Artificial Intelligence. A modern approach*. Pearson.
- Sailer, U. (2020). *Nachhaltigkeitscontrolling: Was Controller und Manager über die Steuerung der Nachhaltigkeit wissen sollten*. UTB.
<https://doi.org/10.36198/9783838553320>
- Salvagno, M., Taccone, F. S. & Gerli, A. G. (2023). Artificial intelligence hallucinations. *Critical Care*, 27(1), 180. <https://doi.org/10.1186/s13054-023-04473-y>
- Sandfuchs, U. (2004). Unterrichtsplanung. In R. W. Keck, U. Sandfuchs & B. Thomas (Hrsg.), *Wörterbuch Schulpädagogik: Ein Nachschlagewerk für Studium und Schulpraxis* (S. 348–350). Klinkhardt.

- Schaltegger, S. (2016). CSR, Nachhaltigkeit und Controlling – Zwischen Praxislücke und Forschungskonzepten. In E. Günther & K.-H. Steinke (Hrsg.), *CSR und Controlling* (S. 55–69). Springer. https://doi.org/10.1007/978-3-662-47702-1_4
- Schartel, S. A. (2012). Giving feedback - an integral part of education. *Best practice & research. Clinical anaesthesiology*, 26(1), 77–87. <https://doi.org/10.1016/j.bpa.2012.02.003>
- Schmaltz, C. (2019). *Heterogenität als Herausforderung für die Professionalisierung von Lehrkräften*. Springer VS.
- Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R. M., Abrami, P. C., Wade, C. A., Surkes, M. A. & Lowerison, G. (2009). Technology's Effect on Achievement in Higher Education: A Stage I Meta-Analysis of Classroom Applications. *Journal of Computing in Higher Education*, 21(2), 95–109. <https://doi.org/10.1007/s12528-009-9021-8>
- Schmidt, J. M.-C. (2024). *Grundlagenwissen zu Künstlicher Intelligenz von angehenden Lehrkräften* [Dissertation]. Universität Leipzig.
- Schöning, S., Mendel, V. & Köse, A. (2020). Mit neuen Controller-Kompetenzen in die Zukunft. *Controlling & Management Review*, 64(1), 58–63. <https://doi.org/10.1007/s12176-019-0073-y>
- Schulz, W. (1980). *Unterrichtsplanung. Praxis und Theorie des Unterrichtens*. Urban & Schwarzenberg.
- Schütt-Sayed, S., Casper, M. & Vollmer, T. (2021). Mitgestaltung lernbar machen – Didaktik der Berufsbildung für nachhaltige Entwicklung. In C. Melzig, W. Kuhlmeier & S. Kretschmer (Hrsg.), *Berufsbildung für nachhaltige Entwicklung: Die Modellversuche 2015-2019 auf dem Weg vom Projekt zur Struktur* (S. 200–227).
- Seifried, J. (2004). *Fachdidaktische Variationen in einer selbstorganisationsoffenen Lernumgebung* [Dissertation]. DUV.
- Seifried, J. (2009). Unterrichtsplanung von (angehenden) Lehrkräften an kaufmännischen Schulen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 105(2), 179–197. <https://doi.org/10.25162/zbw-2009-0012>
- Sembey, R., Hoda, R. & Grundy, J. (2024). Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *Journal of Systems and Software*, 211, 111988. <https://doi.org/10.1016/j.jss.2024.111988>

- Seufert, S., Guggemos, J. & Ifenthaler, D. (2021). Zukunft der Arbeit mit intelligenten Maschinen: Implikationen der Künstlichen Intelligenz für die Berufsbildung: Einleitung zum Beiheft. *Zeitschrift für Berufs- und Wirtschaftspädagogik. Beiheft 31*, 9–27.
- Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P. & Galligan, L. (2022). A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis. *IEEE Access*, 10, 56720–56739.
<https://doi.org/10.1109/ACCESS.2022.3177752>
- Shulman, L. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, 57(1), 1–23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Šíma, J. (1995). Neural expert systems. *Neural Networks*, 8(2), 261–271.
[https://doi.org/10.1016/0893-6080\(94\)00070-3](https://doi.org/10.1016/0893-6080(94)00070-3)
- Simbeck, K., Folkerts, F. & Riaz, S. (2019). Automatisierte Textanalyse bei der Personalauswahl. Potenziale und Grenzen. *Berufsbildung in Wissenschaft und Praxis*, 48(3), 27–29.
- Söll, M. (2021). Die Modellierung von Unterrichtsplanungskompetenz – Ein Ansatz zur Differenzierung wirtschaftsdidaktischer und allgemeiner pädagogischer Begründungsmuster. In J. Klusmeyer & M. Söll (Hrsg.), *Unterrichtsplanung in der Wirtschaftsdidaktik: Aktuelle theorie-, empirie- und praxisbasierte Beiträge* (S. 123–168). Springer.
- Söll, M. & Klusmeyer, J. (2018). Unterrichtsplanung als zentraler Gegenstand der Wirtschaftsdidaktik – Konzeption eines hochschuldidaktischen Ansatzes. In P. T. Tramm, M. Casper & T. Schlömer (Hrsg.), *Didaktik der beruflichen Bildung – Selbstverständnis, Zukunftsperspektiven und Innovationsschwerpunkte* (S. 73–88). Bertelsmann.
- Sormunen, M., Heikkilä, A., Salminen, L., Vauhkonen, A. & Saaranen, T. (2021). Learning Outcomes of Digital Learning Interventions in Higher Education. *Computers, Informatics, Nursing*. <https://doi.org/10.1097/cin.0000000000000797>
- Spener, C. & Schumann, S. (2021). Wissenseffekte des ERP-Einsatzes in der kaufmännischen Berufsschule. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 117(3), 395–430. <https://doi.org/10.25162/zbw-2021-0018>

- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M. & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Stender, A. (2014). *Unterrichtsplanung: Vom Wissen zum Handeln: Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung* [Dissertation]. Christian-Albrechts-Universität, Kiel.
- Stender, A., Brückmann, M. & Neumann, K. (2015). Vom Professionswissen zum kompetenten Handeln im Unterricht: Die Rolle der Unterrichtsplanung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 33(1), 121–133.
- Strzelecki, A. (2023). Students' Acceptance of ChatGPT in Higher Education: An Extended Unified Theory of Acceptance and Use of Technology. *Innovative Higher Education*, 49, 223–245. <https://doi.org/10.1007/s10755-023-09686-1>
- Strzelecki, A. (2024). ChatGPT in higher education: Investigating bachelor and master students' expectations towards AI tool. *Education and Information Technologies*, 1–25. <https://doi.org/10.1007/s10639-024-13222-9>
- Strzelecki, A. & ElArabawy, S. (2024). Investigation of the moderation effect of gender and study level on the acceptance and use of generative AI by higher education students: Comparative evidence from Poland and Egypt. *British Journal of Educational Technology*, 55(3), 1209–1230. <https://doi.org/10.1111/bjet.13425>
- Stütz, S. (2021). Der Rechnungswesenunterricht im retrospektiven Erleben von Studierenden der Wirtschaftspädagogik. *bwp@, Österreich Spezial* 3, 1–23. http://www.bwpat.de/wipaed-at3/stuetz_wipaed-at_2021.pdf
- Stütz, S. (2024). *Charakteristika von Aufgaben in der Domäne Rechnungswesen* [Dissertation]. Springer. <https://doi.org/10.1007/978-3-658-44498-3>
- Stütz, S., Berding, F., Reincke, S. & Scheper, L. (2022). Characteristics of learning tasks in accounting textbooks: an AI assisted analysis. *Empirical Research in Vocational Education and Training*, 14(1), 1–33. <https://doi.org/10.1186/s40461-022-00138-2>
- Sun, Y., Sheng, D., Zhou, Z. & Wu, Y. (2024). AI hallucination: towards a comprehensive classification of distorted information in artificial intelligence-generated content. *Humanities and Social Sciences Communications*, 11(1), 1–14. <https://doi.org/10.1057/s41599-024-03811-x>

- Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C. & Schmid, R. F. (2011). What Forty Years of Research Says About the Impact of Technology on Learning. *Review of Educational Research*, 81(1), 4–28. <https://doi.org/10.3102/0034654310393361>
- Thole, C. (2021). *Berufliche Identitätsarbeit als Bildungsauftrag der Berufsschule: Am Beispiel der dualen Ausbildung im Einzelhandel* [Dissertation]. wbv. <https://doi.org/10.3278/6004730w>
- Traum, A., Müller, C., Hummert, H. & Nerdinger, F. W. (2017). *Digitalisierung – Die Perspektive des arbeitenden Individuums*. https://doi.org/10.18453/rosdok_id00000173
- Tunstall, L., Werra, L. von, Wolf, T. & Géron, A. (2022). *Natural language processing with transformers: Building language applications with hugging face*. O'Reilly.
- Türling, J., Seifried, J., Wuttke, E., Gewiese, A. & Kästner, R. (2011). ‚Typische‘ Schülerfehler im Rechnungswesenunterricht - Empirische Befunde einer Interviewstudie. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 107(3), 390–407.
- Underwood, J. (2008). Effective feedback: guidelines for improving performance. *International Conference of the Learning Sciences*. <https://doi.dx.org/10.22318/icls2008.2.415>
- United Nations. (2023). *Progress towards the Sustainable Development Goals: towards a rescue plan for people and planet*. <https://unstats.un.org/sdgs/files/report/2023/secretary-general-sdg-report-2023--EN.pdf>
- van den Berg, G. & Du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness in Teacher Education. *Education Sciences*, 13(10), 998. <https://doi.org/10.3390/educsci13100998>
- van der Kleij, F. M., Feskens, R. C. W. & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes. *Review of Educational Research*, 85(4), 475–511. <https://doi.org/10.3102/0034654314564881>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). Attention is all you need. In U. von Luxburg, I. Guyon, S. Bengio, H. Wallach, R. Fergus, S. V. N. Vishwanathan & R. Garnett (Hrsg.), *Advances in neural information processing systems 30: 31st Annual Conference on Neural Information Processing Systems (NIPS 2017)* (S. 1–11).
- Vincent-Lancrin, S. & van der Vlies, R. (2020). *Trustworthy artificial intelligence (AI) in education: Promises and challenges*. *OECD Education Working Papers*. <https://dx.doi.org/10.1787/a6c90fa9-en> <https://doi.org/10.1787/19939019>

- Wagner, S., Sibley, L., Weiler, D., Burde, J.-P., Scheiter, K. & Lachner, A. (2024). The more, the better? Learning with feedback and instruction. *Learning and Instruction*, 89, 101844. <https://doi.org/10.1016/j.learninstruc.2023.101844>
- Wang, L. & Zhao, M. (2024). Can artificial Intelligence Technology Promote the Improvement of Student Learning Outcomes? — — Meta Analysis Based on 50 Experimental and Quasi Experimental Studies. In B. V. Santhosh, K. Hemachandran & M. Deepanjali (Hrsg.), *Proceedings of the 3rd International Conference on Educational Innovation and Multimedia Technology*. <https://doi.org/10.4108/eai.29-3-2024.2347685>
- Wang, Z [Zhen], Pei, X., Zhu, H., Gong, S. & Wang, E. (2024). How to Make Computer-Based Feedback More Productive: The Power of Erroneous Solutions. *Journal of Educational Computing Research*, 62(6), 1–21. <https://doi.org/10.1177/07356331241247592>
- Welsch, A., Eitle, V. & Buxmann, P. (2018). Maschinelles Lernen. *HMD Praxis der Wirtschaftsinformatik*, 55(2), 366–382. <https://doi.org/10.1365/s40702-018-0404-z>
- Wernke, S. & Zierer, K. (2017). Die Unterrichtsplanung - Ein in Vergessenheit geratener Kompetenzbereich?! In S. Wernke & K. Zierer (Hrsg.), *Die Unterrichtsplanung - ein in Vergessenheit geratener Kompetenzbereich?! Status Quo und Perspektiven aus Sicht der empirischen Forschung* (S. 7–16). Klinkhardt.
- Wiesböck, F. & Hess, T. (2020). Digital innovations. *Electronic Markets*, 30(1), 75–86. <https://doi.org/10.1007/s12525-019-00364-9>
- Wilbers, K. (2023). *Wirtschaftsunterricht gestalten*. epubli. <https://doi.org/10.25656/01:27824>
- Williams, A. (2024). Delivering Effective Student Feedback in Higher Education: An Evaluation of the Challenges and Best Practice. *International Journal of Research in Education and Science*, 10(2), 473–501. <https://doi.org/10.46328/ijres.3404>
- Winther, E. & Achtenhagen, F. (2008). Kompetenzstrukturmodell für die kaufmännische Bildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 104(4), 511–538. <https://doi.org/10.25162/zbw-2008-0028>
- Wirzberger, M. & Schwarz, M. (2021). Förderung selbstregulierten Lernens durch ein KI-gestütztes Training. *Bildung und Erziehung*, 74(3), 280–295. <https://doi.org/10.13109/buer.2021.74.3.280>

- Wisniewski, B., Zierer, K. & Hattie, J. (2019). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10, 3087. <https://doi.org/10.3389/fpsyg.2019.03087>
- Wöhe, G., Döring, U. & Brösel, G. (2020). *Einführung in die allgemeine Betriebswirtschaftslehre*. Franz Vahlen.
- World Commission on Environment and Development (Hrsg.). (1987). *Our Common Future*. Oxford University Press. <http://www.un-documents.net/wced-ocf.htm>
- Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L. & Tang, Y. (2023). A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122–1136. <https://doi.org/10.1109/JAS.2023.123618>
- Yang, X. & Zhu, C. (2024). Industrial Expert Systems Review: A Comprehensive Analysis of Typical Applications. *IEEE Access*, 12, 88558–88584. <https://doi.org/10.1109/ACCESS.2024.3419047>
- Yoon, Y., Guimaraes, T. & Swales, G. (1994). Integrating artificial neural networks with rule-based expert systems. *Decision Support Systems*, 11(5), 497–507. [https://doi.org/10.1016/0167-9236\(94\)90021-3](https://doi.org/10.1016/0167-9236(94)90021-3)
- Zawacki-Richter, O., Marín, V. I., Bond, M. & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J. & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, 2021(1), Artikel 8812542. <https://doi.org/10.1155/2021/8812542>
- Zhang, J [Jia] & Zhang, Z. (2024). AI in teacher education: Unlocking new dimensions in teaching support, inclusive learning, and digital literacy. *Journal of Computer Assisted Learning*, 40(4), 1871–1885. <https://doi.org/10.1111/jcal.12988>
- Zheng, L., Niu, J., Zhong, L. & Gyasi, J. F. (2023). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interactive Learning Environments*, 31(9), 5650–5664. <https://doi.org/10.1080/10494820.2021.2015693>

Zierer, K., Werner, J. & Wernke, S. (2015). Besser planen? Mit Modell! Empirisch basierte Überlegungen zur Entwicklung eines Planungskompetenzmodells. *Die Deutsche Schule*, 107(4), 375–395.

Anhang

A Liste von Publikationen zum Dissertationsthema

1. Zeitschriftenbeiträge (peer-reviewed)

Pargmann, J., Berding, F., Rebmann, K. & Riebenbauer, E. (im Review bei Empirical Research in Vocational Education and Training). How AI feedback supports lesson planning in vocational teacher education: A longitudinal intervention study using an analytical AI platform.

Pargmann, J., Berding, F., Riebenbauer, E. & Rebmann, K. (eingereicht im *Journal of New Approaches in Educational Research*). Engaging with AI: How student teachers interact with an analytical AI platform for feedback and the effect on their lesson plans – An eye tracking study.

Pargmann, J. & Berding, F. (angenommen). Sustainability content knowledge for teachers in accounting and controlling: A systematic literature review on industry approaches, practices and desired competences. *Frontiers in Education*.

Riebenbauer, E., Rebmann, K., Berding, F., Leube, A., **Pargmann, J.**, Slopinski, A. & Gillert, M. (2025). Zum Planungshandeln von Studierenden der Wirtschaftspädagogik. Teil 2: Empirische Befunde zur Analyse von Unterrichtsentwürfen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 1

Riebenbauer, E., Rebmann, K., Berding, F., Leube, A., **Pargmann, J.**, Slopinski, A. & Gillert, M. (2024). Zum Planungshandeln von Studierenden der Wirtschaftspädagogik. Teil 1: Orientierung bei der Planung mit Unterrichtsentwürfen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 4

Pargmann, J., Leube, A., Berding, F., Riebenbauer, E., Rebmann, K., Slopinski, A. & Gillert, M. (2024). Die elektronisch-didaktische Assistenz (EDDA) zur Unterrichtsplanung auf Basis künstlicher Intelligenz: Funktionsweise, Anwendungsbereiche und Forschungsperspektiven. *Empirische Pädagogik*, 38(1), 118–146.
<https://doi.org/10.62350/WLQO3003>

Pargmann, J., Riebenbauer, E., Leube, A., Berding, F., Slopinski, A. & Rebmann, K. (2024). Der Einsatz von künstlicher Intelligenz zur Unterrichtsplanung in der beruflichen Bildung – exemplarische Möglichkeiten mit der elektronisch-didaktischen Assistenz EDDA. *lernen&lehren*, 39(153), 22–28.

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).

<https://doi.org/10.1186/s40461-023-00141-1>

Pargmann, J., Berding, F., Riebenbauer, E. & Flick-Holtsch, D. (2022). Implikationen der Digitalisierung für die Didaktik des Rechnungswesens – konzeptionelle Überlegungen und Vorstellung eines Learning Cockpits. *bwp@*(43), 1–30.

https://www.bwpat.de/ausgabe43/pargmann_etal_bwpat43.pdf

2. Sammelbandbeiträge (ohne Peer-Review)

Pargmann, J., Berding, F., Riebenbauer, E. & Flick-Holtsch, D. (im Druck). Perspektiven für eine nachhaltigkeitsorientierte Didaktik des Rechnungswesens und Controllings. In J. Schlicht, S. Kaiser & F. Schwehm (Hrsg.), *Innovationen für Nachhaltigkeit durch berufs- und wirtschaftspädagogische Forschung*. wbv.

Pargmann, J. & Berding, F. (2022). Künstliche Intelligenz in der beruflichen Bildung - Utopie oder Dystopie? Chancen und Herausforderungen von KI als Bildungstechnologie. In M. Eckelt, T. J. Ketschau, J. Klassen, J. Schauer, J. K. Schmees & C. Steib (Hrsg.), *Berufsbildungspolitik: Strukturen - Krise - Perspektiven* (S. 215–232). wbv.

3. Monographien

Berding, F. & **Pargmann, J.** (2022). *Iota Reliability Concept of the Second Generation*. Logos.
<https://doi.org/10.30819/5581>

4. Artikel in Praxiszeitschriften

Pargmann, J. & Berding, F. (2023). Kann künstliche Intelligenz (KI) das Lernen von Schüler:innen unterstützen? PädF-Kontrovers. *Pädagogische Führung*, 3, 108.

5. Vorträge

Pargmann, J., Pranger, J. & Hantke, H. (2024). Mehr Nachhaltigkeit im Wirtschaftsunterricht: Eine qualitative Exploration von Nachhaltigkeitsberichten als Grundlage von arbeits- und geschäftsprozessorientierten Lehr-Lern-Arrangements. *Sektionstagung der Berufs- und Wirtschaftspädagogik*, Dresden, 18.09.2024.

Pargmann, J. & Riebenbauer, E. (2024). Förderung der Unterrichtsplanungskompetenz durch KI-basiertes Feedback bei Studierenden der Wirtschaftspädagogik: Eine Interventionsstudie mit EDDA. *Sektionstagung der Berufs- und Wirtschaftspädagogik*, Dresden, 17.09.2024.

- Pargmann, J.** (2023). Der Stand der Nachhaltigkeit in Betrieben und betriebliche Kompetenzanforderungen im Rechnungswesen und Controlling. Ergebnisse eines systematischen Reviews. *Sektionstagung der Berufs- und Wirtschaftspädagogik*, Kiel, 08.09.2023.
- Pargmann, J. & Leube, A.** (2023). Künstliche Intelligenz als Assistenz bei der Unterrichtsplanung. Vorstellung der Lernplattform EDDA. Als Teil des wissenschaftlichen Symposiums zu Unterrichtsplanung und Digitalisierung. Stand und Perspektiven für die Berufs- und Wirtschaftspädagogik. *Sektionstagung der Berufs- und Wirtschaftspädagogik*, Kiel, 07.09.2023.
- Stütz, S. & **Pargmann, J.** (2023). Die Förderung der Modellierungsfähigkeit in der Domäne Rechnungswesen durch Schulbuchaufgaben. Empirische Befunde und Implikationen für die Gestaltung von Lernumgebungen. *Sektionstagung der Berufs- und Wirtschaftspädagogik*, Kiel, 07.09.2023
- Pargmann, J.,** Berding, F., Riebenbauer, E., Rebmann, K., Slopinski, A., Jahncke, H., Leube, A. & Gillert, M. (2022). Die Möglichkeiten künstlicher Intelligenz zur Professionalisierung der Unterrichtsplanungskompetenz. Vorstellung einer KI-basierten Lernplattform. *Sektionstagung der Berufs- und Wirtschaftspädagogik*, Freiburg, 27.09.2022.
- Pargmann, J.,** Berding, F., Riebenbauer, E., Flick-Holtsch, D. (2022). Implikationen der Digitalisierung für die Didaktik des Rechnungswesens. Vorstellung eines neuen didaktischen Konzepts für die berufliche Bildung. *Sektionstagung der Berufs- und Wirtschaftspädagogik*, Freiburg, 27.09.2022.

6. Software

- Berding, F., Tykhonova, Y., **Pargmann, J.,** Riebenbauer, E., Rebmann, K. & Slopinski, A. (2024). *AI for Education (aifeducation). A R package for educators and researchers of the educational and social sciences [Computer software]*.
<https://fberding.github.io/aifeducation/index.html>
- Berding, F., **Pargmann, J.,** Leube, A., Gillert, M., Rebmann, K., Slopinski, A., Riebenbauer, E. (2025). *EDDA – Elektronisch-didaktische Assistenz [Computer software]*. <https://wi-paed.ew.uni-hamburg.de/edda/>

B Beitrag 1

REVIEW

Open Access



Digitalisation in accounting: a systematic literature review of activities and implications for competences

Julia Pargmann^{1*} , Elisabeth Riebenbauer², Doreen Flick-Holtsch³ and Florian Berding¹

*Correspondence:
julia.pargmann@uni-hamburg.de

¹ Institute for Vocational
and Business Education,
University of Hamburg,
Sedanstraße 19, 20146 Hamburg,
Germany

² Institute of Business Education,
University of Graz, Graz, Austria

³ Institute of Education,
University of Zurich, Zurich,
Switzerland

Abstract

The digitalisation of processes is a current topic in accounting. New technologies can change activities which in turn may require different skills from accounting graduates. This paper aims to shed light on the changes that digitalisation brings about in various areas of accounting by assessing the types of activities (non-routine and routine) and corresponding competences in the context of progressing stages of digitalisation. In addition, it is analysed how different technologies are used in these activities and where their execution is placed within the supply chain. The systematic literature review shows a lack of expertise in the field of digitalisation that enables graduates and employees to successfully manage respective processes in the workplace. While routine activities are continuously being automated or digitalised, non-routine activities and the corresponding skills have a similarly increasing importance for employees in accounting as the acquisition of general digital competences.

Keywords: Digitalisation, Expertise, Accounting education, Competences, Activities

Introduction

In the context of digitalisation, it is part of everyday life for customers to purchase products and services with the aid of an application online and in physical stores. The streams of products and finances within and between companies are also highly interconnected when they are based on digitalised processes (Appelfeller and Feldmann 2018). Digitalisation thus encompasses the entire supply chain of a product or service. Accounting is one area of the company that documents these processes with customers and other companies, enables these processes, supports them technically, and connects them with internal and external interfaces (Bleiber 2019; Klein and Küst 2020). In this context, companies nowadays face the challenge to make decisions on the introduction of new technologies and digitalised business processes into the area of accounting, among others (Appelfeller and Feldmann 2018). Digitalisation can address various stages, from substitution (e.g., self-check-out counters in retail shops) to business process innovation (e.g. automated storage and payment with RFID chips; see for these examples e.g. Litfin and Wolfram 2010). It is expected that not only costs

and productivity are essential decision criteria for the introduction of technologies and digitalised business processes (Ashoka et al. 2019; Chen et al. 2012) but also the extent to which employees can master and employ technologies and possess the competence to fulfil new or evolving tasks (Aepli et al. 2017; Bonin et al. 2015; Cong et al. 2018; Egle and Keimer 2017; Seeber and Seifried 2019).

In this context, it is necessary to describe how digitalisation affects the requirement profile regarding the competences of employees in accounting. On the general level, the concept of competence is discussed in many ways and often focused on mastering domain-specific tasks and requirements (Hartig et al. 2008; Weinert 2001). The types of activities themselves range from manual routine tasks (e.g. paper-based bookkeeping) to interactive activities (e.g. solving problems with automated digital bookkeeping) (Seeber et al. 2019; Seeber and Seifried 2019). Hence, in order to learn more about relevant competences in accounting in a digitalised world, we have to analyse the different tasks and actual accounting activities that result in changing competence requirements as a basis. There are insightful studies available that analyse changes and developments of domain-specific requirements and competences as well as types of activities in the context of digitalisation (Aepli et al. 2017; Iten et al. 2016; Sachs et al. 2016; Seeber et al. 2019). However, rather little information is offered by these reports when it comes to accounting and its specific activities as well as technologies as tools for these activities. Thus, we assume that knowledge about these (new) activities in accounting and the corresponding digital technologies provide a good starting point to infer possible changes in necessary competences for acting in professional situations. Thus, our research questions are:

- RQ1: Which activities are concerned with digitalisation in which areas of accounting?
- RQ2: How do accounting activities differ in their respective stages of digitalisation?
- RQ3: Where are these activities placed within the supply chain?
- RQ4: Which technologies are used to perform these activities?
- RQ5: Which advantages and disadvantages are connected to these activities?

The results aim to provoke further curricular and didactic discussions about learning and teaching accounting in vocational education and training on the one hand and higher education on the other.

Starting from a specification of digitalisation along its stages and types of activities, we explain in the following the approach for the systematic analysis based on a review of relevant literature. To actually gain insight into concrete developments of activities in accounting, we provide a systematic overview of the types of activities and the stages of digitalisation. Moreover, we consider the specific stages of digitalisation in combination with the different types of activities, as we expect differences between certain types of activities (e.g. interactive activities) and specific stages of digitalisation (e.g. innovation). In modern accounting, technologies are the main set of tools accountants use to execute their tasks, much like the tools of a craft. When technologies change, so do activities. Therefore, technologies need to be considered in this review as well. In addition, the placement of activities within the supply chain is

included, as different departments require different competences and develop different digitalisation projects since the stakeholders differ. Furthermore, the advantages and disadvantages of digitalisation are to be surveyed since they offer insights why activities are digitalized.

In the German-speaking countries, one's ability to follow an occupation successfully is described by the concept of vocational action competence. Established competence frameworks usually differentiate between professional, social, and self-competences as major dimensions with methodological and technological competences anchored horizontally across all major dimensions. To decide which competences are needed for which occupation, activities are analysed that allow these competences to be fostered. Thus, it is relevant to analyse activities in accounting to derive necessary changes in curricula. Our research systematises digitalised activities that we employ to draw conclusions for the competence development of future accountants.

Framework for categorising digitalisation and types of activities

When discussing the consequences of digitalisation for accounting activities, at least two aspects must be considered in depth: first, the specific stages of digitalisation and, second, the types of activities.

Figure 1 shows an overview of three different stages of digitalisation in the context of accounting, each of which is disruptive to a different degree for activities and thus competences (Bleiber 2019; Hübl 2020): *substitution*, *process change*, and *innovation*. While in stage 1 analogue data or activities are substituted by digital ones, new technologies lead to changes in processes in stage 2. Stage 3 addresses innovation, both processes and their respective outcomes are changed.

An insightful framework of digitalised activities is suggested by Aepli et al. (2017), who adopted the classification from Spitz-Oener (2006) and validated it in expert interviews. The framework was chosen because digitalisation is more likely to change activities in a profession rather than erase it in their entirety (Aepli et al. 2017; Arntz et al. 2016; Autor 2015; Seeber and Seifried 2019; World Economic Forum 2018). In addition, some activities are more prone to digitalisation than others.

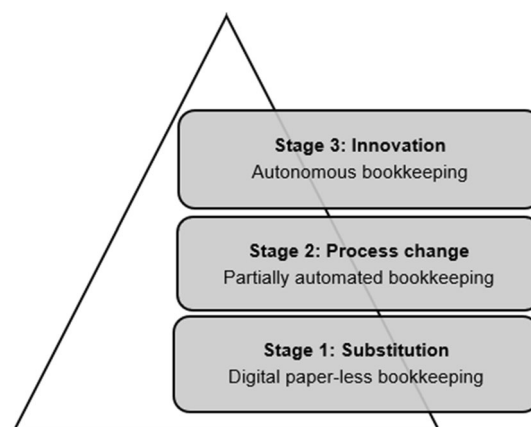


Fig. 1 Stages of digitalisation in accounting

Table 1 Types of activities (Aeppli et al. 2017, 41, examples freely translated and supplemented)

Types of activities	Generic application for different professions	Application in accounting
Manual routine activities	Sort, pack, and ship products and goods	Physically sort and scan invoices and receipts, enter standard invoices and receipts in an accounting program, print analyses
Cognitive routine activities	Data management, check invoices, process transactions	Book incoming invoices, check recorded invoices online, complete incomplete bookings
Analytical non-routine activities	Preparation of teaching and learning processes; planning activities; data preparation, interpretation, and presentation	Analyse monthly and year-end financial reports, develop a plan to improve the dunning process
Manual non-routine activities	Maintain and service machines and equipment	Set up interfaces to customer systems, solve problems with the automatic paper feed of the scanner for the invoices
Interactive non-routine activities	Provide information, teach, advise, instruct	Explain the annual statement to stakeholders, renegotiate payment conditions, internal exchange with the tax advisor

Table 1 illustrates this approach using types of activities (left column) which are intended to provide deeper insight into the facets of potentially digitalised activities with generic examples (middle column) and specific accounting-related examples (right column).

Based on Dengler et al. (2014), we suggest to include automated routine activities, even though they are not performed by professionals (e.g. the automatic import of outgoing invoices from invoicing to accounting). Due to the increasing number of automated workflows in accounting, this adaptation provides the ability to better record and illustrate these changes as they reflect all activities that take place in the different areas of accounting. Thus, the six types of activities provide a comprehensive and differentiated range of activities. This categorization of activities will serve as a framework for our literature review in order to establish comparability and transferability to other (commercial) fields and professions as well.

Research design and methodology

A systematic literature review was conducted to find out how the accounting activities of future accounting professionals have changed due to digitalisation. The four-step process (Fig. 2) started with (1) the definition of up to fourteen keyword groups consisting of technical search terms in both English and German, which for example include “computer-assisted,” “data processing,” “technological speed of change,” “digitalisation rate,” “ERP,” “enterprise-resource-planning,” “accounting software,” “OCR,” “optical character recognition,” “big data,” “blockchain,” or “process mining.” Since the choice of keywords directly impacts the search hits, we used four strategies to ensure that we were able to use as many relevant terms as possible: (a) We conversed with an accounting professor about trends and current developments regarding digitalisation in the field, (b) we scoped the databases that we used for our full search and analysed the most frequently

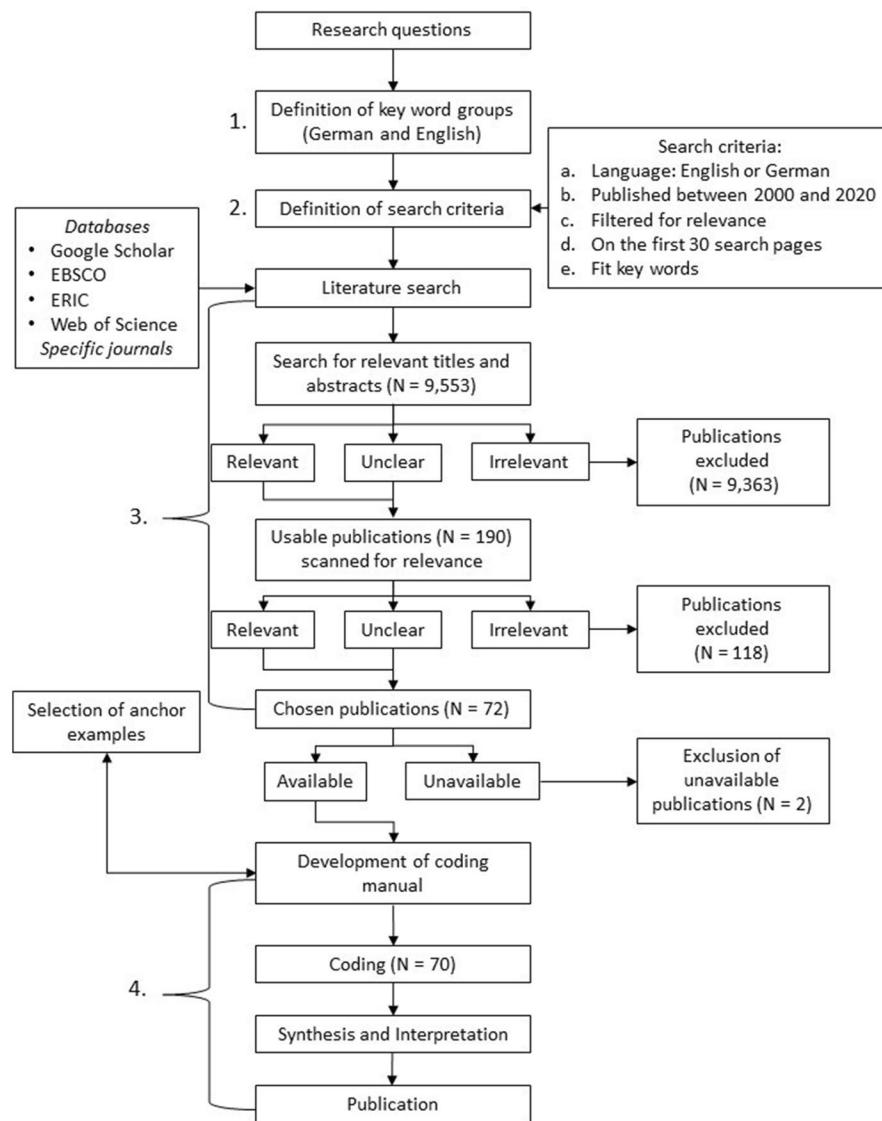


Fig. 2 Flow chart of research design

occurring trends and technologies, (c) we analysed accounting curricula from vocational education and higher education institutions to adjust our concepts of different competence frameworks, and (d) we used a synopsis from different curricula and textbooks to systematise activities. The keywords were grouped because of technical limitations regarding the valid number of connectors, particularly in Google Scholar. To ensure comparability between databases, we used these groups throughout our search.

After the definition of suitable terms, (2) a search was performed in the *Google Scholar*, *ERIC*, *EBSCO*, and *Web of Science* databases and complemented by specific relevant international journals (“Journal of Accounting Education,” “Journal of Emerging Technologies in Accounting,” “Journal of Information Systems”) which were chosen due to their topicality regarding the research questions. They were selected based on their impact and the number of search hits. Due to the increasing speed of digitalisation, the

period between 2000 and 2020 was selected. In the later stages of the study, the research group also decided to include a 2021 article due to its topical suitability.

While performing the content analysis, we applied a strategy that limited the number of hits. Only the first 30 pages of search hits for each keyword group were scoped for relevance. This was necessary because of the broad scope of the search terms. This specific cut-off point was chosen because, in a preliminary scope, the number of potentially useful search hits declined drastically. On average, no more (potentially) relevant hits were identified after 29 to 30 pages, thus setting this number for the full search as well. This process yielded a total of 9,553 potentially relevant hits across all media. A scan of both titles and abstracts resulted in (3) a total of 190 potentially suitable sources of which 72 articles proved relevant to the research questions. Irrelevant publications either referred only to accounting *or* digitalisation, were purely didactic, or had a general IT-orientation without an accounting focus. If the relevance could not be decided, the full text was scanned for the connection between accounting and digitalisation. The last step (4) was the content analysis. A total of 70 articles were retrievable and thus used in the coding process. The coding manual was developed and completed with appropriate anchor examples to illustrate the variety of possible accounting activities and to align our general understanding of the categories. The coding ensued in pairs to promote discussion.

Analytical framework

The coding tree (Fig. 3) consists of (a) publication information, (b) the publication's main foci and (c) types of activities (automated, manual and cognitive routine activities, manual and analytical and interactive non-routine activities) in different areas of accounting (financial accounting, controlling, balancing, financial forecasting, managerial accounting, 'other'). Moreover, the coding tree included (d) respective digitalisation stages of the activity (substitution, process change, innovation) as well as (e) advantages and disadvantages of digitalisation in context of the activities (e.g. financial and time-specific aspects, complexity, required expertise, transparency, data protection).

As a starting point, a few base publications were chosen to be scanned for relevant aspects regarding the research questions. From these notes and discussions, the coding tree was further developed and adapted after the first publications were coded. Advantages and disadvantages were developed inductively while doing the first base

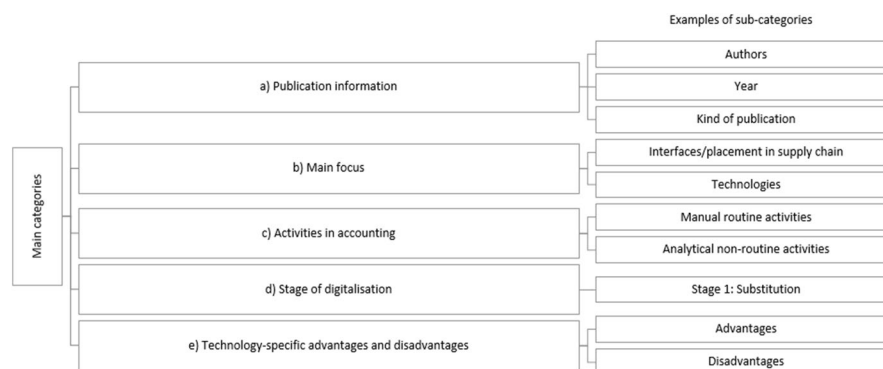


Fig. 3 Coding tree with sample sub-categories

coding. The framework was developed further by using the basic principles of content analysis (Kuckartz 2018).

In the course of a special coding training, the final coding tree in the software MaxQDA was then distributed to two coding teams (three and two people each) along with the coding manual (Fig. 4). The sub-categories in the main category “Activities in Accounting” were further distinguished according to the aforementioned types of activities (Table 1), reaching from *automated routine* to *interactive non-routine activities* (Fig. 9). This more detailed structure allowed interpretations about the manners in which activities might have changed due to digitalisation.

To help the coding process, anchor examples were selected for all categories. As this paper is centred around activities, we are going to illustrate this process for the activity-related categories. For example, an *automatic activity* in accounting is characterised by a fully automated workflow like an automated deduction of worst-case scenarios with algorithms. A *manual routine task* does not require cognitive activation as its execution is part of the employee’s regular working routine, for example sorting accounting documents by date or verifying by hand if the positions on an invoice are complete and filled in correctly by the software. In contrast to this, a *cognitive routine activity* is constituted by a cognitive activation as it would be needed for manual corrections in bookkeeping or the adjustment of false journal entries. An *analytical non-routine task* could, for example, be the preparation and analysis of various financial statements, while *manual non-routine activities* encompass all tasks that are not part of everyday tasks, like solving technical problems or handling issues regarding hardware maintenance. Lastly, *interactive non-routine tasks* require a communicative element, like onboarding processes in the division or consultation with management. Thus, there are different levels of activities that need further specification and presumably undergo different changes through digitalisation.

To identify the extent of changes in the field, activities were assigned to a digitalisation stage. Hence, we applied the three aforementioned stages of digitalisation based on Bleiber (2019) and Hübl (2020) to identify the scope of changes in accounting: *substitution*, *process change*, and *innovation* (Fig. 1).

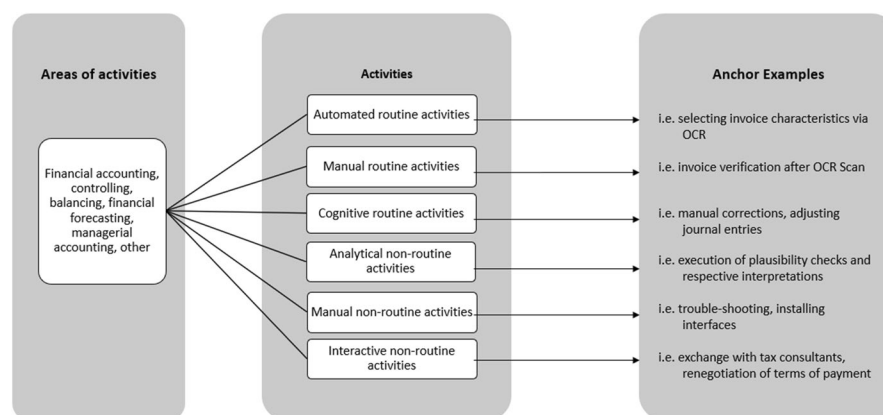


Fig. 4 Overview of the coding manual for accounting activities

- Stage 1, substitution, is completed whenever analogue data or activities are exchanged for automated or digital ones. Neither process nor output is changed, however. This stage applies to activities using, for example, OCR or robotic process automation (RPA) or whenever the publication implies that a manual activity was exchanged for automation.
- In stage 2, process change, workflows, and processes are altered by the use of technology while the output of the process remains the same. This stage applies to all activities that use improved RPA technologies as well as process optimising or mining tools, such as partially automated bookkeeping.
- Stage 3, innovation, implies that both processes and their respective outputs are transformed through the use of technology, like artificial intelligence (AI), (deep) learning systems, machine learning, and neural networks. An example of this stage is a fully autonomous bookkeeping workflow.

Results

Bibliographic information and main foci of publications

The majority of publications are in German ($n=51$) and in English ($n=19$). The dominance of German publications occurred due to the exclusion criteria. In our original set of 190 possibly relevant publications (Fig. 2), the distribution was fairly balanced with 98 publications in German and 92 in English. However, we excluded those publications which dealt with IT in a general manner (missing a direct link to accounting) and those which focused on other aspects of accounting but did not explicitly mention accounting activities. Most of these publications were in English, hence the large difference in the final data set. While most publications are either book chapters ($n=22$) or journal articles ($n=21$), whitepapers ($n=17$), books ($n=5$), and university publications ($n=5$) constitute the minority. Looking at the year of publication, most sources were published between 2015 and 2020 (Fig. 5), fewer sources between 2000 and 2005. This rapid increase in publications indicates that digitalisation in accounting has gained popularity within the past five years and is now a leading topic in accounting publications.

The sub-categories in “main focus” could be coded multiple times within the same publication. Besides the technologies and software systems in use, the two most frequently mentioned aspects regarding the main foci are *developments in accounting/accounting 4.0* ($n=144$) and *implications for accounting education* ($n=108$), a category that describes possible changes in the configuration of activities. (Fig. 6). These two categories are future-oriented and involve both the (technical) developments and the industry’s dynamic requirements for accounting graduates and implications for education.

Regarding the *developments in accounting*, authors primarily describe processes that are improved and standardised to maximise productivity and the accountant’s value to the company (Baier 2019; Müller and Reichmann 2010). Especially technologies like advanced analytics and RPA are becoming increasingly important in accounting 4.0 (Egle et al. 2020; Koch 2017; Losbichler and Gänßlen 2018; Satzger et al. 2018).

Special aspects that were not mentioned in an educational context, such as the related legal requirements ($n=75$) or the generally increasing data availability ($n=39$), are shown in separate categories. Several publications also deal with the concrete

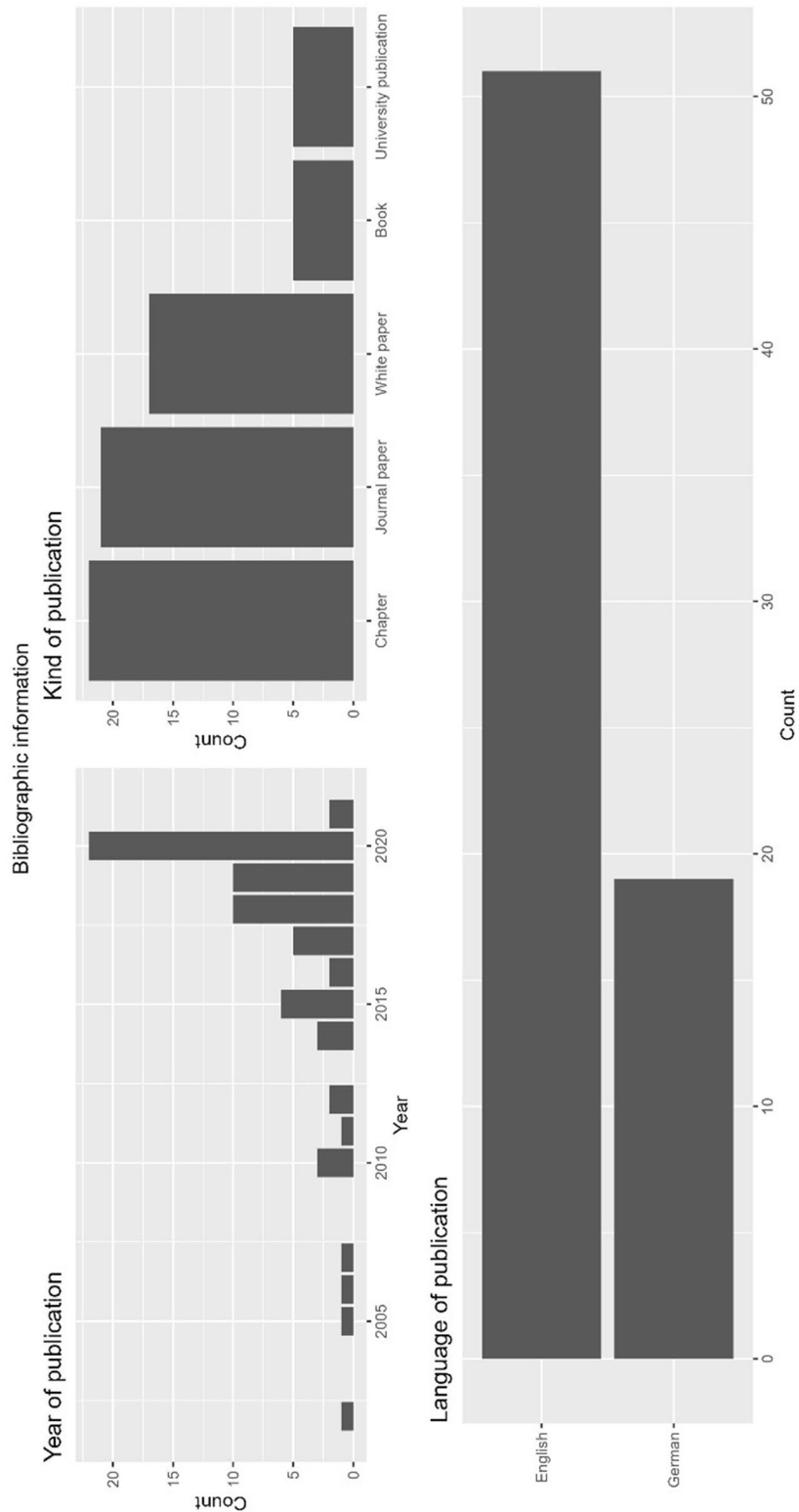


Fig. 5 Bibliographic information (year of publication, kind of publication and language of publication)

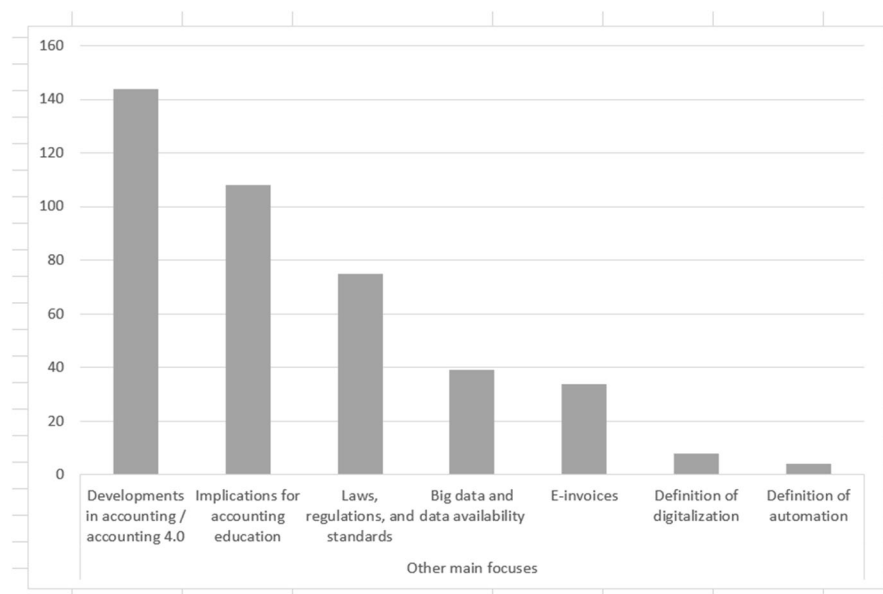


Fig. 6 Main foci of publications (number of codes, multiple codes per paper possible)

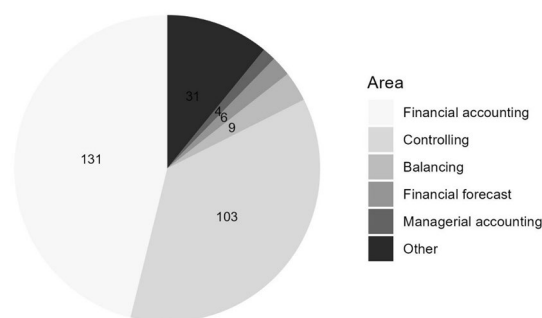


Fig. 7 Areas of activities in accounting divisions

implementation and procedure of electronic invoices ($n=34$). The categories *definition of automation* ($n=4$) and *definition of digitalisation* ($n=8$) are mentioned the least frequently.

RQ 1: Activities concerned with digitalisation in the different areas of accounting

Areas and types of activities in accounting

A total of $n=285$ activities are distributed across six accounting areas including “other” activities (Fig. 7). Most frequently mentioned are activities concerning financial accounting ($n=131$) and controlling ($n=103$). In contrast, all other divisions together make up roughly 25% of the activities described in the publications. The financial accounting code is given to all activities that are rooted in the accounts payable, accounts receivable, banking, payroll, and asset accounting areas (Fig. 8). The activities in the area of financial accounting are dominated by *accounts payable accounting* ($n=98$). Many publications focus on e-invoicing, a technology employed to increase productivity in both accounts payable and *accounts receivable accounting*, the second most mentioned sub-division

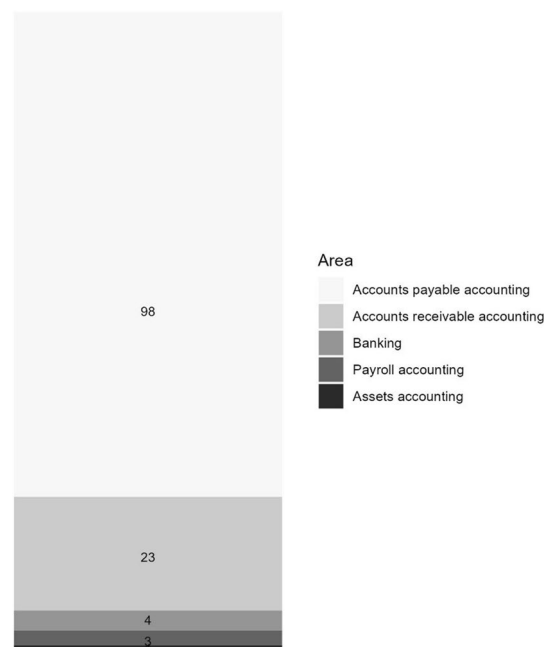


Fig. 8 Areas of activities in financial accounting sub-divisions

($n=23$). In addition, OCR and RPA technologies are often used to increase efficiency in both divisions.

The second most frequently mentioned category regarding activities, controlling, mainly encompasses aspects such as reporting, communication, and interpretation of data (Bär et al. 2019; Egle and Keimer 2017; Heupel and Lange 2019). In contrast to financial accounting, the category of controlling does not focus as much on technologies but rather on specific changes in the job profile (Becker et al. 2020; Losbichler and Gänßlen 2018; Schindera et al. 2018).

The types of activities in the different areas of accounting (Fig. 9A) are differentiated by their mean of action. More specifically, Fig. 9A describes what kinds of activities typically occur in the different areas of accounting according to our analyses. In Fig. 9, those numbers are given as relative numbers because a publication had sometimes multiple codes of the same activity. To account for these duplicates and to maintain proportions, the absolute frequencies are divided by the respective totals per sub-category. As an example, the category *financial accounting* is mentioned 131 times of which 58 mentions are for automatic activities. Thus, there are *automatic activities* without manual dimensions and others that are *routine activities* (manual and cognitive) and *non-routine activities* (manual, interactive and analytical).

In the area of financial accounting, most reported activities are either automatic or manual routine activities. Automatic activities are mentioned frequently, for example the automatic recognition of invoices via RPA or their generation directly from the ERP system, although a fully automated process is not common practice yet (Appelfeller and Feldmann 2018; Jordanski 2020; Kreher 2021; Schömburg and Breitner 2010; Tanner 2016). Instead, some parts of the process are automated and some remain manual. This explains the ratio of manual routine activities, as publications often address the manual

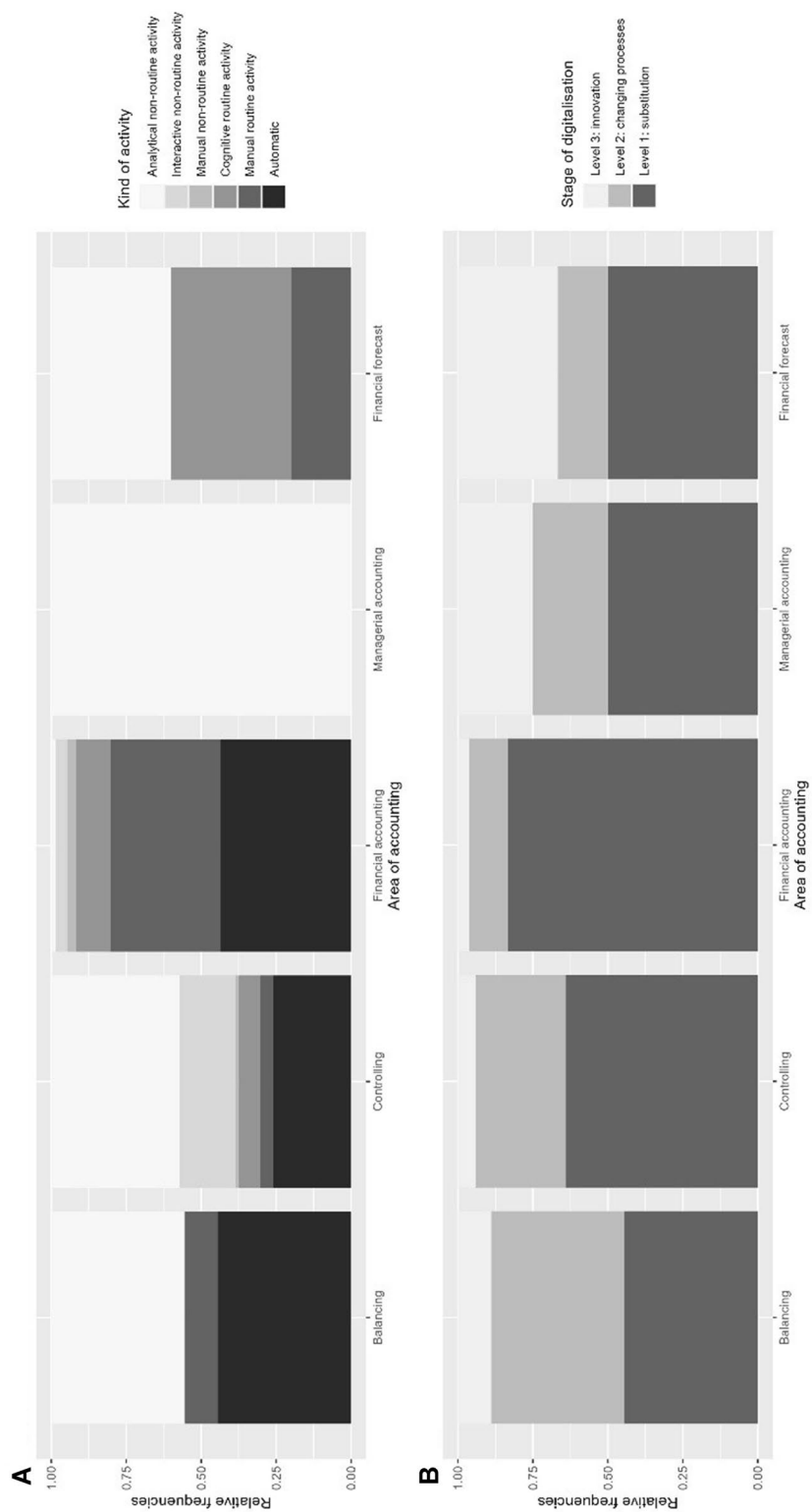


Fig. 9 A–C Types of activities in different accounting areas

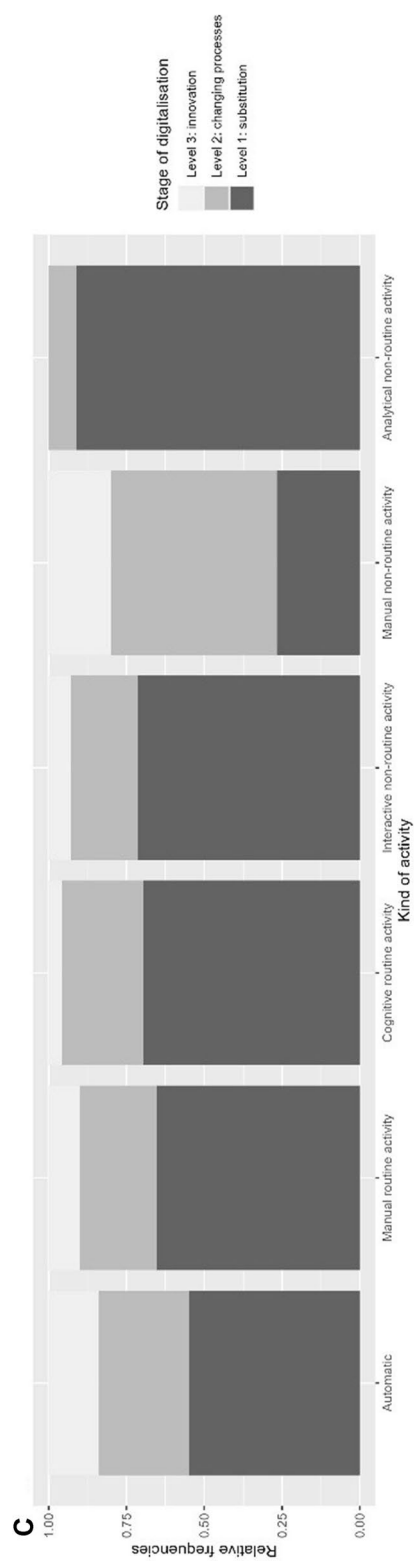


Fig. 9 (continued)

correction or computation of invoices (Bernius and Kreuzer 2014; Koch 2017; Menges 2012; Wilczek 2014). Other areas of accounting like *balancing* or *controlling* display a similar amount of automatic activities that primarily include the use of AI in balancing (Kink 2007; Le Guyader 2020). Another aspect is the automated detection of variances and automated reporting through machine learning in controlling (Alexander et al. 2019; Ashoka et al. 2019; Jonen 2020). In contrast to publications that focus on financial accounting, authors who address managerial accounting topics identify a high proportion of analytical non-routine activities that primarily include the use of RPA to distribute reports and allocate resources which then have to be analysed by human employees (Langmann and Turi 2020).

RQ 2: Accounting activities in different stages of digitalisation

The codes regarding areas of activities in accounting alone, however, do not provide any information regarding the quality of the activities or the respective stages of digitalisation that were presented earlier in this paper. To formulate statements about the quality of accounting activities, it is necessary to analyse the stage of digitalisation and the respective frequencies in publications (Fig. 9B–C).

Stages of digitalisation in the different areas of accounting

Regarding the stages of digitalisation, the first stage (substitution) is assigned the most across all accounting divisions (Fig. 9B). Especially activities that remain in the first stage of substitution are attributed to financial accounting. In contrast to this area of accounting, both financial forecasting and managerial accounting areas show a larger proportion of digitalisation that has progressed to stages 2 and 3, process change and innovation. As to managerial accounting, this extends to the integration of a controlling software that introduces innovations in processes in the respective divisions and the implementation of new processes (Kink 2007; Raschig and Schulze 2020; Selb 2020). In financial forecasting, most processes are still in the substitution stage (Klein and Küst 2020), but some have progressed to integrate process change or innovation, such as the software-based analysis of reporting data for a proactive controlling approach as shorter product lifecycles necessitate shorter planning horizons (Kink 2007; Sledgianowski et al. 2017).

Stages of digitalisation in the different activities

In extent to the types of activities (Fig. 9C), most *routine tasks* (see Table 1 for these categories) are still mainly in the substitution stage. For instance, RPA technologies are used regularly to substitute manual with automated processes (Gadatsch 2020) or electronic invoicing is integrated (Jordanski 2020; Klein and Küst 2020; Pagel 2019; Schömburg and Breitner 2010). Approximately one third of automated activities are in the stage 2 of process change, for example when the invoicing process is adapted to a digital format for both incoming and outgoing transactions (Kreher 2021).

For *non-routine activities*, the trend is mostly different. Interactive non-routine activities consist of a larger proportion of stage 2 digitalisation. Those process changes established in the publications include vendor inquiries through automated online portals (Binkow 2015) or the supervision of RPA systems or AI algorithms (Hmyzo and Muzzu 2020; Klein and Küst 2020). Processes need to be standardised to allocate controllers

more resources to handle more detailed analyses, interpretations, and communications of results and to minimise their routine activities (Keimer and Egle 2020). Schindera et al. (2018) express that, while the chief financial officer slowly transforms into a chief performance officer, employees in the controlling division are confronted with a changing of roles, too—into that of a business partner. This change is conditioned by the use of big data analytics. Thus, analytical non-routine activities gain in importance. Across the areas of accounting, there is a connection between the kind of activity and its respective stage of digitalisation that has not yet been quantified. While the majority of automatic and routine activities remain in stage 1 digitalisation, *interactive non-routine activities* predominantly are assigned to stage 2. This finding can be explained by the successful implementation of software solutions and technologies such as AI leading to more activities involving it in the second stage (Chen et al. 2012; Heupel and Lange 2019; Le Guyader 2020; Losbichler and Gänßlen 2018).

Analytical non-routine activities, however, show a contrasting result with most activities in stage 1 as employees need to make manual corrections (Jordanski 2020). Thus, a human component is required in most processes. However, technologies such as blockchain or AI provide the chance to increase the stage of digitalisation in the future (Chen et al. 2012; Grönke and Heimel 2015; Trachsel and Bitterli 2020). In contrast to manual routine activities (e.g. e-invoicing), the integration of software solutions for analytical non-routine activities requires more resources, hence the lower threshold (Arbeitskreis Externe Unternehmensrechnung der Schmalenbach-Gesellschaft für Betriebswirtschaft e. V. 2018; Egle et al. 2020; Pagel 2019; Weiber et al. 2002).

RQ 3: Placement of activities within the supply chain

Different types of the described activities in different accounting areas can be established and interlinked within a company with the aid of interfaces (Fig. 10). Within a single company, there can be a multitude of interfaces across divisions and along the supply chain. These interfaces are often connected to the activities, for example internal and external financial accounting, which process inbound and outbound invoices or to the procurement division that interacts with vendors. Most papers address one or more relevant interfaces. Most frequently mentioned are *internal interfaces* ($n=46$), especially regarding the adaptation of accounting and controlling division structures to new business processes or ERP implementation (Binkow 2015; Gadatsch 2020; Heupel and

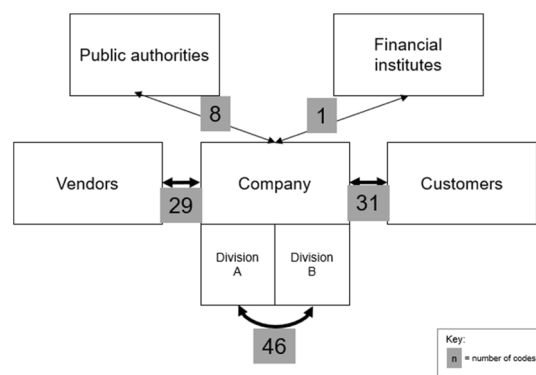


Fig. 10 Interfaces (placement of activities within the supply chain) addressed in the publications

Lange 2019; Najderek 2020; Suden 2010). The rising focus on process management to increase productivity is depicted in the interfaces; more extensive cooperation between financial divisions due to cross-sectional processes gains in importance (Arbeitskreis Externe Unternehmensrechnung der Schmalenbach-Gesellschaft für Betriebswirtschaft e. V. 2018; Bayerl et al. 2020). A functioning flow of information and close cooperation between divisions can aid the success of digitalisation projects. Implementing standardised software across the company is identified as a suitable approach to minimise interface issues (Becker et al. 2020; Gadatsch 2020; Hecht and Scherrer 2020).

Interfaces with *vendors* ($n=29$) or *customers* ($n=31$) are also mentioned. The main context of vendor interfaces lies in the dimension of electronic invoicing and the conditions for successful implementation as well as specific processes (Appelfeller and Feldmann 2018; Klein and Küst 2020; Menges 2012; Najderek 2020; Pagel 2019; Suden 2010; Tanner 2016). It is necessary to harmonise the vendor's system requirements with the company's system requirements (Tanner 2016). Sometimes, companies establish a supplier self-service that enables the vendors to manage processes such as invoicing or the placement of orders within the company's interface (Appelfeller and Feldmann 2018). At the same time, interfaces with customers are mostly focused on improving their experience by distributing invoices digitally or implementing customer self-service (e.g. checkout via smartphone), thus optimising the company's commodity, liquidity, or information flows (Appelfeller and Feldmann 2018; Binkow 2015; Cong et al. 2018; Egle and Keimer 2017; Jonen 2020; Klein and Küst 2020; Nagel 2018).

RQ 4: Technologies used to perform the accounting activities

Most publications address one or more technologies in the context of accounting (Fig. 11). The specific technologies mentioned most frequently are big data analytics ($n=32$), RPA ($n=26$) and AI ($n=26$). Process mining, which allows companies to analyse their processes through digital technology, is mentioned least frequently ($n=8$). This result supports the perception of those technologies as the main disruptors in the industry. *Other technologies* are mentioned in 44 publications. These are mostly generalised statements that refer to basic technologies and software like ERP systems or technology that can be used to process e-invoices (Gadatsch 2020; Koch 2017; Schindera et al. 2018; Schömburg and Breitner 2010; Tanner and Wölflé 2005; Wilczek 2014).

Ashoka et al. (2019) suggest that the use of AI in accounting will free up employees' time capacities and reduce repetitive activities. They also predict a shift towards internet-based reporting. Egle and Keimer (2017) add that the number of repetitive activities will likely be reduced through AI and big data analytic tools, primarily by minimising the number of manual decisions accountants and auditors have to make in their routine activities. These internal factors are supplemented by external aspects that address the customer journey. Especially, AI is used to improve the customer experience (Appelfeller and Feldmann 2018; Bayerl et al. 2020; Egle et al. 2020). In addition, big data is often mentioned in the context of customers (Gadatsch 2020; Losbichler and Gänßlen 2018; Schindera et al. 2018). Yet, certain advantages for other areas of accounting are made visible as analysis tools may aid, for example, auditors in data management or controllers in forecasting (Pan and Seow 2016). This may be the first indication of a shift in the scope of activities in accounting and controlling towards increasing digitalisation.

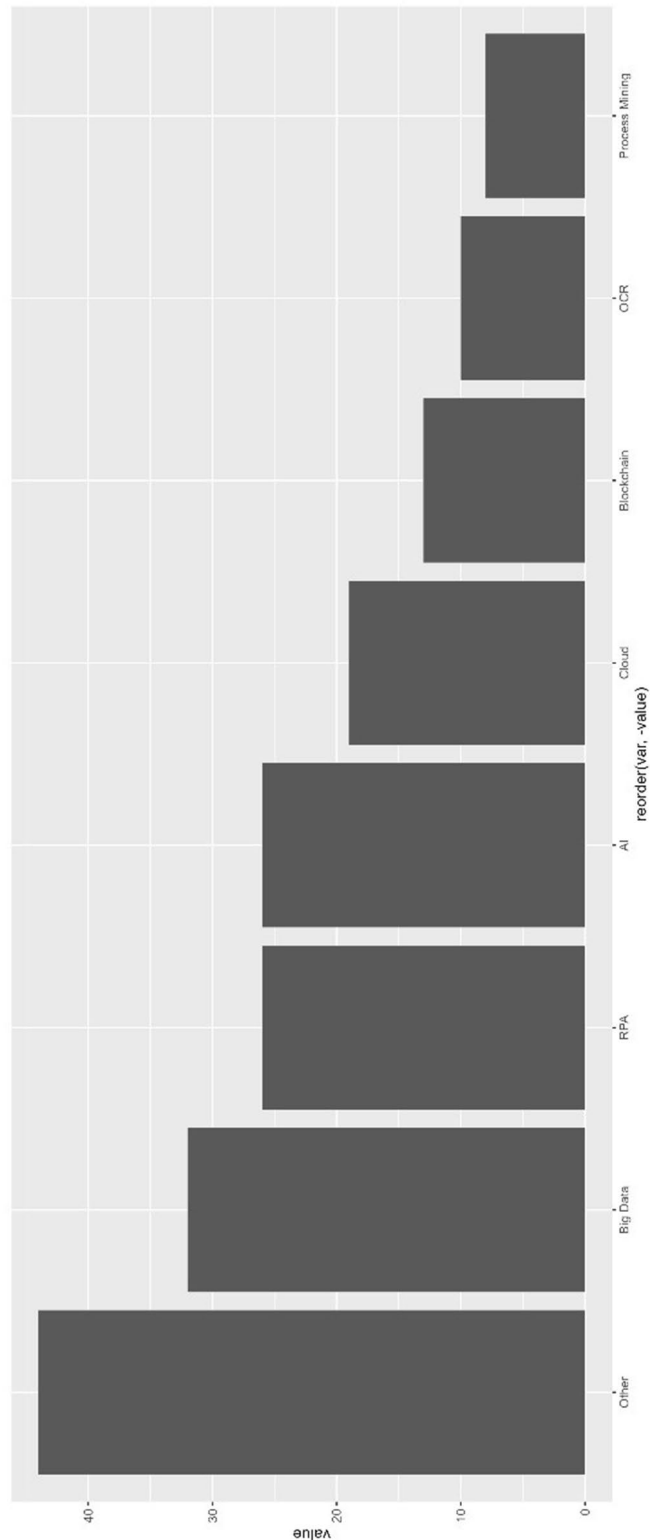


Fig. 11 Frequency of mentioned technologies

Especially, repetitive and physical activities such as monthly reporting or e-invoicing profit from the implementation of RPA (Bowles et al. 2020; Sandner et al. 2020). However, controlling and accounting divisions can be further improved, too (Bayerl et al. 2020; Koch 2017; Langmann and Turi 2020). Through RPA solutions, up to 50% of all back-office tasks—like the synchronisation of supplier and customer accounts—could be automated in the upcoming years (Koch 2017). Additionally, the number of employees necessary to manage the workload can be reduced or shifted to new tasks (Kirchberg 2017). Up to 36% of companies already use RPA or AI for individual tasks in their accounting department (Kreher 2021).

RQ 5: The advantages and disadvantages of digitalised accounting activities

In the review, a variety of advantages and disadvantages of digital technologies in accounting became apparent (Fig. 12). In total, we found almost three times as many advantages ($n=618$) named than disadvantages ($n=249$).

The advantages mentioned most frequently were *saving costs* ($n=90$), *increasing efficiency* ($n=76$), and *saving time* ($n=73$). One example is that digital tools help accountants to manage their workload better and at lower costs (Pan and Seow 2016). Keimer and Egle (2020) state that technologies such as big data analytics or cloud solutions are useful to reduce the susceptibility to errors while simultaneously increasing efficiency in the controlling division. In addition, process automation is perceived as a key aspect of efficiency enhancement in the auditing sector (Werner and Gehrke 2011). The advantage of *saving time* is connected to two time-related opportunities. Firstly, Pagel (2019) mentions the aspect of processes optimised by digital technologies; lead time for invoices was reduced by up to 15 days, leveraging liquidity on the biller's side. The freed-up worktime indicates the second perceived opportunity: The specialised workforce in the controlling division can shift their responsibilities to less administrative tasks, thus extending current activities to strategy-oriented areas (Egle and Keimer 2017).

These productivity-related factors are mentioned in half of the publications. The advantages of *economies of scale* ($n=7$) and *productivity* ($n=11$) are only rarely mentioned. For example, Egle and Keimer (2017) discuss *productivity* as an advantage that results from saved time and error reduction.

“Other” advantages ($n=57$) range from employee engagement and motivation (Langmann and Turi 2020) to simplified auditing processes (Sledgianowski et al. 2017) or improved understanding of financial events (Ashoka et al. 2019). But customer satisfaction (Binkow 2015) and sustainability-driven aspects such as CO₂ reductions (Bernius and Kreuzer 2014) were mentioned as additional aspects as well.

Regarding the disadvantages, content analysis shows the highest risks with digital technologies in accounting in the areas of *interface management* ($n=40$), *financial investments* ($n=42$), and *missing expertise* ($n=31$), mainly with new technologies (Fig. 12). Gadatsch (2020) appoints the central issue of interface management to the complex implementation and its extensive time frame. In addition, data redundancies created by the parallel upkeeping of former and new ERP systems might further increase interface issues, effectively increasing the workload. Thus, there is a need for technological and methodological skills for successful interface management. Schömburg and Breitner's (2010) explanations support this result. They identify the main reasons for companies to

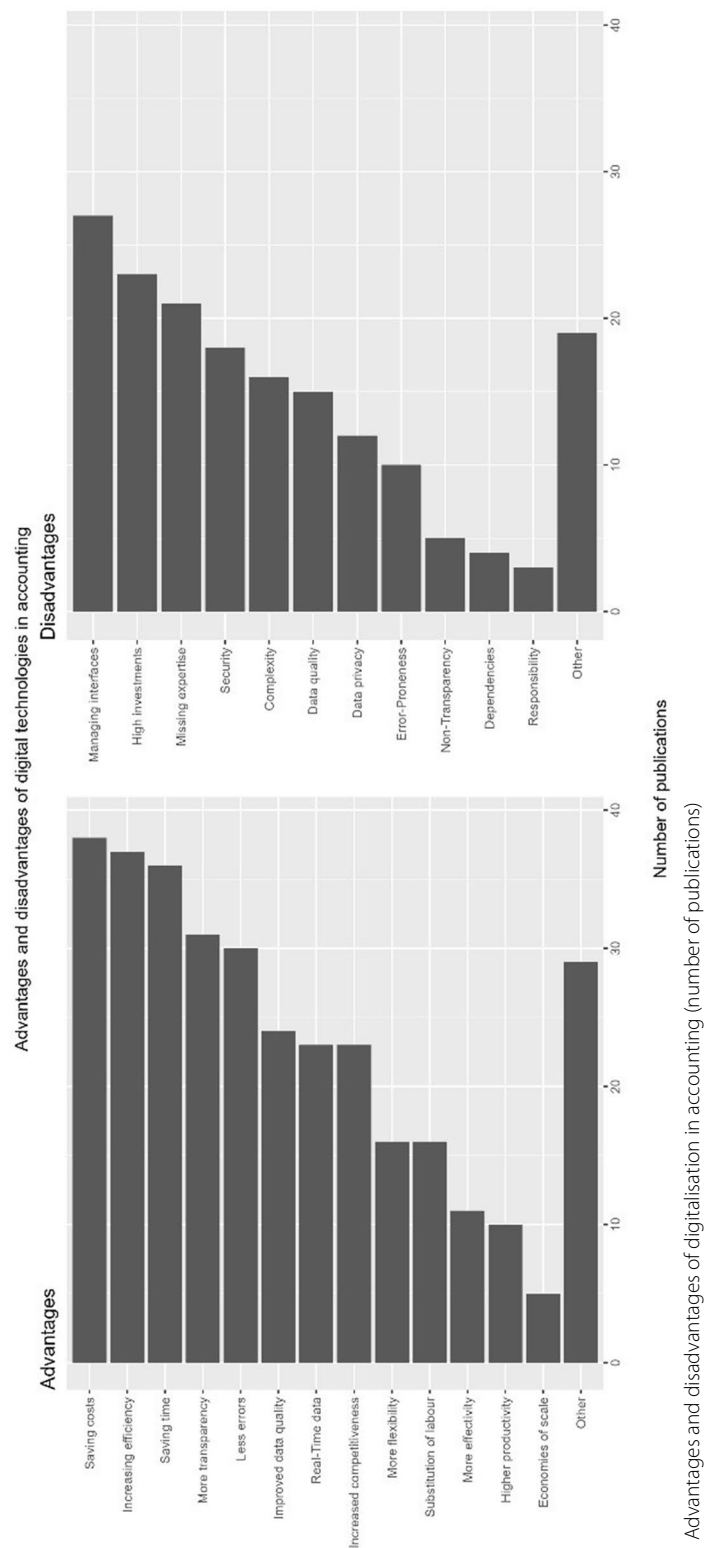


Fig. 12 Advantages and disadvantages of digitalisation in accounting (number of publications)

defer process digitalisation as a lack of expertise, necessary investments, and implementing suitable software. This finding primarily connects to electronic invoicing (Jordan-ski 2020; Koch 2017; Nagel 2018; Najderek 2020; Schömburg and Breitner 2010; Tanner 2016; Tanner and Wölfl 2005) and ERP-specific areas (Appelfeller and Feldmann 2018; Binkow 2015; Chen et al. 2012; Cong et al. 2018; Fuller and Markelevich 2020; Gadatsch 2020; Jonen 2020; Wilczek 2014). The “other” disadvantages ($n=32$) focus on the risks of decentralised innovations that are not implemented company-wide (Schindera et al. 2018) as well as the amount of time needed for software maintenance or to train employees (Gadatsch 2020). In addition, a few technology-specific disadvantages are mentioned: On the one hand, the implementation of RPA does not require the innovation of new processes, thus old, ineffective processes could stall future changes (Gadatsch 2020). On the other hand, blockchain could hinder scalability due to slow transaction speeds and challenging error corrections (Fuller and Markelevich 2020).

Discussion and summary

Summary

Digitalisation is changing numerous processes in the working lives of accounting professionals. This also applies to processes associated with accounting. Against the background of the question of how accounting graduates will have to be trained in the future, we have identified requirements and activities with the help of a systematic literature review. The literature review was conducted based on a conceptual framework of six types of activities and three stages of digitalisation. Figure 13 gives a condensed overview of the most important findings:

Regarding the foci of publications dealing with digitalisation in the field of accounting, we found that developments in accounting and implications for accounting

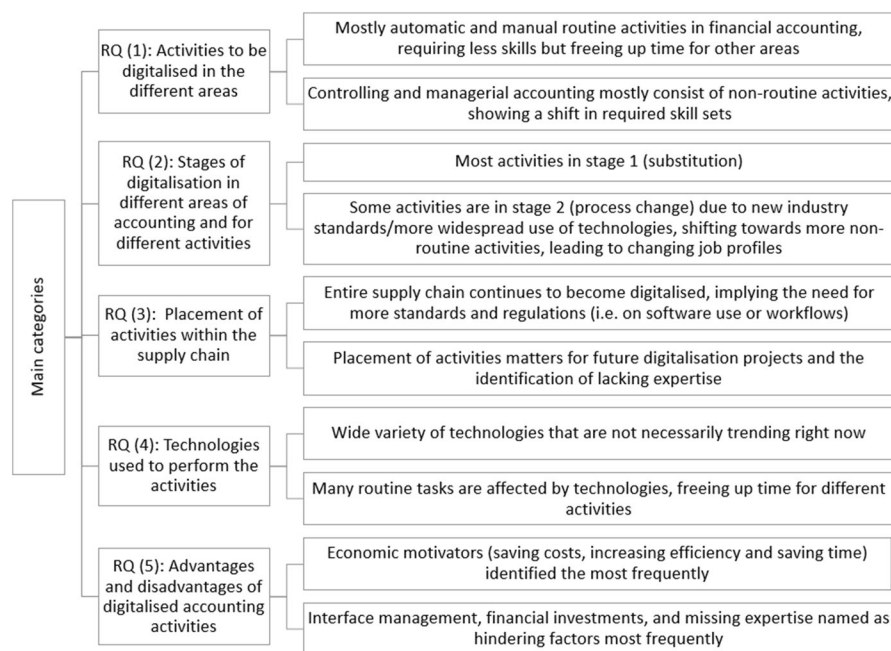


Fig. 13 Overview of findings

education dominate. These findings indicate that the discourse in the field has moved beyond the definitory stages and towards the identification of needs of action, including, among others, the advancement of accounting education and adjustments of curricula.

Results for RQ (1) which deals with activities in the areas of accounting that are subject to digitalisation hint at financial accounting and controlling divisions being the main areas of change. While in financial accounting, there is a focus on technology-based change, the area of controlling mostly encompasses changes in the controller's job profile that are ensuing due to the increasing digitalisation of current tasks. As most activities in financial accounting are based in the area of accounts payable, the focus on technologies could be connected to efficiency-related opportunities that technological updates may implicate. Out of the six types of activities, analytical non-routine, automatic, and manual routine activities are the predominant categories (see Table 1 for an overview of definitions). This finding can be linked to the different degrees of automation in the different areas: In financial accounting and balancing, a large share of tasks has been automated, for example by RPA. In controlling, financial forecasting, and managerial accounting, however, the leading tasks are data analysis (cognitive routine) and consultancy (analytical and interactive non-routine). These results also provide first hints at the stages of digitalisation in the different areas of accounting (RQ 2).

In RQ (2), we identified different stages of digitalisation in the different areas. The results show that the proportion of digitalised activities is largely dependent on the area, yet most activities currently remain in the substitution stage. This indicates that the integration of technologies and process changes is still underway. The combined analysis of areas, activities, and stages was particularly revealing because it allowed differentiated interpretations. The analysis showed, among other things, that types of activities and different accounting areas are digitalised at different stages. Stage 1 (substitution) is to be expected, for example, in financial accounting and analytical non-routine activities. In contrast, interactive non-routine activities can be found more often in the second stage of digitalisation because current processes are supplemented by new technologies. This finding matches the results from a study by Sachs et al (2016) who focused on changing job descriptions in the commercial field. Interaction is required to communicate results and consult with management. This leads to expanding job profiles and further required competences that trace back to updates in accounting education.

All the activities we identified require different skill sets and are placed at different interfaces along the entire supply chain (RQ 3). This implies the supply chain's continuous digitalisation requiring new technical standards and agreements. For example, the increasing popularity of e-invoicing has motivated efforts regarding the standardisation of software and workflows.

Another aspect that shifts competence requirements is the implementation of major technologic trends such as AI and big data analytics in the industry (RQ 4). Frequently mentioned technologies and corresponding business processes need to be integrated into accounting curricula to meet industry needs. On the one hand, technology itself and its evaluation could be addressed. On the other hand, the handling of technology could be prepared when it is integrated into industry practice. Learning materials need to reflect these changes and thus become more digitalised as well.

With RQ (5), we analysed the advantages and disadvantages that the publications under analysis identified during the process of digitalisation in the field of accounting. Even if the advantages mentioned in the publications, such as efficiency reasons, outweighed the disadvantages, risks such as interface management and accounting professionals' lack of skills and competences must still be considered. In consequence, companies lack the expertise to successfully manage digital transformation processes, underlining the importance of specific competence development. Other factors that are main motivators are of a classic economic nature, such as an increase in overall productivity by saving time and financial resources. Both technologies' specific disadvantages and the process of digitalisation in accounting more generally suggest the need for improved accounting graduate education and further training or at least for more guidance for both employers and employees. This need applies to concise abilities, such as the assessment of digital technologies and processes, and the criteria-based evaluation of the potential of software implementation. Depending on how quickly corresponding competences can be developed in the company's workforce, flexibility in competence acquisition for varying technologies and the successful implementation as well as maintenance of technological interfaces is needed. Thus, the field of accounting has reached a level of digitalisation that entails changes in competence profiles to enable both accounting professionals and graduates to manage new demands.

Limitations

Methodological limitations

Concerning limitations, our paper focuses primarily on results from a literature review and not on the analysis of real company situations or accounting professionals' reports. The activities, stages, and combinations described in the articles were evaluated according to the number of codes. Whenever a publication mentioned the same code in a different context, it was coded as a new mention. This could potentially distort results. In addition, we did not analyse the extent to which technologies have actually been established in companies and in particular in accounting. The papers might reflect the past and current states of the discussion.

The next limitations stem from the selection process of publications and their respective language of publication. While there were specific journals in English selected as relevant, we did not select specific German journals. Regarding the dominance of German-speaking publications, it is possible that the databases we used, particularly Google Scholar, produced a bias due to their search algorithms that might include geolocation. To balance this out, we searched for our keywords in both English and German. However, we cannot ignore that this might have implications for the results of our study. In German-speaking countries, accounting methods are rather defensive compared to English-speaking countries, where future-oriented perspectives are more common. In consequence, it is possible that some of our results only pertain to Germany and other German-speaking countries as well as to some to other countries, especially those results that are connected to the flexibility and design of accounting curricula or the stage of digitalisation in the industry.

Content-related limitations

Another limitation is the focus on controlling and financial accounting in the keyword groups as other areas of accounting might have been underrepresented. As for the integration of digitalisation into accounting curricula and learning materials, our paper does not consider external factors that might influence the creation and adaptation of accounting curricula like accountancy professional bodies and other accrediting organisations.

Implications for accounting education and professional practice

The lack of expertise takes a prominent position in our analysis, which is mainly reflected by the prevention of digitalisation processes and innovation of accounting activities. Particularly, technological and methodological skills are needed to integrate new technologies, innovate processes, and, in consequence, produce a shift in activities. The following key implications can be deducted for accounting education and professional practice (see for more details Berding et al. 2022):

- Adaptation of curricula to reflect the change of focus from financial accounting and routine activities to the area of controlling and non-routine activities ("RQ 1: Activities concerned with digitalisation in the different areas of accounting" and "RQ 2: Accounting activities in different stages of digitalisation" section).
- Intensification of the use of technological and analytical tools in accounting education curricula to motivate early contact and contribute to the holistic understanding of processes as the importance of independent reporting and interpretation of financial key figures increases ("RQ 1: Activities concerned with digitalisation in the different areas of accounting" and "RQ 4: Technologies used to perform the accounting activities" section).
- Combined focus on technical and professional competences in the different areas of accounting that recognises the increasing importance of meta-competences such as proactive thinking, self-control, creativity, and interdisciplinary action. This is particularly important to ensure that graduates and employees stay capable to navigate the field of accounting and adapt to the changing parameters flexibly ("RQ 2: Accounting activities in different stages of digitalisation" section).
- Integration of data and process management and new technologies into compulsory curricula and professional development courses to generate a holistic understanding of processes and manage change across the entire supply chain ("RQ 3: Placement of activities within the supply chain" and "RQ 4: Technologies used to perform the accounting activities" section).
- Employees in the areas of accounting and controlling need improved communicative and entrepreneurial skills to market the added value of digitalisation to companies ("RQ 5: The advantages and disadvantages of digitalised accounting activities" section).

Overall, the analysis provides valuable hints for future-oriented accounting education that supports companies in realising a maximum of expertise to manage their

digital transformation. This extends to the opportunity to think about implications for professional practices (How digitalised can the field be and how can digitalisation change job profiles?), for accounting students/employees (Am I willing to engage with changing technologies to improve my skill set and potentially work on entirely different activities within the next few years?), and even for society (How can these changes and their implications on competences help in finding creative solutions to 21st-century challenges?). This review can serve as a starting point for these and similar considerations.

This paper was intended to contribute to the discussion of future (competence) requirements for graduates in accounting. The detailed analyses provide an in-depth insight into six types of activities and the different stages of digitalisation for specific areas of accounting. This systematic and detailed analysis was necessary to gain knowledge on how to initiate concrete improvements in the content and methodology of the education and training of accountants (see for example Author 1 et al. 2022). It became clear, for example, that a deeper look into these areas is valuable as the same technologies are not likely to prevail in all areas at the same stage and for the same types of activities. In the area of financial accounting, for example, the duration of tasks decreases due to the large proportion of routine activities that can be automated or substituted, freeing up time for other activities or areas of accounting. One of these is controlling, where there are more non-routine activities that require new skills. However, it also became clear that interface management in particular is essential for linking the numerous technologies. Against this background, the didactics of accounting are also challenged to educate and train graduates in a differentiated manner.

Appendix I: List of coded publications

Author	Year	Title and publication information	Type	Language
Alexander S, Tiefenbeck F, Sabirzyanova N	2019	Digitale Transformation des Performance Managements: Zielbild und aktuelle Initiativen. [Digital transformation of performance management. Goals and current initiatives]. CON 31 (6): 39–42. https://doi.org/10.15358/0935-0381-2019-6-39	Journal Paper	German
Andiola LM, Masters E, Norman C	2020	Integrating technology and data analytic skills into the accounting curriculum: Accounting department leaders' experiences and insights. Journal of Accounting Education 50:1–18. https://doi.org/10.1016/j.jaccedu.2020.100655	Journal Paper	English

Author	Year	Title and publication information	Type	Language
Appelfeller W, Feldmann C	2018	Stufenweise Transformation der Elemente des digitalen Unternehmens. [Gradual transformation of the elements of a digital company]. In: Appelfeller W, Feldmann C (eds) Die digitale Transformation des Unternehmens: Systematischer Leitfaden mit zehn Elementen zur Strukturierung und Reifegradmessung. Springer, Berlin, pp 19–192	Book	German
Arbeitskreis Externe Unternehmensrechnung der Schmalenbach-Gesellschaft für Betriebswirtschaft e. V	2018	Chancen und Herausforderungen der Digitalisierung für die Effektivität und Effizienz des Rechnungswesens. [Opportunities and challenges of digitalisation for the efficiency and efficacy of accounting]. In: Krause S, Pellens B (eds) Betriebswirtschaftliche Implikationen der digitalen Transformation. Springer Gabler, Wiesbaden, pp 301–17. https://doi.org/10.1007/978-3-658-18751-4_16	Chapter	German
Ashoka, ML, Abhishek N, Divyashree MS	2019	Emerging Trends in Accounting: An Analysis of Impact of Robotics in Accounting, Reporting and Auditing of Business and Financial Information. International Journal of Business Analytics & Intelligence 7(2): 28–34	Journal Paper	English
Baier T	2019	Digitalisation in Management Accounting. Bachelor thesis. Berlin: Hochschule für Wirtschaft und Recht. Available via https://www.theseus.fi/handle/10024/262141 . Accessed 21 Mar 2023	University Publication	English
Bär J, Badura D, Bockshecker A, Hauer L, Karalash M, Nehls S, Neuhaus U, Schröder H, Schulz M, Sharma V, Welter F	2019	Die Mitarbeiter von Morgen: Ergebnisse eines Workshops zu den Kompetenzen künftiger Mitarbeiter im Bereich Business Analytics. [The future's employees: results from a workshop about the skills of future employees in Business Analytics]. Nordblick – Hochschule der Wirtschaft (8):34–49	Whitepaper	German
Bayerl E, Krippner K, Sikora C	2020	Steuerrechtliche Anforderungen und Entwicklungen im Bereich Steuern und Rechnungswesen [Tax law requirements and developments in the scope of taxation and accounting]. In: Rosar W, Krippner K, Setnicka M (eds) Digitalisierung Im Steuer- Und Rechnungswesen. Linde, Wien, pp 277–401	Chapter	German

Author	Year	Title and publication information	Type	Language
Becker W, Nolte M, Schuhknecht F	2020	Die Rolle des Chief Financial Officer im Rahmen der digitalen Transformation von Geschäftsmodellen. [The role of the chief financial officer in the scope of the digital transformation of business models]. In: Keimer I, Egle U (eds) Die Digitalisierung Der Controlling-Funktion: Anwendungsbeispiele Aus Theorie Und Praxis. Springer Gabler, Wiesbaden, pp 373–400	Chapter	German
Bernius S, Kreuzer S	2014	Warum eRechnung? Ökonomische und ökologische Einsparpotenziale in der öffentlichen Verwaltung. [Why e-invoicing? Economic and ecological saving potentials in public administration]. In: Rogall-Grothe C (ed) Leitfaden Elektronische Rechnung in der öffentlichen Verwaltung: Grundlagen, Umsetzungsempfehlungen, Best Practices. Goethe Universität Frankfurt (Main), pp 33–42	Chapter	German
Bernius S, Pfaff D	2014	Mythen der eRechnung – Wie wissenschaftliche Erkenntnisse den Weg zur Umsetzung des elektronischen Rechnungverkehrs zeigen. [Myths of e-invoices. How scientific findings show the way of implementing electronic invoicing]. Wissenschaft trifft Praxis 9(2): 70–80	Journal Paper	German
Binkow P	2015	The Impact of Self-Service Applications on Corporate Accounting and Its Customers. J. Corp. Acct. Fin 26(6): 81–85. https://doi.org/10.1002/jcaf.22084	Journal Paper	English
Bowles M, Ghosh S, Thomas L	2020	Future-Proofing Accounting Professionals: Ensuring Graduate Employability and Future Readiness. Journal of Teaching and Learning for Graduate Employability 11(1): 1–21	Journal Paper	English
Chen H, Yan Huang S, Chiu A, Pa, F	2012	The ERP system impact on the role of accountants. Industr Mngmnt & Data Systems 112(1): 83–101. https://doi.org/10.1108/02635571211193653	Journal Paper	English
Cong Y, Du H, Vasarhelyi MA	2018	Technological Disruption in Accounting and Auditing. Journal of Emerging Technologies in Accounting 15(2): 1–10. https://doi.org/10.2308/jeta-10640	Journal Paper	English
Crookes L, Conway E	2018	Technology Challenges in Accounting and Finance. In: Conway E, Byrne, D (eds) Contemporary Issues in Accounting. Palgrave Macmillan, Cham, pp 61–84	Chapter	English

Author	Year	Title and publication information	Type	Language
Drerup B, Suprano F, Wömpener A	2018	Controller 4.0. Anforderungsprofil des Controllers im digitalen Zeitalter. [Controller 4.0. Profile of requirements for controllers in the digital age]. In: Horváth P, Reichmann T, Baumöl U, Hoffjan A, Möller K, Pedell B (eds) Umbrüche durch VUCA-Umfeld und Digitalisierung. Controlling Sonderheft. München, Vahlen, pp 13–18	Whitepaper	German
Egle U, Keimer I	2017	Digitaler Wandel im Controlling. Schriften aus dem Institut für Finanzdienstleistungen. IFZ, Zug. [Digital change in controlling. Papers from the Institute for financial services Zug. IFZ, Zug.]	Whitepaper	German
Egle U, Frisan A, Steiner M	2020	Digitaler Wandel im Controlling bei der Alpiq Gruppe. [Digital change in controlling of the Alpiq Group]. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 189–98	Chapter	German
Faustino Bauer M, Schulte M, Schwab JB	2019	Was Blockchain für das Accounting bedeutet. [What blockchain means for accounting]. <i>Control Manag Rev</i> 63(5):40–45. https://doi.org/10.1007/s12176-019-0025-6	Whitepaper	German
Fordham DR, Hamilton CW	2019	Accounting Information Technology in Small Businesses: An Inquiry. <i>Journal of Information Systems</i> 33(2):63–75. https://doi.org/10.2308/isy-51982	Journal Paper	English
Fuller SH, Markelevich A	2020	Should accountants care about blockchain? <i>J. Corp. Acct. Fin</i> 31(2):34–46. https://doi.org/10.1002/jcaf.22424	Journal Paper	English
Gadatsch A	2020	Grundkurs Geschäftsprozess-Management. [Business process management 101]. Springer, Wiesbaden	Chapter	German
Grönke K, Heimel J	2015	Big Data im CFO-Bereich—Kompetenzanforderungen an den Controller. [Big Data in the scope of the CFO. Competence requirements for controllers]. <i>CON</i> 27(4–5):242–48. https://doi.org/10.15358/0935-0381-2015-4-5-242	Whitepaper	German
Groß, C, Pfennig R	2019	Branchenübergreifende Anwendungen. [Applications across industry sectors]. Gabler, Wiesbaden	Book	German

Author	Year	Title and publication information	Type	Language
Hecht N, Scherrer P	2020	Nutzen und Stolpersteine bei der Einführung einer Business Intelligence-Lösung für KMU am Beispiel der Firma SIGA. [Benefits and stumbling blocks for the introduction of business intelligence solutions in small and medium-sized businesses by the example of the SIGA company]. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 83–102	Chapter	German
Heupel T, Lange W	2019	Wird der Controller zum Data Scientist? Herausforderungen und Chancen in Zeiten von Big Data, Predictive Analytics und Echtzeitverfügbarkeit. [Does the controller become a data scientist? Challenges and opportunities in times of big data, predictive analytics and real-time availability]. In: Hermeier B, Heupel T, Fichtner-Rosada, S (eds) Arbeitswelten der Zukunft: Wie die Digitalisierung unsere Arbeitsplätze und Arbeitsweisen verändert. FOM-Edition. Springer Gabler, Wiesbaden, pp 201–221. https://doi.org/10.1007/978-3-658-23397-6_12	Chapter	German
Hmyzo E Muzzu A	2020	Technologie im Rechnungswesen – Wenn die Maschine besser und schneller bucht. [Technology in accounting. When the machine transfers better and faster]. In: Berding F, Jahncke H, Slopinski A (eds) Moderner Rechnungswesenunterricht 2020: Status Quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive. Springer, VS, Wiesbaden, pp 99–113	Chapter	German
Janvrin DJ Weidenmier Watson M	2017	"Big Data": A new twist to accounting. Journal of Accounting Education 38:3–8. https://doi.org/10.1016/j.jaccedu.2016.12.009	Journal Paper	English
Jonen A	2020	Aktuelle Trends und zukünftige Potenziale der Digitalisierung im Beschaffungscontrolling. [Current trends and future potentials of digitalisation in procurement controlling]. In: Keimer I, Egle U (eds) Die Digitalisierung Der Controlling-Funktion: Anwendungsbeispiele Aus Theorie Und Praxis. Springer Gabler, Wiesbaden, pp 349–72. https://doi.org/10.1007/978-3-658-29196-9_20	Chapter	German

Author	Year	Title and publication information	Type	Language
Jordanski G	2020	Kaufmännische Steuerung und Kontrolle im 4.0 Arbeitsumfeld – Anforderungen an duale Ausbildungsberufe. [Commercial management and controlling in a 4.0 workfield]. In: Berding F, Jahncke H, Slopinski A (eds) <i>Moderner Rechnungswesenunterricht 2020: Status Quo und Entwicklungen aus wissenschaftlicher Und praktischer Perspektive</i> . Springer VS, Wiesbaden, pp 59–82	Chapter	German
Keimer I, Egle U	2020	Digital Controlling – Grundlagen für den erfolgreichen digitalen Wandel im Controlling. [Digital controlling. Fundamentals of a successful digital transformation in controlling]. Springer Gabler, Wiesbaden	Book	German
Kink N	2007	Umfassende Controllingunterstützung. [Extensive controlling support]. <i>Z Control Manag</i> 51(5): 303–5. https://doi.org/10.1007/s12176-007-0083-z	Whitepaper	German
Kirchberg A	2017	Harmonisierung des externen und des internen Rechnungswesens aus aufbau- und ablauforganisatorischer Sicht. [Harmonisation of external and internal accounting from a workflow management and organisational structure point of view]. In: Klein A, Gräf A (eds) <i>Reporting und Business Intelligence.. 3rd edn</i> Haufe Gruppe, Freiburg, München, Stuttgart, pp 85–102	Chapter	German
Klein J, Küst C	2020	Wie die Digitalisierung im Rechnungswesen die Aufgaben und Anforderungen an die Mitarbeiter/-innen verändert [How digitalisation in accounting changes the tasks and requirements of employees]. In: Berding F, Jahncke H, Slopinski A (eds) <i>Moderner Rechnungswesenunterricht 2020: Status Quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive</i> . Springer VS, Wiesbaden, pp 83–97	Chapter	German
Koch B	2017	Die E-Rechnung steht im Zeichen großer Marktveränderungen. Studie E-Rechnung. [E-invoices will bring about changes in the market]. Available via https://www.billentis.com/Marktstudie2017_elektronische_Rechnung.pdf . Accessed 21 Mar 2023	Whitepaper	German
Kokina J, Gilleran R, Blanchette S, Stoddard D	2021	Accountant as Digital Innovator: Roles and Competencies in the Age of Automation. <i>Accounting Horizons</i> 35(1):153–84. https://doi.org/10.2308/HORIZONS-19-145	Journal Paper	English

Author	Year	Title and publication information	Type	Language
Kreher M	2021	Digitalisierung im Rechnungswesen 2021. [Digitalisation in accounting 2021]. KPMG. Available via https://home.kpmg/de/home/themen/2021/09/digitalisierung-im-rechnungswesen-2021.html . Accessed 05 Oct 2022	Whitepaper	German
Langmann C, Turi D	2020	Robotic Process Automation (RPA) – Digitalisierung und Automatisierung von Prozessen. Voraussetzungen, Funktionsweise und Implementierung am Beispiel des Controllings und Rechnungswesen. [Robotic process automation (RPA) – Digitalisation and automation of processes. Prerequisites, functions and implementation by the example of controlling and accounting]. Springer Gabler, Wiesbaden	Book	German
Le Guyader LP	2020	Artificial intelligence in accounting: GAAP's "FAS133". J. Corp. Acct. Fin 31(3):185–89. https://doi.org/10.1002/jcaf.22407	Journal Paper	English
Losbichler H, Gänßlen S	2018	Performance Measurement in Zeiten von Big Data. Auswirkungen auf Kennzahlen und deren Reporting. [Performance measurement in times of big data. Implications on operating figures and their reporting]. In: Horváth P, Reichmann T, Baumöl U, Hoffjan A, Möller K, Pedell B (eds) Umbrüche durch VUCA-Umfeld und Digitalisierung. Controlling Sonderheft. München, Vahlen, pp 31–37	Whitepaper	German
Menges T	2012	Implementierung eines automatisierten Rechnungseingangs [Implementation of automated invoice reception]. In: Deickert F (ed) Erfolgsfaktor strategisches Management, Controlling und Personal: Zukunft des Gesundheitswesens. Mannheimer Schriften zur Gesundheitswirtschaft 3. Centaurus, Freiburg (Breisgau), pp 115–125. https://doi.org/10.1007/978-3-86226-855-9_8	Chapter	German
Mertens P	2015	Industrie 4.0 – Herausforderungen auch an Rechnungswesen und Controlling im Überblick. [Industry 4.0. Overview of challenges for accounting and controlling]. CON 27 (8–9):452–54. https://doi.org/10.15358/0935-0381-2015-8-9-452	Whitepaper	German
Müller A, Graumann M, Weiß HJ	2020	Innovationen für eine digitale Wirtschaft. Wie Unternehmen den Wandel meistern. [Innovations for a digital economy. How companies succeed in the transformation]. Springer Gabler, Wiesbaden	Journal Paper	German

Author	Year	Title and publication information	Type	Language
Müller R, Reichmann T	2010	Rechnungswesen und Controlling: Integrierte Softwarelösungen für den Mittelstand. [Accounting and controlling: integrated software solutions for small and medium-sized businesses]. CON 22(3):204–206. https://doi.org/10.15358/0935-0381-2010-3-204	Whitepaper	German
Nagel L	2018	Die Einführung der “E-Rechnung” bei den Kommunen gemäß EU-Richtlinie. Bachelorarbeit. [The introduction of e-invoices in local communities according to EU law. Bachelor thesis]. Hochschule für öffentliche Verwaltung und Rechtspflege. Available via https://opus.bsz-bw.de/hsf/files/119/bachelorarbeit_lisa_nagel.pdf . Accessed 21 Mar 2023	University Publication	German
Najderek A	2020	Auswirkungen der Digitalisierung im Rechnungswesen – ein Überblick. [Implications of digitalisation for accounting—an overview]. In: Müller A, Graumann M, Weiß H-J (eds) Innovationen für eine digitale Wirtschaft: Wie Unternehmen den Wandel meistern. Springer Gabler, Wiesbaden, pp 127–45	Chapter	German
Nwokike FO, Eya GM	2015	A Comparative Study of the Perceptions of Accounting Educators and Accountants on Skills Required of Accounting Education Graduates in Automated Offices. World Journal of Education 5(5):64–70. https://doi.org/10.5430/wje.v5n5p64	Journal Paper	English
Pagel P	2019	Bei der Übermittlung und Verarbeitung elektronischer Rechnungen müssen die geltenden Anforderungen an Datenschutz und Datensicherheit erfüllt sein. [The transmission and handling of electronic invoices needs to follow current regulations of data protection and safety]. Wirtsch Inform Manag 11(5):342–46. https://doi.org/10.1365/s35764-019-00201-w	Journal Paper	German
Pampel JR	2018	Digitale Horizonterweiterung. Begleitung der Innovation von Geschäftsmodellen durch das Controlling. [Broadening one’s horizon. Accompanying innovative business models through controlling]. In: Horváth P, Reichmann T, Baumöl U, Hoffjan A, Möller K, Pedell B (eds) Umbrüche durch VUCA-Umfeld und Digitalisierung. Controlling Sonderheft. Vahlen, München, pp 21–29	Whitepaper	German

Author	Year	Title and publication information	Type	Language
Pan G, Seow P-S	2016	Preparing Accounting Graduates for Digital Revolution: A Critical Review of Information Technology Competencies and Skills Development. Singapore Management University, Singapore	University Publication	English
Peters P	2006	Standardisierung im Rechnungswesen – Leistungssprung durch Shared Services. [Standardizing accounting. A performance leap through shared services]. Z. Control. Manag. 50(S8):94–100. https://doi.org/10.1365/s12176-006-0606-z	Whitepaper	German
Raschig S, Schulze M	2020	Weiterentwicklung des Finanz-Forecasts im Rahmen der digitalen Transformation am Beispiel der SAP SE. [Development of the financial forecast as part of digital transformation by the example of SAP SE]. In Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 25–42	Chapter	German
Sandner P, Lange A, Schulden P	2020	The Role of the CFO of an Industrial Company: An Analysis of the Impact of Blockchain Technology. Future Internet 12(128):1–16. https://doi.org/10.3390/fi12080128	Journal Paper	English
Satzger G, Holtmann C, Peter S	2018	Advanced Analytics im Controlling. Potenzial und Anwendung für Umsatz- und Kostenprognosen. [Advanced analytics in controlling. Potentials and applications for sales and cost projections]. In: Horváth P, Reichmann T, Baumöl U, Hoffjan A, Möller K, Pedell B (eds) Umbrüche durch VUCA-Umfeld und Digitalisierung. Controlling Sonderheft. Vahlen, München, pp 47–53	Whitepaper	German
Schindera F, Jux M, Glustin O	2018	Banken-Controlling im digitalen Zeitalter. [Controlling in financial institutes in the digital age]. In: Horváth P, Reichmann T, Baumöl U, Hoffjan A, Möller K, Pedell B (eds) Umbrüche durch VUCA-Umfeld und Digitalisierung. Controlling Sonderheft. Vahlen, München, pp 61–67	Whitepaper	German

Author	Year	Title and publication information	Type	Language
Schömburg H, Breitner MH	2010	Elektronische Rechnungen zur Optimierung der Financial Supply Chain: Status Quo, empirische Ergebnisse und Akzeptanzprobleme. [Electronic invoices to optimize the financial supply chain: current state, empirical results and acceptance issues]. MKWI 2010 – E-Commerce und E-Business:1253–64. Available via http://webdoc.gwdg.de/univerlag/2010/mkwi/03_anwendungen/e-commerce_und_e-business/07_elektronische_rechnungen_zur_optimierung_der_financial_supply_chain.pdf . Accessed 21 Mar 2023	Journal Paper	German
Selb M	2020	Von der Erfolgssicherung zur Produktentwicklung – Datenanalyse bei Gebrüder Weiss im Fachbereich Corporate Logistics. [From securing long-term success to product development—Data analysis of Gebrüder Weiss in the area of corporate logistics]. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 43–64	Chapter	German
Shah H, Jiles L	2020	Management accountants and other professionals are leveraging data analytics to react effectively to the COVID-19 pandemic. A data-driven approach to the pandemic. Strategic Finance, 26–33	Whitepaper	English
Sinzig W	2015	In-memory Technik für Rechnungswesen und Controlling. [In-memory technique for accounting and controlling]. CON 27 (4–5):236–241. https://doi.org/10.15358/0935-0381-2015-4-5-236	Whitepaper	German
Sledgianowski D, Gomaa M, Tan C	2017	Toward integration of Big Data, technology and information systems competencies into the accounting curriculum. Journal of Accounting Education 38:81–93. https://doi.org/10.1016/J.JACCE DU.2016.12.008	Journal Paper	English
Spraakman G, O'Grady W, Askary D, Akroyd C	2015	Employers' Perceptions of Information Technology Competency Requirements for Management Accounting Graduates. Accounting Education 24(5):403–422	Journal Paper	English
Suden PT	2010	Einführung elektronischer Rechnungen. [Integration of electronic invoices]. Gabler, Wiesbaden	Book	German

Author	Year	Title and publication information	Type	Language
Tanner C	2016	E-Invoicing – ein erster Schritt zur digital vernetzten Wirtschaft. [E-invoicing. A first step to a digitally cross-linked economy]. Die Fachzeitschrift für erfolgreiche Unternehmer und Top-Manager (KMU Magazin) 19(4):44–48. https://doi.org/10.26041/fhnw-649	Whitepaper	German
Tanner C, Wölfle R	2005	Elektronische Rechnungsstellung zwischen Unternehmen. Konzentriertes Wissen aus der swissDIGIN-Initiative zur Förderung des elektronischen Rechnungsaustausch. [Electronic invoicing between companies. Concentrated knowledge from the swissDIGIN initiative for the promotion of electronic invoicing]. Available via https://unece.org/fileadmin/DAM/cefact/forum_grps/tbg/tbg15/project_docs/swissDIGIN_Leitfaden.pdf . Accessed 21 Mar 2023	University Publication	German
Trachsel V, Bitterli C	2020	Controller-Profil in der Schweiz. Bedeutung der Digitalisierung. [Profiles of controlling employees in Switzerland. The meaning of digitalisation]. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 199–210	Chapter	German
Weiber R, Gassler H, Meyer J	2002	Qualifizierungsanforderungen im E-Business [Qualification requirements in e-business]. In: Weiber R (ed) Handbuch Electronic Business: Informationstechnologien – Electronic Commerce – Geschäftsprozesse. 2 nd edn. Gabler, Wiesbaden, pp 261–77	Chapter	German
Werner M, Gehrke N	2011	Potentiale und Grenzen automatisierter Prozessprüfungen durch Prozessrekonstruktionen. [Potentials and boundaries of automated process verification through process reconstruction]. In: Plate G (ed) Forschung für die Wirtschaft. pp 99–120	Chapter	German
Wettstein K, Caderas R	2020	Die Digitale Transformation des Reportings beim Schweizer Radio und Fernsehen (SRF). [The digital transformation of reporting at the Swiss Radio and Television organisation (SRF)]. In: Keimer I, Egle, U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 65–82	Chapter	German

Author	Year	Title and publication information	Type	Language
Wilczek T	2014	Elektronische Rechnungseingangsbearbeitung. Eine Make-or-Buy-Studie. [Electronic reception of invoices and their processing. A make-or-buy study]. Disserta, Hamburg	University Publication	German

Abbreviations

AI	Artificial intelligence
ERP	Enterprise resource planning
OCR	Optical character recognition
RPA	Robotic process automation
RQ	Research question

Acknowledgements

Not applicable.

Author contributions

JP analysed and interpreted the data and was a major contributor in writing the manuscript. ER, DFH and FB initiated the project, drafted the theoretical background and methodology and were contributors in writing and editing the manuscript. All authors read and approved the final manuscript.

Funding

No funding has been received for this research.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 24 October 2022 Accepted: 27 March 2023

Published online: 04 April 2023

References

- Aeppli M, Angst V, Iten R, Kaiser H, Lüthi I, Schweri J (2017) Die Entwicklungen der Kompetenzenanforderungen auf dem Arbeitsmarkt im Zuge der Digitalisierung: Schlussbericht. Eidgenössisches Hochschulinstitut für Berufsbildung, Zollikofen
- Alexander S, Tiefenbeck F, Sabirzyanova N (2019) Digitale transformation des performance managements: zielbild und aktuelle initiativen. CON 31:39–42. <https://doi.org/10.15358/0935-0381-2019-6-39>
- Appelfeller W, Feldmann C (2018) Stufenweise Transformation der Elemente des digitalen Unternehmens. In: Appelfeller W, Feldmann C (eds) Die digitale Transformation des Unternehmens: Systematischer Leitfaden mit zehn Elementen zur Strukturierung und Reifegradmessung. Springer, Berlin, pp 19–192
- Arbeitskreis Externe Unternehmensrechnung, der Schmalenbach-Gesellschaft für Betriebswirtschaft e. V. (2018) Chancen und Herausforderungen der Digitalisierung für die Effektivität und Effizienz des Rechnungswesens. In: Krause S, Pellens B (eds) Betriebswirtschaftliche Implikationen der digitalen Transformation. Springer Gabler, Wiesbaden, pp 301–317
- Ashoka ML, Abhishek N, Divyashree MS (2019) Emerging trends in accounting: an analysis of impact of robotics in accounting, reporting and auditing of business and financial information. Int J Business Anal Intell 7:28–34
- Autor DH (2015) Why are there still so many jobs? the history and future of workplace automation. J Econ Perspect 29:3–30. <https://doi.org/10.1257/jep.29.3.3>
- Baier T (2019) Digitalisation in Management Accounting. Bachelorarbeit. Hochschule für Wirtschaft und Recht, Berlin
- Bär J, Badura D, Bockshecker A, Hauer L, Karalash M, Nehls S, Neuhaus U, Schröder H, Schulz M, Sharma V, Welter F (2019) Die Mitarbeiter von Morgen: Ergebnisse eines Workshops zu den Kompetenzen künftiger Mitarbeiter im Bereich Business Analytics. Nordblick Hochschule Der Wirtschaft 8:34–49
- Bayerl E, Krippner K, Sikora C (2020) Steuerrechtliche Anforderungen und Entwicklungen im Bereich Steuern und Rechnungswesen. In: Rosar W, Krippner K, Setnicka M (eds) Digitalisierung im Steuer- und Rechnungswesen. Linde, Wien, pp 277–401
- Becker W, Nolte M, Schuhknecht F (2020) Die Rolle des Chief Financial Officer im Rahmen der digitalen Transformation von Geschäftsmodellen. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 373–400

- Berding F, Riebenbauer E, Doreen Flick-Holtsch (2022) Fachwissenschaftliche Implikationen für die Aus- und Weiterbildung von Lehrkräften zur Digitalisierung im Rechnungswesen. In: Gerholz K.-H., Schlottmann P, Slepcevic-Zach P, Stock M (eds) *Digital Literacy in der beruflichen Lehrerinnen- und Lehrerbildung. Didaktik, Empirie und Innovation*. WBV, Bielefeld, pp 199–212.
- Bernius S, Kreuzer S (2014) Warum eRechnung? Ökonomische und ökologische Einsparpotenziale in der öffentlichen Verwaltung. In: Rogall-Grothe C (ed) *Leitfaden Elektronische Rechnung in der öffentlichen Verwaltung: Grundlagen, Umsetzungsempfehlungen, Best Practices*, Frankfurt (Main), pp 33–42.
- Binkow P (2015) The Impact of Self-Service Applications on Corporate Accounting and Its Customers. *J Corp Acct Fin* 26:81–85. <https://doi.org/10.1002/jcaf.22084>
- Bleiber R (2019) Digitalisierung in der Finanzbuchhaltung: Vom Status quo in die digitale Zukunft. Haufe, Freiburg
- Bowles M, Ghosh S, Thomas L (2020) Future-proofing accounting professionals: ensuring graduate employability and future readiness. *J Teach Learn Grad Employability* 11:1–21
- Chen H-J, Yan Huang S, Chiu A-A, Pai F-C (2012) The ERP system impact on the role of accountants. *Industr Mngmnt & Data Systems* 112:83–101. <https://doi.org/10.1108/02635571211193653>
- Cong Y, Du H, Vasarhelyi MA (2018) Technological disruption in accounting and auditing. *J Emerg Technol Accounting* 15:1–10. <https://doi.org/10.2308/jeta-10640>
- Egle U, Frisan A, Steiner M (2020) Digitaler Wandel im Controlling bei der Alpiq Gruppe. In: Keimer I, Egle U (eds) *Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis*. Springer Gabler, Wiesbaden, pp 189–198
- Fuller SH, Markelevich A (2020) Should accountants care about blockchain? *J Corp Acct Fin* 31:34–46. <https://doi.org/10.1002/jcaf.22424>
- Gadatsch A (2020) *Grundkurs Geschäftsprozess-Management*. Springer, Wiesbaden
- Grönke K, Heimel J (2015) Big Data im CFO-Bereich – Kompetenzanforderungen an den Controller. *CON* 27:242–248. <https://doi.org/10.15358/0935-0381-2015-4-5-242>
- Hartig J, Klieme E, Leutner D (2008) Assessment of competences in educational contexts. Hogrefe, Hogrefe eLibrary
- Hecht N, Scherrer P (2020) Nutzen und Stolpersteine bei der Einführung einer Business Intelligence-Lösung für KMU am Beispiel der Firma SIGA. In: Keimer I, Egle U (eds) *Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis*. Springer Gabler, Wiesbaden, pp 83–102
- Heupel T, Lange V (2019) Wird der Controller zum Data Scientist? Herausforderungen und Chancen in Zeiten von Big Data, Predictive Analytics und Echtzeitverfügbarkeit. In: Hermeier B, Heupel T, Fichtner-Rosada S (eds) *Arbeitswelten der Zukunft: Wie die Digitalisierung unsere Arbeitsplätze und Arbeitsweisen verändert*. Springer Gabler, Wiesbaden, pp 201–221
- Hmyzo E, Muzzu A (2020) Technologie im Rechnungswesen – Wenn die Maschine besser und schneller bucht. In: Berding F, Jahncke H, Slopinski A (eds) *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive*. Springer VS, Wiesbaden, pp 99–113
- Hübl L (2020) Kritische Erfolgsfaktoren bei der Digitalisierung von Wirtschaftstreuhandkanzleien. In: Setnicka M, Krippner K, Rosar W (eds) *Digitalisierung im Steuer- und Rechnungswesen*. Linde, Wien, pp 191–203
- Iten R, Peter M, Gschwend E, Angst V, Lachenmeier P, Heinemann E (2016) Offshoring und Wandel der Berufsbilder. Aktuelle Trends und Konsequenzen für kaufmännische Berufe: Schlussbericht. Eidgenössisches Hochschulinstitut für Berufsbildung, Zollikofen
- Jonen A (2020) Aktuelle Trends und zukünftige Potenziale der Digitalisierung im Beschaffungscontrolling. In: Keimer I, Egle U (eds) *Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis*. Springer Gabler, Wiesbaden, pp 349–372
- Jordanski G (2020) Kaufmännische Steuerung und Kontrolle im 4.0 Arbeitsumfeld – Anforderungen an duale Ausbildungsberufe. In: Berding F, Jahncke H, Slopinski A (eds) *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive*. Springer VS, Wiesbaden, pp 59–82
- Keimer I, Egle U (2020) Digital Controlling - Grundlagen für den erfolgreichen digitalen Wandel im Controlling. In: Keimer I, Egle U (eds) *Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis*. Springer Gabler, Wiesbaden, pp 1–16
- Kink N (2007) Umfassende Controllingunterstützung *Z Control Manag* 51:303–305. <https://doi.org/10.1007/s12176-007-0083-z>
- Kirchberg A (2017) Harmonisierung des externen und des internen Rechnungswesens aufbau- und ablauforganisatorischer Sicht. In: Klein A, Gräf J (eds) *Reporting und Business Intelligence*, 3rd edn. Haufe Gruppe, Freiburg, München, Stuttgart, pp 85–102.
- 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 (eds) *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive*. Springer VS, Wiesbaden, pp 83–97
- Koch B (2017) *Die E-Rechnung steht im Zeichen großer Marktveränderungen: Studie E-Rechnung*. Billentis, Wil.
- Kreher M (2021) *Digitalisierung im Rechnungswesen 2021*. KPMG, München
- Kuckartz U (2018) *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*, 4th edn. Beltz, Weinheim.
- Langmann C, Turi D (2020) *Robotic Process Automation (RPA). Digitalisierung und Automatisierung von Prozessen: Voraussetzungen, Funktionsweise und Implementierung am Beispiel des Controllings und Rechnungswesens*. Springer Gabler, Wiesbaden.
- Le Guyader LP (2020) Artificial intelligence in accounting: GAAP's "FAS133." *J Corp Acct Fin* 31:185–189. <https://doi.org/10.1002/jcaf.22407>
- Litfin T, Wolfram G (2010) New Automated Checkout Systems. In: Krafft M, Mantrala M (eds) *Retailing in the 21st Century*. Springer, Berlin, Heidelberg, pp 189–203
- Losbichler H, Gänßlen S (2018) Performance Measurement in Zeiten von Big Data: Auswirkungen auf Kennzahlen und deren Reporting. In: Horváth P, Reichmann T, Baumöl U, Hoffman A, Möller K, Pedell B (eds) *Transformation im Controlling: Umbrüche durch VUCA-Umfeld und Digitalisierung*. Vahlen, München, pp 31–37

- Menges T (2012) Implementierung eines automatisierten Rechnungseingangs. In: Deickert F (ed) Erfolgsfaktor strategisches Management, Controlling und Personal: Zukunft des Gesundheitswesens. Centaurus, Freiburg (Breisgau), pp 115–125.
- Müller R, Reichmann T (2010) Rechnungswesen und Controlling: Integrierte Softwarelösungen für den Mittelstand. CON 22:204–206. <https://doi.org/10.15358/0935-0381-2010-3-204>
- Nagel L (2018) Die Einführung der „E-Rechnung“ bei den Kommunen gemäß EU-Richtlinie: Bachelorarbeit, Hochschule für öffentliche Verwaltung und Rechtspflege. Bachelorarbeit. Meißen.
- Najderek A (2020) Auswirkungen der Digitalisierung im Rechnungswesen – ein Überblick. In: Müller A, Graumann M, Weiß H-J (eds) Innovationen für eine digitale Wirtschaft: Wie Unternehmen den Wandel meistern. Springer Gabler, Wiesbaden, pp 127–145
- Pagel P (2019) Bei der Übermittlung und Verarbeitung elektronischer Rechnungen müssen die geltenden Anforderungen an Datenschutz und Datensicherheit erfüllt sein. Wirtsch Inform Manag 11:342–346. <https://doi.org/10.1365/s35764-019-00201-w>
- Pan G, Seow P-S (2016) Preparing accounting graduates for digital revolution: a critical review of information technology competencies and skills development. Management University, Singapore
- Raschig S, Schulze M (2020) Weiterentwicklung des Finanz-Forecasts im Rahmen der digitalen Transformation am Beispiel der SAP SE. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 25–42
- Sachs S, Meier C, Mccorley V, Schweiz KV, Thurgau TT (2016) Digitalisierung und die Zukunft kaufmännischer Berufsbilder – eine explorative Studie. Schlussbericht. <https://www.swissleaders.ch/uploads/media/page/0001/04/898a7462b82f51ab749f0e3ce678aaac0be495eb.pdf>. Accessed 21 Mar 2023.
- Sandner P, Lange A, Schulden P (2020) The Role of the CFO of an industrial company: an analysis of the impact of Blockchain technology. Future Internet 12:1–16. <https://doi.org/10.3390/fi12080128>
- Satzger G, Holtmann C, Peter S (2018) Advanced Analytics im Controlling: Potenzial und Anwendung für Umsatz- und Kostenprognosen. In: Horváth P, Reichmann T, Baumöl U, Hoffjan A, Möller K, Pedell B (eds) Transformation im Controlling: Umbrüche durch VUCA-Umfeld und Digitalisierung. Vahlen, München, pp 47–53
- Schindera F, Jux M, Glustin O (2018) Banken-Controlling im digitalen Zeitalter. In: Horváth P, Reichmann T, Baumöl U, Hoffjan A, Möller K, Pedell B (eds) Transformation im Controlling: Umbrüche durch VUCA-Umfeld und Digitalisierung. Vahlen, München, pp 61–67
- Schömburg H, Breitner MH (2010) Elektronische Rechnungen zur Optimierung der Financial Supply Chain: Status Quo, empirische Ergebnisse und Akzeptanzprobleme. MKWI 2010 – E-Commerce und E-Business:1253–1264.
- Seeber S, Seifried J (2019) Herausforderungen und Entwicklungsperspektiven der beruflichen Bildung unter veränderten Rahmenbedingungen. Z Erzieh 22:485–508
- Seeber S, Weber S, Geiser S, Zarnow S, Hackenberg T, Hiller F (2019) Effekte der Digitalisierung auf kaufmännische Tätigkeiten und Sichtweisen ausgewählter Akteure. Berufsbildung 73:2–7
- Selb M (2020) Von der Erfolgssicherung zur Produktentwicklung – Datenanalyse bei Gebrüder Weiss im Fachbereich Corporate Logistics. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 43–64
- Sledgianowski D, Gomaa M, Tan C (2017) Toward integration of Big Data, technology and information systems competencies into the accounting curriculum. J Account Educ 38:81–93. <https://doi.org/10.1016/J.JACCEDU.2016.12.008>
- Spitz-Oener A (2006) Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. J Law Econ 24:235–270. <https://doi.org/10.1086/499972>
- Suden P (2010) Einführung elektronischer Rechnungen. In: Suden P (ed) Die elektronische Rechnung in Handels- und Steuerrecht: Einführung, Signatur, Dokumentation. Gabler, Wiesbaden, pp 17–47
- Tanner C (2016) E-Invoicing – ein erster Schritt zur digital vernetzten Wirtschaft. Die Fachzeitschrift Für Erfolgreiche Unternehmer Und Top-Manager (KMU Magazin) 19:44–48. <https://doi.org/10.26041/fhnw-649>
- Tanner C, Wölfl R (2005) Elektronische Rechnungsstellung zwischen Unternehmen: Konzentriertes Wissen aus der swissDIGIN-Initiative zur Förderung des elektronischen Rechnungsaustauschs. Fachhochschule beider Basel Nordwestschweiz, Basel.
- Trachsel V, Bitterli C (2020) Controller-Profil in der Schweiz. Bedeutung der Digitalisierung. In: Keimer I, Egle U (eds) Die Digitalisierung der Controlling-Funktion: Anwendungsbeispiele aus Theorie und Praxis. Springer Gabler, Wiesbaden, pp 199–210
- Weiber R, Gassler H, Meyer J (2002) Qualifizierungsanforderungen im E-Business. In: Weiber R (ed) Handbuch Electronic Business: Informationstechnologien – Electronic Commerce – Geschäftsprozesse, 2nd edn. Gabler, Wiesbaden, pp 261–277
- Weinert F (2001) Concept of competence: A conceptual clarification. In: Rychen D, Salganik LH (eds) Defining and selecting key competencies. Hogrefe & Huber, Ashland, pp 45–65.
- Werner M, Gehrke N (2011) Potentiale und Grenzen automatisierter Prozessprüfungen durch Prozessrekonstruktionen. In: Plate G (ed) Forschung für die Wirtschaft. Shaker, Aachen, pp 99–120.
- Wilczek T (2014) Elektronische Rechnungseingangsbearbeitung: Eine Make-or-Buy-Studie. Disserta, Hamburg.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

C Beitrag 2

Sustainability content knowledge for teachers in accounting and controlling: A systematic literature review on industry approaches, practices and desired competences

Julia Pargmann^{1*}, Florian Berding¹,

¹ Faculty of Education, Department for Vocational Education and Lifelong Learning, Institute of Vocational and Business Education, University of Hamburg, Hamburg, Germany

* Correspondence:

Julia Pargmann

julia.pargmann@uni-hamburg.de

Keywords: sustainability, accounting, controlling, content knowledge , review

Abstract

Sustainable transformation is critical in the face of today's challenges, with business at the center of sustainable development. Students in vocational business education need to integrate sustainability into decision-making processes, which are often based on accounting and controlling. These form a significant operational foundation, which is why sustainability should play an essential role here. Teachers need appropriate content knowledge and knowledge of industry needs. This study uses a PRISMA-led systematic review and qualitative content analysis to explore the current discourse. The aim is to derive starting points for adequate content knowledge. 79 full texts were analyzed that link sustainability to accounting and/or controlling. The literature shows a shift towards a comprehensive approach where sustainability permeates legal and profit-orientated operational aspects. Challenges exist in the integration of non-monetary information, with proposed solutions such as social and environmental accounting. Essential competences include sustainability awareness and ethical principles. The study emphasizes the need to integrate sustainability into education and recommends a process-oriented perspective to promote a multidimensional understanding.

1 Introduction

Accounting and controlling are fundamental elements for managing entrepreneurial success and promoting economic thinking (Chiapello, 2009; Preiß and Tramm, 1996). Seifried (2004) pointed out twenty years ago that accounting lessons should be expanded to include more and more controlling elements. Current global developments support this call: traditional accounting is increasingly being questioned, particularly as a result of digitalization (Klein and Küst, 2020; Pargmann et al., 2023). However, little seems to have changed; there is still a need in practice for more comprehensive controlling knowledge that goes beyond accounting and offers deeper insights into the management and control of companies (Schöning et al., 2020). In order to do justice to the operational relevance that accounting and management control have in theory, they must face modern challenges such as sustainability.

The concept of sustainability in accounting and controlling was previously understood to primarily entail the assurance of the company's long-term viability. However, the advent of integrated non-financial reporting has also prompted a growing recognition of sustainability as a core aspect of

accounting and control (Meeh-Bunse & Luer, 2016; Osburg, 2013). Presently, this is predominantly observed in the form of discrete sustainability management and controlling departments, rather than being integrated in a more comprehensive manner (Ghosh et al., 2019; Schaltegger, 2016). The implementation of sustainability efforts also necessitates the presence of suitably qualified employees, who facilitate corporate transformation towards sustainability (e.g. Berding et al., 2020; Bliesner-Steckmann, 2018; Gallagher et al., 2020). For a variety of reasons, not all companies demonstrate a commitment to pursuing sustainability. Some companies cite a lack of expertise, financial resources, or industry interest as reasons for their inaction. Others attempt to disguise unsustainable practices through green marketing and communications. Conversely, other organizations have already transitioned to more environmentally sustainable business models.

Accounting and controlling take on a vital role in vocational business education, as they facilitate students' comprehension of economic correlations, causal chains and the ramifications of their own economic decisions. This process-oriented approach facilitates the application of knowledge in novel contexts and situations. The curriculum for accounting and controlling typically emphasizes an understanding of double entry bookkeeping as well as with traditional control instruments, such as SWOT analysis and various key performance indicators (KPIs), including return on investment (ROI) and ABC analysis. It is uncommon for students to complete the control cycle, omitting a central component designed to instruct them in the process of making economic decisions (Stütz, 2024). The control cycle is comprised of six steps that enable students to translate the theoretical domain of accounting into practice. The six steps are 1) orientation in the material, 2) translation into the real world, 3) working in the systems of accounting, 4) interpret results, 5) validate solutions and 6) decision making. In an analysis of 2,000 accounting and controlling tasks, Stütz (2024) shows that the last two steps that connect the results back to reality are often omitted. This might be due to a lack of action-orientation. In addition, few didactical concepts exist for the integration of sustainability into accounting and controlling lessons. This points to the issue of missing real world examples of how sustainability can be incorporated into accounting and controlling classes.

In recent years, there has been a notable surge in the number of publications on sustainability in accounting and controlling. In addition to examining the processes of sustainable transformation, these publications also address the competences that employees must possess in order to contribute to these changes. Other current reviews either focused on sustainability in accounting and controlling in general (i.e., Broccardo et al., 2024), on higher education accounting and controlling curricula (i.e., Frizon & Eugénio, 2022) or on sustainability accounting as a separate concept (i.e., Gil-Marín et al., 2022). The present review adds a process-oriented approach that makes the results useable for teaching practices, especially in German vocational education contexts. In order to successfully integrate sustainability into schoolwide accounting and controlling curricula, educators must first ascertain which types of content are relevant to their students' vocation and then determine the manner in which these contents are connected to the various competence dimensions. The key problem in this article is the current scope of sustainability in business literature to derive teacher content knowledge.

To shed light on the relevant content knowledge for teachers in vocational accounting and controlling education, a systematic review of n=79 publications was carried out using PRISMA guidelines (Page et al., 2021) and evaluated using qualitative content analysis (Kuckartz, 2018). Contributions to the following research questions will be made:

1. How is sustainability discussed in accounting and controlling literature? Which tools, processes and practices are relevant in this context?

2. Which sustainability-focused competences are discussed in accounting and controlling research and practice?

The aim of this article is a) to gain insights into the state of the discussion on sustainability in accounting and controlling literature. Through this we aim to discover which content knowledge teachers who would like to incorporate sustainability into lessons (or sequences) focused on controlling or accounting need. Aim b) is to derive the sustainability-related professional, methodological and social competence requirements teachers need to foster in their classes. As a result of this critical analysis, we extract implications for accounting and controlling classroom activities.

The remainder of this paper is structured as follows: First, the methodology is presented and the categories for the content analysis are deducted as a backdrop for the critical analysis. Then, the results will be reported, structured by the analytical categories. Lastly, the findings will be discussed and implications for vocational accounting and controlling education will be derived to help both research and practice to sharpen their efforts regarding vocational sustainability education.

2 Methods

When designing this literature review, the preferred reporting items for systematic reviews and meta-analyses (PRISMA) (Page et al., 2021) were used to ensure a replicable search. We developed selection criteria, a consistent search strategy and a framework for data analysis. In the later stages of the analysis, we also performed a qualitative content analysis following Kuckartz (2018) to identify trends. Thus, this review follows PRISMA standards for record identification and retrieval but a qualitative content analysis framework for analysis.

2.1 Search strategy

We searched five electronic databases (Web of Science, PEDOCS, Google Scholar, EBSCO and WISO) for the years 2000 to 2022. The databases were chosen because they provide literature from both the technical content side of business education as well as the pedagogical-didactical side. While Web of Science, Google Scholar and EBSCO are well known international databases, WISO and PEDOCS are only established in Germany. Since the German vocational education system has a unique structure that combines workplace learning and vocational schools, the curricula encompass a distinct connection to the current developments in businesses and industries (Fürstenau et al., 2014). To honor this connection, we included the WISO database, which offers more practice-related publications. This matches well with our goal to derive implications for competences, as we get to analyze developments from the industry that are specific to the German market. PEDOCS on the other hand is a database that offers publications from educational science. Internationally, accounting education is mostly a topic for higher education institutions (Kleinberg Neimark, 1996). Since this paper focuses on developing implications for vocational accounting education, it was appropriate to extend our search to databases that focus on educational science and are often published with the German educational system in mind. This makes it easier to connect the implications to the use cases of this paper. PEDOCS has been used in reviews before, such as by Hermann and Gerlach (2020). Same is true for the WISO database, which was used by Niesen and Houy (2015).

We still searched for both English and German articles to reflect international discussions. Two researchers searched the databases using the same search terms to make aligned decisions on exclusion and inclusion. The search terms were grouped to search for sustainability and controlling/accounting literature as well as literature that includes competence, skills, or

knowledge/expertise in this context. Table 1 presents an overview of keywords used. The search included titles, abstracts and full texts.

2.2 Article selection criteria

To identify as many suitable articles as possible, we determined the following selection criteria. Articles were excluded if they did not connect sustainability and controlling/accounting. Publications about accounting and controlling classes in higher education or case studies on higher education were also excluded, even if they had a distinct sustainability focus. Additionally, the screening excluded all publications which did not use a suitable definition of sustainability, i.e. in the sense of longevity. The review only includes original research or reports from practice (grey literature, book chapters and working papers), no literature reviews or meta-analyses. The search strategy yielded 5,131 records for screening of which 4,671 were excluded right away as they did not connect the two perspectives of sustainability and controlling/accounting. We derived a total of n=79 full texts for coding. Figure 1 shows the PRISMA flow diagram for this paper. A detailed list of the publications can be found in Appendix A.

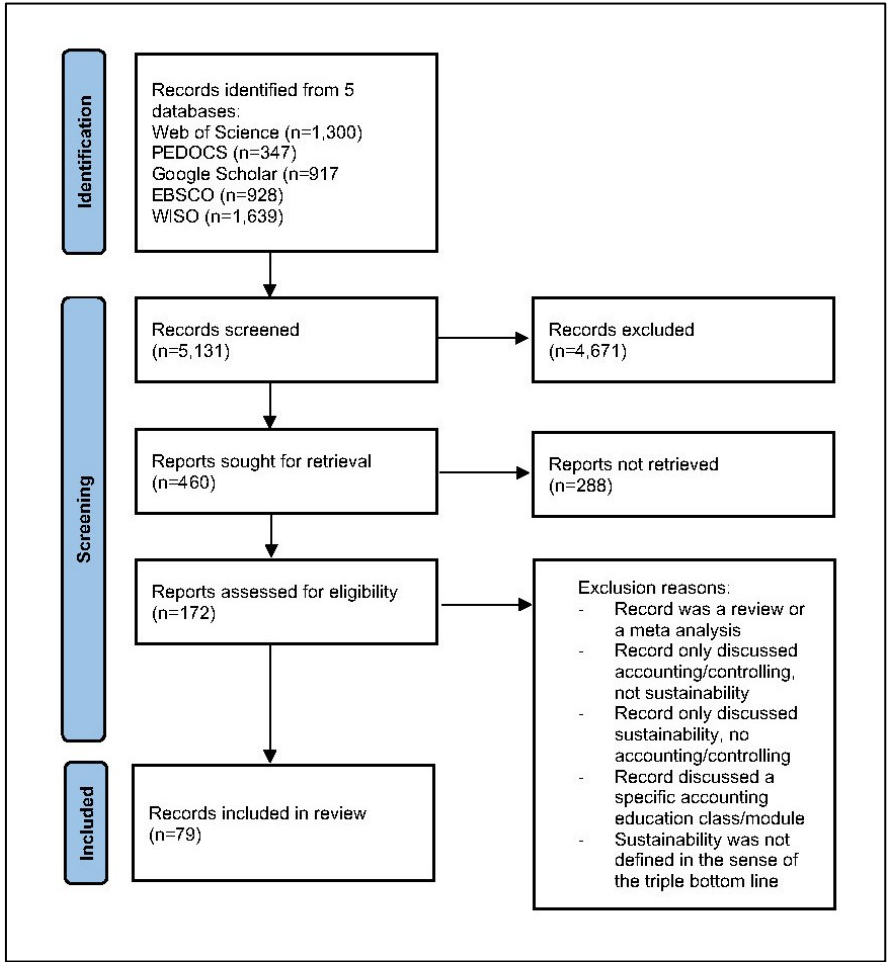


Figure 1: PRISMA Flow Diagram

2.3 Qualitative content analysis

To gain structured insight into the topic, we performed a qualitative content analysis using inductive and deductive categories. The coding manual was developed using guidelines for qualitative content

analysis (Kuckartz, 2018) and was carried out using the software MaxQDA (Release version 22.8.0). The full texts were coded using the analytical framework (Table 2). After baseline coding for alignment, the remainder was coded in pairs with regular discussion regarding inconsistencies.

Most of the analytical framework was developed deductively, as the purpose of this paper is not to create a new systemization but rather to see how far the discourse on sustainability in controlling and accounting has progressed. The following categories were developed to enable a critical analysis:

- **Approaches towards sustainability, sustainability dimensions, strategies, and conflicts:** To identify where corporate sustainability efforts are located and how well companies have integrated sustainability into their supply chain, the approaches towards sustainability were coded by the relevant business process. The differentiation between controlling, core and support processes helps match sustainability contents with accounting and controlling activities (Gadatsch, 2017). The areas of activities were added because not every company, especially not small- and medium-sized enterprises (SME), are organized in a process-oriented manner (Wöhe et al., 2020). Here, sustainability efforts are closely intertwined between different areas of the business. The last three categories were included to develop an understanding of how companies conceptualize their sustainability efforts, how they define them and how long-term they think about the topic (Perridon et al., 2017). The most common definition of sustainability is a balance between the ecological, social and economic dimensions (Schmidt, 2012) and a sustainable development requires a transformation that stops the exploitation of resources (World Commission on Environment and Development, 1987). Generally, sustainability efforts can follow one of three strategies: efficiency, consistency or sufficiency. Each of these strategies has its own risks and benefits and their implementation largely depends on the business model. In consequence, there are some conflicts that can be derived from the triple bottom line and the sustainability strategies, i.e. rebound effects. As these also affect controlling and accounting, they were included. We also inductively coded the motivation of companies to become more sustainable.
- **Sustainability reporting, certificates, frameworks, tools and methods:** Corporate sustainability efforts usually include some notion of corporate social responsibility and thus integrated reporting. Due to legislative developments, companies registered in the European Union that exceed the size of 500 employees and are of public interest are now obligated to report on their efforts in a suitable way, regulations for other enterprises are to follow in the coming years (European Union, 2022). For the purpose of this paper, it is of interest to identify reporting standards and certificates in which businesses have a particular interest in (Scholz and Pastoors, 2018). As we focus on controlling and accounting, the ‘tools and methods’ category also includes key performance indicators (KPIs) used by companies. This is especially interesting because in vocational business education in Germany, the curricula use KPIs as means to illustrate impact relationships to students.
- **Competences:** In the German-speaking countries, the promotion of action competence is a crucial part of vocational education programs. This concept depicts the ability to solve occupational tasks successfully and to execute routine and non-routine activities. Typically, frameworks distinguish between professional, social and self-competences, with methodological and technological competences as interdisciplinary dimensions (Roth, 1971; Winther and Achtenhagen, 2008). Thus, we included the sub-category of “Focus on sustainability” to each competence dimension.

3 Results

3.1 The sample

Out of the 79 publications, 25 were journal articles, 33 were magazines, whitepapers or grey literature, three were conference proceedings, 15 were book chapters and three were monographs. Regarding the publishing years, three peaks can be identified in the years 2011, 2016 and in 2020/2021. Nineteen publications are from the period between 2001 and 2011, 60 were published between 2012 and 2022. Language-wise, most publications (n=55) are German, the remainder (n=24) are English. The specific characteristics of each publication are provided in the supplemental material.

The next section presents the results for each of the main categories.

3.2 Content analysis results

3.2.1 Approaches towards sustainability

Few texts focus on the economic dimension of the triple bottom line (Fig. 2, left). Instead, the ecological dimension is discussed most frequently, followed by the social dimension. The right side of Fig. 2 shows that conflicts between the social and ecological dimension seem to be less important, as conflicts between economic and ecological decisions and economic and social decisions dominate.

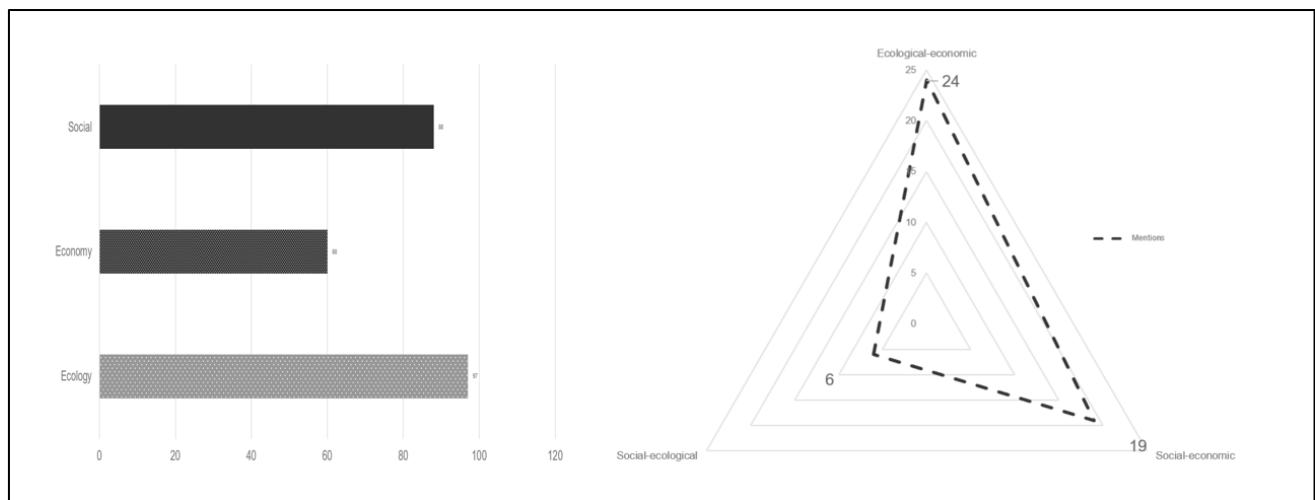


Fig. 2.: Dimensions of the triple bottom line mentioned (left) and conflicts arising (right)

In the material, the departments most frequently discussed to be responsible for integrating sustainability into practice were sales and marketing, finance and the sustainability department (Fig. 3). Less frequently, procurement and HR were mentioned.

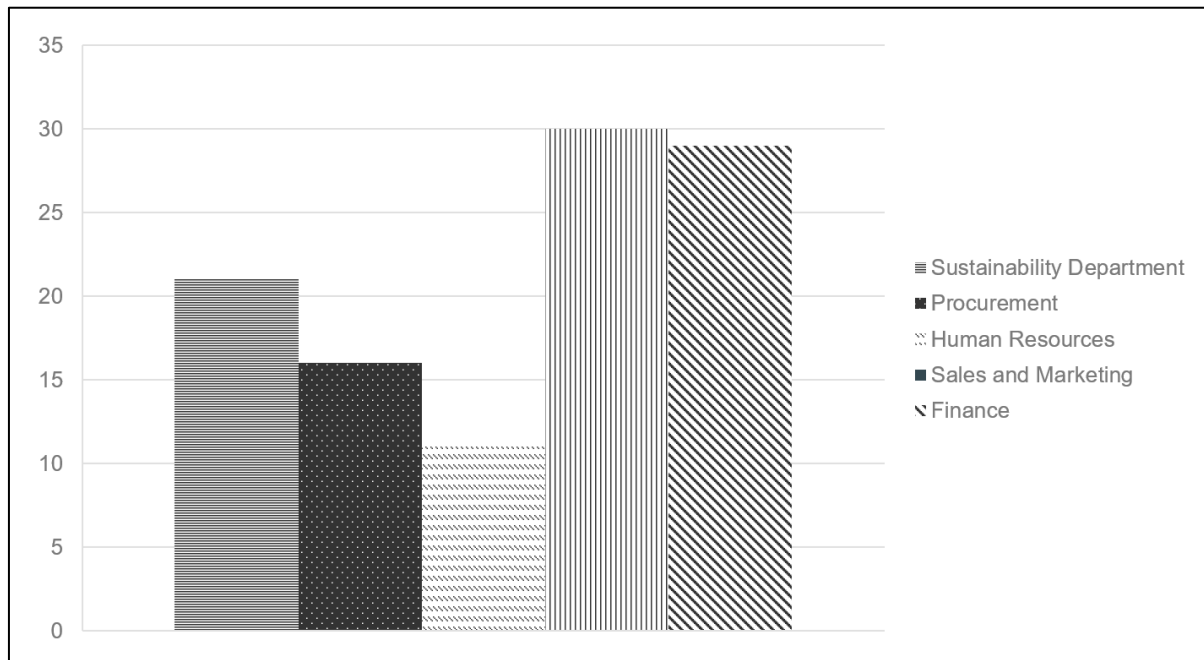


Fig. 3: Integration of sustainability on a departmental level (frequency of mentions)

The sustainability department was mentioned as a possible point of interaction with controlling/accounting because they aggregate corporate information from different areas of the business, generating an overview of environmental, social and economic sustainability progress (Isensee and Michel, 2011). However, the degree of cooperation between the interfaces varies. While not all businesses have a sustainability department, those that do usually are the main locus of sustainability information for all other business areas (Michel et al., 2014). In contrast, employees in controlling and accounting do not see themselves as relevant stakeholders of the sustainability department and rarely keep up with their developments, unless they have a personal interest in the topic (Kämmeler-Burraak and Bauer, 2022; Stehle and Stelkens, 2018).

In procurement, the sustainability efforts that are connected to controlling and accounting usually focus on the corporate supply chain's emissions and resource consumption (Ferreira et al., 2010; Glanze, 2022; Krause, 2016). However, some publications mention progress in the sense of design for sustainability, procurement for sustainability or the development of more efficient and environmentally friendly means of production (Greiling and Ther, 2010; Isensee and Michel, 2011; Schaltegger, 2020; Schaltegger and Zvezdov, 2011).

Sustainability in HR mostly focuses on the social dimension through employee development and training, as well as the controlling of employee KPIs (Greiling and Ther, 2010; Weber et al., 2010). They are a stakeholder of the sustainability department because for an increasing number of potential employees, a company's sustainability efforts are relevant in the application process. HR needs sustainability information to market the roles to applicants (Braun, 2019). It is also in the company's interest that employees act in line with corporate strategy. This extends to sustainability-specific aspects. It is therefore necessary for HR to also offer further training in this area in cooperation with the sustainability department (Lacy et al., 2009). Regarding controlling, HR also records data relevant to the social dimension of sustainability, such as absence and sickness rates, workplace incidents and personal development hours, amongst others (Schwarzmaier, 2015).

The development of an ‘honest’ marketing mix dominates the codings in the field of sales and marketing (Gebauer, 2013; Schaltegger and Zvezdov, 2011; Schwarzmaier, 2013). Braun (2019) even calls some CSR reports a mere marketing tool that is used strategically as a means of corporate communications. These reports provide little value to stakeholders as they lack substance (Gaggl, 2021). To enforce truly sustainable marketing efforts, reporting should be seen as an opportunity to realize the corporate sustainability strategy, i.e. by commenting on internal sustainability efforts and failures (Schwarzmaier, 2013; Weigand, 2020).

In the finance department, most codings relate to the area of controlling (Fig. 4). These were coded whenever the area was mentioned in a sustainability context. The other areas like external and internal accounting (the latter represented by the different areas of cost accounting and investment accounting) were mentioned far less frequently, although explicitly included in the search terms (Table 1).

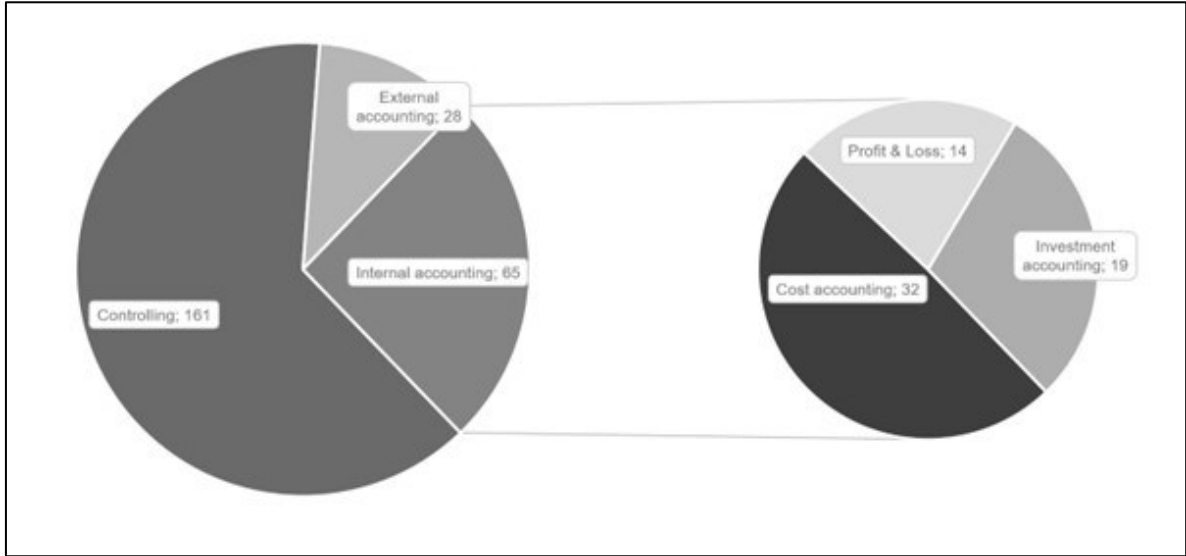


Fig. 4: Coding distribution across different accounting and controlling departments with an overview on the left and a zoom into internal accounting on the right

Most texts did not discuss new sustainability-oriented instruments for controlling but rather applied traditional tools to the new context (Wellbrock et al., 2020). Some traditional tools that are mentioned are ecology-oriented SWOT analysis, portfolio analysis, ABC-scoring, eco-efficiency indicators or ecological life cycle assessments (Greiling and Ther, 2010). Some publications specify controlling tools that are unique to the issue of sustainability, for example the ‘sustainable value’ or the ‘sustainability balanced scorecard’ (Figge and Hahn, 2004; Gaggl, 2021; Heimel and Momberg, 2021; Hubbard, 2009; Wellbrock et al., 2020). Fischer et al. (2010) and Michel et al. (2014) highlight the need of new KPIs that include the triple bottom line approach. Thus, new KPI systems need to be developed to ensure number-based reporting (Lühn et al., 2022). According to Aras and Crowther (2009) and Rönz and Ryba (2012) however, companies need guidance to develop these KPIs, as many sustainability aspects are difficult to quantify for accounting purposes.

Some publications discuss the connection between controlling and CSR departments, as both collect relevant data that requires profound sustainability expertise (Rönz and Ryba, 2012). Since most controlling employees are not sustainability experts, cooperation is needed to successfully develop and control the corporate sustainability strategy (Endenich and Trapp, 2019).

In both external and internal accounting, sustainability aspects were mostly discussed as supplemental to the regulations companies already follow (Günther and Stechemesser, 2011; Sulaiman and Mokhtar, 2012). In external accounting, most codings related to hardships that sustainable companies face when it comes to financial accounting and reporting. Sulaiman and Mokhtar (2012) specify that professional accounting bodies should require monetary information to be reported in the CSR reports, because only this will motivate companies to be serious about the integration of environmental issues into their accounting and control systems. Schwarzmaier (2015) adds that companies that report in Germany can use accounting instruments such as social balance sheets in addition to the traditional management reports. In cost accounting, it is an option to internalize costs, e.g. CO2 compensation. Some companies go a step further and prepare an ecology-oriented profit and loss statement that internalizes effects on ecology and values them at an internal cost rate (Michel et al., 2014).

In addition to sustainability efforts on a departmental level, some aspects affect some or all parts of a business' supply chain. The results for this category are presented in Fig. 5. Most frequently, management and control processes as well as the entire supply chain were coded. Key and support processes were not coded as often.

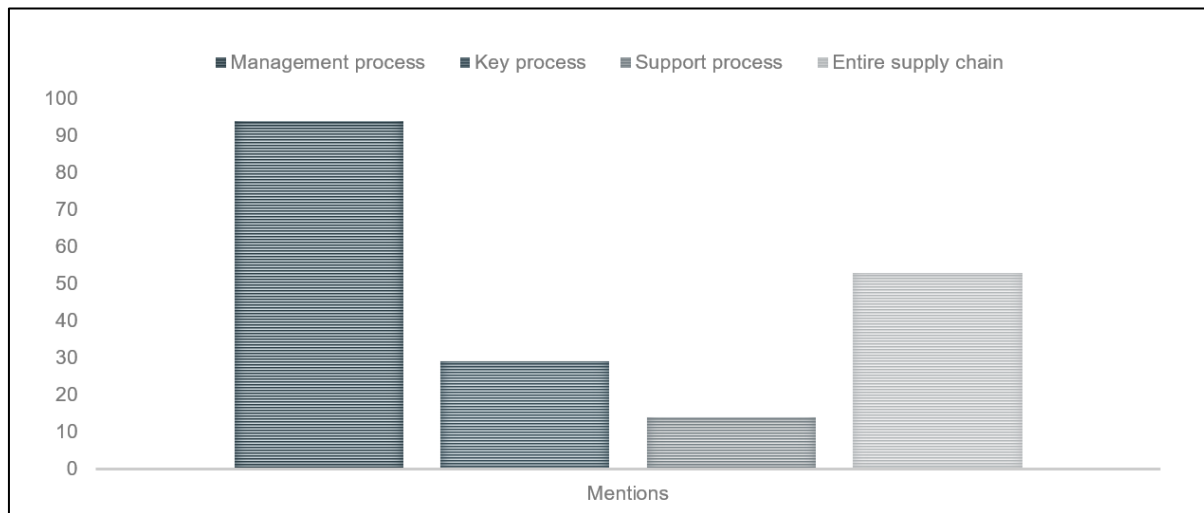


Fig. 5: Distribution of codings across different business processes

Publications that discuss sustainability in management processes suggest that the topic can be integrated in multiple directions (Lacy et al., 2009; Olbert-Bock et al., 2015; Stehle and Stelkens, 2018). Sustainability should be integrated both top-down and bottom-up by engaging in continuous stakeholder dialogue. However, a significant amount of data is needed for adequate sustainability performance management within these processes. Another option is to incentivize sustainable decision-making (Schäffer, 2022). Additionally, non-financial information should be as relevant in management decisions as financial information and thus discussed as such in corresponding contexts (Sulaiman and Mokhtar, 2012). Some publications also discuss key processes such as production, logistics and procurement as appropriate options to promote sustainability (Ferreira et al., 2010; Fischer et al., 2010; Weber et al., 2010). This could be the reduction of scrap products and materials, resource and energy consumption, as well as the extension of product life. In support processes, sustainability takes on the form of employee habit change and the sustainability of the company itself (Krause, 2016; Weigand, 2020). This extends to aspects such as using recycled printing paper (or not printing at all), being conscious of cloud data and server energy usage, recycling in-house, using

regenerative energy to power production sites and offices and donating to charity. Publications discussing a company's entire supply chain mostly focus on the social dimension of sustainability by describing the impact on local labor laws and even environmental protection laws at production sites (Gaggl, 2021; Hubbard, 2009; Klute-Wenig and Refflinghaus, 2015; Rodriguez-Olalla and Aviles-Palacios, 2017; Schulz, 2014).

Additionally, most texts discuss a strategic integration of sustainability (n=81 codings). Operative and tactical timeframes are mentioned less often (Fig. 6).

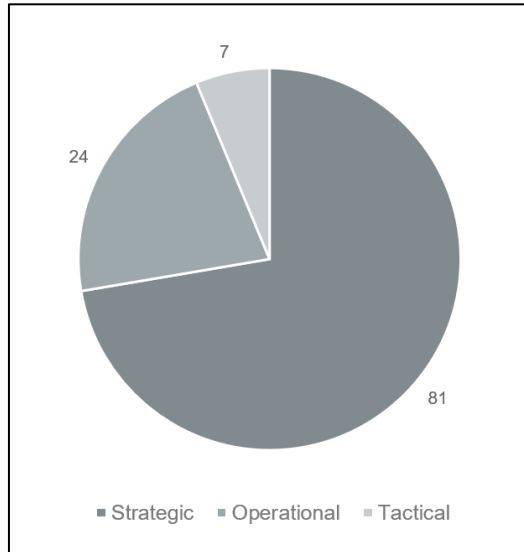


Fig. 6: Mentioned timeframe for sustainable transformation

3.2.2 Sustainability reporting, certificates, frameworks, tools and methods

Regarding the non-financial data that companies reported on, a few trends can be identified in the publications (Fig. 7). Most publications mention that companies use the GRI standards and focus on emissions, the exploitation of resources and work conditions (Gawenko et al., 2020; Pandit and Rubenfield, 2016; Rönz and Ryba, 2012; Weber et al., 2010). Less emphasis is put on creating equal opportunities, keeping human rights or fighting lobbyism and corruption (Schier and Bonnländer, 2018; Schmitz, 2021).

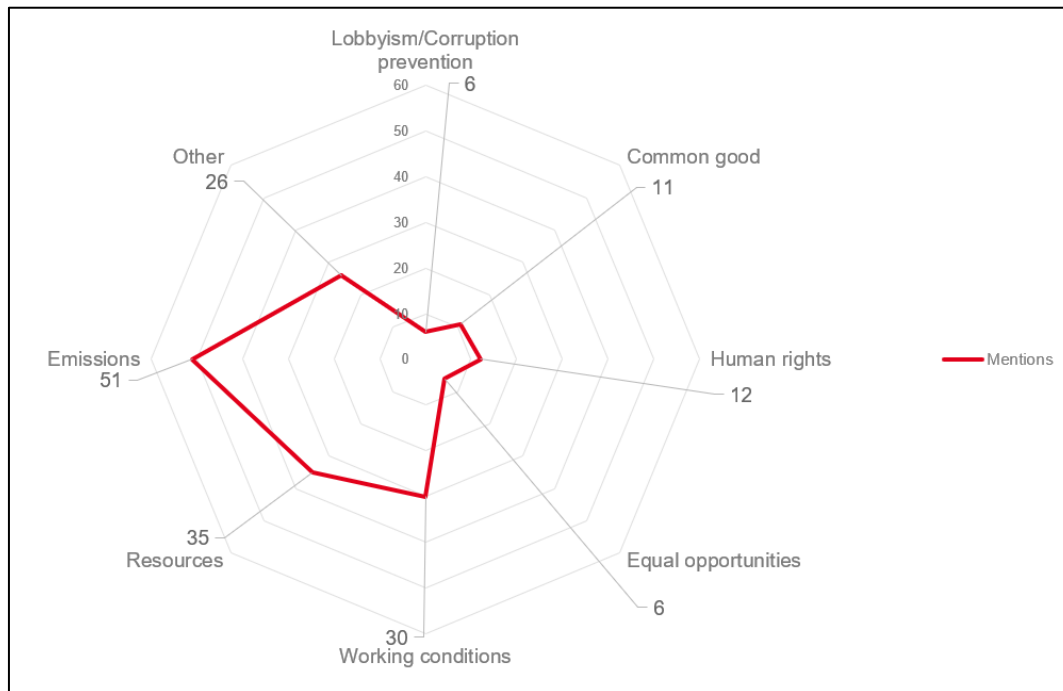


Fig. 7: CSR reporting items mentioned by publications

The tools that are mentioned in controlling and accounting codings go beyond more general calculations like KPIs with (n=19) and without a sustainability focus (SKPI, n=16) to more specific ones, such as eco efficiency analysis (n=3), social return on investment (SROI, n=2), sustainability balanced scorecard (SBSC, n=19) and sustainable value (n=9). The SBSC was frequently discussed from a goal-setting perspective, thus more useful as a starting point.

The general KPIs mentioned were rather specific control systems like economic value added (EVA) and total return on investment (Lux and Olbert-Bock, 2016; Schulze and Thomas, 2012; Schwarzmaier, 2013). Newer publications suggested SKPIs such as the recording of CO2 emissions, circular transition indicators or sustainable value added (Gaggl, 2021; Kämmler-Burrak and Bauer, 2022; Lühn et al., 2022). Additionally, newer publications also stressed the issue of data availability and standardization (Kämmler-Burrak and Bauer, 2022). In recent years, there has been an influx of KPIs and tools in sustainability controlling and accounting, most of which companies have developed themselves through benchmarking (Heimel and Momberg, 2021; Rönz and Ryba, 2012).

3.2.3 Competences

The competences discussed in the texts were presented in a general manner or with a focus on sustainability, since some competences are easily adaptable for different contexts (e.g. researching alternatives for products or developing new strategies), whereas others are only needed in specific sustainability-focused situations (e.g. the calculation of CO2 emissions or the corporate carbon footprint). In the n=79 publications, professional and methodological competences were discussed the most frequently (Fig. 8, left), social competences the least.

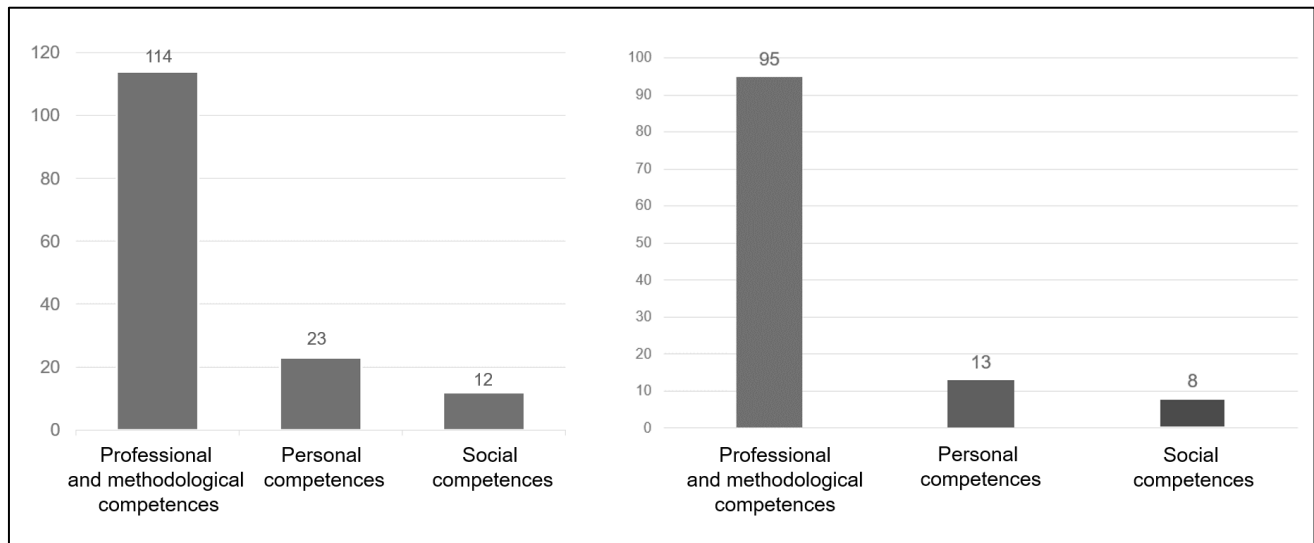


Fig. 8: Distribution of the mentioned competence dimensions as a total (left) and with a sustainability focus (right)

Roughly 83% of the codings for professional and methodological competences are connected to sustainability in some way. In contrast, only 56% of personal competences and 66% of social competences can be connected to sustainability (Fig. 8, right) (Hoekstra, 2020; Kämmler-Burraak and Bauer, 2022; Sneidere and Bumane, 2019; Tsiligiris and Bowyer, 2021).

Results from the professional and methodological dimension can be clustered into three key aspects: 1) sustainability knowledge, 2) knowledge about tools and KPIs and 3) data and interface management. The first aspect touches on the development of professional knowledge about sustainability. Employees in controlling and accounting need to identify intersections between their area of expertise and potential sustainability aspects to analyze information potentially relevant for sustainability controlling processes (Gaggl, 2021). It is not required for them to become sustainability experts, but there is nuance: employees in controlling and accounting should be open to include sustainability-oriented tools and processes into their repertoire and be available to take on learning opportunities (Lühn et al., 2022; Schulz, 2014; Stehle and Stelkens, 2018). Especially the methodological skill can be developed to include the transformation of ecological and social effects into financial units (Weber et al., 2010). To aid this, companies need to train controllers as additional ‘sustainability managers’ by developing their general skill of preparing management information (Schulz, 2014; Schulze and Thomas, 2012). Lastly, these employees need to be able to communicate the performance of management accounting reports or other regular external reporting and realize their value for controlling processes (Weber et al., 2010).

Regarding the second aspect, publications mention three relevant perspectives. Those concepts that are currently used need to be reevaluated for their suitability regarding a sustainable transformation. Thus, employees need to be able to measure, analyze, report and evaluate financial data to quantify non-financial information, i.e. by adapting existing KPIs and/or tools (Clarke and O’Neill, 2005). In doing so, they need to apply legal frameworks and keep up with industry standards and best practices. When adaptation is not an option, new concepts need to be developed by defining new relevant KPIs that gather the corporate environmental and social impacts (Stehle and Stelkens, 2018). Necessary competences include benchmarking, risk analysis, stakeholder surveys and external ratings (Kämmler-Burraak and Bauer, 2022). When finally, all tools are up to date, they need to be competent

in the use of the new tools, thus broadening their competences in measuring sustainability-relevant information (Hartmann et al., 2016). Then, employees should consider the fit with the corporate sustainability strategy and control accordingly. This might require new workflows (Schulze and Thomas, 2012).

The last aspect, data and interface management, considers the support of other potentially relevant departments and processes. Not all information relevant to sustainability is aggregated in a single department. Thus, departmental cooperation is required. Some publications suggest collaboration between the sustainability experts and those employees who now need to include sustainability in their activities (Fischer et al., 2010; Isensee and Michel, 2011). This can be done through an additional or integrated sustainability department. This reduces specialized technical knowledge that employees in controlling and accounting need to develop (Schaltegger, 2020).

Regarding social competences, publications primarily focus on cooperation, collaboration and communication. Employees are required to use high-quality data as a basis and keep it up to date by collaborating with other departments through empathy and initiative (Gebauer, 2013; Schubert and Gerhardt, 2021). This extends to the collection of relevant data from other departments, including the sustainability department (Fischer et al., 2010).

The personal competences can be divided into ethical guidelines and respecting values as well as identifying complexities and being willing to learn. Regardless of their personal opinion on sustainability, employees need to respect company values and guidelines in their recommendations (Walinska and Dobroszek, 2021) and produce comparable, reliable and consistent data and reports (Arora et al., 2022). Additionally, they need to put problems into larger contexts and derive implications for their actions (D'Amato et al., 2009).

The competence requirements clearly show that controllers need to expand their existing competences: a certain level of sustainability knowledge is required as well as the ability to navigate sustainability-oriented tools and key figures by adapting them or developing new ones. Data management is also important, as non-financial data may need to be transformed or communicated. Since sustainability data occurs everywhere in the company, controllers also have to manage interfaces. Overall, controllers do not need to be experts, but they do need to cooperate.

4 Discussion

4.1 Sustainability in the Occupational Reality of Employees in Controlling and Accounting

The exploration of sustainability in accounting and controlling literature reveals a dynamic shift from a traditional perspective, which primarily aimed at securing the long-term existence of companies, to an integrated approach aligned with the three-pillar model. Notably, sustainability discussions have expanded beyond separate sustainability management and controlling departments, indicating a need for a more integrative approach within organizations. The literature highlights the challenges faced by companies, including a lack of expertise, financial resources, and industry interest, emphasizing the multifaceted nature of sustainability integration. Relevant tools and processes in controlling predominantly focus on KPIs, cost management, and sustainability KPIs, underscoring the need for a comprehensive understanding of both financial and non-financial aspects. Furthermore, the literature suggests the importance of sustainability-oriented information and tools that integrate sustainability into mainstream controlling processes rather than segregating it into a distinct sustainability controlling function.

The identified integration of sustainability initiatives spans across various organizational levels, ranging from dedicated sustainability departments to functions within sales, marketing, PR, and finance, here most often it is sustainability controlling. The predominant motivations driving this integration are legal compliance and profit optimization. The strategic implementation of sustainability measures is dominated by the economic dimension, emphasizing the prevalence of economic goal conflicts. The ecological dimension is affected through the reduction of emissions and resources, while the social dimension mostly considers working conditions, often interconnected with legal obligations such as labor laws.

The primary objective of controlling activities is to monitor and evaluate key performance indicators (KPIs) with and without a sustainability focus, and to assess their relationship with cost management. It is notable that there is an emphasis on capturing sustainability-relevant management information across the entire company. This is to be achieved by utilizing tools that consider sustainability holistically, rather than relying on separate sustainability controlling entities. In the domain of accounting, there are a number of challenges to be overcome. These include the incorporation of non-monetary information, such as the accrual of financial information, and the external reporting of non-financial aspects. A number of proposed solutions have been put forward. These include the amortisation of costs through funds, the utilisation of social balance sheets, or environmental investment calculations as complementary tools alongside traditional instruments like the income statement.

The discussion on sustainability-focused competences in accounting and controlling research and practice underscores the critical role of accounting and controlling in VET in driving corporate transformation towards sustainability. Students require general competences such as sustainability-related learning readiness, collaboration, identification, and adaptation of relevant tools and KPIs, understanding and addressing complexities and conflicts, and weighing ethical principles. The literature suggests that even in organizations with dedicated sustainability departments, controllers need to possess sustainability-related competences, indicating the pervasive nature of sustainability across different facets of business operations.

The identified competences extend beyond technical proficiency to include sustainability-related learning readiness, collaboration, and cooperation. Professionals need to identify and potentially develop or adapt relevant tools and metrics, recognizing the evolving nature of sustainability metrics. The literature focuses on the importance of understanding complexities and conflicts to enable students to propose viable solutions as change agents. In addition, personal competence should balance personal and organizational (ethical) principles, which are both crucial for navigating sustainability decision-making.

Synthesizing these insights, the literature points towards a paradigm shift in organizational practices, where sustainability is a strategic imperative while considering economic implications. The competences highlighted serve as a starting point to develop critical implications for vocational education. These indicate the need for a comprehensive and dynamic curriculum that not only imparts technical knowledge but cultivates a holistic perspective on the complexities and ethical dimensions of sustainable decision-making as change agents.

4.2 Implications for teacher content knowledge and curriculum design

Combining the results described above and the central arguments from the discussion with the critical perspective on sustainability in accounting and controlling developed in the introduction, the

following implications should be considered to include a sustainability-forward approach in accounting and controlling lessons:

- The integration of sustainability in **learning scenarios** to connect the increasing relevance of sustainability in decision-making processes. Instead of a selective approach, teachers need to view sustainability in controlling as an intersectional topic that affects students in multiple dimensions.
- Connecting to this, educators need to be aware of **alternative controlling and accounting tools and KPIs** that they can add to their repertoire. They need to be able to research relevant additions to the school-wide curriculum and integrate them into learning scenarios. In this context, controlling educators should also consider the importance of non-financial KPIs and their potential for learning opportunities as they may aid to illustrate conflicts and foster valuable learning opportunities as they connect the dimensions of economics, business management and legal.
- Educators should include the premise of **economic sustainability as the primary goal** to employ sustainability conflicts in a didactically beneficial manner, rather than to make the students feel demotivated. Students will need to develop a set of competences that helps them drive organizational change as part of their vocational action competence. For this, it might be helpful to discuss alternatives and best practice examples.
- A deepened **business process orientation** that focuses on management processes to enable students to take on a multidimensional approach. German vocational business education is, in theory, inherently process-oriented; thus, the development of sustainability integrated processes may uncover potential starting points for lesson planning. When discussing the entire supply chain, sustainability-oriented aspects need to be considered to depict the corporate situation holistically. In addition, discussing business processes helps illustrate conflicts between sustainability dimensions because it illustrates cause-and-effect chains.

4.3 Limitations

There are some limitations to this review that need to be discussed. Although with PRISMA, we used a standardized strategy, we cannot control ‘intelligent’ search algorithms that use geolocalization (particularly with Google Scholar) and smart indexing. This might also explain the uneven language distribution. Although we searched for the same keywords in English and German, we retrieved more German than English publications. We prioritized publications that would give us insights into relevant content knowledge for vocational business education and, more specifically, the area of accounting and controlling. It is very likely that we missed publications that showcase sustainability efforts but do not do so in the context of accounting and controlling. Additionally, no case studies about higher education contexts were included as we focused on developments in the relevant academic field. These two aspects might have limited the scope of the papers that we analyzed. Another limitation is the uneven distribution of controlling and accounting. Controlling literature was predominant as external accounting often faces legal constraints. Thus, we naturally identified more implications for this area. In addition, there is another content-related limitation as in German concepts, controlling only comprises internal accounting and is decision-oriented, whereas internationally, a more integrated approach is common (Beckers, 2010). Thus, the data basis controllers use is different. These variations have grown historically and hence contributed to different views on the roles of controllers. Through this we might have unintentionally excluded papers using different vocabulary than us. As a consequence, this literature review is neither exhaustive nor representative for the discourse but rather serves as a starting point for discussions about relevant teacher content knowledge.

5 Conclusion

This research paper sheds light on the role of sustainability in accounting and controlling, with a particular emphasis on its implications for vocational education. The study provides valuable insights into how sustainability is discussed in accounting and controlling literature and the competences essential for practitioners in this field:

In the context of vocational education, the study proposes implications for accounting and controlling teachers. It advocates for the integration of sustainability into learning scenarios to align with the increasing relevance of sustainability in decision-making processes. The paper emphasizes the need for teachers to be aware of alternative tools and KPIs, fostering a proactive approach to researching and integrating relevant additions into the curriculum. Additionally, ethical considerations are highlighted, acknowledging the potential dissonance between personal values and corporate practices that students may encounter in their future careers. The study encourages a deepened business process orientation to enable students to take on a multidimensional approach and emphasizes the development of sustainable controlling and accounting action competences. In essence, the research not only contributes to understanding where companies are currently at regarding sustainability in accounting and controlling literature but also provides practical guidance for vocational education, urging educators to adapt their approaches to equip students with the necessary competences for a future where sustainability is integral to corporate decision-making. The implications drawn from this research paper hold particular relevance within the context of the vocational education system in Germany. German vocational business education has long been renowned for its emphasis on practical training and its early and thorough integration into the workforce. As the study highlights the topic's increasing importance, it becomes imperative to consider how these findings can be effectively incorporated into research for accounting and controlling didactics in vocational business education. New concepts need to be developed, trialed and established to develop the fitting content knowledge.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

JP: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft

FB: Conceptualization, Writing – review & Editing, Methodology, Supervision

Funding

No funds, grants, or other support was received.

Data Availability Statement

The datasets analyzed for this study can be made available from the corresponding author on reasonable request.

References

- 537 Alsdorf, H., Kosmol, L., and Rivera Fernandez, Z. (2023). Relevanz und Bedeutung der
538 Nachhaltigkeit – Green IT/IS als Nischenthema in KMU der DACH-Region. *HMD* 60:162–72.
539 doi: 10.1365/s40702-022-00863-y
- 540 Aras, G., and Crowther, D. (2009). Corporate Sustainability Reporting: A Study in Disingenuity? *J*
541 *Bus Ethics* 87:279–88. doi: 10.1007/s10551-008-9806-0
- 542 Arora, M. P., Lodhia, S., and Stone, G. (2022). Enablers and barriers to the involvement of
543 accountants in integrated reporting. *MEDITARI ACCOUNTANCY RESEARCH* 30:676–709. doi:
544 10.1108/MEDAR-11-2020-1102
- 545 Beckers, B. (2010). *Controlling Berufsfeldanalyse Deutschland und USA*. Igel Verlag.
- 546 Braun, S. (2019). Nachhaltigkeitsberichterstattung: Im Spannungsfeld unterschiedlicher
547 Anforderungen und Interessen. *Ökologisches Wirtschaften*, 22–24.
- 548 Broccardo, L., Giordino, D., Yaqub, M. Z., and Alshibani, S. (2024). Implementing sustainability:
549 What role do knowledge management and management accounting play? Agenda for
550 environmentally friendly businesses. *Corporate Social Responsibility and Environmental*
551 *Management*, 32:383–403. doi: 10.1002/csr.2936
- 552 Bundesinstitut für Berufsbildung. (2023). Datenreport zum Berufsbildungsbericht 2023:
553 Informationen und Analysen zur Entwicklung der beruflichen Bildung. Bonn.
- 554 Chiapello, E. (2009). “Die Konstruktion der Wirtschaft durch das Rechnungswesen,” in *Diskurs und*
555 *Ökonomie: Diskursanalytische Perspektiven auf Märkte und Organisationen*, ed. R. Diaz-Bone and
556 G. Krell (Wiesbaden: VS), 125–49.
- 557 Clarke, K., and O’Neill, S. (2005). Is the Environmental Professional ... an Accountant? *Greener*
558 *Management International*, 111–24.
- 559 D’Amato, A., Roome, N., and Lenssen, G. (2009). Toward an integrated model of leadership for
560 corporate responsibility and sustainable development: a process model of corporate responsibility
561 beyond management innovation. *Corporate Governance: The International Journal of Effective*
562 *Board Performance* 9:421–34. doi: 10.1108/14720700910984972
- 563 Endenich, C., and Trapp, R. (2019). “Nachhaltigkeitscontrolling in Klein- und Mittelunternehmen,”
564 in *Controlling – Aktuelle Entwicklungen und Herausforderungen: Digitalisierung, Nachhaltigkeit*
565 *und Spezialaspekte*, ed. B. Feldbauer-Durstmüller and S. Mayr (Wiesbaden: Springer Fachmedien
566 Wiesbaden), 229–46.
- 567 European Union. (2022). Directive (EU) 2022/2464 of the European Parliament and of the Council.
- 568 Ferreira, A., Moulang, C., and Hendro, B. (2010). Environmental management accounting and
569 innovation: an exploratory analysis. *Accounting, Auditing & Accountability Journal* 23:920–48.
570 doi: 10.1108/09513571011080180
- 571 Figge, F., and Hahn, T. (2004). Sustainable Value Added—measuring corporate contributions to
572 sustainability beyond eco-efficiency. *Ecological Economics* 48:173–87. doi:
573 10.1016/j.ecolecon.2003.08.005
- 574 Fischer, T. M., Huber, R., and Sawczyn, A. (2010). Nachhaltige Unternehmensführung als
575 Herausforderung für das Controlling. *CON* 22:222–30. doi: 10.15358/0935-0381-2010-4-5-222
- 576 Frizon, J., and Eugénio, T. (2022). Recent developments on research in sustainability in higher
577 education management and accounting areas. *The International Journal of Management Education*,
578 20:100709. doi: 10.1016/j.ijme.2022.100709

- 579 Fürstenau, B., Pilz, M., and Gonon, P. (2014). "The Dual System of Vocational Education and
580 Training in Germany – What Can Be Learnt About Education for (Other) Professions," in
581 International Handbook of Research in Professional and Practice-based Learning, ed. S. Billett, C.
582 Harteis and H. Gruber (Wiesbaden: Springer). 427-460. doi: 10.1007/978-94-017-8902-8_16
- 583 Gadatsch, A. (2017). Grundkurs Geschäftsprozess-Management: Analyse, Modellierung,
584 Optimierung und Controlling von Prozessen. Wiesbaden: Springer Vieweg.
- 585 Gaggl, P. (2021). "Nachhaltigkeitscontrolling: Wie nichtfinanzielle Informationen zum Werttreiber
586 werden," in Funktions-Controlling: Praxishandbuch für Unternehmen, Non-Profit-Organisationen
587 und die öffentliche Verwaltung, ed. R. Eschenbach, J. Baumüller and H. Siller (Wiesbaden:
588 Springer Fachmedien Wiesbaden), 327–64.
- 589 Gawenko, W., Richter, F., Hinz, M., and Götze, U. (2020). Interne Ansätze zur
590 Nachhaltigkeitsbewertung in der externen Berichterstattung. Konzeptionelle und empirische
591 Analyse der DAX-Unternehmen. Internal Approaches for Sustainability Assessment in the
592 External Reporting—A Conceptual and Empirical Analysis of DAX-Companies 74:264–84. doi:
593 10.5771/0042-059X-2020-3-264
- 594 Gebauer, J. (2013). Die Zukunft der Nachhaltigkeitsberichterstattung III und IV/IV: Glaubwürdigkeit
595 können nur die Unternehmen selbst herstellen. Ökologisches Wirtschaften, 8–9.
- 596 Gil-Marín, M., Vega-Muñoz, A., Contreras-Barraza, N., Salazar-Sepúlveda, G., Vera-Ruiz, S., &
597 Losada, A. V. (2022). Sustainability Accounting Studies: A Metasynthesis. *Sustainability* 14:
598 9533. doi:10.3390/su14159533
- 599 Glanze, E. (2022). Kompass für Green Change: Das Reifegradmodell des betrieblichen
600 Nachhaltigkeitsmanagements. OrganisationsEntwicklung, 103–7.
- 601 Glen, J., Hilson, C., and Lowitt, E. (2009). The emergence of green talent. *Business Strategy Review*,
602 52-56.
- 603 Greiling, D., and Ther, D. (2010). "Leistungsfähigkeit des Sustainable Value-Ansatzes als Instrument
604 des Sustainability Controlling," in Corporate Sustainability: Der Beitrag von Unternehmen zu einer
605 nachhaltigen Entwicklung in Wirtschaft und Gesellschaft., ed. H. K. Prammer (Wiesbaden:
606 Gabler), 37–68.
- 607 Günther, E., and Stechemesser, K. (2011). Instrumente des Green Controllings: ein Blick zurück, ein
608 Blick nach vorn. *Controlling* 23:417–23. doi: 10.15358/0935-0381-2011-8-9-417
- 609 Hartmann, F., Maas, K., and Perego, P. (2016). "Den Wald vor lauter Bäumen nicht sehen:
610 Controller auf der Suche nach Nachhaltigkeit," in CSR und Controlling: Unternehmerische
611 Verantwortung als Gestaltungsaufgabe des Controlling, ed. E. Günther and K.-H. Steinke (Berlin,
612 Heidelberg: Springer Berlin Heidelberg), 71–82.
- 613 Heimel, J., and Momberg, M. (2021). "Sustainable Finance: Nachhaltigkeitscontrolling zur
614 Steuerung des sozialen und ökologischen Wirtschaftens von Unternehmen," in Nachhaltiger
615 Konsum: Best Practices aus Wissenschaft, Unternehmenspraxis, Gesellschaft, Verwaltung und
616 Politik, ed. W. Wellbrock and D. Ludin (Wiesbaden: Springer Fachmedien Wiesbaden), 83–106.
- 617 Herrmann, C., and Gerlach, E. (2020). Unterrichtsqualität im Fach Sport – Ein Überblicksbeitrag
618 zum Forschungsstand in Theorie und Empirie. *Unterrichtswissenschaft*, 48:361-384. doi:
619 10.1007/s42010-020-00080-w
- 620 Hoekstra, R. (2020). SNA and beyond: Towards a Broader Accounting Framework That Links the
621 SNA, SDGs and Other Global Initiatives. *Statistical Journal of the IAOS* 36, 657–75.

- Hubbard, G. (2009). Measuring organizational performance: beyond the triple bottom line. *BUSINESS STRATEGY AND THE ENVIRONMENT* 18:177–91. doi: 10.1002/bse.564
- Isensee, J., and Michel, U. (2011). Green Controlling - Die Rolle des Controllers und aktuelle Entwicklungen in der Praxis. *Controlling* 23:436–42. doi: 10.15358/0935-0381-2011-8-9-436
- Kämmeler-Burrak, A., and Bauer, R. (2022). Nachhaltigkeit wird Standardaufgabe im Controlling. *CONTROLLER Magazin*, 24–29.
- Klein, J., and Küst, C. (2020). “Wie die Digitalisierung im Rechnungswesen die Aufgaben und Anforderungen an die Mitarbeiter/-innen verändert,” in *Moderner Rechnungswesenunterricht 2020: Status quo und Entwicklungen aus wissenschaftlicher und praktischer Perspektive*, ed. F. Berding, H. Jahncke and A. Slopinski (Wiesbaden: Springer VS), 83–97.
- Kleinberg Neimark, M. (1996). Caught in the squeeze: An essay on higher education in accounting. *Critical Perspectives on Accounting*, 7:1-11. doi: 10.1006/cpac.1996.0001
- Klute-Wenig, S., and Refflinghaus, R. (2015). Integrating sustainability aspects into an integrated management system. *TQM Journal* 27:303–15. doi: 10.1108/TQM-12-2013-0128
- Krause, H.-U. (2016). Controlling-Kennzahlen für ein nachhaltiges Management: Ein umfassendes Kompendium kompakt erklärter Key Performance Indicators. De Gruyter.
- Kuckartz, U. (2018). Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. Weinheim: Beltz.
- Lacy, P., Arnott, J., and Lowitt, E. (2009). The challenge of integrating sustainability into talent and organization strategies: investing in the knowledge, skills and attitudes to achieve high performance. *Corporate Governance: The International Journal of Effective Board Performance* 9, 1-0.
- Lühn, M., Nuzum, A.-K., Petersen, H., and Schaltegger, S. (2022). Controller und Nachhaltigkeit im Kontext neuer Berichterstattungspflichten. *Rethinking Finance*, 25–33.
- Lux, W., and Olbert-Bock, S. (2016). Strategisches Controlling als Teil des Sustainability Performance Managements - auch für KMU. *CONTROLLER Magazin*, 4–12.
- Martins, A., Branco, M. C., Melo, P. N., and Machado, C. (2022). Sustainability in Small and Medium-Sized Enterprises: A Systematic Literature Review and Future Research Agenda. *SUSTAINABILITY* 14:6493. doi: 10.3390/su14116493
- Michel, U., Isensee, J., and Stehle, A. (2014). “Sustainability Controlling: Planung, Steuerung und Kontrolle der Realisierung der Nachhaltigkeitsstrategie,” in *CSR und Finance*, ed. T. Schulz and S. Bergius (Berlin, Heidelberg: Springer Berlin Heidelberg), 97–112.
- Niesen, Tim and Houy, Constantin (2015). Zur Nutzung von Techniken der Natürlichen Sprachverarbeitung für die Bestimmung von Prozessmodellähnlichkeiten – Review und Konzeptentwicklung. *Wirtschaftsinformatik Proceedings* 122.
- Olbert-Bock, S., Bussmann, C., Lux, W., and Garrels, S. (2015). Nachhaltigkeitsmanagement erfolgreich im Unternehmen professionalisieren. *IM+io Fachmagazin*, 50–57.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 10:89. doi: 10.1186/s13643-021-01626-4
- Pandit, G. M., and Rubenfield, A. J. (2016). The Current State of Sustainability Reporting by Smaller S&P 500 Companies. *CPA Journal* 86, 52–57.

- 664 Pargmann, J., Riebenbauer, E., Flick-Holtsch, D., and Berding, F. (2023). Digitalisation in
665 accounting: a systematic literature review of activities and implications for competences. *Empirical*
666 *Research in Vocational Education and Training* 15.
- 667 Perridon, L., Rathgeber, A. W., and Steiner, M. (2017). *Finanzwirtschaft der Unternehmung*.
668 München, Germany: Verlag Franz Vahlen.
- 669 Preiß, P., and Tramm, T. (1996). “Die Göttinger Unterrichtskonzeption des
670 wirtschaftsinstrumentellen Rechnungswesens,” in *Rechnungswesenunterricht und ökonomisches*
671 *Denken*, ed. P. Preiß and T. Tramm (Wiesbaden: Gabler), 222–323.
- 672 Rodriguez-Olalla, A., and Aviles-Palacios, C. (2017). Integrating Sustainability in Organisations: An
673 Activity-Based Sustainability Model. *SUSTAINABILITY* 9. doi: 10.3390/su9061072
- 674 Rönz, M., and Ryba, M. (2012). Kennzahlen in Nachhaltigkeitsberichten. *CONTROLLER Magazin*,
675 87–91.
- 676 Roth, H. (1971). *Pädagogische Anthropologie, Band 2. Entwicklung und Erziehung. Grundlagen*
677 *einer Entwicklungspädagogik*. Hannover.
- 678 Schäffer, U. (2022). Nachhaltigkeit und Controlling: Ein Aufruf zum Handeln. *CONTROLLER*
679 *Magazin*, 6–10.
- 680 Schaltegger, S. (2020). Nachhaltigkeit als Treiber des Unternehmenserfolgs: Ein
681 mehrperspektivischer Controllingansatz. *Rethinking Finance*, 34–39.
- 682 Schaltegger, S., and Zvezdov, D. (2011). Konzeption und Praxis des Nachhaltigkeitscontrollings.
683 *Controlling* 23:430–35. doi: 10.15358/0935-0381-2011-8-9-430
- 684 Schier, M., and Bonnländer, A. (2018). Nachhaltigkeit als Teil der Geschäftsstrategie. *Die Bank*, 38–
685 41.
- 686 Schmidt, O. (2012). Von den Wurzeln der Nachhaltigkeit: Vor drei Jahrhunderten erfand die
687 Forstwirtschaft das Prinzip der Nachhaltigkeit. *LWF aktuell*, 50–51.
- 688 Schmitz, H. (2021). GRI-Nachhaltigkeitskennzahlen für das Controlling im Mittelstand.
689 *CONTROLLER Magazin*, 60–61.
- 690 Scholz, U., and Pastoors, S. (2018). “Betriebliche Nachhaltigkeit,” in *Praxishandbuch Nachhaltige*
691 *Produktentwicklung*, ed. U. Scholz, S. Pastoors, J. H. Becker, D. Hofmann and R. van Dun (Berlin,
692 Heidelberg: Springer Berlin Heidelberg), 11–21.
- 693 Schöning, S., Mendel, V., and Köse, A. (2020). Mit neuen Controller-Kompetenzen in die Zukunft.
694 *Control Manag Rev* 64:58–63. doi: 10.1007/s12176-019-0073-y
- 695 Schubert, R., and Gerhardt, N. (2021). Praxisanforderungen mittelständischer Unternehmen an
696 Controlling-Nachwuchskräfte. *CONTROLLER Magazin*, 31–35.
- 697 Schulz, T. (2014). “CFO-Agenda: Gute Gründe, Nachhaltigkeit auf die Tagesordnung zu setzen,” in
698 *CSR und Finance*, ed. T. Schulz and S. Bergius (Berlin, Heidelberg: Springer Berlin Heidelberg),
699 3–34.
- 700 Schulze, M., and Thomas, S. (2012). Strategisches Nachhaltigkeits-Controlling bei der Deutschen
701 Telekom AG Corporate Responsibility Key Performance-Indikatoren von Mike Schulze und Silke
702 Thomas. *CONTROLLER Magazin*, 58–63.

- 703 Schwarzmaier, U. (2013). Entwicklungstendenzen des Controllings unter besonderer
704 Berücksichtigung der Veränderungen durch die Nachhaltigkeitsdiskussion. *CONTROLLER*
705 *Magazin*, 29–36.
- 706 Schwarzmaier, U. (2015). Die Rolle des Controllings im Sustainable Management. *Der Betriebswirt*,
707 22–27.
- 708 Seifried, J. (2004). Fachdidaktische Variationen in einer selbstorganisationsoffenen Lernumgebung.
709 [Dissertation].
- 710 Sneider, R., and Bumane, I. (2019). “THE CHALLENGES FOR THE PROFESSION OF
711 ACCOUNTANT IN THE CHANGING GLOBAL ECONOMIC ENVIRONMENT,” in *NEW*
712 *CHALLENGES OF ECONOMIC AND BUSINESS DEVELOPMENT - 2019: INCENTIVES*
713 *FOR SUSTAINABLE ECONOMIC GROWTH (19 RAINA BLVD, RIGA, LV 1586, LATVIA:*
714 *UNIV LATVIA)*, 761–71.
- 715 Stehle, A., and Stelkens, V. (2018). Green Controlling - eine Aufgabe für Controller und
716 Nachhaltigkeitsexperten! *CONTROLLER Magazin*, 4–10.
- 717 Stütz, S. (2024). Charakteristika von Aufgaben in der Domäne Rechnungswesen. Wiesbaden:
718 Springer Fachmedien Wiesbaden.
- 719 Sulaiman, M., and Mokhtar, N. (2012). Ensuring Sustainability: A Preliminary Study of
720 Environmental Management Accounting in Malaysia. *International Journal of Business &*
721 *Management Science* 5, 85–102.
- 722 Tsiligiris, V., and Bowyer, D. (2021). Exploring the impact of 4IR on skills and personal qualities for
723 future accountants: a proposed conceptual framework for university accounting education.
724 *ACCOUNTING EDUCATION* 30:621–49. doi: 10.1080/09639284.2021.1938616
- 725 Walinska, E., and Dobroszek, J. (2021). The Functional Controller for Sustainable and Value Chain
726 Management: Fashion or Need? A Sample of Job Advertisements in the COVID-19 Period.
727 *SUSTAINABILITY* 13. doi: 10.3390/su13137139
- 728 Weber, J., Georg, J., and Janke, R. (2010). Nachhaltigkeit: Relevant für das Controlling? *Controlling*
729 *& Management* 54:395–400. doi: 10.1007/s12176-010-0101-4
- 730 Weigand, H. (2020). “Green Marketing - nachhaltig erfolgreich,” in *Die 10 wichtigsten*
731 *Zukunftsthemen im Marketing: Buzzwords die bleiben*, ed. M. Stumpf (Haufe-Lexware GmbH &
732 Co. KG), 47–69.
- 733 Wellbrock, W., Ludin, D., and Krauter, S. (2020). Nachhaltigkeitscontrolling: Instrumente und
734 Kennzahlen für die strategische und operative Unternehmensführung. Wiesbaden: Springer
735 Fachmedien Wiesbaden.
- 736 Winther, E., and Achtenhagen, F. (2008). Kompetenzstrukturmodell für die kaufmännische Bildung.
737 *zbw* 104:511–38. doi: 10.25162/zbw-2008-0028
- 738 Wöhe, G., Döring, U., and Brösel, G. (2020). Einführung in die allgemeine Betriebswirtschaftslehre.
739 München: Verlag Franz Vahlen.
- 740 World Commission on Environment and Development. (1987). Report of the world commission on
741 environment and development: Our common future. UN Documents Gathering a body of global
742 agreements.
- 743

744

745

746

747 Table 1. List of Keywords

748

Group No.	List of Keywords
1	Sustainability, ecology, sustainable development goals, non-financial reporting, corporate social responsibility, integrated reporting, sustainability controlling, accounting for sustainability, financial accounting, cost accounting, investment accounting, external accounting, sustainable key performance indicators
2	competence, skills, activity, education, training, learning process, accounting, controlling competence, skills, activity, education, training, professional competence, action competence, methodological competence, social competence, self-competence, accounting, controlling, sustainability controlling

749

750 Table 2: Analytical framework with sub-categories

751

Main category	Level 1 sub-category	Level 2 sub-category	Selected examples
Approach to sustainability	Level of business areas	Procurement, HR, PR, marketing, finance	Marketing/HR: “Our business is growing at a rate that we have a significant demand for energy engineers, for mechanical technicians, for systems technicians. And one of the big reasons that we continue to talk about and promote our sustainability initiatives is to recruit people into these roles.” (Glen et al., 2009) Control process: “Accordingly, many of the companies (11) surveyed have not integrated environmental information in their decision-making process.” (Sulaiman & Mokthar, 2012) Consistency: “Cradle-to-cradle” approaches (Weigand, 2020)
	Level of business processes	Control, core, support processes, entire supply chain	
	Time frame of integration	Strategic (long-term), operational (medium-term), tactical (short-term)	
	Sustainability dimensions	Ecology, Economy, Social	
	Sustainability strategies	Efficiency, consistency, sufficiency	
	Sustainability conflicts	Ecologic-social, economic-social, economic-ecologic	
	Activities	Controlling	
		Integrated reporting	
		External accounting	
Relevant certificates and frameworks,	Certificates	Internal accounting	Corruption prevention: “The remaining companies discussed environ-mental
		SA8000, ISO5000, ISO26000, EMASII, ISO9000, others	

tools, and methods	Reporting standards Tools, Concepts, KPIs	Corruption prevention, human rights, public welfare, equal opportunities, working conditions, resources, emissions Sustainable Value, Sustainability Balanced Scorecard, Social Return on Investment, eco-efficiency analysis, sustainability key performance indicators	impact at various locations through-out the report or disclosure. The information provided in this section varied from simply reporting on compliance efforts undertaken to meet the requirements of local, regional, or national agencies to providing a year-to-year comparison of the “goals and results” achieved in these areas“ (Pandit & Rubenfield, 2016) Sustainable Value: “Sustainable Value Added is the extra value created when the overall level of environmental and social impacts is kept constant.” (Figge & Hahn, 2004)
Competence requirements	Professional and methodological competence Social competence Self-competence Focus on sustainability		Professional competence with focus on sustainability: “Applying strategies to sustainability-related goals, create a balance between short term financial results and long term sustainability goals” (Kämmeler-Burraak & Bauer, 2022, freely translated) Self-competence with focus on sustainability: “Wish to learn, create ecological learning processes and keep themselves motivated” (Schaltegger, 2020)

752

753

Appendix A. List of coded publications

Author	Year	Title	Journal/Publisher
Altenburger, Reinhard	2022	Strategisches CSR-Controlling in Familienunternehmen	Springer Nature
Anderson, Kai; Sommer, Carina; Fassino, Gina	2022	Kann HR Nachhaltigkeit?	Personalmagazin
Aras, Güler; Crowther, David	2009	Corporate Sustainability Reporting: A Study in Disingenuity?	Journal of Business Ethics
Arora, Mitali Panchal; Lodhia, Sumit; Stone, Gerard	2022	Enablers and barriers to the involvement of accountants in integrated reporting	Mediterranean Accountancy Research
Asenkerschbaumer, Stefan; Watterott, Richard	2011	Controlling-Steuerungsimpulse für ein Mehr an Nachhaltigkeit am Beispiel Bosch	Controlling
Bernatzky, Simone; Endenich, Christoph; Wömpener, Andreas	2018	Zur Integration von Nachhaltigkeit in das Controlling - Eine empirische Analyse	Betriebswirtschaftliche Forschung und Praxis
Biel, Alfred	2020	Nachhaltigkeit ein Controlling-Thema?	CONTROLLER Magazin
Biswas, Sharlene; O'Grady, Winnie	2016	Using external environmental reporting to embed sustainability into organisational practices	Accounting Research Journal

Bogicevic, Jasmina; Domanovic, Violeta; Krstic, Bojan	2016	The role of financial and non-financial performance indicators in enterprise sustainability evaluation	Ekonomika
Braun, Sabine	2019	Nachhaltigkeitsberichterstattung: Im Spannungsfeld unterschiedlicher Anforderungen und Interessen	Ökologisches Wirtschaften
Burritt, Roger; Schaltegger, Stefan	2001	Eco-efficiency in corporate budgeting	Environmental Management and Health
Clarke, Kevin; O'Neill, Sharron	2005	Is the Environmental Professional ... an Accountant?	Greener Management International
D'Amato, Alessia; Roome, Nigel; Lenssen, G.	2009	Toward an integrated model of leadership for corporate responsibility and sustainable development: a process model of corporate responsibility beyond management innovation	Corporate Governance: The International Journal of Effective Board Performance
Dumitru, Mădălina; JINGA, Gabriel	2015	Integrated Reporting Practice for Sustainable Business: A Case Study	Audit Financiar
Endenich, Christoph; Trapp, Rouven	2022	Nachhaltigkeitscontrolling in Klein- und Mittelunternehmen	Springer Nature
Faupel, Christian; Stremmel, Florian	2011	Berücksichtigung von Nachhaltigkeit im Rahmen einer wertorientierten Unternehmensführung	Controlling & Management
Ferreira, Aldónio; Moulang, Carly; Hendro, Bayu	2010	Environmental management accounting and innovation: an exploratory analysis	Accounting, Auditing & Accountability Journal

Figge, Frank; Hahn, Tobias	2004	Sustainable Value Added—measuring corporate contributions to sustainability beyond eco-efficiency	Ecological Economics
Figge, Frank; Hahn, Tobias; Schaltegger, Stefan; Wagner, Marcus	2002	The Sustainability Balanced Scorecard - linking sustainability management to business strategy	Business Strategy and the Environment
Fischer, Stephan; Schmitz, Anja; Knepel, Kirke	2014	Auswahl von nachhaltig handelnden Mitarbeitern - vom Kompetenzmodell zur Auswahlmethode	Wirtschaftspsychologie
Fischer, Thomas M.; Huber, Robert; Sawczyn, Angelika	2010	Nachhaltige Unternehmensführung als Herausforderung für das Controlling.	Controlling
Gaggl, Philipp	2021	Nachhaltigkeitscontrolling: Wie nichtfinanzielle Informationen zum Werttreiber werden	Springer Gabler
Gänßlen, Siegfried; Kraus, Udo; Ette, Daniel	2011	Green Controlling - Green Profit - Nachhaltigkeitscontrolling bei Hansgrohe	Controlling
Gawenko, Wladislav; Richter, Fanny; Hinz, Michael; Götze, Uwe	2020	Interne Ansätze zur Nachhaltigkeitsbewertung in der externen Berichterstattung. Konzeptionelle und empirische Analyse der DAX-Unternehmen	Die Unternehmung
Gebauer, Jana	2013	Die Zukunft der Nachhaltigkeitsberichterstattung III und IV/IV: Glaubwürdigkeit können nur die Unternehmen selbst herstellen	Ökologisches Wirtschaften
Glanze, Eva	2022	Kompass für Green Change	OrganisationsEntwicklung

Greiling, Dorothea; Slacik, Johannes	2022	Nachhaltigkeitsberichterstattung von Elektrizitätsversorgungsunternehmen: Ein internationaler Vergleich	Springer Nature
Greiling, Dorothea; Ther, Daniela	2010	Leistungsfähigkeit des Sustainable Value-Ansatzes als Instrument des Sustainability Controlling	Springer
Günther, Edeltraud; Stechemesser, Kristin	2011	Instrumente des Green Controllings: ein Blick zurück, ein Blick nach vorn	Controlling
Günther, Thomas	2014	Planungs- und Kontrollinstrumente zur unternehmenswertorientierten Führung in mittelständischen Unternehmen	Controlling
Hartmann, Frank; Maas, Karen; Perego, Paolo	2016	Den Wald vor lauter Bäumen nicht sehen: Controller auf der Suche nach Nachhaltigkeit	Springer Gabler
Heimel, Jana; Momberg, Martin	2021	Sustainable Finance: Nachhaltigkeitscontrolling zur Steuerung des sozialen und ökologischen Wirtschaftens von Unternehmen	Springer Gabler
Hoekstra, Rutger	2020	SNA and beyond: Towards a Broader Accounting Framework That Links the SNA, SDGs and Other Global Initiatives	Statistical Journal of the IAOS
Horváth, Péter; Pütter, Judith M.; Dagilienė, Lina; Dimante, Dzineta; Haldma, Toomas; Kochalski, Cezary; Král, Bohumil; Labaš, Davor; Lääts, Kertu; Bedenik, Nidžara Osmanagić; Pakšiová,	2017	Status Quo and Future Development of Sustainability Reporting in Central and Eastern Europe	Journal for East European Management Studies

Renáta; Petera, Petr; Ratajczak, Piotr; Buhovac, Adriana Rejc; Sava, Adriana; Sucală, Voicu-Ion; Tirnitz, Tamás József; Wagner, Jaroslav			
Hubbard, Graham	2009	Measuring organizational performance: beyond the triple bottom line	Business Strategy and the Environment
Isensee, Johannes; Michel, Uwe	2011	Green Controlling - Die Rolle des Controllern und aktuelle Entwicklungen in der Praxis	Controlling
Kämmeler-Burak, Andrea; Bauer, Rainer	2022	Nachhaltigkeit wird Standardaufgabe im Controlling	CONTROLLER Magazin
Klute-Wenig, Sandra; Refflinghaus, Robert	2015	Integrating sustainability aspects into an integrated management system	TQM Journal
Krause, Hans-Ulrich	2016	Controlling-Kennzahlen für ein nachhaltiges Management: Ein umfassendes Kompendium kompakt erklärter Key Performance Indicators	DeGruyter
Krivačić, Dubravka; Janković, Sandra	2021	Sustainability reporting during the pandemic: current state and expectations for the future	Journal of Accounting & Management
Kuttner, Michael	2022	Corporate Social Responsibility-Controlling: Eine instrumentelle Perspektive	Springer Nature

Lühn, Michael; Nuzum, Anne-Katrin; Petersen, Holger; Schaltegger, Stefan	2022	Controller und Nachhaltigkeit im Kontext neuer Berichterstattungspflichten	Rethinking Finance
Lux, Wilfried; Olbert-Bock, Sibylle	2016	Strategisches Controlling als Teil des Sustainability Performance Managements - auch für KMU	CONTROLLER Magazin
Michel, Uwe; Isensee, Johannes; Stehle, Alexander	2014	Sustainability Controlling: Planung, Steuerung und Kontrolle der Realisierung der Nachhaltigkeitsstrategie	Springer
Mödritscher, Gernot; Wall, Friederike	2021	Nachhaltiger Konsum und seine Verankerung im Controlling	Springer
Muller, Roger; Vesper, Mark	2020	The Current State of Nonfinancial Reporting in Switzerland and Beyond	Internal Approaches for Sustainability Assessment in the External Reporting -- A Conceptual and Empirical Analysis of DAX-Companies
Nikolic, Gordana; Vitenberg, Valentina; Karanovic, Bisera	2020	Adoption of Sustainable Accounting Practices Among Croatian SME's	Economies of the Balkan and Eastern European Countries
Noodt, Andreas; Grede, Jessica	2013	Die Welt ändert sich - die Rechnungslegung auch	Der Betrieb

Olbert-Bock, Sibylle; Bussmann, Christian; Lux, Wilfried; Garrels, Sven	2015	Nachhaltigkeitsmanagement erfolgreich im Unternehmen professionalisieren	IM+io Fachmagazin
Pandit, Ganesh M.; Rubenfield, Allen J.	2016	The Current State of Sustainability Reporting by Smaller S&P 500 Companies	CPA Journal
Prammer, Heinz Karl	2010	Wie lässt sich die operative Umweltleistung von Unternehmen messen? – Streiflichter auf ausgewählte Konzepte und Normen	Gabler
Preiß, Marlene	2019	Methoden zur Steigerung der Ressourceneffizienz und ihr Einsatz im Unternehmen	Hampp
Raschke, Christian	2021	Das F in CFO steht für Future	Creditreform
Rodriguez-Olalla, Ana; Aviles-Palacios, Carmen	2017	Integrating sustainability in organisations: an activity-based sustainability model	Sustainability
Rönz, Malte; Ryba, Michael	2012	Kennzahlen in Nachhaltigkeitsberichten	CONTROLLER Magazin
Schäfer, Henry; Kröner, Fabian; Seeberg, Barbara	2011	Zielgruppengerechte Nachhaltigkeitsberichterstattung mit Hilfe des Socially Responsible Investment (SRI)-Factsheets	Controlling
Schäffer, Utz	2022	Nachhaltigkeit und Controlling: Ein Aufruf zum Handeln	CONTROLLER Magazin
Schäffer, Utz; Weber, Jürgen	2015	Controlling im Wandel: Die Veränderung eines Berufsbilds im Spiegel der 2. WHU-Zukunftsstudie	Controlling

Schaltegger, Stefan	2020	Nachhaltigkeit als Treiber des Unternehmenserfolgs	Rethinking Finance
Schaltegger, Stefan; Zvezdov, Dimitar	2011	Konzeption und Praxis des Nachhaltigkeitscontrollings	Controlling
Schier, Michael; Bonnländer, Anton	2018	Nachhaltigkeit als Teil der Geschäftsstrategie	Die Bank
Schmidpeter, René; Günther, Edeltraud	2013	CSR — Innovationen im Rechnungswesen	Controlling & Management Review
Schmitz, Hans	2021	GRI-Nachhaltigkeitskennzahlen für das Controlling im Mittelstand	CONTROLLER Magazin
Schneider, Georg; Müllner, Thomas	2021	CSR-Berichterstattung in Europa - Quo vadis?	Der Konzern
Schubert, René; Gerhardt, Nadine	2021	Praxisanforderungen mittelständischer Unternehmen an Controlling-Nachwuchskräfte	CONTROLLER Magazin
Schulz, Thomas	2014	CFO-Agenda: Gute Gründe, Nachhaltigkeit auf die Tagesordnung zu setzen	Springer
Schulze, Mike; Thomas, Silke	2012	Strategisches Nachhaltigkeits-Controlling bei der Deutschen Telekom AG Corporate Responsibility Key Performance-Indikatoren von Mike Schulze und Silke Thomas	CONTROLLER Magazin
Schwarzmaier, Ulrich	2013	Entwicklungstendenzen des Controllings unter besonderer Berücksichtigung der Veränderungen durch die Nachhaltigkeitsdiskussion	CONTROLLER Magazin

Schwarzmaier, Ulrich	2015	Die Rolle des Controllings im Sustainable Management	Der Betriebswirt
Sneidere, Ruta; Bumane, Inga	2019	The challenges for the profession of accountant in the changing global economic environment	University of Latvia
Stehle, Alexander; Stelkens, Vera	2018	Green Controlling - eine Aufgabe für Controller und Nachhaltigkeitsexperten!	CONTROLLER Magazin
Sulaiman, Maliah; Mokhtar, Norsyahida	2012	Ensuring Sustainability: A Preliminary Study of Environmental Management Accounting in Malaysia	International Journal of Business & Management Science
Vuković, Renata; Vukić, Nikolina Markota; Sesar, Darija	2020	Non-Financial Reporting as Part of Sustainability Accounting. With the Examples of Good Practices.	Journal of Accounting & Management
Walinska, Ewa; Dobroszek, Justyna	2021	The Functional Controller for Sustainable and Value Chain Management: Fashion or Need? A Sample of Job Advertisements in the COVID-19 Period	Sustainability
Weber, Jürgen; Georg, Johannes; Janke, Robert	2010	Nachhaltigkeit: Relevant für das Controlling?	Controlling & Management
Weber, Jürgen; Schäffer, Utz	2016	Nachhaltigkeit - Modewelle oder ein neues Arbeitsfeld für Controller?	Springer
Weigand, Heiner	2020	Green Marketing - nachhaltig erfolgreich	Haufe

Wellbrock, Wanja; Ludin, Daniela; Krauter, Sina	2020	Nachhaltigkeitscontrolling: Instrumente und Kennzahlen für die strategische und operative Unternehmensführung	Springer
Wördenweber, Martin	2017	Nachhaltigkeitsmanagement	Schäffer-Poeschel

D Beitrag 3

How AI feedback supports lesson planning in vocational teacher education: A longitudinal intervention study using an analytical AI platform

Julia Pargmann^{1,*}, Florian Berding¹, Karin Rebmann², Elisabeth Riebenbauer³

¹ Universität Hamburg, Sedanstr. 19, 20146 Hamburg, Germany

² University of Oldenburg, Ammerländer Heerstraße 114-119, 26129 Oldenburg, Germany

³ University of Graz, Universitätsstraße 15, 8010 Graz, Austria

* Corresponding author; Email: julia.pargmann@uni-hamburg.de

Abstract

This study explores the use of analytical artificial intelligence (AI) to enhance lesson planning in teacher education, particularly in vocational business education. Providing detailed, individualized feedback for lesson planning is challenging for university instructors. An analytical AI platform that offers expert-level feedback on lesson plans could be a suitable supplement to traditional feedback as it balances out critical aspects of generative AI systems. In a longitudinal study with 103 participants, an experimental group received AI feedback on the lesson plans they developed throughout a semester, while a control group received human feedback. Results indicate that AI feedback can match or surpass human feedback in some dimensions, improving planning aspects like certain content-related sustainability aspects, action orientation, relatedness or even technology use, lesson structure, and learner engagement. Besides the integration of didactic planning dimensions into the lesson plans, this study looked at student motivation, self-reported AI literacy, AI attitudes, self-efficacy with digital technologies and digital competences. Results of these meta-dimensions show that the mere use of AI does not automatically enhance AI literacy, change attitudes or improve motivation. The study suggests that as a consequence, analytical AI platforms can effectively supplement traditional feedback but should be used within a holistic framework.

Keywords: lesson planning, teacher education, artificial intelligence, AI feedback

Declarations

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due to students' copyrights but are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding was received for this study.

Authors' contributions

JP conceptualized the study, analyzed the data, interpreted the data, wrote the original draft and did the review and editing. FB conceptualized the study, helped with data analysis and provided commentary and revision and supervised the study. KR provided commentary and revisions in the editing process and provided resources for the data analysis. ER provided commentary and revisions in the editing process and provided resources for the data analysis. All authors read and approved the final manuscript.

Acknowledgements

We thank all participating students and our student assistants who were involved in coding the data.

Abbreviations

AI	Artificial Intelligence
VET	Vocational Education and Training
LLM	Large Language Model
ML	Machine Learning
NLP	Natural Language Processing
MLM	Masked Language Modeling
BERT	Bidirectional Encoder Representations
GPT	Generative Pre-trained Transformer
GUI	Graphical User Interface
EDDA	Electronic Didactical Assistance
SDG	Sustainable Development Goals
ZPD	Zone of Proximal Development

1 Introduction

Lesson planning is an integral part of teacher education (Reinisch 2021). Through lesson planning, students train their didactical and methodological thinking, learn how to process the necessary content knowledge and derive decisions on technology and materials (Koehler and Mishra 2009). A variety of lesson planning models is available. No matter which model student teachers follow, they have to practice making certain decisions and look at the content and goals from different perspectives, as didactical decisions often directly impact each other. In doing so, they practice decision-making and prepare for teaching, where spontaneous didactical decisions will also be necessary. They become competent in the unification of theoretical, content-specific foundations, planning conditions/environments and didactical decisions, which is a highly complex process (Sandfuchs 2009; Riebenbauer 2021).

Given these complex considerations, the planning process often results in rather long and detailed written documents. As novices, to become better at lesson planning, students need feedback that fosters learning. Feedback in this context follows the definition given by Hattie and Timperley (2007); as such it is “information provided by an agent regarding aspects of one’s performance or understanding”. Although feedback itself is not automatically conducive to learning, it can be under the right circumstances (Shute 2008). Research by Beaumont et al. (2011), Lemmrich et al. (2024), Narciss et al. (2021) or Wisniewski et al. (2019) has shown that feedback should be provided formatively, timely and in a certain format, follow communicated criteria and provide action items for improvement. Jonsson (2013) adds that feedback needs to be deemed useful by the receiver, should be specific, detailed and individualized, be on eye-level with the students, and agrees that criteria should be communicated and explained.

In an ideal world, students would receive frequent, high-quality feedback on their decisions to develop their planning competences. However, providing this kind of detailed, forward-oriented feedback is often not possible for university teachers, i.e., due to time restrictions or large learning groups (Henderson et al. 2019). Thus, finding a way to optimize this process might promote student competence development.

There is some evidence that artificial intelligence applications might be an appropriate alternative (Hooda et al. 2022; Shaik et al. 2022; Hsia et al. 2024; Lee and Moore 2024; Sembey et al. 2024). This conclusion builds on the effects of digital feedback on learning, which was analyzed in meta analyses and systematic reviews by Maier and Klotz (2022), Mertens et al. (2022) or Sembey et al. (2024). However, the use of generative AI for feedback poses some problems for in-class use, for example a lack of reproducibility, untransparent training data, the reproduction of bias or issues with data protection and ethics (Banihashem et al. 2022). Gerlich (2025) shows that the use of generative AI tools reduces

students' critical thinking skills. Also, domain-specific knowledge is not available in the most common tools, as generative tools need to be trained specifically for this purpose (Liu et al. 2023). Furthermore, solutions are produced at random, and the user has to be able to maintain some level of quality control (Liu et al. 2023). Additionally, generative AI tools require a lot of energy (Khowaja et al. 2024). Luccioni et al. (2024) show that multi-purpose applications consume more energy than specialized models, even when the number of parameters is constricted.

Analytical AI applications, in contrast, use AI algorithms to classify uploaded data into categories and can then be linked to pre-produced feedback that is based on relevant scientific literature (Beheshti 2023; Pargmann et al. 2024b). These systems are smaller, thus run on less energy, and can be specialized on both domain and didactics. As they classify the work into an existing framework, the results can be reproduced easily, ensuring comparability in formative learning scenarios. Since students have to incorporate the feedback into their lesson plan themselves, which might be beneficial for their cognition. As a consequence, it might be useful to develop analytical AI tools that give context-sensitive, domain-specific expert feedback for the different specialized didactics.

As shown above, there is a growing amount of research on the suitability and effectiveness of generative AI tools for feedback processes, however, little is known about analytical AI tools that provide individualized expert-level feedback. To allow assumptions on the suitability of AI for educational settings, especially feedback situations, research has to go beyond generative tools as both means can supplement each other. Especially in vocational education and training (VET) with focus on business education, where a plethora of didactical principles needs to be considered, complex feedback situations could be improved through the help of AI.

In this study, the authors contribute to the closure of this gap with an analytical AI platform that analyzes lesson plans and materials and provides individualized, elaborate expert-level text as well as visual feedback. This platform is called EDDA, short for *electronic didactical assistance for lesson planning*. EDDA was developed through a cooperation between the University of Hamburg, Germany, the University of Oldenburg, Germany and the University of Graz, Austria. In a longitudinal intervention setting, students (in total N=103) planned a 90-minute lesson. In the experimental group, students (n=55) received feedback from the platform, supplemented by feedback from their university teacher if needed. The control group consisted of students (n=48) who received human feedback only. The lesson plans were collected at the beginning and at the end of the semester and then rated for their execution of the planning principles and dimensions through content analysis.

We address the following research questions:

1. How do the two groups differ in their implementation of the didactical and planning principles?

2. How does the use of the AI platform influence the students' motivation, their AI literacy, the perceived usefulness of AI and their self-efficacy with technology?

The remainder of this paper is structured as follows. First, the relevant literature is reviewed to give context on the state of research on AI in teacher education, then a short introduction into lesson planning for vocational business education is given. The theoretical part concludes with a summary of the state of research on feedback, both human and digital/AI to derive what high-quality feedback can look like. In the method section, the AI platform and the study design are introduced. After that, results are reported and discussed. After considering the limitations of the study and the platform, the paper ends with a conclusion and an outlook for further research.

2 Theoretical Contexts

2.1 AI in Teacher Education

McCarthy (2004) defines AI as the “science and engineering of making intelligent machines, especially intelligent computer programs” by using methods that are not necessarily visible to humans, i.e. neural networks. Through machine learning (ML), computers can make predictions based on likelihood (Buxmann and Schmidt 2021). Generative AI systems dominate current research, although there are alternatives: cognitive and analytical systems (Beheshti 2023). Analytical AI systems detect patterns in datasets and produce the desired output based on these patterns, i.e., they return specific feedback, detect whether criteria are fulfilled or prognose results. To function properly, they require large datasets. Cognitive systems simulate human cognitive processes to support decision-making. The benefit of analytical AI results is their reproducibility of the texts they return, which are often pre-produced. Particularly in educational sciences, where imbalanced data is omnipresent (Stütz et al. 2022), the required large datasets pose a significant limitation that needs to be considered when aiming to employ an analytical AI system. As popular AI tools often lack domain-specificity, suitable tools need to be developed to fit the specific needs of education. AI systems should thus acquire the domain-specific structures needed for professional decision-making (Pargmann et al. 2024a).

Regarding the classification of AI applications, three main types are available today, according to Luckin et al. (2016): personal tutors, intelligent support for collaborative learning and intelligent virtual reality. As the discussion on AI in education is gaining speed rapidly, this literature review focuses on its application in teacher education (for a discussion of the general opportunities in education, see for example Zhai et al. 2021). Previous studies on AI in teacher education found that future teachers are generally interested in the technology and use it during their studies (Celik et al. 2022), but some are hesitant due to a lack of skill (Kaplan-Rakowski et al. 2023). In a systematic review, Celik et al. (2022) identify three core areas where AI assists (future) teachers: lesson planning and development, implementation

of feedback for learners as well as assessment of learning conditions. Zhang and Zhang (2024) detail these and name teaching support regarding classroom management, the creation of learning environments, digital literacy, individualization to learning styles and the improvement of teacher-student relationships as the main aspects. While personalized and individualized learning opportunities pose a chance to improve learning outcomes as they support teachers in their diagnostics, adequate literacy is needed to be able to design, implement and assess the collected data (Long and Magerko 2020). In addition, this data is highly sensitive and requires special protection, i.e. through respective policies (Pargmann and Berding 2022). Regarding lesson planning, (future) teachers should be cautious about the results of generative AI and critically evaluate them, implying the necessity of critical thinking in teacher education (van den Berg and Du Plessis 2023). To keep up with the technological changes, future teachers need to develop AI literacy. Multiple competence frameworks and scopings have been developed and analyzed (Sperling et al. 2024), from a holistic approach by Long and Magerko (2020) to an integration of AI into the TPACK model (Ng et al. 2021). Long and Magerko (2020) define AI literacy as a “set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace”. Since generative AI systems have far-reaching implications for ethics, sustainability and education, it could be a sensible alternative to use analytical AI systems.

In summary, there are multiple ways in which AI can be beneficial in teacher education and for practicing teachers, primarily regarding lesson planning and feedback. In addition, teacher education should promote AI literacy through multiple dimensions: Student teachers need professional and technological skills as well as practical AI experiences in order to understand and be able to reflect the implications of the topic. However, technology should be used with caution. A common issue of AI tools is their lack of understanding the necessary know-how, or the performative aspect of a skill (Lebovitz et al. 2021). Applied to lesson planning, this implies a poor understanding of the didactical implications and opportunities. In addition, available large language models/ are weak in understanding complex context for specialized topics (Spreitzer et al. 2024). The development of lesson planning competence in student teachers is thus still crucial.

2.2 Lesson Planning

One core area where AI can support (future) teachers is lesson planning (Celik et al. 2022). Lesson planning itself can be defined as all activities that precede the lesson itself and are supposed to improve learning and teaching processes (Sandfuchs 2004). Lesson planning competence manifests itself through a person’s professional, didactical and pedagogical competence (Wernke and Zierer 2017). Especially for student teachers and teachers in training, the careful considerations of planning decisions contribute to their professional competence (Rusznyak and Walton 2011).

Two general planning models dominate the discourse in Germany: Klafki's (1963) Model of Didactical Analysis and the Berlin approach by Heimann et al. (1965). Certain regional variations were developed as well, such as the Hamburg approach (Schulz 1980). While each planning model is based on different learning theories or develops a pre-existing model further, they all (some more than others) cover the dimensions of learning conditions and environments, goals, methods and media (Schmaltz 2019). In addition to general planning models, specific models are available for specialized didactics, such as vocational business didactics. These also cover similar prerequisites as the general models (see for an overview Wilbers 2023) but extend further to consider domain-specific principles, such as action orientation. The planning process generally starts with a competence goal or learning outcome that is derived from the relevant curriculum (Heimann et al. 1965). Then, a first idea for the lesson is developed. After that, the learning group is considered to make individualized planning decisions. For this, diversity factors as well as motivational and learning preferences need to be analyzed. Now the first lesson idea can be modified to suit the learning group and the other prerequisites, such as the availability of technologies, the size of the room or other media and methods. When a comprehensive plan is developed, suitable material needs to be created or researched and adapted. Whether the person planning the lesson is a novice or an expert, they consider similar aspects, but in varying depth and with different reasoning. As they play with different ideas and consider different didactical consequences and implications, they get used to think in alternatives (Rusznayak and Walton 2011).

In addition to using a planning model, there are certain didactical planning principles that need to be considered in vocational business education: action orientation (Aebli 2011; Hacker and Sachse 2014), process orientation (Bloemen 2011; Stütz et al. 2022) and problem orientation (Piaget 1977; Dörner 1979). In addition, there are some planning dimensions that are empirically proven to foster learning, such as individualization and motivation (Deci and Ryan 1993), or basic dimensions like learner activation, lesson and content structure and cross-linked learning (Riebenbauer 2021). Furthermore, the German Federal Institute for Vocational Education (BIBB) has published a framework of complementary competences that students should develop over the course of their education (Bundesinstitut für Berufsbildung 2021). These include sustainability and digitalization, to be set for future vocational challenges.

Throughout the professionalization of student teachers, lesson plans are often the medium of choice to foster this competence development (Söll and Klusmeyer 2018; Reinisch 2021). The ability to plan lessons professionally thus develops from continuous improvement, as lesson planning is highly complex, and only little routine can be developed as learning groups change frequently. In their transformation model of lesson planning, Stender et al. (2015) allocate a special meaning to the planning

process, as it helps student teachers to develop professional competences. In consequence, Stender et al. (2015) assume that high-quality planning entails high-quality teaching.

So far, few studies are available on the consideration of didactical principles and other planning factors in vocational business education (Riebenbauer et al. 2024, 2025). Existing literature focuses on the use of planning models (Seifried 2009) and the competence development throughout the academic education (Riebenbauer 2021). A recent study by Riebenbauer et al. implies that there are certain planning dimensions that are particularly underrepresented in student teachers' lesson plans (i.e. sustainability, digitalization, certain aspects of action, process and problem orientation as well as motivation). As a consequence, the authors demand increased support and feedback for student teachers in their lesson planning efforts.

2.3 (Digital and AI) Feedback

Good feedback focuses on performance rather than the individual, is specific, based on observations, formative and delivered using neutral language (Schartel 2012; Narciss 2013). It should compare actual performance to expected levels and provide information for improvement (Geyskens et al. 2012). To be effective for the student, feedback must be noticed, accepted and understood (Underwood 2008). In addition, the qualification of the person providing feedback is an important dimension, as expert-level feedback can lead to bigger improvements than peer feedback alone (Prilop and Weber 2023). Wisniewski et al. (2019) derive that feedback needs to be informative, appropriate, timely and, if possible, self-regulated.

As digital feedback tools are becoming an increasing research interest, there is clear evidence that digital feedback has similar or even better effects on student learning than traditional feedback (van der Kleij et al. 2015; Riedel and Kleppsch 2021; Cai et al. 2023) and that elaborated digital feedback is more effective than corrective feedback (van der Kleij et al. 2015; Mertens et al. 2022; Wagner et al. 2024). The advantage of digital feedback mainly lies in the timeliness and the self-regulation; students can request it whenever necessary, not only when the university teacher offers it.

In early digital feedback applications, teachers had to anticipate mistakes and pre-program appropriate elaborate feedback. As generative artificial intelligence is becoming more widely available, an increasing number of researchers has focused on its opportunities and risks for feedback in educational contexts (Hooda et al. 2022; Shaik et al. 2022; Hsia et al. 2024; Lee and Moore 2024; Sembey et al. 2024). In traditional analogue feedback, there are three main areas for improvement regarding feedback in learning processes: consistency, effectiveness and timeliness (Williams 2024). All three might benefit from the use of AI; as consistency can be improved through communicated criteria, effectiveness through the adaptation of an empirically sound feedback structure and timeliness through either the

easier generation of feedback and thus shorter timelines or through personalized applications that provide feedback whenever the student requests it. Existing research on AI feedback mainly focuses on large language models (LLM) and generative AI (Jeon and Lee 2023; Dai et al. 2024; Jacobsen and Weber 2024; Steiss et al. 2024). Steiss et al. (2024) show that human feedback barely outperforms ChatGPT with only small differences when including the overall feedback quality, i.e. time-savings. Jacobsen and Weber (2024) demonstrate that well-prompted ChatGPT feedback is of higher quality in some dimensions than both novice and expert traditional feedback. Sembey et al. (2024) show in their meta-analysis that generally, using AI for feedback provision seems to be effective. According to Tubino and Adachi (2022), using AI in feedback processes not only reduces teacher workload but also has the potential to develop students' feedback literacy skills.

Lee and Moore (2024) point out that the role of the instructor is changing when implementing AI feedback, as the instructor now has to support students in the use of the tool, go into further details or answer questions about the feedback, make final assessment decisions and stimulate learning through the feedback.

3 Hypotheses

There is evidence that digital and AI feedback promotes learning as well as or even better than traditional feedback. Thus, we expect that the experimental group (using the AI platform) receives the same or higher ratings for their lesson plans. We also expect that students with higher initial ratings show less growth than students with a lower initial rating, as the room for development with standardized feedback might be limited. There is also evidence that technology use can promote motivation, digital and AI literacy and respective self-efficacy and attitudes. As a consequence, we expect these meta dimensions to develop positively throughout the semester, the experimental group scores higher than the control group (with traditional feedback only). Based on these theoretical and empirical findings, we derive the following hypotheses:

H1: The experimental group receives higher ratings for their lesson plans.

H2: Students with a higher initial rating generally show less growth than students with a lower initial rating.

H3: The meta dimensions develop positively throughout the semester. The experimental group scores higher than the control group.

4 Method

4.1 The AI Platform

Following the implications from Chapter 2, we developed an analytical AI platform called EDDA based on unsupervised machine learning that can analyze lesson plans and give individualized expert feedback. The platform uses a BERT (bidirectional encoder representations) transformer model as means for natural language processing (NLP) to analyze the textual representation of lesson plans. The BERT model uses masked language modeling (MLM) to identify missing words in context (Tunstall et al. 2022). Although GPT (generative pre-trained transformer) models are better at generating texts, BERT models can use better context and use less resources (Tunstall et al. 2022). We trained our transformer model with the German BERT base (Chan et al. 2020). In the application, the transformer heads solve classification tasks based on the data set they were trained with. The platform provides feedback on ten main and 57 sub-categories – or classes – in total. It was trained through pairwise coding with $n=1,304$ lesson plans and materials from three different universities in Germany and Austria as well as 4,679 coded textbook tasks. The transformer then classifies the uploaded plan and/or material and gives feedback and recommendations in a graphical user interface (GUI). Figure 1 gives an overview of this process:

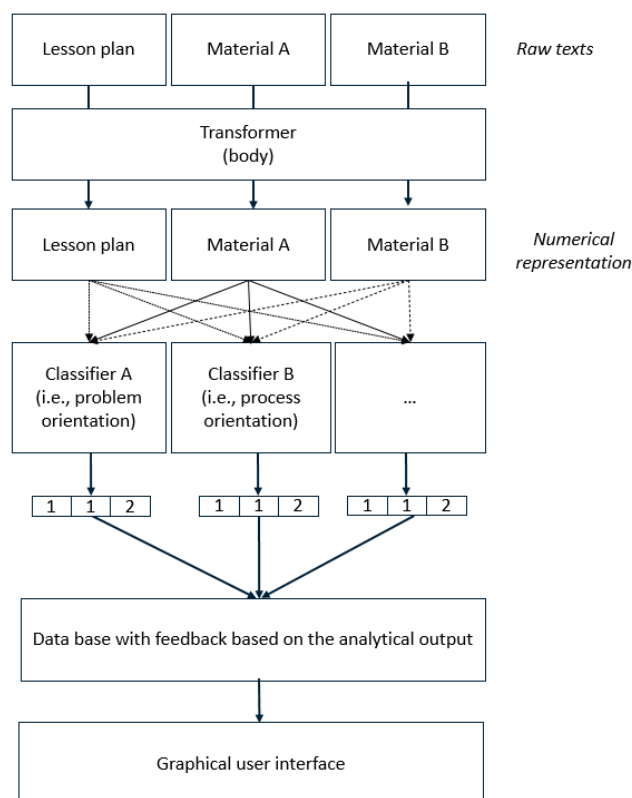


Figure 1. Schematic image of EDDA's analytical process

The reliability of the training data (Appendix A) was calculated using Percentage Agreement, Krippendorff's α , Gwet's AC coefficient (Gwet 2001) and the Iota reliability concept (Berding and Pargmann 2022). As discussed in Riebenbauer et al. (2025), the reliability of most categories is satisfactory when

judged across the different coefficients. Only three categories do not provide sufficient reliability (process orientation, plan and multiple solution strategies, see 4.4.1 for a definition). These results impact the present study in a sense that it poses limitations on the recommendations that students receive from EDDA, as it increases the potential degree of error.

4.2 Study Design

The base of the study are lesson plans and accompanying materials that student teachers developed in a master's didactics class in small groups. The overall design of the study is presented in Figure 2.

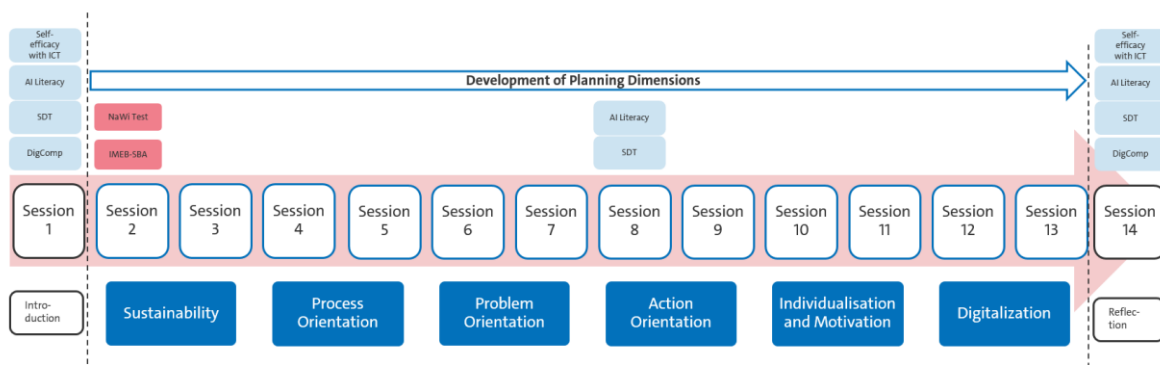


Figure 2. Study Design

Throughout the semester, students worked on six main planning dimensions or principles: sustainability, process orientation, problem orientation, action orientation, individualization and digitalization (dark blue boxes). Each of the six main dimensions or principles was focused on for two weeks. In the first week, students would get a short theoretical input on the topic to activate prior knowledge, get to know relevant research and discuss its relevance to business education. Then, they would start their design decisions and build a base for the upcoming session where they would incorporate the principle or dimension into their lesson plan and use the AI platform to receive feedback. This structure continued until all dimensions and principles were incorporated into the same lesson plan.

Connecting to Section 2.2, the goal was not to simulate lesson planning under realistic conditions (where not every single dimension or principle would be incorporated into the same lesson), but rather to allow students room for multidimensional contemplation and critical consideration of different planning options and their respective didactical and practical consequences. To allow for enough room for development and creativity, the general planning conditions were rather broad:

- 90-minute plan: The majority of business education classes in Germany are 90 minutes long. Students thus need to learn how to structure their plans accordingly. No phasing model was given.
- Focus on accounting or controlling: These areas of business education often lack didactical innovation and action orientation but are essential for the understanding of business (Chiapello 2009).

Students were free to decide which level of controlling they chose (operative or strategic) or if they wanted to focus on internal or external accounting.

- Compulsory integration of sustainability: German VET frameworks urge educators for a holistic integration of sustainability (Bundesinstitut für Berufsbildung 2021).
- The use of a model company to situate the lesson: To allow for process orientation and foster situated learning, the students had to integrate a model company.
- Choice between eight authentic learning groups to allow for individualized decisions: Didactical considerations should always be focused on the learner. All considerations regarding motivation and individualization would be void without a learner in mind.

Both the experimental and the control groups received the same theoretical input from the same lecturer, completed the same preparatory assignments and planned lessons within the same general planning conditions. The only difference between the two groups was the feedback medium. While the control group received human feedback throughout the semester, the experimental group consulted the AI platform first and then requested human feedback to discuss the AI.

The light blue boxes in Figure 2 show the study for the meta dimensions which aimed to capture developments in the students' AI literacy three times throughout the semester (Session 1, 8 and 14) and their digital competences twice (Session 1 and 14). In addition, we controlled for their prior knowledge on sustainability as well as their attitudes and epistemic beliefs towards sustainability (red boxes in Figure 2) to differentiate their effects from the intervention effects.

4.3 Description of the sample

Regarding grouping, the study followed a two-group structure with an experimental and a control group. The first cycle of the study took place during the winter semester of 2023/2024, a second cycle took place during the winter semester of 2024/2025 to recruit more participants for the experimental group. Assignment to the experimental groups was not randomized in a traditional sense as the study was integrated into a mandatory masters' module. The control group was recruited through an elective masters' module at a different university. There were no known contact points between the two groups.

Data was collected in Germany from October of 2023 until February of 2025. We draw on a sample of $N=103$ students enrolled in a master's program for business education, with $n=55$ students in the experimental group (58.2% female, 41.8% male, no diverse) and $n=48$ students in the control group (47.9% female, 52.1% male, no diverse). The lesson plans were developed in small groups. The experimental group consists of 15 groups with a minimum of three members and a maximum of five, the control group consists of 13 groups with a minimum of three members and a maximum of four. Age

wise, the median in the experimental group is at 28.53 years ($SD=4,56$), in the control group it is at 26,46 years ($SD=2,73$). In the experimental group, 85.5% of participants have a completed apprenticeship or equivalent practical experience, in the control group this number is at 85.4%.

Participation in the study for the meta dimensions was voluntary; the lesson planning process was mandatory content of the seminar. Students were informed about the study and gave consent to the use of the lesson plans. They did not receive grades for their plans as those were part of their active participation and they could drop their participation without any consequences. Due to their course of study, some prior knowledge on lesson planning could be expected.

4.4 Measures

4.4.1 Measures for the assessment of the lesson plans

The analysis of the lesson plans has two dimensions: First, the lesson plans were coded for the implementation of didactical criteria and planning dimensions. Second, students' prior knowledge on sustainability and epistemic beliefs on sustainable business were recorded to differentiate between feedback effects and the effects of prior knowledge and attitudes. Thus, these measures are reported with the assessment of the lesson plans and not with the meta dimensions.

The coding manual for the lesson plans

The coding manual was developed for the training process of EDDA (see Section 4.1). It consists of the following 10 main and 57 subcategories that encompass didactical criteria and planning dimensions for lesson planning in vocational business education:

- action orientation: the lesson promotes the ability to act through 1) authentic situations (Brown et al. 1989) (4 subcategories) and 2) by fulfilling the complete action cycle (Hacker and Sachse 2014) (6 subcategories)
- process orientation: 3) connecting the content knowledge of different business areas and activities within these business processes (Stütz et al. 2022) (4 subcategories)
- problem orientation: 4) the creation or adaptation of cognitive schemes (Piaget 1977; Dörner 1979) (1 subcategories)
- motivation: 5) stimulation of the three dimensions for self-determined motivation (competence, relatedness, autonomy) (Deci and Ryan 2000) (5 subcategories)
- digitalization: the inclusion of 6) content-related (Krämer et al. 2017) and 7) methodological digitalization (i.e., Bernard et al. 2009; Tamim et al. 2011) (6 and 4 subcategories, respectively)

- sustainability: 8) the promotion of competences for sustainable change agents (Berding et al. 2020) and 9) the inclusion of the sustainable development goals (SDG) (United Nations 2015) (7 and 17 subcategories, respectively)
- basic dimensions: 10) learner activation (students actively engage with the topic), lesson and content structure (lesson phasing and structuring of the teaching process) as well as cross-linked learning (the development of connections between topics and contents) (Riebenbauer 2021) (3 subcategories)

A more detailed discussion and evaluation on these (sub-)categories can be found in Riebenbauer et al. (2024), Pargmann et al. (2024a) and Riebenbauer et al. (2025).

Sustainability knowledge

The NAWI test by Berding et al. (2019b) was used to operationalize prior knowledge on sustainability to differentiate those effects from the intervention effects. The NAWI test is a single choice test with 15 items on knowledge about sustainable management. Students have to answer these knowledge-based questions on sustainability and sustainable management by selecting one of three answer options. They are rated on the total of correct answers. The following Table 1 shows some sample questions.

Table 1. Sample questions for the NAWI test

Scale	Sample questions
NAWI test	“How can the relationship between economy, human and nature be described in sustainable management?” ☐ Nature must be subordinated to human interests. The aim is to use the available resources as efficiently as possible. ☐ Nature is worth protecting. Economic activities must therefore not use natural resources to produce goods or provide services. ☐ Nature provides the resources for economic activities. At the same time, it is a habitat that is worth protecting.

Epistemic beliefs towards sustainability

We used the IMEB-SBA, the ‘Instrument for Measuring Epistemic Beliefs in Sustainable Business Administration’ by Berding et al. (2019a) to operationalize these beliefs in order to differentiate their effects from the intervention effects. The IMEB-SBA consists of 35 items in five scales: a) structure, b) source, c) practicability, d) security and e) justification of knowledge on sustainability in business. Sample items are shown in Table 2. The IMEB-SA uses a 7-point Likert scale (0-6) with 0 meaning “not true at all” and 6 meaning “very true”. The framework was applied without adaptations.

Table 2. Sample items for the IMEB-SBA

Scale	Sample items
Structure	"I am convinced that sustainable business requires the integration of economic, social and ecological goals"
Source	"I am convinced that the right measures for sustainable HR management can only be decided by experts"
Practicability	"I am convinced that the knowledge about the reconciliation of family and work can help companies overcome current and future challenges"
Security	"I am convinced that we already know everything about sustainable procurement"
Justification	"I am convinced that statements about sustainable logistics can only be seen as knowledge when they are verified and justified"

4.4.2 Measures for the meta dimensions

The study was framed with a questionnaire-based study (see blue boxes in Figure 2). The aim was to measure certain meta dimensions of the intervention: the development of the students' AI literacy, their attitudes towards AI, their digital competences, their self-efficacy with digital technologies and their motivation. These are not dimensions of the lesson plan itself but rather dimensions that we aimed to impact through the use of the AI platform.

AI Literacy

AI Literacy was operationalized through the Meta AI Literacy Scale (MAILS) by Carolus et al. (2023). The scale consists of 72 items and measures: a) AI literacy through the facets of *Use & apply AI*, *Understand AI*, *Detect AI*, and *AI Ethics*, b) the ability of AI creation in *Create AI*, c) AI self-efficacy through the faces *problem solving* and *learning* and d) AI Self-competency through *persuasion literacy* and *emotion regulation*. As the creation of AI was not part of this study, we only used the dimensions of AI literacy (a), AI self-efficacy (c) and AI self-competency (d). Sample items are given in Table 3. The first and the last dimensions are rated on an 11-point Likert scale (0 to 10), 0 meaning the ability is hardly or not at all pronounced, 10 meaning the ability is very well pronounced. The second dimension is rated on a 5-point Likert scale. For consistency reasons, we decided to change this to the same 11-point Likert scale.

Table 3. Sample items for AI literacy

Scale	Sample items
AI literacy	"I can operate AI applications in everyday life"
AI self-efficacy	"I can handle most problems in dealing with artificial intelligence well on my own"
AI self-competency	"I realise if artificial intelligence is influencing me in my everyday decisions"

AI Attitude

To measure AI attitude, we combined four scales with 14 items on attitudes towards digital technologies and technology acceptance and applied them to artificial intelligence platforms: a) Perceived usefulness, b) Perceived ease of use and c) Job relevance by Venkatesh and Davis (2000) and d) Attitude toward using technology by Venkatesh et al. (2003). All items were measured on a 7-point Likert scale (1=strongly disagree and 7=strongly agree). Sample items from the original scales are shown in Table 4.

Table 4. Sample items for AI attitude

Scale	Sample items
Perceived usefulness	"Using the system improves my performance in my job"
Perceived ease of use	"I find it easy to get the system to do what I want it to do"
Job relevance	"In my job, usage of the system is relevant"
Attitude towards using technology	"Working with the system is fun"

Digital Competences

The digital competences were operationalized through the DigKomp2.2.de framework by Krempkow (2022). We chose this scale because it uses the digital competence framework from the European Union. It is based on self-reports on skills and use of digital media and technologies. Five facets with a total of 15 closed and one open item are part of the scale: a) Data processing and evaluation, b) Communication and cooperation, c) Content creation, d) Safety and e) Problem solving. All facets are rated on a 5-point Likert scale from "very well" to "not at all". Sample items are shown in Table 5. We used this framework without adaptations.

Table 5. Sample items for digital competences

Scale	Sample items
Data processing and evaluation	"I can use criteria to explain the reliability of information from the internet"
Communication and cooperation	"I can decide which tool is useful to work collaboratively"
Content creation	"I know how to use licenses and copyrights"
Safety	"I know how to make the right security settings on my devices"
Problem solving	"I can find the right application for my problem"

SWIT – Self-efficacy of teachers regarding the integration of digital technologies

The SWIT framework by Doll and Meyer (2021) measures the self-efficacy of teachers on the integration of digital technologies into their classes. It uses the following four subscales with a total of 12 items on a 6-point Likert scale (1=not convinced at all to 6=totally convinced): a) Knowledge and learning, b) Technical knowledge, c) Digital diagnostics and d) Digital teaching. Items are phrased similarly across

all four subscales, following the structure of “How convinced are you, that...”. Sample items are shown in Table 6. We used this framework without adaptations.

Table 6. Sample items for SWIT

Scale	Sample items
Knowledge and learning	“... you have the necessary knowledge to judge the quality of a software product for teaching and learning?”
Technical knowledge	“... you have the necessary skills to produce digital learning materials yourself that fit your teaching goals?”
Digital diagnostics	“... can assess plagiarism that was caused by the use of digital technologies?”
Digital teaching	“... you can make suitable adaptations on the contents of your subject?”

Basic needs satisfaction at work (motivation)

To measure motivation through the self-determination theory, we used the German translation of the Basic Need Satisfaction at Work scale (Deci and Ryan 2000; Gagné 2003). It contains 21 items across the three subscales autonomy, competence and relatedness and is rated on a 7-point Likert scale (1=not true at all, 7=very true). We adapted the wording to reflect the seminar setting and changed terms from “coworkers” to “classmates” and “work” to “seminar”. Sample items from the original scale are shown in Table 7.

Table 7. Sample items for basic needs satisfaction at work

Scale	Sample items
Autonomy	“I can freely express my ideas and opinions at work”
Competence	“I do not feel like I do good work”
Relatedness	“I like my coworkers”

4.5 Data Analysis

4.5.1 Measurement model for the lesson plans

To analyze the development of the planning dimensions and principles, we first coded all lesson plans and teaching materials for all categories with the coding manual described in Section 4.4.1 (frequencies of coded dimensions are available in Appendix E). Mplus software (version 8.11) was used to analyze the data. The model was estimated using the BAYES estimator for a two-level probit regression. Because the lesson plans were created in groups, the reliability of the observation is impacted. Thus, a two-level regression is needed (Hox and van de Schoot 2013). One level is the individual level vs. the group level, the other is the control group vs. the experimental group. Although the BAYES estimator uses

standardized values, it often distorts the results with increasing missing cases (Heo et al. 2024). The main reason to use BAYES in this study was the small sample size, as BAYES is more reliable in estimating complex models in these settings. Additionally, BAYES estimation is robust against non-normality of the data (Hox and van de Schoot 2013). While we chose the BAYES estimator due to the small sample size, we also calculated the model fits with a maximum likelihood estimator, as these estimators handle a larger share of missing data better than the BAYES estimator, although this mostly applies to larger samples. The algorithm was programmed so that only scales that fit both models were used for the final analysis. The maximum likelihood model was calculated and optimized to reach scalar measurement invariance, as suggested in the literature to be optimal (Schmitt and Kuljanin 2008). Common fit indices such as the CFI or RMSEA were not used to assess the fit, as they are not as suitable for smaller samples (Depaoli et al. 2024). However, to allow for statistical re-evaluation with new methods in the future, CFI and RMSEA are still reported in Appendix B and C. For the final measurement model we employed a BAYES estimator. Here, the model fit was assessed by the posterior predictive p-value (PPP), as this is a good indicator for model fit with BAYES (Depaoli et al. 2024). While generally, a PPP value around 0.5 indicates an excellent fit, Cain and Zhang (2019) discuss 0.05 for bigger samples and 0.15 for smaller samples. Due to the sample size, this study uses a cut-off point of 0.15. Most PPP values indicate a good fit of the model (see the Results section), meeting the 0.5 benchmark. No category falls under the cut-off of 0.15, thus, all dimensions are included in the results.

To detect whether the participants' prior sustainability knowledge and attitudes (IMEB-SBA) had an effect on their planning performance, we included these variables in the evaluation of the lesson plans. We first calculated the reliability of all dimensions with Cronbach's alpha. To optimize the values, we removed items from the calculations. These optimized values are shown in Table 8. Generally, Cronbach's Alpha values are deemed sufficient from 0.7. All dimensions except practicability exceeded this value. Since some literature suggests that values from 0.6 are sufficient (Taber 2018), the scale was included in the analysis. As the NAWI test result is based on points awarded for correct answers, no reliability measure is calculated.

Table 8. Cronbach's alpha values for the IMEB-SBA

Scale	Optimized Cronbach's Alpha Values
Structure	0.83
Source	0.89
Practicability	0.69
Security	0.82
Justification	0.86

Note: Cronbach's alpha values are satisfactory from 0.7

To estimate the influence of the sustainability attitudes on the sustainability-related planning dimensions, a measurement model was calculated based on the maximum likelihood estimator. Table 9 shows the model fit. More detailed information on the fit can be found in Appendix D. Due to the small sample size, CFI and RMSEA are not considered to be suitable indices for this study. Thus, Chi Square is used to derive usable scales. Only the dimensions Structure and Practicability show a non-significant Chi Square value. Thus, only those dimensions are used in the final model.

Table 9. Model fit for IMEB-SBA with maximum likelihood estimator

Scale	n	Chi ²	p	CFI
Security	94	37.13*	0.00	0.73
Structure	94	13.59	0.26	0.97
Justification	94	42.12*	0.00	0.71
Source	94	23.11*	0.02	0.92
Practicability	94	7.77	0.17	0.91

Note: * indicates that $p < 0.05$

4.5.2 Measurement model for the meta dimensions

Next, the measurement model for the meta dimensions was developed. The meta dimensions and their development across the semester were calculated with a latent growth model (Muthén and Curran 1997). Using the latent growth model has several advantages: changes across time can be analyzed more easily and individual differences as well as latent variables can be considered. In addition, latent growth models handle missing data well. This makes them superior for this application than other methods, i.e., an ANOVA (Fan 2003). Then, the measurement model was validated for both time and group. To develop a measurement model that fits the small sample size, we needed to reduce the number of parameters. To reach scalar invariance across all scales, the model needed to be optimized. Table 10 shows the alpha values for the two survey periods for both the normal and the optimized model. More detailed data on the optimization is available in Appendix D. Table 11 shows the remaining scales that were suitable for the latent growth model.

Table 10. Cronbach's alpha values for the meta dimensions

	t0 normal	t0 optimized	t1 normal	t1 optimized
<i>SWIT</i>				
Knowledge on teaching	0.79	0.79	0.78	0.78
Technical knowledge	0.65	0.65	0.78	0.78
Digital diagnostics	0.66	0.66	0.69	0.69
Digital teaching	0.65	0.65	0.69	0.69
<i>Digital competences</i>				
Data processing and evaluation	0.69	0.69	0.80	0.80

Communication and cooperation	0.74	0.74	0.75	0.75
Content creation	0.80	0.80	0.85	0.85
Safety	0.85	0.85	0.86	0.86
Problem solving	0.72	0.72	0.75	0.75
<i>AI literacy</i>				
Use and apply AI	0.94	0.95	0.96	0.95
Understand AI	0.82	0.79	0.91	0.90
Detect AI	0.68	0.68	0.85	0.85
AI ethics	0.75	0.75	0.89	0.89
AI problem solving	0.67	0.67	0.91	0.91
AI learning	0.69	0.69	0.97	0.97
AI persuasion literacy	0.38	0.38	0.85	0.85
AI emotion regulation	0.79	0.79	0.86	0.86
<i>AI attitude</i>				
Perceived usefulness	0.93	0.93	0.98	0.98
Perceived ease of use	0.92	0.90	0.97	0.96
Attitude towards using technology	0.86	0.86	0.93	0.93
Job relevance	0.83	0.83	0.87	0.87
<i>Self-determination framework</i>				
Competence	0.41	0.51	0.68	0.70
Autonomy	0.60	0.72	0.72	0.78
Relatedness	0.79	0.77	0.77	0.89

Note: normal=with all items, optimized=optimized model fit with reduced items.

Table 11. Final scales for the final model with the BAYES estimator

Scale	n	CFI	RMSEA	Chi ²	PPP
Digital teaching	102	0.95	0.06	-18.24	0.23
Data processing and evaluation	102	1	0	-26.22	0.43
Safety	102	0.97	0.07	-16.20	0.18
Detect AI	102	0.98	0.05	-18.75	0.23
AI emotion regulation	101	1	0	-24.66	0.38
Job relevance	102	1	0	-19.7	0.44
Autonomy	101	1	0	-41.08	0.52

5 Results

In the following, the results are presented. First, we present the results for the didactical criteria and planning dimensions which we derived in 4.4.1, including the effect of the IMEB-SBA and NAWI test for the sustainability dimension. Then, the results for the meta dimensions are presented.

5.1 Results of the lesson plans

Sustainability

The first planning dimension is the implementation of the sustainable development goals and the competences for sustainable change agents. Table 12 shows the R^2 values and the p-values for prior knowledge (base), group and gender as well as sustainability knowledge (NAWI) and the dimensions of the IMEB-SBA that reached sufficient reliability to check for effects of content-specific prior knowledge. The differences between the experimental group and the control group can be seen in the column *group*. Here, a positive significant value indicates that the experimental group received higher ratings than the control group. The non-standardized probit regression results for all dimensions are available in Appendix F.

Table 12. R-squared and Standardized probit regression results for sustainability

Variable	PPP	R ²	Base	Group	Gender	NAWI	Structure	Practi- cability
<i>Sustainable Development Goals</i>								
SDG 2	0.521	0.665	-0.067	-0.057	0.065	0.087	0.090	-0.353
SDG 7	0.480	0.769	-0.075	0.402	-0.023	-0.218	0.179	0.109
SDG 8	0.555	0.780	-0.614*	-0.088	-0.003	0.147	-0.287	-0.113
SDG 9	0.458	0.739	-0.315	0.103	0.438	-0.054	0.274	-0.273
SDG 11	0.522	0.757	-0.052	0.377	0.170	0.050	0.038	-0.351
SDG 12	0.455	0.800	-0.556*	-0.181	0.209	0.075	-0.235	-0.025
SDG 13	0.497	0.835	-0.589*	-0.202	-0.037	0.170**	0.019	-0.437*
SDG 14	0.508	0.762	-0.012	0.388	0.177	0.051	0.031	-0.345
SDG 15	0.459	0.782	-0.056	0.447**	-0.155	-0.235**	-0.032	0.451**
<i>Sustainable Change Agent Competences</i>								
Content knowledge	0.347	0.923	-0.864*	-0.023	0.040	0.046	-0.063	0.026
Network knowledge	0.511	0.881	-0.781*	0.006	-0.177	-0.110	-0.019	0.309**
Methodological competences	0.950	0.959	-0.926*	0.057	0.026	0.023	0.014	0.023
Social competences	0.491	0.906	-0.837*	0.449*	-0.001	0.093	0.007	-0.222
Decision-making	0.419	0.898	-0.816*	0.046	0.087	0.099	-0.144	-0.030
Open and critical perspective	0.557	0.920	-0.797*	0.157	-0.125	-0.134**	0.132	0.247**
Reflexivity	0.504	0.818	-0.634*	0.219**	0.144	-0.100	0.335**	-0.125

Note: $N=103$, $*=p<0.05$; $**=p<0.10$. PPP cut off is at 0.15; No changes for SDG 1, 3, 4, 5, 6, 10, 16 and 17. Base = improvement from starting level.

The R^2 values of all variables in this dimension are fairly high, indicating that the model explains the variables well. This holds true when drawing on the PPP values, which are all sufficient. For the SDGs, the starting values have a significant effect ($p<0.05$) on the development of SDG 8, 12 and 13, but not on the other SDGs. The intervention has a significant positive effect ($p<0.10$) on the implementation of SDG 15, *life on land*. There are no significant gender effects. The prior knowledge on sustainability

(NAWI) has a significant positive effect ($p < 0.10$) on SDG 13 and a significant negative effect ($p < 0.10$) on SDG 15. This indicates that students who scored higher points in the NAWI test showed higher ratings regarding SDG 13, but lower ratings regarding SDG 15. The attitude of practicability towards sustainability has a significant negative effect ($p < 0.05$) on the implementation of SDG 13, but a significant positive effect ($p < 0.10$) on SDG 15. This indicates that students who think that sustainability knowledge has practical use tend to integrate SDG 15 better, those who find it hard to apply struggle with the integration of SDG 13 into their lesson plans.

There are significant effects for some of the *competences for sustainable change agents*; the development based on prior knowledge is significant ($p < 0.05$) on all dimensions, indicating that students with a high starting point showed less growth across the semester than students with a lower starting point. The intervention has a significant positive effect ($p < 0.05$) regarding the promotion of *social competences for sustainable development*, i.e. through the promotion of communicative strategies to overcome conflict in the context of a sustainability-focused company setting. The results for *social competences* indicate that students in the experimental group received higher ratings on their implementation of the category than the control group. The same is true for *reflexivity*. Here, the intervention also showed a significant positive effect ($p < 0.10$). Thus, students in the experimental group received higher ratings for their integration of opportunities to reflect on sustainability.

Prior knowledge on sustainability (NAWI) is significant ($p < 0.10$) in the dimension of an *open and critical perspective*, i.e. when established perspectives are criticized or deconstructed. Students who scored higher on the NAWI-test did not integrate an open and critical perspective as frequently as students with lower ratings did. Gender has no significant effect on the development. Regarding the attitudes towards sustainable business, *structure* has a significant positive effect ($p < 0.10$) on reflexivity, indicating that students who conceptualize sustainability knowledge as a network integrate the dimension better. *Practicability* has a positive effect ($p < 0.10$) on both *network knowledge* and *open critical perspective*. Thus, students who see practical use for their sustainability knowledge implement the dimension better than the others.

Action orientation, process orientation and problem orientation

Table 13 shows the regression results for the categories that are connected to authentic learning scenarios: *action orientation*, *process orientation* and *problem orientation*. The R^2 values indicate a good explanation of variances through the regression for most categories, except for the category *financial streams*. All results for the improvement from the starting values (base) except for *financial streams* are significant. This indicates that students with an already high starting value show less improvement throughout the semester. The intervention has significant effects on the categories *inform*, *plan*, *decide*

and *check* as well as for *information streams*. This result indicates that students from the experimental group slightly outperformed the control group. Gender again has no significant effect on the development.

Table 13. R-squared and probit regression results for action orientation, process orientation and problem orientation

Variable	PPP	R ²	Base	Group	Gender
<i>Action orientation</i>					
Mental planning	0.271	0.80	-0.576*	0.381	0.004
Actual execution	0.412	0.82	-0.900*	0.137	0.016
Kind of action	0.358	0.69	-0.763*	0.037	0.184
Inform	0.47	0.85	-0.772*	0.328*	0.046
Plan	0.478	0.79	-0.755*	0.271*	0.037
Decide	0.371	0.81	-0.621*	0.391*	0.11
Act	0.437	0.80	-0.877*	-0.053	0.033
Check	0.463	0.84	-0.809*	0.384*	-0.165
Reflect	0.533	0.60	-0.702*	0.17	0.134
Context	0.491	0.65	-0.751*	0.055	0.015
<i>Process orientation</i>					
Process orientation	0.355	0.62	-0.669*	-0.144	0.319
Value streams	0.452	0.77	-0.718*	-0.249	0.205
Financial streams	0.545	0.21	-0.213	0.258	0.13
Information streams	0.506	0.69	-0.557*	0.452*	-0.123
<i>Problem orientation</i>					
Problem orientation	0.497	0.82	-0.89*	-0.128	0.156

Note: $N=103$, $*=p<0.05$; $**=p<0.10$. PPP cut off is at 0.15; Base = improvement from starting level.

Motivation

Table 14 shows the results for the dimension of motivation. The R-squared values are sufficient. Regarding the promotion of motivation in the lesson plans, the model explains the variances fairly well. Students who had high starting values again improved not as much as students with lower starting values, as all estimates for base are significant. The intervention has a significant positive effect on the dimensions of competence and relatedness. Gender has no significant effect on the development.

Table 14. R-squared and probit regression results for motivation

Variable	PPP	R ²	Base	Group	Gender
Contains problems	0.523	0.743	-0.738*	0.174	0.076
Multiple solution strategies	0.521	0.778	-0.794*	0.142	0.201
Multiple solutions possible	0.519	0.684	-0.724*	0.112	0.053

Competence	0.519	0.636	-0.643*	0.433*	0.085
Relatedness	0.536	0.599	-0.620*	0.414*	0.202

Note: $N=103$, $*$ = $p<0.05$; $**$ = $p<0.10$. PPP cut off is at 0.15; Base = improvement from starting level.

Digitalization

Table 15 shows the results for the dimension of digitalization. The R-squared values indicate that for methodological digitalization, the results are explained well. For the scale of content-related digitalization, the values are lower, except for the category condition. Thus, there could be other factors explaining the variances that are not present in the model. Since the PPP-values are around 0.5 for these dimensions, we still assume a good enough fit.

Table 15. R-squared and probit regression results for digitalization

Variable	PPP	R ²	Base	Group	Gender
<i>Methodological digitalization</i>					
Technology use	0.523	0.890	-0.954*	0.269*	0.060
Engagement	0.539	0.618	-0.696*	0.176	0.180
<i>Content-related digitalization</i>					
General technologies	0.513	0.375	-0.493*	-0.043	-0.138
Specific technologies	0.524	0.250	-0.177	-0.071	-0.282
Verbalization	0.506	0.321	-0.415*	-0.115	-0.112
Topic	0.498	0.267	-0.371**	0.113	-0.034
Condition	0.499	0.602	-0.059	-0.537**	-0.370**

Note: $N=103$, $*$ = $p<0.05$; $**$ = $p<0.10$. PPP cut off is at 0.15; Base = improvement from starting level. No changes for cooperation/collaboration, feedback and development.

On the scale of methodological digitalization, starting values for both *technology use* and *technology engagement* have a significant effect on the improvement of students. As in the previous categories, students with high starting values do not improve as much as those with lower starting values. This is also true for the categories *general technologies*, *verbalization* and *topic* from the content-related digitalization scale. The intervention has a significant positive effect on the implementation of *technology use* as students who used the AI platform reached higher final values. There are no other significant positive effects for this dimension. However, students from the control group slightly outperformed the experimental group in the category of *condition*, where the implementation of the conditions for technology use is coded. Gender has a small but significant negative effect ($p<0.10$) on the category *condition*, thus male students implemented the category better than female students. There are no gender effects on the other dimensions.

Basic dimensions

Table 16 shows the results for the basic planning dimensions. The R-squared values indicate a good explanation of the variances. The regression results for the improvement from high starting values are all significant and indicate smaller improvements for high starting values. The intervention has significant positive effects on all three dimensions of basic lesson planning. There is also a small significant effect ($p < 0.10$) of gender on *learner activation*, indicating that female students integrated this dimension better into their lesson plans than male students.

Table 16. R-squared and probit regression results for the basic dimensions

Variable	PPP	R ²	Base	Group	Gender
Lesson structure	0.537	0.767	-0.732*	0.553*	0.006
Learner activation	0.501	0.614	-0.531*	0.520*	0.314**
Cross-linked learning	0.539	0.581	-0.525*	0.607*	0.103

Note: $N=103$, $*$ = $p < 0.05$; $**$ = $p < 0.10$. PPP cut off is at 0.15; Base = improvement from starting level.

Application of the results

As a consequence of these results, we reject hypothesis 1, as the experimental group sometimes receives higher ratings for their lesson plans, but not always. The experimental group performs better than the control group on the dimensions of social competences in sustainability, inform, plan, decide, check, information streams, competence, relatedness, technology use, lesson structure, learner activation and cross-linked learning. However, the control group performs better for conditions of technology use. In all other categories, there are no significant differences between the groups.

We accept hypothesis 2, as students with higher starting values improve less well than students with lower starting values across all dimensions except SDGs 2, 7, 9, 11, 14, 15, financial streams and conditions of technology use.

The unstandardized results of the probit regression can be plugged into a probability calculation where the regression results are weighted and used to calculate the likelihood that a specific event becomes true. In our case, this would be the likelihood that a person with certain characteristics can improve their lesson plans using EDDA. For example, a female student with a medium high starting value in the category *technology use* has a 87% chance of improvement using EDDA. However, the same student only has a 4% chance to improve her results through the AI feedback in the category *condition*.

5.2 Results of the meta dimensions

The latent growth model for the effect of the intervention on the meta dimensions was calculated with the constructs with a good model fit from Table 11. Table 17 shows the standardized estimations for these constructs on level 3 scalar invariance. Both *data processing and evaluation* and *detect AI* fall

under the PPP cut off point of 0.15. All other dimensions can be interpreted. Non-standardized values are available in Appendix G.

Table 17. Standardized latent growth model results

	Control group			Experimental group			
Scale	Initial	Growth	Treatment	Initial	Growth	Treatment	PPP
<i>SWIT</i>							
Digital teaching	6.137*	0.139	0	5.937*	0.139	2.219	0.18
<i>DigComp</i>							
Data processing and evaluation	6.786*	0.364	0	6.736*	0.364	-1.256	0.13
Safety	5.194*	0.349	0	4.627*	0.349	0.6	0.21
<i>AI literacy</i>							
Detect AI	3.308*	0.021	0	2.809*	0.021	28.935*	0.07
AI emotion regulation	6.134*	-0.848*	0	5.817*	-0.848*	26.887*	0.35
<i>AI attitude</i>							
Job relevance	6.774*	0.246	0	6.395*	0.246	1.916	0.51
<i>Self-determination framework</i>							
Autonomy	8.462*	0.556*	0	9.143*	0.556*	-0.121	0.30

Note: *= $p < 0.05$; **= $p < 0.10$. PPP cut off is at 0.15.

Both groups show significant growth in the latent variables of AI emotion regulation and the subscale of autonomy from the self-determination framework. The direction of the effect in AI emotion regulation indicates a decrease throughout the semester, the development on autonomy however is positive. The treatment shows a significant positive effect on two subscales of AI literacy and AI emotion regulation.

Due to these results, we reject hypothesis 3. The meta dimensions do not all show positive growth throughout the semester and the experimental group does not have significantly better results than the control group. This is only true for two dimensions of the AI literacy construct, detect AI and AI emotion regulation.

6 Discussion

This study investigated how the use of an analytical AI platform for feedback influences the lesson planning process and the implementation of didactic and planning principles in a pre-post intervention setting. In addition, it analyzed how the use of the platform influences students' motivation, AI literacy, attitudes towards AI and their sense of technological self-efficacy in comparison to students who do not use the platform.

Our results for the planning process show that the intervention has a positive effect on the lesson plans in some, but not all dimensions. This indicates that both human and AI feedback contribute to student

learning, and each have a time to be given or used. These findings extend the findings by Sembey et al. (2024) or Steiss et al. (2024), that AI feedback generally is effective, but that human-in-the-loop-strategies promise even better learning effectiveness. We found that there are some dimensions, most of which are already challenging to the student teachers, that needed to be discussed further. Thus, we support the demand for AI feedback frameworks and extend this demand to analytical AI platforms. In this first study, the AI's feedback quality seems to be sufficient in some, but not all dimensions.

In a recent study, Riebenbauer et al. pointed out that lesson plans from student teachers lack the transformation of some didactical and planning principles. Especially sustainability and digitalization were integrated poorly. Our results indicate that the AI platform can improve these results for the dimensions social competence and technology use. Our results show that analytical AI feedback is a suitable supplement to human feedback in many areas of lesson planning. Since providing feedback is a time-consuming task for (university) teachers, the results support the idea to (additionally) use AI for feedback purposes.

Action orientation is a core principle for vocational business education (KMK 2021). While older evidence on action-oriented teaching methods indicate that more teacher-focused methods are preferred for reasons of efficiency and controllability, even though the dominating classroom arrangements (group work, single work, partner work) differs (Pätzold et al. 2003; Seifried et al. 2006; Götzl et al. 2013; Kozina and Pilz 2018), there is no research on the representation of the phases of complete action in VET lessons. However, the study by Riebenbauer et al. (2025) suggests that only the checking phase is integrated sufficiently in student teachers' lesson plans. Our results show that students who used EDDA, implemented the early phases of the complete action cycle better (inform, plan, decide) and also include the checking phase.

In a longitudinal study on lesson planning in accounting education, Riebenbauer (2021) showed that most students develop a good, but superficial understanding of the basic dimensions of learner activation, lesson structure and cross-linked learning. The findings from this study suggest, that the AI platform could be a suitable medium to improve this understanding, as the experimental group showed significantly better results than the control group across all three categories. This is particularly true for weaker learners or those with less prior knowledge.

A significant finding of this study is that students who already had high starting scores at the beginning showed less progress over time than their peers with lower starting scores. This finding points at the need for more differentiated and personalized education in higher education to promote competence development for all students (Felder and Brent 2005). According to the Zone of Proximal Development (ZPD) theory (Vygotsky 1981), students at the extremes of the achievement scale are in particular need

of customized and challenging learning opportunities to remain in their ideal learning zone and reach their full potential. Digital technologies are promising for this cause (Maier and Klotz 2022), thus AI platforms should allow a certain degree of personalization.

While some research suggests that men have more positive attitudes towards digital technologies and have a higher self-efficacy (Cai et al. 2017; Borokhovski et al. 2018), our findings indicate that these differences are becoming less significant, as the authors projected. We add to these findings that there are no significant differences regarding the methodological aspects of digitalization, but men do discuss the conditions for the use of technology slightly better than women.

Regarding the meta dimensions, only six scales could be included in the final model, as the others did not perform well enough in the longitudinal setting. While we used established scales to avoid this, the ones we chose seem to either not work for the longitudinal setting or for our sample. Regarding the impact of digital technologies on motivation, similarly to Rubach and Lazarides (2021), we did not find significant differences in motivation in the experimental group. Present literature shows that learning about AI in theory does not necessarily improve literacy or competences and neither does the mere use, without knowing the purpose or functionality (Hornberger et al. 2023; Schiavo et al. 2024). Thus, this study could not significantly increase the general AI skills, which would be desirable (Kaplan-Rakowski et al. 2023). As a consequence, students need to use the technology and use it productively with guidance on specific AI literacy dimensions. This study adds to these findings as our results show that students who use the AI platform also judge their abilities to detect AI better than students who do not use the platform. But the mere use of an AI platform does not improve self-efficacy with digital technologies on digital teaching, the competences in data processing and evaluation or the ability to judge a product's safety, as there are no statistical differences. Neither does it influence the perceptions of job relevance or foster autonomy. Thus, the mere use of a technology does not change perceptions, competences or is more motivating than traditional settings. Specific instruction is still necessary. Interestingly, the literacy dimension AI emotion regulation shows a significant decline throughout the semester for both groups, but the treatment has a positive effect on the experimental group. This means that the intervention could have helped to attenuate or even reverse the decline in AI emotion regulation compared to the control group. Without the intervention, the AI emotion regulation declines throughout the semester. As students from the experimental group still have to engage with the tool, it decreases this decline as students are continuously confronted with the technology and cannot necessarily remove themselves from it.

Limitations

There are some limitations to this study. Firstly, students only planned these lessons in theory, they did not get to realize this planning in practice. Thus, the critical reflection of their planning processes and decisions could not take place. Theoretical sufficiency of the categories does not necessarily imply practicability. Secondly, we did not analyze the lesson plans qualitatively, we only coded the transformation of each category. Thus, reasonings or nuances for and against the implementation of content-related aspects such as sustainability or content-related digitalization were not considered. This was not part of the study, as we focused on AI feedback. Students learned about the EDDA algorithm early on in the semester, where the platform and its strengths and weaknesses were explained to them. Since the accuracy of EDDAs recommendations is not always the same, this might have negatively affected their motivation. In addition, formulation and frequency of human feedback was not controlled in either group as it was not the goal to substitute human feedback but to supplement it.

Another limitation to this study is the sample size. Intervention studies in educational science tend to face this issue, inflating the effectiveness of interventions (Slavin and Smith 2009; Tipton et al. 2017). We tried to manage this issue with the estimators we used. There is always a risk of bias when using the BAYES estimator, however, we tried to minimize it as much as possible by double-checking the fit with the maximum likelihood estimator. Data should be re-evaluated when more advanced statistical methods are available.

Lastly, the poor model fit for the meta dimensions indicates that the scales we used should be carefully evaluated when applied in a longitudinal setting. Only few scales were useable for the final analysis. This implies that researchers performing longitudinal studies who use the same scales should cautiously consider the model fit and verify the measurement levels. If this is not done, results should be interpreted carefully. The bad model fit for most meta dimensions could be due to the block seminar setting in the control group, where students were surveyed more frequently (timing-wise) than in the experimental group. In addition, there are certain effects that could have been affecting the results of the experimental group, such as stress towards the end of the semester.

7 Conclusion

This study has shown that analytical AI feedback on lesson planning dimensions can supplement human feedback in a way that it matches or even exceeds the results of lesson plans that received only human feedback. Many planning principles were integrated significantly better in the experimental group such as technology use, lesson structure, learner activation and motivation. In line with current research, the mere use of a technology does not automatically improve the user's literacy or impede their attitudes towards it. It also is not necessarily more motivating. It helps, however, in the detection of AI platforms.

These results indicate that not only generative AI platforms are a suitable feedback provider, as these have significant disadvantages and even risks. But an analytical AI platform that provides expert-level, structured feedback with recommendations can serve as a suitable feedback option in the lesson planning process. The feedback itself, however, needs to be more individualized to avoid ceiling effects for already strong learners. Institutions interested in using such a feedback option should provide more opportunities for improvement, for example by implementing reflection questions or a personalized profile where more detailed or more critical feedback can be requested.

Future research in this direction should investigate further to which degree human feedback can be supplemented or even substituted with AI feedback to further individualize feedback processes. The more time instructors have for individualized feedback, the more the ceiling effects of the analytical AI can be reduced, as the AI feedback can serve as a starting point for more complex changes, if necessary. In addition, the lesson plans could also be coded for their goal formulation and the realization of their learning goals, as this was not part of this study but could provide helpful suggestions for extensions of the AI platform.

References

- Aebli H (2011) Zwölf Grundformen des Lehrens: Eine Allgemeine Didaktik auf psychologischer Grundlage. Medien und Inhalte didaktischer Kommunikation, der Lernzyklus, 14th edn. Klett-Cotta, Stuttgart
- Banihashem SK, Noroozi O, van Ginkel S, Macfadyen LP, Biemans H (2022) A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review* 37:100489. <https://doi.org/10.1016/j.edurev.2022.100489>
- Beaumont C, O'Doherty M, Shannon L (2011) Reconceptualising assessment feedback: a key to improving student learning? *Studies in Higher Education* 36:671–687. <https://doi.org/10.1080/03075071003731135>
- Beheshti A (2023) Empowering Generative AI with Knowledge Base 4.0: Towards Linking Analytical, Cognitive, and Generative Intelligence. In: 2023 IEEE International Conference on Web Services (ICWS). IEEE, pp 763–771. <https://doi.org/10.1109/ICWS60048.2023.00103>
- Berding F, Pargmann J (2022) Iota Reliability Concept of the Second Generation. Logos Verlag Berlin. <https://doi.org/10.30819/5581>
- Berding F, Slopinski A, Gebhardt R, Heubischl S, Rebmann K, Schlömer T (2019a) IMEB-SBA. Instrument for Measuring Epistemic Beliefs in Sustainable Business Administration [Verfahrensdokumentation und Fragebogen]. <https://doi.org/10.23668/psycharchives.4608>
- Berding F, Slopinski A, Gebhardt R, Heubischl S, Rebmann K, Schlömer T (2019b) NaWi-Test -Test über das Wissen im nachhaltigen Wirtschaften. [Verfahrensdokumentation und Fragebogen mit Auswertungshinweisen]. <https://doi.org/10.23668/psycharchives.4609>
- Berding F, Slopinski A, Frerichs R (2020) Auszubildende als zukünftige Change Agents for Sustainable Innovations. *Betriebswirtschaftliche Forschung und Praxis* 72:313–337
- Bernard RM, Abrami PC, Borokhovski E, Wade CA, Tamim RM, Surkes MA, Bethel EC (2009) A Meta-Analysis of Three Types of Interaction Treatments in Distance Education. *Review of Educational Research* 79:1243–1289
- Bloemen A (2011) Lernaufgaben in Schulbüchern der Wirtschaftslehre. Analyse, Konstruktion und Evaluation von Lernaufgaben für die Lernfelder industrieller Geschäftsprozesse. Schriften zur Berufs- und Wirtschaftspädagogik. Hampp, München
- Brown JS, Collins A, Duguid P (1989) Situated Cognition and the Culture of Learning. *Educational Researcher* 18:32. <https://doi.org/10.2307/1176008>
- Bundesinstitut für Berufsbildung (ed) (2021) Vier sind die Zukunft. Digitalisierung. Nachhaltigkeit. Recht. Sicherheit: Die modernisierten Standardberufsbildpositionen anerkannter Ausbildungsberufe. Verlag Barbara Budrich, Leverkusen

- Buxmann P, Schmidt H (2021) Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens. In: Buxmann P, Schmidt H (eds) Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg, 2nd edn. Springer, Berlin, pp 3–26
- Cai Z, Fan X, Du J (2017) Gender and attitudes toward technology use: A meta-analysis. *Computers & Education* 105:1–13. <https://doi.org/10.1016/j.compedu.2016.11.003>
- Cai Z, Gui Y, Mao P, Wang Z, Hao X, Fan X, Tai RH (2023) The effect of feedback on academic achievement in technology-rich learning environments (TREs): A meta-analytic review. *Educational Research Review* 39:100521. <https://doi.org/10.1016/j.edurev.2023.100521>
- Cain MK, Zhang Z (2019) Fit for a Bayesian: An Evaluation of PPP and DIC for Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal* 26:39–50. <https://doi.org/10.1080/10705511.2018.1490648>
- Celik I, Dindar M, Muukkonen H, Järvelä S (2022) The Promises and Challenges of Artificial Intelligence for Teachers: a Systematic Review of Research. *TechTrends* 66:616–630. <https://doi.org/10.1007/s11528-022-00715-y>
- Chan B, Schweter S, Möller T (2020) German’s Next Language Model. In: Scott D, Bel N, Zong C (eds) *Proceedings of the 28th International Conference on Computational Linguistics*. International Committee on Computational Linguistics, Stroudsburg, PA, USA, pp 6788–6796. <https://doi.org/10.18653/v1/2020.coling-main.598>
- Chiapello E (2009) Die Konstruktion der Wirtschaft durch das Rechnungswesen. In: Diaz-Bone R, Krell G (eds) *Diskurs und Ökonomie: Diskursanalytische Perspektiven auf Märkte und Organisationen*. VS, Wiesbaden, pp 125–149. https://doi.org/10.1007/978-3-531-91914-0_6
- Dai W, Tsai Y-S, Lin J, Aldino A, Jin F, Li T, Gašević D (2024) Assessing the Proficiency of Large Language Models in Automatic Feedback Generation: An Evaluation Study. *Computers and Education: Artificial Intelligence* 7:1–10. <https://doi.org/10.35542/osf.io/s7dvy>
- Deci EL, Ryan RM (1993) Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik* 39:223–238
- Deci EL, Ryan RM (2000) The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*:227–268
- Depaoli S, Winter SD, Liu H (2024) Under-Fitting and Over-Fitting: The Performance of Bayesian Model Selection and Fit Indices in SEM. *Structural Equation Modeling: A Multidisciplinary Journal* 31:604–625. <https://doi.org/10.1080/10705511.2023.2280952>
- Dörner D (1979) Problemlösen als Informationsverarbeitung, 2nd edn. Kohlhammer, Stuttgart
- Fan X (2003) Power of Latent Growth Modeling for Detecting Group Differences in Linear Growth Trajectory Parameters. *Structural Equation Modeling: A Multidisciplinary Journal* 10:380–400. https://doi.org/10.1207/S15328007SEM1003_3

- Felder RM, Brent R (2005) Understanding Student Differences. *Journal of Engineering Education* 94:57–72. <https://doi.org/10.1002/j.2168-9830.2005.tb00829.x>
- Gagné M (2003) The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation and Emotion*:199–223
- Gerlich M (2025) AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies* 15:6. <https://doi.org/10.3390/soc15010006>
- Geyskens J, Donche V, Petegem P (2012) Towards effective feedback in higher education: bridging theory and practice. *Reflecting education* 8:132–147
- Götzl M, Jahn R, Held G (2013) Bleibt alles anders!? Sozialformen, Unterrichtsphasen und echte Lernzeit im kaufmännischen Unterricht. *bwp@* 24:1–22
- Gwet K (2001) *Handbook of Inter-Rater Reliability*. STATAXIS, Gaithersburg
- Hacker W, Sachse P (2014) *Allgemeine Arbeitspsychologie: Psychische Regulation von Tätigkeiten*. Hogrefe, Göttingen
- Hattie J, Timperley H (2007) The Power of Feedback. *Review of Educational Research* 77:81–112. <https://doi.org/10.3102/003465430298487>
- Heimann P, Otto G, Schulz W (1965) *Unterricht - Analyse und Planung*. Schroedel, Hannover
- Henderson M, Ryan T, Phillips M (2019) The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education* 44:1237–1252. <https://doi.org/10.1080/02602938.2019.1599815>
- Heo I, Jia F, Depaoli S (2024) Performance of Model Fit and Selection Indices for Bayesian Piecewise Growth Modeling with Missing Data. *Structural Equation Modeling: A Multidisciplinary Journal* 31:455–476. <https://doi.org/10.1080/10705511.2023.2264514>
- Hooda M, Rana C, Dahiya O, Rizwan A, Hossain MS (2022) Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education. *Mathematical Problems in Engineering* 2022:1–19. <https://doi.org/10.1155/2022/5215722>
- Hornberger M, Bewersdorff A, Nerdel C (2023) What do university students know about Artificial Intelligence? Development and validation of an AI literacy test. *Computers and Education: Artificial Intelligence* 5:100165. <https://doi.org/10.1016/j.caeai.2023.100165>
- Hox J, van de Schoot R (2013) Robust methods for multilevel models. In: Scott MA, Simonoff JS, Marx BD (eds) *The SAGE handbook of multilevel modeling*. SAGE, Los Angeles, London, New Delhi, Singapore, Washington DC, pp 387–402
- Hsia L-H, Hwang G-J, Hwang J-P (2024) AI-facilitated reflective practice in physical education: an auto-assessment and feedback approach. *Interactive Learning Environments* 32:5267–5286. <https://doi.org/10.1080/10494820.2023.2212712>

- Jeon J, Lee S (2023) Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Educ Inf Technol* 28:15873–15892.
<https://doi.org/10.1007/s10639-023-11834-1>
- Jonsson A (2013) Facilitating productive use of feedback in higher education. *Active Learning in Higher Education* 14:63–76. <https://doi.org/10.1177/1469787412467125>
- Kaplan-Rakowski R, Grotewold K, Hartwick P, Papin K (2023) Generative AI and Teachers' Perspectives on Its Implementation in Education. *Journal of Interactive Learning Research* 34:313–338
- Khowaja SA, Khuwaja P, Dev K, Wang W, Nkenyereye L (2024) ChatGPT Needs SPADE (Sustainability, PrivAcy, Digital divide, and Ethics) Evaluation: A Review. *Cogn Comput* 16:2528–2550.
<https://doi.org/10.1007/s12559-024-10285-1>
- Klafki W (1963) *Studien zur Bildungstheorie und Didaktik*. Beltz, Weinheim
- KMK (2021) Handreichung für die Erarbeitung von Rahmenlehrplänen der Kultusministerkonferenz für den berufsbezogenen Unterricht in der Berufsschule und ihre Abstimmung mit Ausbildungsordnungen des Bundes für anerkannte Ausbildungsberufe. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2021/2021_06_17-GEP-Handreichung.pdf
- Koehler M, Mishra P (2009) What is Technological Pedagogical Content Knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education* 9:60–70
- Kozina A, Pilz M (2018) Handlungsorientierung im Wirtschaftsunterricht: Befunde zur Realisierung von Sozialformen 33:23–36
- Lebovitz S, Levina N, Lifshitz-Assa H (2021) Is AI Ground Truth Really True? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What. *MISQ* 45:1501–1526.
<https://doi.org/10.25300/MISQ/2021/16564>
- Lee SS, Moore RL (2024) Harnessing Generative AI (GenAI) for Automated Feedback in Higher Education: A Systematic Review. *Online Learning* 28:82–106
- Lemmrich S, Ehmke T, Reusser K (2024) Adaptive Lernunterstützung durch fachliche Präzision und interaktionale Qualität. *Zeitschrift für Schul- und Professionsentwicklung* 6:6–23.
<https://doi.org/10.11576/pflb-6862>
- Liu Y, Han T, Ma S, Zhang J, Yang Y, Tian J, He H, Li A, He M, Liu Z, Wu Z, Zhao L, Zhu D, Li X, Qiang N, Shen D, Liu T, Ge B (2023) Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology* 1:100017. <https://doi.org/10.1016/j.met-rad.2023.100017>
- Long D, Magerko B (2020) What is AI Literacy? Competencies and Design Considerations. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, pp 1–16. <https://doi.org/10.1145/3313831.3376727>

- Luccioni S, Jernite Y, Strubell E (2024) Power Hungry Processing: Watts Driving the Cost of AI Deployment? In: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency. Association for Computing Machinery, Erscheinungsort nicht ermittelbar, pp 85–99.
<https://doi.org/10.1145/3630106.3658542>
- Luckin R, Holmes W, Griffiths M, Forcier LB (2016) Intelligence Unleashed. An argument for AI in Education. Pearson, London
- Maier U, Klotz C (2022) Personalized feedback in digital learning environments: Classification framework and literature review. Computers and Education: Artificial Intelligence 3:100080.
<https://doi.org/10.1016/j.caeai.2022.100080>
- McCarthy J (2004) What is artificial intelligence?
- Mertens U, Finn B, Lindner MA (2022) Effects of computer-based feedback on lower- and higher-order learning outcomes: A network meta-analysis. Journal of Educational Psychology 114:1743–1772. <https://doi.org/10.1037/edu0000764>
- Muthén B, Curran P (1997) General Longitudinal Modeling of Individual Differences in Experimental Designs: A Latent Variable Framework for Analysis and Power Estimation. Psychological Methods 2:371–402
- Narciss S (2013) Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the interactive feedback model. Digital Education Review 23:7–26
- Narciss S, Hammer E, Damnik G, Kisielski K, Körndle H (2021) Promoting Prospective Teacher Competencies for Designing, Implementing, Evaluating, and Adapting Interactive Formative Feedback Strategies. Psychology Learning & Teaching 20:261–278.
<https://doi.org/10.1177/1475725720971887>
- Ng DTK, Leung JKL, Chu SKW, Qiao MS (2021) Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence 2:100041.
<https://doi.org/10.1016/j.caeai.2021.100041>
- Pargmann J, Berding F (2022) Künstliche Intelligenz in der beruflichen Bildung - Utopie oder Dystopie? Chancen und Herausforderungen von KI als Bildungstechnologie. In: Eckelt M, Ketschau TJ, Klassen J, Schauer J, Schmees JK, Steib C (eds) Berufsbildungspolitik: Strukturen - Krise - Perspektiven, 1st edn. wbv Publikation, Bielefeld, pp 215–232
- Pargmann J, Riebenbauer E, Leube A, Berding F, Slopinski A, Rebmann K (2024a) Der Einsatz von künstlicher Intelligenz zur Unterrichtsplanung in der beruflichen Bildung – exemplarische Möglichkeiten mit der elektronisch-didaktischen Assistenz EDDA. lernen&lehren 39:22–28
- Pargmann J, Leube A, Berding F, Riebenbauer E, Rebmann K, Slopinski A, Gillert M (2024b) Die elektronisch-didaktische Assistenz (EDDA) zur Unterrichtsplanung auf Basis künstlicher Intelligenz:

- Funktionsweise, Anwendungsbereiche und Forschungsperspektiven. EP 38:118–146.
<https://doi.org/10.62350/WLQO3003>
- Pätzold G, Klusmeyer J, Wingels J, Lang M (2003) Lehr-Lern-Methoden in der beruflichen Bildung: Eine empirische Untersuchung in ausgewählten Berufsfeldern. BIS, Oldenburg
- Piaget J (1977) The development of thought: Equilibration of cognitive structures. Viking
- Prilop CN, Weber KE (2023) Digital video-based peer feedback training: The effect of expert feedback on pre-service teachers' peer feedback beliefs and peer feedback quality. Teaching and Teacher Education 127:104099. <https://doi.org/10.1016/j.tate.2023.104099>
- Reinisch H (2021) Unterrichtsplanung im Kontext bildungstheoretischer Grundannahmen. In: Klusmeyer J, Söll M (eds) Unterrichtsplanung in der Wirtschaftsdidaktik: Aktuelle theorie-, empirie- und praxisbasierte Beiträge. Springer, Wiesbaden, pp 17–34
- Riebenbauer E (2021) Kompetenzentwicklung im Masterstudium Wirtschaftspädagogik. Längsschnittstudie zur Unterrichtsplanung im Rechnungswesen. wbv, Bielefeld
- Riebenbauer E, Rebmann K, Berding F, Leube A, Pargmann J, Slopinski A, Gillert M (2024) Zum Planungshandeln von Studierenden der Wirtschaftspädagogik - Teil 1: Orientierung bei der Planung mit Unterrichtsentwürfen. Zeitschrift für Berufs- und Wirtschaftspädagogik 4
- Riebenbauer E, Rebmann K, Berding F, Leube A, Pargmann J, Slopinski A, Gillert M (2025) Zum Planungshandeln von Studierenden der Wirtschaftspädagogik – Teil 2: Empirische Befunde zur Analyse von Unterrichtsentwürfen. Zeitschrift für Berufs- und Wirtschaftspädagogik 1
- Riedel J, Kleppsch J (2021) Wie bereit sind Studierende für die Nutzung von KI-Technologien? Eine Annäherung an die KI-Readiness Studierender im Kontext des Projektes „tech4comp“. In: Wollersheim H-W, Karapanos M, Pengel N (eds) Bildung in der digitalen Transformation. Waxmann, Münster, pp 283–292. <https://doi.org/10.25656/01:26644>
- Rubach C, Lazarides R (2021) Bedingungen für die Umsetzung motivationsförderlicher Unterrichtsstrategien durch digitale Medien. In: Lazarides R, Raufelder D (eds) Motivation in unterrichtlichen fachbezogenen Lehr-Lernkontexten, vol 10. Springer Fachmedien Wiesbaden, Wiesbaden, pp 427–453. https://doi.org/10.1007/978-3-658-31064-6_15
- Rusznayak L, Walton E (2011) Lesson planning guidelines for student teachers: A scaffold for the development of pedagogical content knowledge. Education as Change 15:271–285.
<https://doi.org/10.1080/16823206.2011.619141>
- Sandfuchs U (2004) Unterrichtsplanung. In: Keck RW, Sandfuchs U, Thomas B (eds) Wörterbuch Schulpädagogik: Ein Nachschlagewerk für Studium und Schulpraxis. Klinkhardt, Bad Heilbrunn, pp 348–350
- Sandfuchs U (2009) Grundfragen der Unterrichtsplanung. In: Arnold K-H, Sandfuchs U, Wiechmann J (eds) Handbuch Unterricht. UTB, Stuttgart, Bad Heilbrunn, pp 512–518

- Schartel SA (2012) Giving feedback - an integral part of education. *Best Pract Res Clin Anaesthesiol* 26:77–87. <https://doi.org/10.1016/j.bpa.2012.02.003>
- Schiavo G, Businaro S, Zancanaro M (2024) Comprehension, apprehension, and acceptance: Understanding the influence of literacy and anxiety on acceptance of artificial Intelligence. *Technology in Society* 77:102537. <https://doi.org/10.1016/j.techsoc.2024.102537>
- Schmaltz C (2019) Heterogenität als Herausforderung für die Professionalisierung von Lehrkräften. Springer VS, Wiesbaden
- Schmitt N, Kuljanin G (2008) Measurement invariance: Review of practice and implications. *Human Resource Management Review* 18:210–222. <https://doi.org/10.1016/j.hrmr.2008.03.003>
- Schulz W (1980) Unterrichtsplanung. Praxis und Theorie des Unterrichtens. Urban & Schwarzenberg, München, Wien
- Seifried J (2009) Unterrichtsplanung von (angehenden) Lehrkräften an kaufmännischen Schulen. *zbw* 105:179–197. <https://doi.org/10.25162/zbw-2009-0012>
- Seifried J, Grill L, Wagner M (2006) Unterrichtsmethoden in der kaufmännischen Unterrichtspraxis. *Wirtschaft und Erziehung* 58:236–241
- Sembey R, Hoda R, Grundy J (2024) Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *Journal of Systems and Software* 211:111988. <https://doi.org/10.1016/j.jss.2024.111988>
- Shaik T, Tao X, Li Y, Dann C, McDonald J, Redmond P, Galligan L (2022) A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis. *IEEE Access* 10:56720–56739. <https://doi.org/10.1109/ACCESS.2022.3177752>
- Shute VJ (2008) Focus on Formative Feedback. *Review of Educational Research* 78:153–189. <https://doi.org/10.3102/0034654307313795>
- Slavin R, Smith D (2009) The Relationship Between Sample Sizes and Effect Sizes in Systematic Reviews in Education. *Educational Evaluation and Policy Analysis* 31:500–506. <https://doi.org/10.3102/0162373709352369>
- Söll M, Klusmeyer J (2018) Unterrichtsplanung als zentraler Gegenstand der Wirtschaftsdidaktik – Konzeption eines hochschuldidaktischen Ansatzes. In: Tramm PT, Casper M, Schlömer T (eds) *Be-richte zur beruflichen Bildung: Bd. 22. Didaktik der beruflichen Bildung – Selbstverständnis, Zu-kunftsperspektiven und Innovationsschwerpunkte*. Bertelsmann, Bielefeld, pp 73–88
- Sperling K, Stenberg C-J, McGrath C, Åkerfeldt A, Heintz F, Stenliden L (2024) In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open* 6:100169. <https://doi.org/10.1016/j.caeo.2024.100169>

- Spreitzer C, Straser O, Zehetmeier S, Maaß K (2024) Mathematical Modelling Abilities of Artificial Intelligence Tools: The Case of ChatGPT. *Education Sciences* 14:698.
<https://doi.org/10.3390/educsci14070698>
- Steiss J, Tate T, Graham S, Cruz J, Hebert M, Wang J, Moon Y, Tseng W, Warschauer M, Olson CB (2024) Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction* 91:101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Stender A, Brückmann M, Neumann K (2015) Vom Professionswissen zum kompetenten Handeln im Unterricht: Die Rolle der Unterrichtsplanung. *Beiträge zur Lehrerinnen- und Lehrerbildung* 33:121–133
- Stütz S, Berding F, Reincke S, Scheper L (2022) Characteristics of learning tasks in accounting textbooks: an AI assisted analysis. *Empirical Res Voc Ed Train* 14:1–33.
<https://doi.org/10.1186/s40461-022-00138-2>
- Taber KS (2018) The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res Sci Educ* 48:1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tamim RM, Bernard RM, Borokhovski E, Abrami PC, Schmid RF (2011) What Forty Years of Research Says About the Impact of Technology on Learning. *Review of Educational Research* 81:4–28
- Tipton E, Hallberg K, Hedges LV, Chan W (2017) Implications of Small Samples for Generalization: Adjustments and Rules of Thumb. *Eval Rev* 41:472–505.
<https://doi.org/10.1177/0193841X16655665>
- Tubino L, Adachi C (2022) Developing feedback literacy capabilities through an AI automated feedback tool. In: ASCILITE (ed) *Reconnecting relationships through technology : Proceedings of the 2022 ASCILITE Conference*, 1-5. <https://doi.org/10.14742/apubs.2022.39>
- Tunstall L, Werra L von, Wolf T, Géron A (2022) *Natural language processing with transformers: Building language applications with hugging face*. O'Reilly, Sebastopol
- Underwood J (2008) *Effective feedback: guidelines for improving performance*. International Conference of the Learning Sciences
- United Nations (2015) *Transforming our World: The 2030 Agenda for Sustainable Development*.
<https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>
- van den Berg G, Du Plessis E (2023) ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness in Teacher Education. *Education Sciences* 13:998. <https://doi.org/10.3390/educsci13100998>
- van der Kleij FM, Feskens RCW, Eggen, T. J. H. M. (2015) Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes. *Review of Educational Research* 85:475–511. <https://doi.org/10.3102/0034654314564881>

- Venkatesh V, Davis FD (2000) A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science* 46:186–204.
<https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, Morris, Davis (2003) User Acceptance of Information Technology: Toward a Unified View. *MISQ* 27:425. <https://doi.org/10.2307/30036540>
- Vygotsky LS (1981) *Mind in Society: The development of higher psychological processes*. Harvard University Press, Cambridge, Mass.
- Wagner S, Sibley L, Weiler D, Burde J-P, Scheiter K, Lachner A (2024) The more, the better? Learning with feedback and instruction. *Learning and Instruction* 89:101844.
<https://doi.org/10.1016/j.learninstruc.2023.101844>
- Wernke S, Zierer K (2017) Die Unterrichtsplanung - Ein in Vergessenheit geratener Kompetenzbereich?! In: Wernke S, Zierer K (eds) *Die Unterrichtsplanung - ein in Vergessenheit geratener Kompetenzbereich?!: Status Quo und Perspektiven aus Sicht der empirischen Forschung*. Klinkhardt, Bad Heilbrunn, pp 7–16
- Wilbers K (2023) *Wirtschaftsunterricht gestalten*. epubli, Berlin
- Williams A (2024) Delivering Effective Student Feedback in Higher Education: An Evaluation of the Challenges and Best Practice. *IJRES* 10:473–501. <https://doi.org/10.46328/ijres.3404>
- Wisniewski B, Zierer K, Hattie J (2019) The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology* 10:3087.
<https://doi.org/10.3389/fpsyg.2019.03087>
- Zhai X, Chu X, Chai CS, Jong MSY, Istenic A, Spector M, Liu J-B, Yuan J, Li Y (2021) A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity* 2021.
<https://doi.org/10.1155/2021/8812542>
- Zhang J, Zhang Z (2024) AI in teacher education: Unlocking new dimensions in teaching support, inclusive learning, and digital literacy. *J Comput Assist Learn* 40:1871–1885.
<https://doi.org/10.1111/jcal.12988>

Appendix A. Reliability of the training data

Dimension	Category	Levels	N Coder	N Units		t		%		α		AC		Alpha Values per level									
				1	2	1	2	1	2	1	2	1	2	0		1		2		3		4	
														1	2	1	2	1	2	1	2	1	2
Basic dimensions	Lesson structure	0-4	2	58	144	0.51	0.88	44.8	88.9	0.34	0.97	0.23	0.86	0.36	n.a.	0.69	0.59	0.51	0.95	0.49	0.95	0.67	0.98
	Learner activation	0-4	2	58	144	0.56	0.97	39.7	97.9	0.22	0.98	0.34	0.98	0.5	0.5	n.a.	0.83	0.5	0.98	0.49	1	0.76	0.99
	Cross-linked learning	0-4	2	58	144	0.37	0.97	24.1	97.2	0.13	0.99	0.13	0.97	0.5	0.97	0.28	1	0.37	0.97	0.52	0.99	0.27	1
Authenticity of actions	Mental planning	0-1	3	10	10	0.54	0.61	50	60	0.29	0.27	0.38	0.59	0.78	1	1	0.85						
	Actual execution	0-1	3	10	10	0.59	0.88	50	90	-0.18	.76	0.54	0.91	0.78	1	1	0.85						
	Kind of action	0-1	3	10	40	0.78	0.52	80	60	-0.05	0.14	0.85	0.61	0.93	0.88	0.5	0.51						
	Context of action	0-1	3	10	41	0.75	0.55	80	63	0.43	0.27	0.83	0.63	1	0.91	0.63	0.58						
Complete action	Inform	0-1	3	10	41	0.13	0.88	30	90	0.08	0.86	0.07	0.88	0.52	0.92	1	1						
	Plan	0-1	3	10	10	0.66	0.28	70	17	0.45	-0.22	0.69	0	0.91	0.66	0.78	0.5						
	Decide	0-1	3	10	10	n.a.	0.36	100	50	1	0.29	1	0.38	n.a.	1	n.a.	0.56						
	Act	0-1	3	10	10	0.47	0.88	50	90	0.34	0.83	0.33	0.89	0.69	0.89	1	1						
	Check	0-1	3	10	10	0.77	0.49	80	60	0.63	0.38	0.79	0.54	1	0.6	0.77	0.95						
Process orientation	Reflect	0-1	3	10	10	0.78	0.69	80	70	-0.05	0.39	0.85	0.71	0.93	0.89	0.5	1						
	Process orientation	0-3	3	10	41	0.5	0.4	30	50	0.5	0	0.27	0.12	0.88	0.33	0.4	0.5	0.67	0.55	0.36	0.52		
	Value streams	0-2	3	10	41	0.66	0.65	70	61	0.39	0.55	0.71	0.65	1	0.91	0.63	0.72	n.a.	0.5				
	Financial streams	0-2	3	10	10	0.57	0.63	40	70	0.42	0.58	0.46	0.63	0.86	0.75	0.5	0.95	0.5	n.a.				
Problem orientation	Information streams	0-2	3	10	10	0.58	0.59	70	50	0.5	-0.17	0.67	0.56	1	0.43	0.63	0.8	n.a.	0.4				
	Kinds of problems	0-2	2	50	24	0.65	n.a.	76	100	0.34	1	0.72	1	0.9	1	0.33	n.a.	0.4	n.a.				
Motivation	Contains problems	0-3	2	50	48	0.66	0.85	60	92	0.25	-0.02	0.53	0.91	0.8	0.96	0.55	0.5	0.5	n.a.	0.5	n.a.		
	Multiple solution strategies	0-3	2	50	48	0.46	0.33	28	38	-0.15	-0.42	0.1	-0.1	0.33	0.69	0.5	0.5	0.48	n.a.	0.5	n.a.		
	Multiple solutions possible	0-2	2	50	48	0.52	n.a.	28	100	0.29	1	0.09	1	0.5	n.a.	0.62	1	0.58	n.a.				
	Competence	0-2	2	50	48	0.62	0.65	66	80	0.52	0.48	0.53	0.66	0.69	0.87	0.88	0.94	0.67	n.a.				
	Relatedness	0-2	2	50	48	0.77	1	84	100	0.88	1	0.78	1	0.95	1	0.8	1	0.97	n.a.				
Content-related digitalization	General technologies	0-1	2	60		.92		95		.70		.94		1		0.76							
	Specific technologies	0-1	2	60		n.a.		100		1		1		1		n.a.							
	Verbalization	0-1	2	60		n.a.		100		1		1		1		n.a.							
	Topic	0-1	2	60		.83		90		-0.04		0.89		0.95		0.5							
	Condition	0-1	2	60		n.a.		100		1		1		1		n.a.							
	Development	0-1	2	60		n.a.		100		1		1		1		n.a.							
Methodological digitalization	Technology use	0-2	2	41		0.91		95		0.48		0.95		0.98		0.9		n.a.					
	Cooperation and collaboration	0-1	2	60		.94		97		-0.01		0.97		0.98		0.5							

Dimension	Category	Levels	N Coder	N Units		t		%		α		AC		Alpha Values per level											
														0		1		2		3		4			
				1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2		
Sustaina- bility com- petences	Engagement	0-1	2	60		.85		92		.40		.90		0.96		0.94									
	Feedback	0-3	2	55		.90		95		-0.02		.94		0.97		0.5		n.a.		n.a.					
	Content knowledge	0-2	3	10	9	0.42	1	10	100	-0.03	1	0.21	1	0.5	n.a.	0.67	1	0.5	1						
	Network knowledge	0-2	3	10	9	0.5	0.73	50	67	-0.02	0.6	0.51	0.7	0.8	0.89	0.5	1	n.a.	0.83						
	Methodological competences	0-2	3	10	9	0.88	0.48	90	33	0.87	0.28	0.87	0.36	1	0.62	0.93	0.87	n.a.	0.63						
	Social competences	0-2	3	10	9	0.36	0.66	40	56	0.15	0.3	0.25	0.6	0.56	1	1	0.89	n.a.	0.65						
	Decision-making	0-2	3	10	9	0.55	0.71	70	67	0.29	0.29	0.72	0.74	1	0.98	0.5	0.63	n.a.	0.5						
	Open and critical perspective	0-2	3	10	9	0.55	0.72	70	67	0.29	0.65	0.72	0.7	1	0.95	0.5	1	n.a.	0.65						
	Reflexivity	0-2	3	10	9	n.a.	n.a.	100	100	1	1	1	1	n.b	n.a.	n.b	n.a.	n.a.	n.a.						
Sustainable develop- ment goals	SDG 1	0-1	3	10		n.a.		100																	
	SDG 2	0-1	3	10		n.a.		100																	
	SDG 3	0-1	3	10		n.a.		100																	
	SDG 4	0-1	3	10		n.a.		100																	
	SDG 5	0-1	3	10		n.a.		100																	
	SDG 6	0-1	3	10		n.a.		100																	
	SDG 7	0-1	3	10		n.a.		100																	
	SDG 8	0-1	3	10		n.a.		100																	
	SDG 9	0-1	3	10		n.a.		100																	
	SDG 10	0-1	3	10		n.a.		100																	
	SDG 11	0-1	3	10		n.a.		100																	
	SDG 12	0-1	3	10		0.88		90		-0.02		0.93		0.97		0.5									
	SDG 13	0-1	3	10		n.a.		100																	
	SDG 14	0-1	3	10		n.a.		100																	
	SDG 15	0-1	3	10		n.a.		100																	
	SDG 16	0-1	3	10		n.a.		100																	
	SDG 17	0-1	3	10		n.a.		100																	

Appendix B. Model fit information with BAYES estimator

Scale	n	CFI	RMSEA	PPP	Invariance
Knowledge on teaching	102	0.95	0.09	0.13	metric
Technical knowledge	102	0.90	0.11	0.07	metric
Digital diagnostics	102	0.97	0.05	0.26	metric
Digital teaching	102	0.93	0.07	0.18	scalar
Data processing and evaluation	102	0.90	0.10	0.13	scalar
Communication and cooperation	102	0.96	0.09	0.23	metric
Content creation	102	0.93	0.11	0.07	metric
Safety	102	0.98	0.06	0.21	scalar
Problem solving	102	0.75	0.16	0.00	configural
Use and apply AI	102	0.95	0.12	0.00	configural
Understand AI	102	0.92	0.12	0.02	configural
Detect AI	102	0.91	0.11	0.07	scalar
AI ethics	101	0.95	0.10	0.09	metric
AI problem solving	102	0.96	0.10	0.13	configural
AI learning	101	1.00	0.02	0.35	metric
AI persuasion literacy	101	0.98	0.05	0.27	metric
AI emotion regulation	101	1.00	0.00	0.35	scalar
Perceived usefulness	102	0.95	0.11	0.00	configural
Perceived ease of use	102	0.98	0.08	0.14	configural
Attitude towards using technology	102	0.89	0.18	0.00	configural
Job relevance	102	1.00	0.00	0.51	scalar
Competence	96	0.85	0.10	0.02	configural
Autonomy	101	1.00	0.00	0.30	scalar
Relatedness	100	0.96	0.09	0.05	configural

Appendix C. Model fit information with maximum likelihood estimator

Scale	n	Chi ²	p	CFI	RMSEA	SRMR	level
Knowledge on teaching	102	36.91	0.06	0.93	0.10	0.19	metric
Technical knowledge	102	46.60	0.00	0.81	0.19	0.08	configural
Digital diagnostics	102	33.71	0.01	0.87	0.15	0.10	configural
Digital teaching	102	44.95	0.05	0.88	0.09	0.24	scalar
Data processing and evaluation	102	28.97	0.57	1.00	0.00	0.20	scalar
Communication and cooperation	102	11.80	0.16	0.96	0.10	0.19	metric
Content creation	102	35.30	0.00	0.91	0.15	0.05	configural
Safety	102	39.70	0.14	0.96	0.07	0.13	scalar
Problem solving	102	110.34	0.00	0.69	0.19	0.12	configural
Use and apply AI	102	119.49	0.00	0.94	0.12	0.06	configural
Understand AI	102	69.33	0.00	0.90	0.13	0.08	configural
Detect AI	102	38.63	0.16	0.96	0.07	0.13	scalar
AI ethics	101	30.10	0.02	0.92	0.13	0.10	configural
AI problem solving	102	27.58	0.04	0.95	0.12	0.09	configural
AI learning	101	33.55	0.12	0.97	0.08	0.20	metric
AI persuasion literacy	101	29.70	0.24	0.96	0.06	0.17	metric
AI emotion regulation	101	34.51	0.10	0.93	0.09	0.17	metric
Perceived usefulness	102	182.43	0.00	0.94	0.12	0.05	configural
Perceived ease of use	102	58.26	0.02	0.97	0.10	0.07	configural
Attitude towards using technology	102	103.02	0.00	0.88	0.18	0.07	configural
Job relevance	102	12.81	0.31	0.99	0.06	0.27	scalar
Competence	96	72.70	0.00	0.79	0.14	0.15	configural
Autonomy	101	66.14	0.24	0.96	0.05	0.16	scalar
Relatedness	100	73.11	0.00	0.90	0.14	0.15	configural

Appendix D. Model fit information for IMEB-SBA with the maximum likelihood estimator

Scale	n	Chi ²	p	CFI	RMSEA	SRMR	level
Security	94	5.21	0.27	0.99	0.08	0.03	configural
Security	94	30.24	0.00	0.77	0.24	0.28	metric
Security	94	37.13	0.00	0.73	0.23	0.28	scalar
Security	94	237.19	0.00	0.00	0.54	3.97	residual
Structure	94	2.53	0.64	1.00	0.00	0.03	configural
Structure	94	9.04	0.34	0.99	0.05	0.18	metric
Structure	94	13.59	0.26	0.97	0.07	0.21	scalar
Structure	94	176.38	0.00	0.00	0.46	3.70	residual
Justification	94	53.22	0.00	0.53	0.51	0.08	configural
Justification	94	49.20	0.00	0.61	0.33	0.11	metric
Justification	94	42.12	0.00	0.71	0.25	0.12	scalar
Justification	94	146.40	0.00	0.00	0.42	1.61	residual
Source	94	18.22	0.00	0.90	0.28	0.04	configural
Source	94	19.34	0.01	0.92	0.17	0.05	metric
Source	94	23.11	0.02	0.92	0.15	0.07	scalar
Source	94	140.80	0.00	0.16	0.41	1.45	residual
Practicability	94	0.00	0.00	1.00	0.00	0.00	configural
Practicability	94	2.14	0.54	1.00	0.00	0.15	metric
Practicability	94	7.77	0.17	0.91	0.11	0.15	scalar
Practicability	94	135.19	0.00	0.00	0.55	3.17	residual

Appendix E. Frequencies of coded dimensions from both groups at start and finish of the semester

	Experimental t0						Control t0						Experimental t5						Control t5					
	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4
SDG 1		100	0					100.0	0					100.0	0					100.0	0			
SDG 2		100	0					91.7	8.3					92.7	7.3					85.4	14.6			
SDG 3		92.7	7.3					100.0	0					82.7	7.3					100.0	0			
SDG 4		100	0					100.0	0					92.7	7.3					100.0	0.0			
SDG 5		100	0					100.0	0					100.0	0					100.0	0.0			
SDG 6		100	0					100.0	0					92.7	7.3					91.7	8.3			
SDG 7		90.9	9.1					100.0	0.0					92.7	7.3					100.0	0.0			
SDG 8		54.5	45.5					75.0	25.0					32.7	67.3					62.5	37.5			
SDG 9		80	20					91.7	8.3					76.4	23.6					91.7	8.3			
SDG 10		100	0					100.0	0.0					100.0	0					100.0	0			
SDG 11		94.5	5.5					100.0	0.0					92.7	7.3					100.0	0			
SDG 12		25.5	74.5					6.3	93.8					29.1	70.9					8.3	91.7			
SDG 13		45.5	54.5					68.8	31.3					52.7	47.3					37.5	62.5			
SDG 14		100	0					91.7	8.3					92.7	7.3					100.0	0			
SDG 15		96.4	3.6					100.0	0					87.3	12.7					100.0	0			
SDG 16		100	0					100.0	0					100	0					100.0	0			
SDG 17		100	0					100.0	0					100	0					100.0	0			
Content knowledge		12.7	0	87.3				20.8	0	79.2				0	0	100.0				0	0	100.0		
Network knowledge		76.4	0	23.6				68.8	0	31.3				45.5	0	54.5				31.3	0	68.8		
Methodological competences		40.0	7.3	52.7				43.8	0	56.3				14.5	0	85.5				6.3	0	93.8		
Social competences		47.3	10.9	41.8				66.7	0	33.3				0	0	100.0				27.1	0	72.9		

	Experimental t0						Control t0						Experimental t5						Control t5					
	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4
Decision-making		83.6	0	16.4				68.8	0	31.3				7.3	0	92.7				8.3	0	91.7		
Open and critical perspective		76.4	0	23.6				58.3	0	41.7				20	0	80.0				29.2	0	70.8		
Reflexivity		89.1	0	10.9				83.3	0	16.7				27.3	0	72.7				60.4	0	39.6		
Mental planning		74.5	25.5					12.5	87.5					20.0	80.0					37.5	62.5			
Actual execution		40.0	60.0					60.4	39.6					0	100.0					16.7	83.3			
Kind of action		67.3	32.7					68.8	31.3					23.6	76.4					33.3	66.7			
Inform		54.5	41.8					27.1	72.9					0	100.0					12.5	87.5			
Plan		70.9	29.1					43.8	56.3					7.3	92.7					37.5	62.5			
Decide		92.7	7.3					62.5	37.5					0	100.0					22.9	77.1			
Act		70.9	29.1					45.8	54.2					7.3	92.7					8.3	91.7			
Check		61.8	38.2					58.3	41.7					0.0	100.0					45.8	54.2			
Reflect		74.5	25.5					75.0	25.0					30.9	69.1					58.3	41.7			
Context		81.8	18.2					77.1	22.9					9.1	90.9					20.8	79.2			
Process orientation		45.5	54.5	0	0			22.9	45.8	22.9	8.3			0	45.5	49.1	5.5				8.3	43.8	47.9	
Value streams		38.2	47.3	14.5				60.4	31.3	8.3				0	18.2	81.8				8.3	14.6	77.1		
Financial streams		40.0	52.7	7.3				77.1	22.9	0.0				12.7	30.9	56.4				52.1	14.6	33.3		
Information streams		29.1	56.4	14.5				14.6	62.5	22.9				0	5.5	94.5				31.3	14.6	54.2		
Contains problems		56.4	43.6	0	0			27.1	58.3	14.6	0			0	34.5	65.5	0			6.3	68.8	25.0	0	

	Experimental t0						Control t0						Experimental t5						Control t5					
	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4
Multiple solution strategies		38.2	54.5	7.3	0.0			35.4	41.7	22.9	0.0			0	16.4	76.4	7.3			66.7	25.0	8.3	0.0	
Multiple solutions possible			76.4	16.4	7.3				43.8	16.7	39.6				3.6	34.5	61.8				14.6	85.4	0.0	
Competence		54.5	38.2	7.3	0			58.3	33.3	8.3	0.0			0	40.0	60.0	0.0			12.5	45.8	41.7	0.0	
Relatedness		38.2	38.2	23.6	0			41.7	50.0	8.3	0			0	36.4	63.6	0			27.1	47.9	25.0	0	
Problem orientation		49.1	43.6	7.3				27.1	50.0	22.9				0	78.2	21.8				6.3	60.4	33.3		
Technology use		54.5	0	45.5				75.0	0	25.0				0	0.0	100				20.8	0	79.2		
Cooperation and collaboration		100.0	0.0					100.0	0.0					74.5	25.5					83.3	16.7			
Feedback		100.0	0.0	0.0	0.0			100.0	0.0	0.0	0.0			94.5	0.0	0.0	5.5			100.0	0.0	0.0	0.0	
Engagement		83.6	16.4					91.7	8.3					23.6	76.4					52.1	47.9			
General technologies		92.7	7.3					83.3	16.7					45.5	54.5					37.5	62.5			
Specific technologies		100.0	0.0					91.7	8.3					81.8	18.2					60.4	39.6			
Verbalization		92.7	7.3					91.7	8.3					45.5	54.5					29.2	70.8			
Topic		89.1	10.9					100.0	0.0					52.7	47.3					68.8	31.3			
Condition		100.0	0.0					91.7	8.3					100.0	0.0					83.3	16.7			
Development		100.0	0					100.0	0					100.0	0					100.0				
Lesson structure	27.3	5.5	0.0	14.5	21.8	30.9	41.7	0.0	8.3	0	8.3	41.7	0.0	0	0.0	0.0	21.8	78.2	20.8	0.0	6.3	6.3	8.3	58.3

	Experimental t0						Control t0						Experimental t5						Control t5					
	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4	-1	0	1	2	3	4
Learner activation	20	12.7	0.0	7.3	32.7	27.3	35.4	0.0	8.3	16.7	31.3	8.3	0.0	0	0.0	7.3	23.6	69.1	27.1	0.0	0.0	16.7	47.9	8.3
Cross-linked learning	20	5.5	12.7	10.9	40.0	10.9	35.4	0.0	8.3	25.0	31.3	0	0	0	0	14.5	36.4	49.1	20.8	8.3	8.3	22.9	39.6	0

Appendix F. Non-standardized probit regression results

Variable	Base	Group	Gender	NAWI	Structure	Practicability
SDG 2	-0.662	-0.199	0.318	0.289	0.301	-1.477
SDG 7	-0.617	1.612	-0.136	-0.876	0.745	0.531
SDG 8	-2.711*	-0.379	-0.019	0.648	-1.263	-0.573
SDG 9	-1.585	0.395	2.428	-0.215	1.087	-1.275
SDG 11	-0.595	1.477	0.990	0.203	0.151	-1.695
SDG 12	-3.033	-0.788	1.341	0.338	-1.060	-0.134
SDG 13	-2.816*	-0.977	-0.252	0.880	0.095	-2.454
SDG 14	-0.136	1.542	1.038	0.209	0.131	-1.694
SDG 15	-0.671	1.857**	-0.939	-1.082**	-0.140	2.219**
Content knowledge	-4.047*	-0.162	0.418	0.334	-0.456	0.220
Network knowledge	-2.478*	0.037	-1.491	-0.665	-0.119	2.093**
Methodological competences	-4.249*	0.497	0.338	0.212	0.126	0.250
Social competences	-2.854*	2.907*	-0.007	0.639	0.045	-1.704
Decision-making	-2.932	0.287	0.792	0.580	-0.932	-0.227
Open and critical perspective	-2.908*	1.104	-1.272	-1.001**	0.948	1.988**
Reflexivity	-2.070*	1.036**	0.997	-0.459	1.646**	-0.705
Mental planning	-2.552	1.718	0.040			
Actual execution	-4.272	0.639	0.161			
Kind of action	-2.928	0.135	1.314			
Inform	-4.016	1.660	0.497			
Plan	-3.320	1.189	0.346			
Decide	-3.476	1.753	1.005			
Act	-4.047	-0.236	0.315			
Check	-4.121	1.904	-1.577			
Reflect	-2.466	0.555	0.880			
Context	-3.111	0.191	0.108			
Process orientation	-1.410	-0.464	2.021			

Variable	Base	Group	Gender	NAWI	Structure	Practicability
Value streams	-2.233	-1.037	1.743			
Financial streams	-0.424	0.583	0.587			
Information streams	-1.576	1.650	-0.876			
Contains problems	-2.361	0.695	0.627			
Multiple solution strategies	-2.481	0.605	1.716			
Multiple solutions possible	-1.581	0.397	0.388			
Competence	-1.711	1.439	0.577			
Relatedness	-1.421	1.314	1.248			
Problem orientation	-2.956	-0.595	1.472			
Technology use	-2.957	1.618	0.768			
Engagement	-2.424	0.590	1.181			
General technologies	-2.020	-0.111	0.724			
Specific technologies	-1.120	-0.167	-1.300			
Verbalization	-1.965	-0.282	-0.559			
Topic	-1.689	0.269	-0.168			
Condition	-0.536	-1.636	-2.230			
Lesson structure	-0.698	2.279	0.058			
Learner activation	-0.443	1.674	1.886			
Cross-linked learning	-0.455	1.879	0.651			

Note: *=p<0.05; **=p<0.10. Table only contains categories where regression results could be calculated, thus where there was a change throughout the semesters.

Appendix G. Non-standardized latent growth model results

	Control group			Experimental group		
Scale	Initial	Growth	Treatment	Initial	Growth	Treatment
<i>SWIT</i>						
Digital teaching	4.221*	0.077	0	5.283*	0.237	0
<i>DigComp</i>						
Data processing and evaluation	3.749*	0.114	0	3.718*	0.114	-0.04
Safety	3.23*	0.138	0	2.873*	0.138	0.019
<i>AI Literacy</i>						
Detect AI	5.493*	0.021	0	4.666*	0.021	0.915*
AI emotion regulation	8.836*	-1.37*	0	8.367*	-1.37*	0.85*
<i>AI Attitude</i>						
Job relevance	5.283*	0.237	0	4.984*	0.237	0.061
<i>Self-determination framework</i>						
Autonomy	5.014*	0.244*	0	5.418*	0.244*	-0.004

Note: *= $p < 0.05$; **= $p < 0.10$.

Engaging with AI: How student teachers interact with an analytical AI platform for feedback and the effect on their lesson plans – An eye tracking study

Julia Pargmann^{1,*}, Florian Berding¹, Karin Rebmann², Elisabeth Riebenbauer³

¹ Universität Hamburg, Sedanstr. 19, 20146 Hamburg, Germany

² University of Oldenburg, Ammerländer Heerstraße 114-119, 26129 Oldenburg, Germany

³ University of Graz, Universitätsstraße 15, 8010 Graz, Austria

* Corresponding author; Email: julia.pargmann@uni-hamburg.de

Abstract

Despite the increasing use of AI platforms in the education sector, the understanding of their interactions and their impact on learning processes and the development of expertise remain insufficiently researched, particularly in the area of teacher training and vocational education. This study examines the types of interactions of N=32 student teachers through the example of the AI platform EDDA, which provides feedback for lesson planning. With the help of eye tracking technology, we analyzed how these interactions change over the course of the semester and what influence they have on the quality of the lesson plans created. lesson plans. Three general types of interaction were identified: feedback-oriented, balanced and results-oriented. The analysis showed that students with feedback-oriented or balanced approaches tend to develop better lesson plans. This typology also reflects preferences for different feedback representations, written or visual. The results contribute to a better understanding of the interaction with AI platforms by demonstrating the link between visual attention and performance levels in lesson planning. They emphasize the importance of detailed, written feedback for the integration of didactic criteria. The results provide valuable insights into the interactions with AI and the effective use of AI feedback in teacher education.

Keywords: eye tracking, visual strategies, teacher education, artificial intelligence, lesson planning

Declarations

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available but will be made available by the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding was received for this study.

Human Ethics and Consent to Participate and Publish

Written informed consent to evaluate the lesson plans retrospectively was obtained from the participants at the start of the semester. Written informed consent to record, analyze and publish the eye tracking data was obtained from all participants at every eye tracking instance.

Authors' contributions

Julia Pargmann conceptualized the study, analyzed the data, interpreted the data, wrote the original draft and did the review and editing. Florian Berding conceptualized the study, helped with data analysis and provided commentary and revision and supervised the study. Karin Rebmann provided critical commentary and revisions in the editing process and provided resources for the data analysis. Elisabeth Riebenbauer provided critical commentary and revisions in the editing process and provided resources for the data analysis. All authors read and approved the final manuscript.

Acknowledgements

We thank all participating students and our student assistants who were involved in coding the data.

1 Introduction

As learning processes continue to move into the digital spheres, the evaluation of digital environments and particularly AI platforms is necessary. Different approaches exist towards these environments; while some researchers employ existing, more generic platforms (Jacobsen & Weber, 2024; Liu et al., 2023; Steiss et al., 2024; van den Berg & Du Plessis, 2023), others develop their own, more domain-specific ones (Pargmann, Riebenbauer, et al., 2024; Rausch et al., 2021). Research on interactions with AI platforms and their effects on learning as well as expertise (or competence) development, however, is minimal (see for example Ebner & Sageder, 2023 or Frenzke–Shim et al., 2024). Despite the increasing popularity of digital learning environments and the growing use of AI systems, the understanding of the specific interactions and their impact on learning outcomes remains underdeveloped, particularly in teacher and vocational education and training (VET). Existing research from other domains suggests that the interaction with digital tools changes with growing user expertise (C. W. Mayer et al., 2023). Proficient users interact differently with digital tools than novice users. Hahn and Klein (2022) and Kok et al. (2023) show that more proficient users focus on the relevant content of the tools, while novices do not interact with tools in a structured or goal-oriented way. With a growing body of research on the opportunities and challenges of AI platforms in education, the interaction with the platforms itself becomes more relevant to gain insights into expertise development with AI. These so-called visual strategies provide information on the ways in which someone looks at and interacts with different components of a platform. However, little is known about the development of visual strategies regarding AI platforms in teacher education.

One of the specialized learning platforms in the domain of vocational business education is EDDA, electronic-didactical assistance for lesson planning. EDDA provides users with feedback on the integration of 57 general and domain-specific didactical planning principles. The feedback uses established criteria and provides personalized visual representations of the integration. Thus, the platform is a multimedia tool in accordance with R. E. Mayer and Fiorella (2021), which are tools that combine multiple ways of content representation, preferably written and visual. The effectiveness of EDDA's feedback has recently been validated in a longitudinal intervention study by

Pargmann et al. (in review). It shows that EDDA's feedback has a positive impact on some planning dimensions, but not on all.

This paper aims to contribute to this gap by using EDDA as a sample AI platform to establish a typology of visual strategies or interaction types for expertise development through AI platforms that can potentially be applied to other feedback platforms as well. As Brückner and Zlatkin-Troitschanskaia (2024a) argue based on Gegenfurtner et al. (2023), the effect of visual perceptions on expertise development can be uncovered through the tracking of eye movements. Integrating eye-tracking technologies into research on interactions with AI platforms could provide specific insights into which visual strategies are successful and how they evolve over time. Eye trackers illuminate the eye, causing corneal reflections which are then captured by a camera. Through algorithms, this data can then be aggregated, i.e., into a reconstruction of the scan path (Beach & McConnel, 2019).

Studies on eye-tracking in VET show that this technology can clarify visual strategies in simulations and computer-based learning. Browning et al. (2016) and O'Meara et al. (2015) report that simulations with eye tracking in healthcare improve situational awareness and clinical performance. Eklund et al. (2020) show that eye tracking improves the transfer of tacit knowledge when training sea pilots. Li et al. (2023) show that experienced surgeons in simulation focus more on task-relevant areas, and Vijayan et al. (2018) describe systematic scan paths that relate to process improvements in engineering. A study by Schweinberger et al. (2023) found that regarding the use of feedback tools in physics teacher education, steering students' gaze in a specific direction can aid their feedback reception. In vocational business education and its respective teacher education, only few eye tracking studies are known (Brückner & Zlatkin-Troitschanskaia, 2024b; C. W. Mayer et al., 2023), although the technology seems generally promising to gain deeper insight into noticing processes and the development of tacit knowledge in other fields and stages of teacher education (Beach & McConnel, 2019; Beach & Willows, 2014; Faiella et al., 2019; Jarodzka et al., 2021; Langner et al., 2022; McIntyre et al., 2019; Schweinberger et al., 2023; Stürmer et al., 2017; Wyss et al., 2021). This gap is remarkable, as interactions with digital learning environments are becoming an integral part of technology use and design.

Eye tracking thus seems to offer great potential for educational science to investigate various effects and processes. Jarodzka et al. (2017) derive three research areas that have previously used eye tracking technology on a more general level. These are the improvement of instructional designs in digital environments, the development of expertise and the use of tacit knowledge and lastly visual expertise, i.e., in the sense of triangulation or pattern recognition. Studies in vocational business education have been focusing on the first two, those in health and medical education on the latter two.

Overall, there is a gap regarding the visual strategies of student teachers in VET in the engagement with digital technologies, such as AI platforms. While there is evidence that eye trackers can uncover connections between visual interaction strategies and expertise development, more research is needed on vocational teacher education and the interaction with AI platforms, for example through the analysis of interaction types (i.e., individual strategies to engage with the technology). A better understanding of the connections between the improvement of digital learning platforms and learning experiences could aid in more individualized solutions and optimize AI tools for learning. The present paper contributes to the gap on interaction strategies with AI and their effects on learning outcomes using the AI platform EDDA as a use case. Thus, the research questions are:

RQ1: Which interaction types can be identified using EDDA? How do these interaction types change throughout the semester?

RQ2: Which interaction types produce better lesson plans?

The remainder of this paper is structured as follows: In the following chapter, we establish the theoretical context of eye tracking technology and gaze behavior and give a short overview over eye tracking studies in educational sciences as well as vocational (business) education. Then, we quickly introduce the AI platform and its structure. After that, the hypotheses are derived, followed by the methods that we applied. This includes the sample, the apparatus, the procedure and the measurement models. Then we provide the results for the data analysis and discuss them. The paper finishes with a short conclusion with implications for further research and practice.

2 Context

2.1 Eye tracking technology and gaze behavior

The most common eye tracking events are fixations and saccades (Holmqvist et al., 2011). While fixations occur when the eye remains still, saccades are rapid movements from one fixation to another. By recording fixations and saccades, the point of gaze is developed. Eye trackers record the reflection of the cornea to aggregate the eye movements. In a review on the use of eye tracking in educational science, Lai et al. (2013, p. 90) find that there are seven general themes: “patterns of information processing, effects of instructional design, reexamination of existing theories, individual differences, effects of learning strategies, patterns of decision making, and conceptual development”. This conceptualization is more detailed than that of Jarodzka et al. (2017), who suggest three main directions: the improvement of instructional designs in digital environments, the development of expertise and the use of tacit knowledge and lastly visual expertise. Lai et al. (2013) expand these areas in regards to the reexamination of existing theories and conceptual developments, however the other aspects can generally be assigned to the improvement of instructional design and the development of expertise. Thus, this paper follows the general structure of Jarodzka et al. (2017).

Regarding general learning behavior and expertise development, Kok et al. (2023) found in a study on learners’ comprehension of video lectures that the gaze point of the students correlates with their score on a comprehension test. Hahn and Klein (2022) conclude that in physics education, most studies point to a link between visual attention and levels of performance. In a scoping review, C. W. Mayer et al. (2023) take stock that higher performing participants and experts focus their attention on relevant areas and also exhibit longer fixation durations. This matches the findings of Li et al. (2023), who indicate that experts show longer fixation durations, shorter saccade durations and focus more on areas relevant to the task. In addition, van Gog and Jarodzka (2013) establish that eye tracking can be useful in educational settings to adapt both instructions and the format of the task based on user characteristics. Strohmaier et al. (2020) expand these findings with regards to mathematics education and establish that eye tracking in their domain is more suitable for process documentation, mental representations and subconscious aspects of mathematical thinking. Regarding domain-specificity of the present study, Tunga and Cagiltay (2023) only found two studies related to business education. The first study by Emhardt et al. (2022) looked at the benefits of using the teacher’s gaze as a cursor in an online lecture vs. the mouse cursor to guide learners’ attention. They did not identify significant effects of the use of either cursor on learning, although students reported to prefer the gaze cursor, where they would see where their instructor looked. The second study by Winter et al. (2021) discusses

how expert gaze can be used to guide how novices look at business process maps, as these can be challenging to understand. They found that by using these expert paths as an example, novices can increase their understanding of the process maps.

Most eye tracking studies in teacher education do not explicitly focus on aspiring teachers for vocational business education (Beach & McConnel, 2019; Faiella et al., 2019; Langner et al., 2022; Schweinberger et al., 2023; Wyss et al., 2021). For example, Wyss et al. (2021) show differences in the professional vision of novices and experts. Brückner and Zlatkin-Troitschanskaia (2024b) find for the domain of business education that students' self-confidence affects their performance in the eye tracking study. In another study, they show that the visual attention of the participants during task solving depends on the study progress (Brückner & Zlatkin-Troitschanskaia, 2024a).

Based on the current body of research, there are indicators that link interaction and expertise. As a consequence, it seems relevant to gain insights into the interaction process itself and not only focus on the outcome. The findings support the application of eye tracking in diverse educational settings to assess visual attention, guide feedback, and inform strategy development from expert performance. To establish the use case for this study, the next section introduces the AI platform EDDA.

2.2 The AI platform EDDA

EDDA is an analytical AI platform that provides individualized feedback for lesson plans and materials (currently only available in the context of business education) (Pargmann, Leube, et al., 2024). The feedback is aimed at student teachers, teachers in training and more experienced teachers as well. The aim of EDDA is to support users in their lesson planning and to promote their ability to design targeted lessons. Users should further develop their AI literacy by critically examining the suggestions and possibly using them for their lessons. Through machine learning, EDDA classifies uploaded documents into 57 didactical planning dimensions and rates their integration. The algorithm was trained with a total of $N= 5,983$ domain-specific lesson plans, materials and coded textbook tasks. Examples of the reliability of the replication of human ratings through the algorithm for unknown data are shown in Table 1 (for an elaboration on the categories, see Section 4.4.2). Based on the classification results, EDDA provides feedback that defines the category, explains the rating result, suggests changes and offers questions for reflection.

Table 1: Metrics for the reproduction of human ratings through EDDA

Construct	Dynamic Iota Index	Krippendorffs Alpha	Percentage Agreement	Gwet's AC
<i>Complete action</i>				
Inform	.421	.312	.743	.588
Plan	.596	.098	.813	.762
Decide	.517	.147	.772	.687
Act	.512	.386	.799	.700
Check	.443	.519	.764	.538
Reflect	.819	.019	.926	.920
<i>Competences for sustainable change agents</i>				
Content knowledge	.659	.474	.826	.789

Network knowledge	.649	.631	.831	.787
Methodological competences	.816	.508	.911	.903
Social competences	.769	.282	.874	.862
Decision-making	.878	.196	.926	.922
Open and critical perspective	.857	.209	.921	.917
Reflexivity	.649	.670	.837	.786

Please note that currently, EDDA is only available in German. For transparency reasons, the following screenshots were translated using the browser's automatic translation feature. We chose the best translations for the following examples, but some inaccuracies in the translation may occur. Figures 3-5 were manually translated.

When starting the application, users are directed to the upload page where they can choose their lesson plan and up to four files of supplementary materials (Figure 1). They then select the areas they would like to receive feedback on. After accepting the data protection and privacy statements, the analysis can start. Once the analysis begins, a pop-up opens that shows a progress bar.

The screenshot displays the EDDA landing page with four main sections:

- 1. Uploading the design:** Instructions state the draft must be a .pdf file. It includes a file selection area for 'Educational design' with a 'Browse...' button and a file named 'OL_0024.pdf'. An 'Upload complete' button is visible.
- 2. Select the areas to be analyzed:** A list of six areas with checkboxes: Action orientation, Basic dimensions, Digitalisation, Motivation, Problem orientation, and Sustainability. All are checked.
- 3. Uploading complementary materials:** Instructions state all material must be uploaded as .pdf. It shows four material slots (Material 1 to Material 4). Material 1 has a file 'OL_0024_M.pdf' uploaded. Materials 2, 3, and 4 show 'No file selected'.
- 4. Start of analyses:** Instructions state to click 'Start analysis'. A checkbox is checked: 'I have explained the data processing statement and the Disclaimer read and agree to the use of the software and data processing.' An 'Analyses begin' button is at the bottom.

Fig. 1 Landing page (automatically translated)

Once the upload is finished, users are directed to an overview, where the results for all dimensions and principles are shown (Figure 2). From there they can navigate to the desired dimensions.

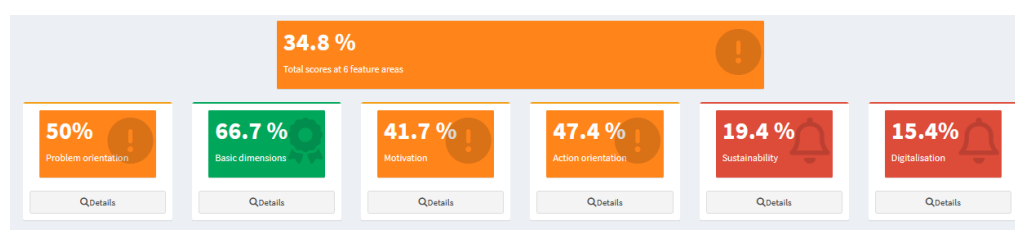


Fig. 2 Overview page (automatically translated)

The specific pages for the dimensions are always structured in the same way. At the top, the percentage for the dimension is given as the median across the sub-dimensions. Next, three images provide visual detail for the results (Figures 3–5). The result analysis shows the ratings for the different sub-dimensions, the coherence analysis shows the differences between the plan and the materials, and the security analysis shows how sure EDDA is about the ratings.

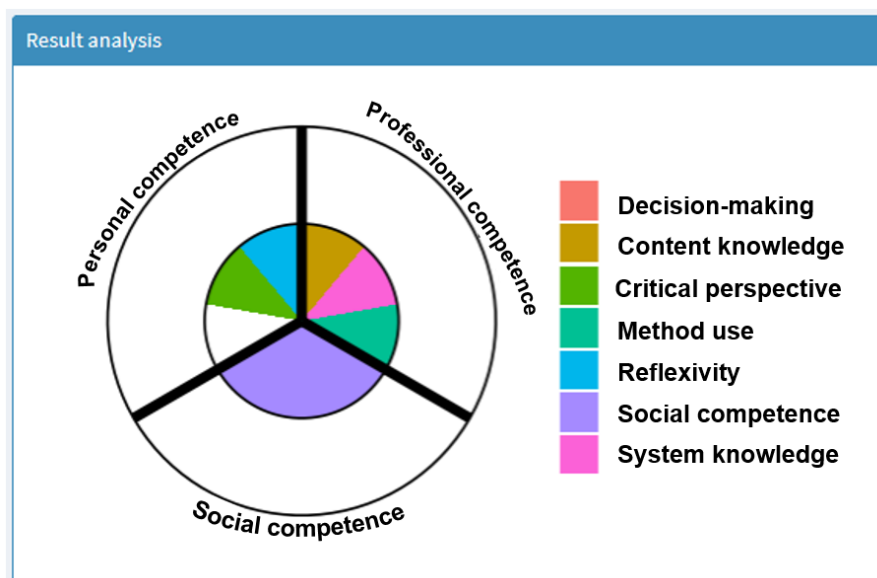


Fig. 3 Example for a result analysis

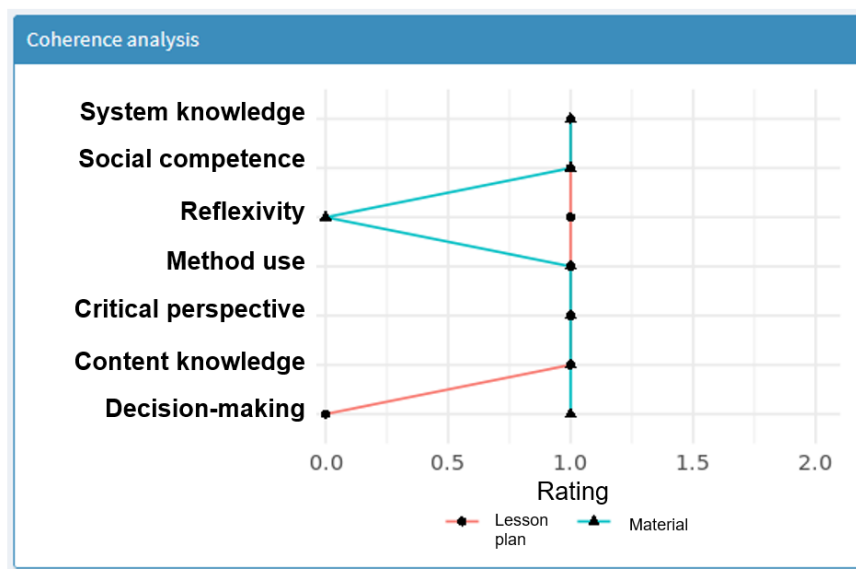


Fig. 4 Example for a coherence analysis

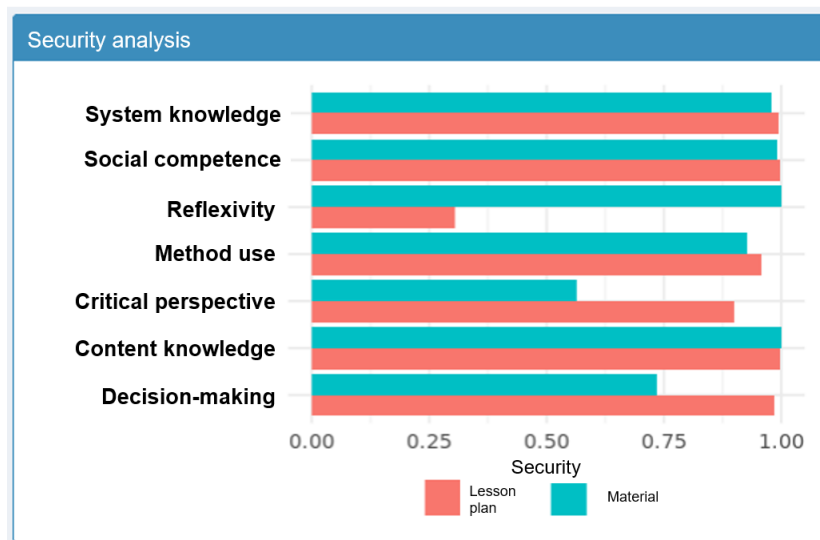


Fig. 5 Example for a security analysis

Following the visual results, textual feedback is given to the user. Here, users can decide which sub-dimensions they would like to read the feedback and open the respective box (Figure 6). Figure 7 shows an example for how feedback in EDDA is constructed. The feedback provides a short definition of the category/dimension, gives the rating and then reasons for the rating. After that, recommendations for improvement are given and supplemental sources can be used to gain a deeper understanding of the topic.

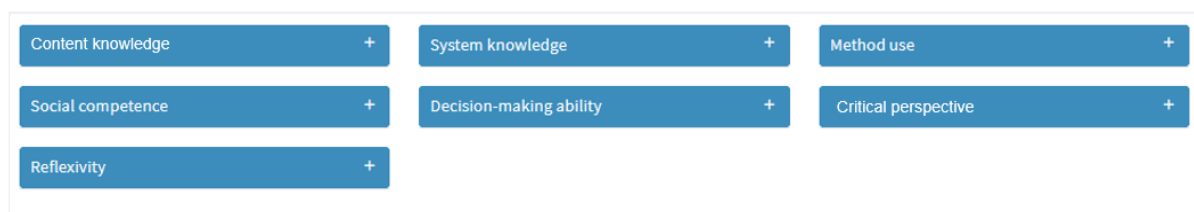


Fig. 6 Overview of feedback boxes

Decision-making ability

General description

The decision-making ability of learners plays an important role in promoting sustainable transformation, because it supports learners in making business decisions binding and representing them to other persons/stakeholder groups. They also need to take the room for manoeuvre to implement their decisions by decision-makers: to be able to use them. This may be associated with setbacks. It is therefore necessary that the learners already learn to make decisions in vocational training and to justify them in such a way that they are logically comprehensible to stakeholder groups and decision-makers.

My recommendations for you

Unfortunately, I cannot find such a promotion of decision-making in your draft teaching. This may be because the students either do not have to prepare or make decisions.

You can promote this important competence by integrating the opportunity for decisions in your tasks and problems if it is thematically appropriate. Due to the high relevance of decisions in professional action, this possibility can often be created if only small changes, such as more room for manoeuvre, are made.

In the second step, you can then build up references to sustainability, by the learners, for example, Make decisions about new sustainable products, marketing strategies, suppliers or controlling key figures and have to represent them to other persons, e.g. in discussions.

Do you need to plan a sustainability-oriented lesson in accounting or controlling? Here, too, I have some suggestions for you:

▼ Sustainability-oriented decision-making skills in accounting and controlling

General description

Controlling is the “economic conscience” of a company (Michel et al. 2014). However, they can also give more importance to sustainability in the company and argue the rationality of sustainability measures. They also need to adapt, develop or revise suitable key figures to manage the company's sustainability performance. For this purpose, it is necessary to anchor sustainability in the **company's mission** statement. This facilitates decision-making for controlling and management can strengthen environmental and social aspects. They make decisions about the key figures and valuation bases used, e.g. whether qualitative or quantitative environmental aspects such as CO2 emissions are taken into account. This is done by actual, plan and forecast scenarios, which require a pronounced decision-making ability (cf. 2014, Schaltegger 2020 and Schmitz 2021).

Notes for teaching

For controlling lessons, this means that the learners should deal intensively with different decision-making options and reflect on them. Try to anchor decision-making more closely in your lessons by giving rise to change using authentic materials such as e-mails, receipts, company data or survey results. Perhaps the company's sustainability strategy is also changing or there is dissatisfaction with the current key performance systems because they do not agree with the company's mission? Decisions can be changed or even obsolete by means of small changes in the scenario. Show this complexity to your students.

Further sources

Fig. 7 Example for feedback on “decision-making competence”

As explained above, students can freely choose which boxes they click on and which texts they would like to expand. Thus, a plethora of different interactions is possible. To examine whether this has an effect on the integration of didactical dimensions into their own lesson plans, an eye tracking study was performed.

3 Hypotheses

The theoretical context established above leads to the following two hypotheses for this eye tracking study:

Since EDDA provides individualized feedback opportunities, it seems likely that students interact differently with the software. Thus, we hypothesize that:

H1: There are different interaction types in the sample.

C. W. Mayer et al. (2023) and Hahn and Klein (2022) found that knowledge retention depends on the way users interact with a tool. Thus, we hypothesize that:

H2: Students who engage more with EDDA’s content receive higher ratings on their lesson plans.

4 Method

4.1 Participants

$N=32$ students ($n=15$ female, $n=17$ male) from a master’s program were recruited for this eye tracking study, which took place twice, once at the beginning (t_0) and once at the end (t_1) of the semester. Participation in the eye tracking study was voluntary. Ten of those 32 students participated in both settings, nine only in the first and 13 only in the second setting. Thus, there are 19 datasets for the first instance and 23 for the second, with a total of 42 datasets (22 datasets from those who only participated in one instance and 20 datasets from those who participated in both

instances). All participants study to become business education teachers. Approximately 81% of the participants have done an apprenticeship before studying. Two participants wore glasses. One participant was color-blind. The median age is 28.34 years (min=24, max=44).

4.2 Apparatus

In this study, we use the reporting guidelines for eye tracking studies by Dunn et al. (2024) to ensure a transparent procedure. The Smart Eye AI X portable eye tracking devices were mounted on the bottom side of the frame of laptops with a 33.6 cm by 59.8 cm screen and a 60 Hz refresh rate. The Smart Eye AI X is a video-based eye tracker that works with dark pupils and corneal reflections through infrared. Eye positions were recorded for both eyes. The sampling rate is 60 Hz, and the tracking accuracy is approximately 0.5 degrees. The latency is 25ms from camera exposure with a recovery time of one frame.

Participants sat in chairs and could move their upper bodies within certain limits. Using the software iMotions (Version 9.3.30921.6), the eye tracker and the screen recording were used to capture both the students' eye movement and all on-screen activities. Data was collected in a large windowless room with large fluorescent lighting from above. The devices were calibrated for every participant through the iMotions software, using a 9-point calibration with two cycles. The calibration background was grey, participants had to follow a large white dot across the screen with their eyes. Regarding calibration, in t0, 12 students had an excellent calibration result and seven were good. In t1, 18 were excellent, five were good.

Data quality was at an average of 82% in t0 (min=65%, max=95%) and an average of 88,25 % (min=64%, max=98%) in t1. Regarding the data processing, all data was transferred through an I-VT script for processing. The velocity threshold (I-VT) was at 30 degrees/second and the minimum fixation duration was set to 60ms. The participants sat at a comfortable distance of 70 cm away from the screen, as they had to still use the mouse and the keyboard. All recordings were supervised by the test administrator or a trained assistant.

4.3 Design and procedure

The eye tracking study was part of a semester-long intervention study (14 weeks) where students used the AI platform for feedback on their lesson plans. Eye tracking opportunities were provided twice during the semester, once at the beginning (second week) and once at the end (13th week). Both instances followed the same procedure: First, the experiment and the functionality of the eye tracking devices were explained to the students. This was necessary so they could make an informed decision whether or not they wanted to participate in the study. Study participation was voluntary. We also gave them a short introduction to the AI platform, so they were aware of the basic functionalities. They then signed a waiver, which explained the study, the data that was recorded and the data archiving policies. The students then imported a PDF of their lesson plans to the laptop and started the calibration. When the calibration was sufficient, the experiment started. Students saw a welcome message on the screen, explaining their task. They were supposed to upload their lesson plans to EDDA and to select the didactical principle. In t0, this was sustainability, in t1 it was digitalization, as these were the lesson plan dimensions that they worked on in the respective weeks. In t0, students had their first real interaction with EDDA, in t1 they already used it for feedback for the most part of the semester. Students were then asked to use EDDA for feedback on their lesson plans and interact with the platform like they normally would. The recording ended automatically when the browser window with EDDA was closed.

4.4 Data analysis

4.4.1 Analyzing the interaction types

To classify interaction types, the raw data was labelled with areas of interest (AOIs). AOIs are “defined regions in the stimulus that the researcher is interested in gathering data about” (Holmqvist et al., 2011, p. 187). Thus, the AOIs are not necessarily what a participant looks at, but rather what the researcher defines as interesting to analyze. Holmqvist et al. (2011) describe that expert-set AOIs can lack objectivity. This risk was counteracted in this study by pre-determining the AOIs, which were decided to be all boxes, menus and illustrations that could be interacted with. They were labelled and assigned to the video in post-production using iMotions (Version 9.3.30921.6). The AOI was activated once the appropriate screen was loaded. Dynamic AOIs were used to gather all fixations on the AOI, even when it was zoomed in or out, the participant used the scrollbar or boxes were interacted with. A total of 22 AOIs was determined, consisting of ten general AOIs, six sustainability-related AOIs and six digitalization-related AOIs (Figure 8, detailed AOIs). These were used for detailed interaction types for both t0 (sustainability) and t1 (digitalization). For a more general analysis, the 22 AOIs were then aggregated into seven (Figure 8, aggregated AOIs). We decided to aggregate the AOIs to allow for a more general analysis regarding the impact of interaction types on the lesson plans. This was necessary as two different topics were interacted with, including different AOIs.

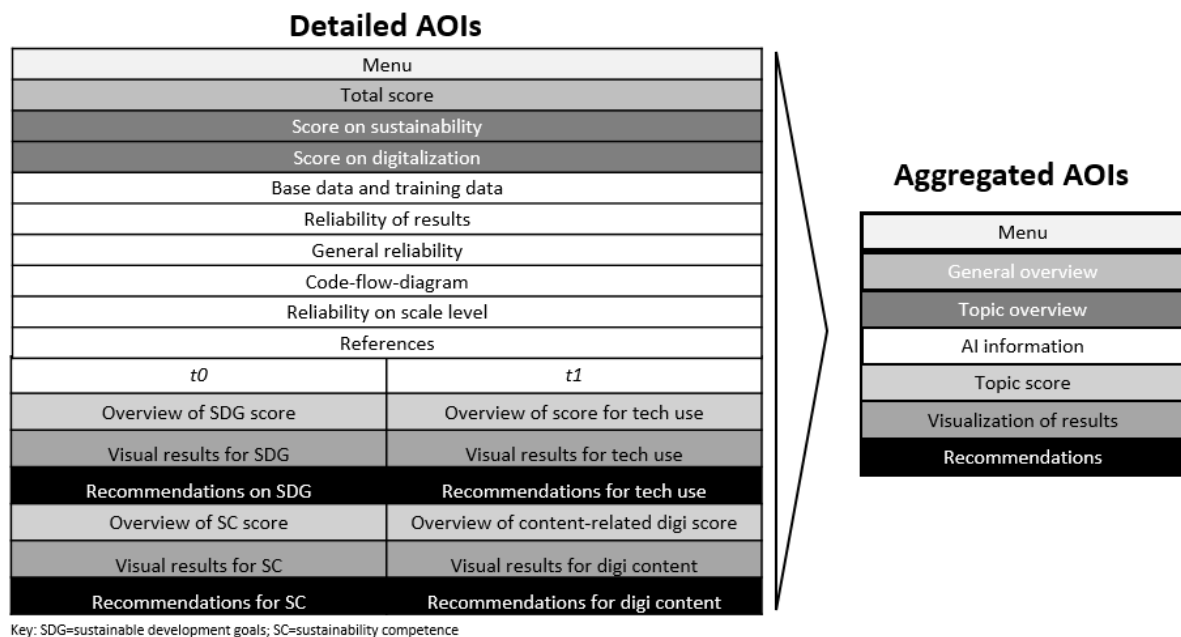


Fig. 8 Detailed and aggregated AOIs

To develop interaction types, we performed a sequence analysis, which can reflect sequential data in clustering tasks (Abbott, 1995). A latent class analysis, which is more common in educational science, was not suitable for this setting because it neither reflects the sequential aspect nor the duration of the interaction. Sequence analysis can be used with categorical sequence data that reflects states (Gabadinho et al., 2011). To perform the sequential analysis, we used the R package TraMineR by Gabadinho et al. (2011) with R (Version 4.4.3) in RStudio (Version 2024.12.1+563). The package provides a framework for the different steps of the sequence analysis. First, the data is strung into sequences. To reflect the fixation durations, we used the minimum fixation duration of the study

(60ms, see Section 4.2) and used it to create a ratio for the fixation on each AOI. A single sequence thus is the combination of all AOIs a participant looked at, weighted by the respective fixation duration. Then, we calculated a distance matrix that was then used for a cluster analysis using Ward's method of hierarchical grouping (Ward, 1963). The ward method has several advantages over other clustering methods: It minimizes variance within clusters, it is more robust against outliers and the clusters are easier to interpret because they are more homogeneous. It also outperforms other linking methods (Vijaya et al., 2019). We performed this analysis for the detailed AOIs separated by t0 and t1 and the aggregated AOIs. Since the number of clusters is supposed to help with the reconstruction of differences, there is no established formula that can be employed. We based our decisions on the dendrograms. For t0, the three-cluster solution produced the most consistent outcome. For t1, these were four clusters and for the aggregated analysis, we again chose three clusters.

After the sequence analysis, we performed a more descriptive analysis of the different types by accumulating the total seconds spent in an AOI across all participants, separated by type. By doing so, we can provide more insights into the different sequences and the characteristics of the interaction types.

4.4.2 *Analyzing the effects of the interaction types on the lesson plans*

After establishing the interaction types, we then aimed to establish the effects of each interaction type on the lesson planning process. The eye tracking study was embedded into a larger intervention study on the effects of EDDA on the lesson planning process Pargmann et al. (im Review). The students who participated in the eye tracking sessions also developed lesson plans for which they used and integrated EDDA's feedback. These lesson plans were coded once at the beginning and once at the end of the semester using the same coding manual that was used to rate EDDA's training material. The pre-evaluated manual included 57 didactical criteria and planning dimensions (subcategories) across ten main categories (Pargmann et al., im Review; Riebenbauer et al., 2024, 2025). These categories are:

- action orientation: the lesson promotes the ability to act through 1) authentic situations (Brown et al., 1989) (4 subcategories) and 2) by fulfilling the complete action cycle (Hacker & Sachse, 2014) (6 subcategories)
- process orientation: 3) connecting the content knowledge of different business areas and activities within these business processes (Stütz et al., 2022) (4 subcategories)
- problem orientation: 4) the creation or adaptation of cognitive schemes (Dörner, 1979; Piaget, 1977) (1 subcategories)
- motivation: 5) stimulation of the three dimensions for self-determined motivation (competence, relatedness, autonomy) (Deci & Ryan, 2000) (5 subcategories)
- digitalization: the inclusion of 6) content-related (Krämer et al., 2017) and 7) methodological digitalization (i.e., Bernard et al., 2009; Tamim et al., 2011) (6 and 4 subcategories, respectively)
- sustainability: 8) the promotion of competences for sustainable change agents (Berding et al., 2020) and 9) the inclusion of the sustainable development goals (SDG) (United Nations, 2015) (7 and 17 subcategories, respectively)
- basic dimensions: 10) learner activation (students actively engage with the topic), lesson and content structure (lesson phasing and structuring of the teaching process) as well as cross-linked learning (the development of connections between topics and contents) (Riebenbauer, 2021) (3 subcategories)

To analyze the effects of the interaction types on the categories of the lesson plans, we calculated a probit regression using the BAYES estimator with *mpls* (Version 8.11). We mainly used BAYES due to the small sample size, as it is more reliable in these situations (Hox & van de Schoot, 2013). The interaction types were included as variables in the regression using $k-1$ dummy variables with one interaction type as the reference type (Canela et al., 2019; Suits, 1957). The model fit was assessed with the posterior predictive p-value (PPP), which is a good indicator for model fit with BAYES (Depaoli et al., 2024). Generally, a PPP value around 0.5 points at an excellent fit for bigger samples and 0.15 for smaller samples (Cain & Zhang, 2019). No categories fell below the cut-off point in this study. Instead of a probit regression, another possible analysis would have been a variance analysis (ANOVA). Since the lesson plans were rated and we do not compare questionnaire data, calculating meaningful means is not possible.

5 Results

This section presents the results for the two analyses introduced in 4.4.1 and 4.4.2. First, the interactions of the students with the AI platform are analyzed through a sequence analysis to identify whether there are different interaction types or not. These results are presented in Section 5.1 for both instances, t_0 and t_1 , as well as an aggregated typology that is not specific to sustainability or digitalization. Then, in Section 5.2., we present the results for the probit regression using the total interaction type at the end of the semester.

5.1 *Results of the sequence analysis and analysis of the interaction types*

5.1.1 *Topic specific types for sustainability – t_0*

For t_0 with focus on sustainability, the sequence analysis produced three interaction types as the most viable result. Figure 9 shows bar diagrams for the representative sequences for all three types. 19 datasets were analyzed for this typology. We chose a coverage rate of 100% for all plots in this section to identify outliers. This results in a high number of representative sequences per type and instance, which are then analyzed for their characteristics. Type 1 and type 3 each only consist of three sequences; type 2 is the most common with 13 sequences. The number of sequences indicates the number of students for each type. In Figure 9, the legend for the Y-axis indicates how many representative sequences were identified, the number in the brackets indicates the total number of students in each type. The height of the bar indicates the representativeness of the sequence relative to the total number of students in each type. For example, in type 1, there are three cluster members and three representative sequences. Thus, one third of the cluster belongs to each sequence. In contrast, type 2 has 13 cluster members and 13 representative sequences. Thus, one bar only represents $1/13^{\text{th}}$ of cluster members, hence the smaller bar.

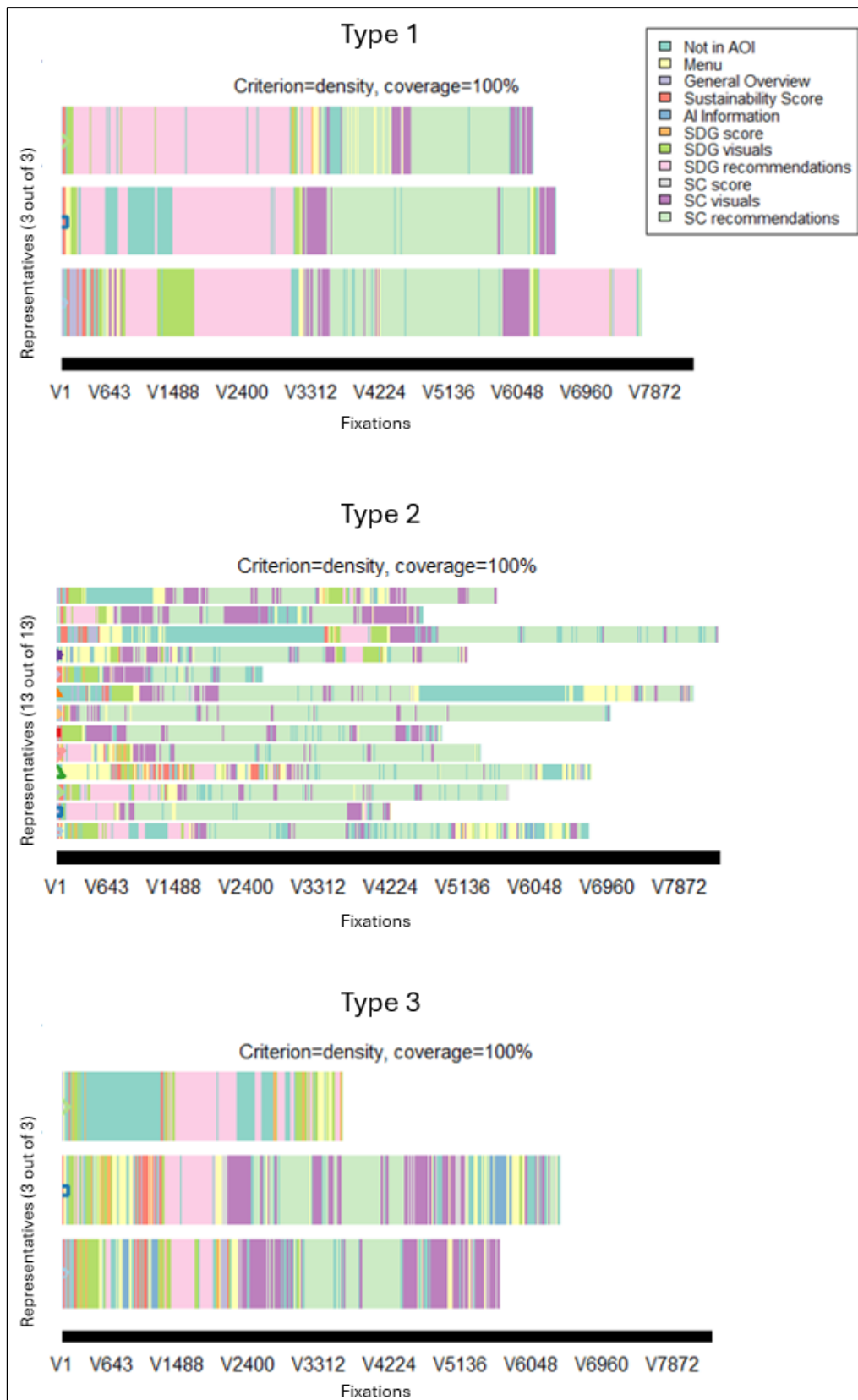


Fig. 9 Representative types for t0

The sequences for type 1 show a similar structure: They start out with general information on the AI platform but move onto the SDG score quickly. After that, they spend longer time in the recommendations for the implementation of the SDG. After checking their score again, they move onto the visual results for the sustainability

competences (SC), before they spend longer periods with the recommendations for the SC. Overall, students of type 1 spend longer durations in both recommendation AOIs. This pattern applies to all three representative sequences for type 1.

For type 2, students again start out with the general results for their lesson plan. They then move onto visual results and the recommendations. However, the periods spent in the recommendations are shorter and the students are more focused on the SC recommendations. The SDG recommendations are only fixated at briefly. The characteristic pattern for this type is a longer time in the visual results for both SDG and SC, and a longer time in the recommendations for SC. The recommendations for SDG are not fixated at as much as with type 1.

Students of type 3 visit the visual results for SDG and SC more frequently and spend a shorter duration in the recommendations. They also fixate on the different scores and overviews more frequently and for longer periods of time. Thus, they prioritize the visual results and scores higher than the recommendations.

The distribution of average total fixation time in percent for every type in t0 is shown in Figure 10. The times in milliseconds that are referred to below can be found in Appendix A.

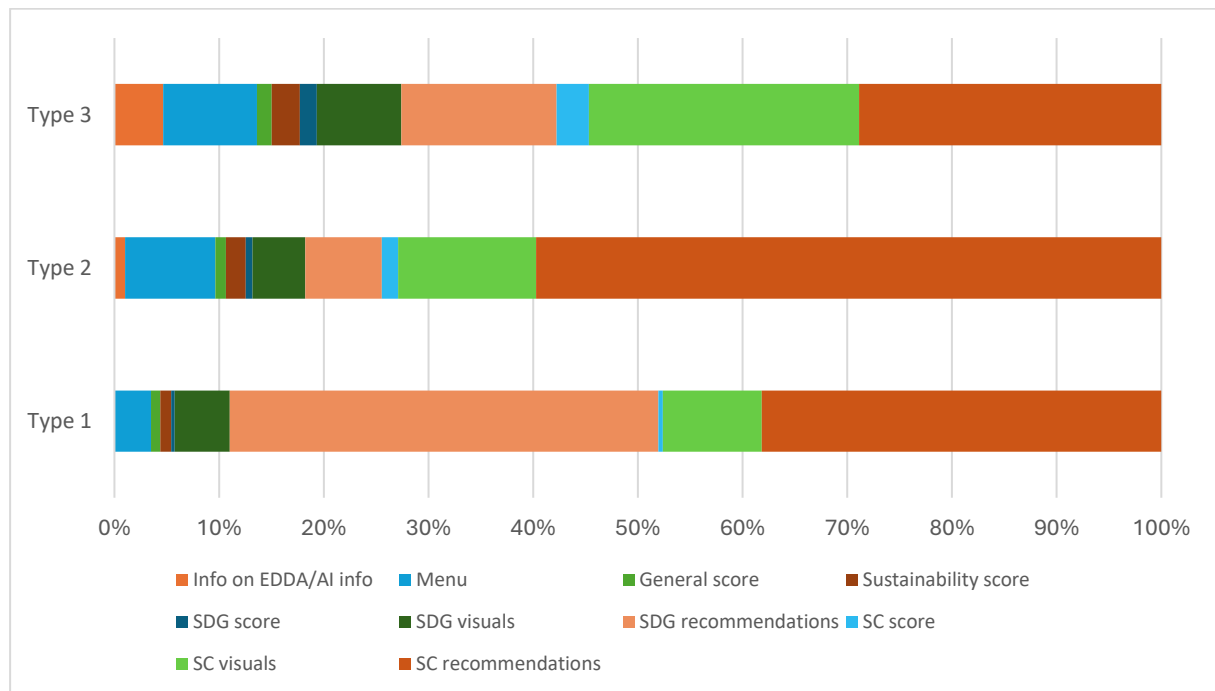


Fig. 10 Average total fixation duration in each AOI for three topic-specific types in t0 in percent

Considering the distribution, type 1 can be characterized through long fixations at both the SDG and SC recommendations. Roughly 79% of their fixation time can be assigned to the recommendations. In contrast, type 2 spends about 60% of their fixations at the recommendations for SC and only about 7% on the SDG recommendations. Type 3 only spends a total of about 44% of the time on recommendations. Regarding the menu, scores and overviews, type 1 spends the least time there, with only about 6.2% of their fixations, type 2 spends 13.7% of the time there and type 3 is the longest with 17.7%. Type 1 does not visit the information page on EDDA, type 2 and 3 do, with type 3 fixating there the longest.

There are clear overall tendencies for each type: While type 1 mainly focuses on the recommendations (79.12%), type 2 considers the visuals (18.23% in total) and only one recommendation site (59,71%). Type 3 focuses on the scores and visuals (52,08% in total) and spends the least time on the recommendations.

5.1.2 Topic specific types for digitalization – t1

The sequence analysis for digitalization (t1) proved four types to be the most appropriate result. Figure 11 shows an overview of the representative sequences. Twenty-three datasets were analyzed for this instance. Type 1 consists of six sequences, five cases belong to type 2, nine to type 3 and three to type 4.

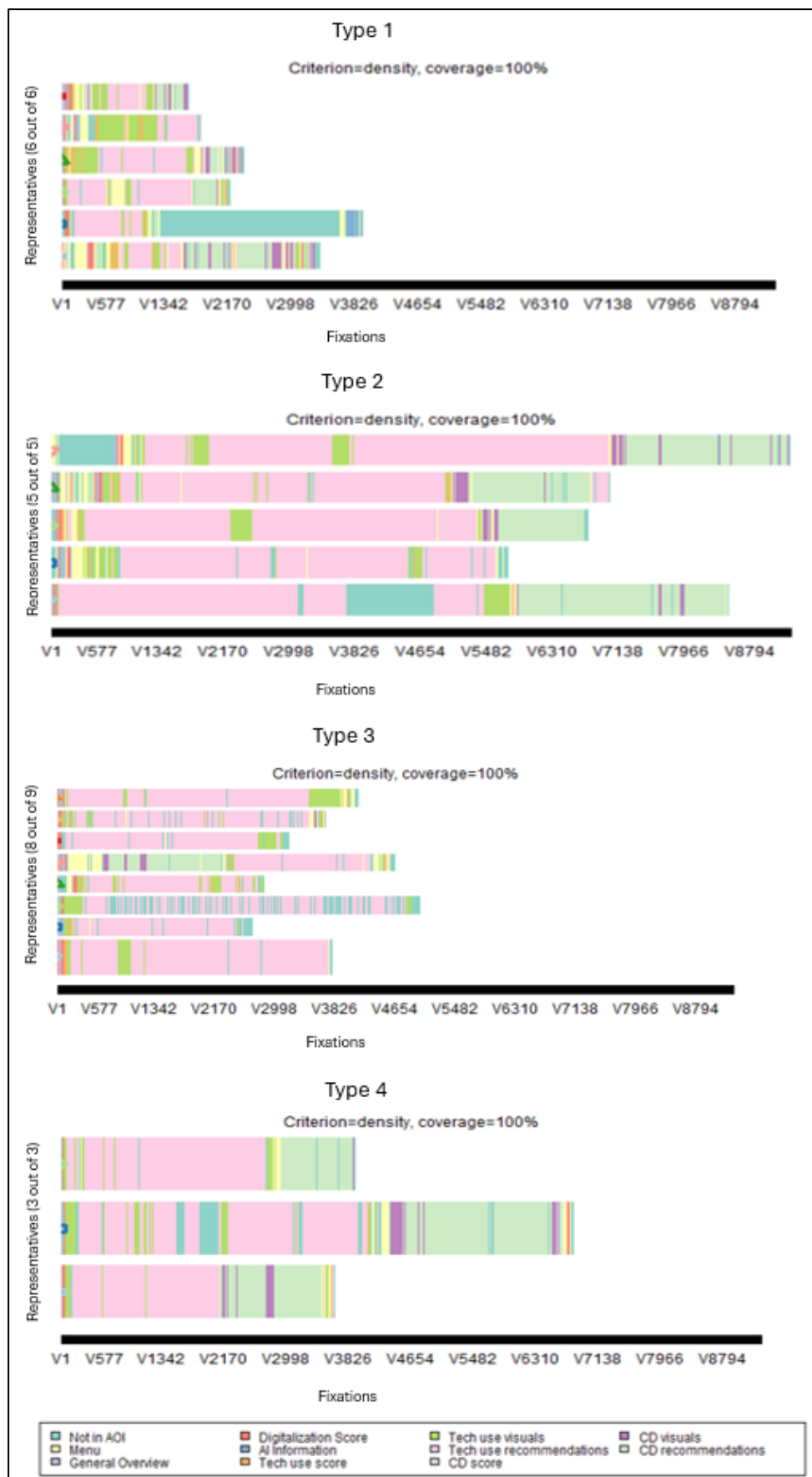


Fig. 11 Representative types for t1

Looking at the sequences for type 1, they all have in common that they focus on the recommendations for tech use. While they all start out with a general look at the menu and the overview, most of them move onto the visual results for tech use, before spending a longer time looking at the recommendations for that category. This seems to be the characterizing pattern for this type. Some sequences then proceed to look at the visual results for content-related digitalization but only briefly focus on the recommendations for that. There is one outlier sequence, where the student looks somewhere else than the AOI and then browses the AI information, however, this case still shares the initial characteristics of the type.

For type 2, students again briefly start out with the general overview, the menu and the scores and then move onto the visual results for tech use. They then fixate for a substantial amount of time on the recommendations for that type before moving on to the visual results and recommendations for content-related digitalization, switching between fixating on the visual results and the recommendations. Type 2 also shows the longest sequences compared to the other three types.

The representative patterns for type 3 show a focus on the overview and general digitalization score at the beginning, with a quick look over the visual results. Then, students fixate for longer durations on the recommendations for tech use. There is one outlying sequence where the student first fixates on the visual results and recommendations for content-related digitalization, but this person shares the initial interaction structure of the other sequences for the type and spends longer periods in the recommendations for tech use.

Regarding type 4, students of this type move on to the recommendations for tech use fairly slowly after looking at the visual results for that category for a longer duration. They then show long sequences of fixations on the recommendations, sometimes going back to the visual results. After that, they check the visual results for content-related digitalization and proceed to fixate on the respective recommendations for an extended period of time. The characteristic sequence for this type seems to be a check of the visual results and then a focus on the recommendations, sometimes checking back with the visual results in-between the recommendations.

The distribution of average total fixation time in percent for every type in t0 is shown in Figure 12. The times in milliseconds that are referred to below can be found in Appendix B.

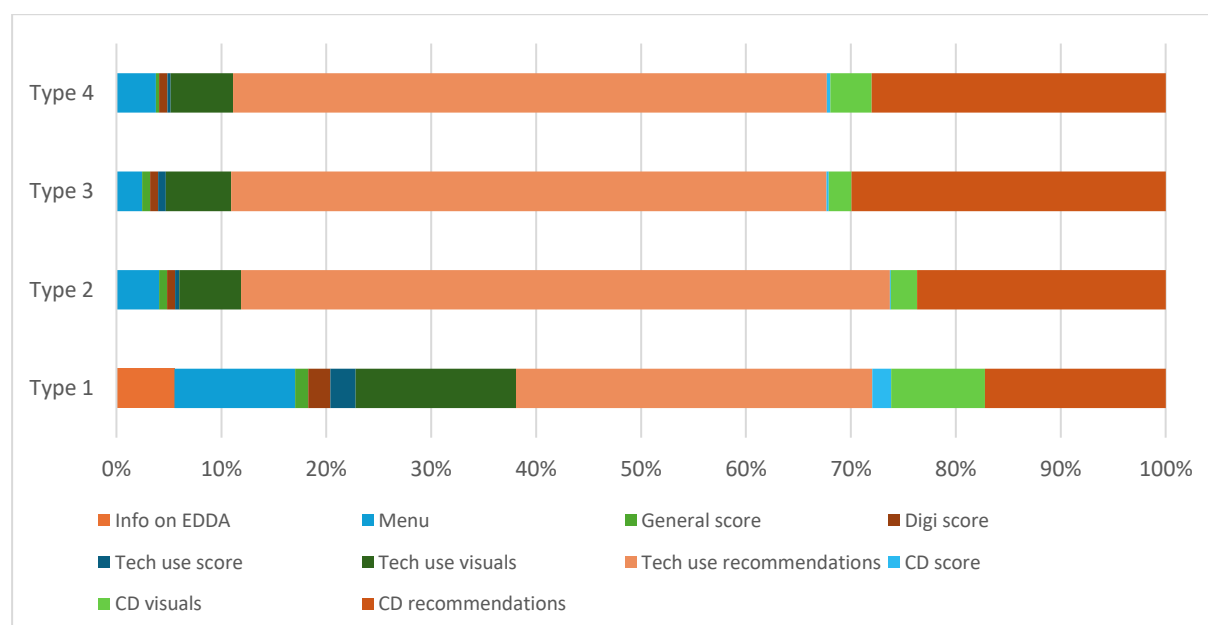


Fig. 12 Average total fixation duration in each AOI for the three topic-specific types in t1 in percent

Considering Figure 12, Type 1 spends the least amount of time fixating on the recommendations (51,18%), but the longest on the visuals (24,24%). All three other types spend roughly 85% of their fixations on the recommendations and only around 8–9% on the visual results. Thus, type 1 seems the most balanced between the different AOIs. Type 1 is also the only type that considers AI information (5,54%). They also fixate the longest on the menu (11,49%) and on the scores and overviews (7,55%). In addition, Type 1 focuses more on the recommendations and visuals of tech use. This is a general trend among the different types, as all of them spend more time fixating on the AOIs in that subcategory compared to content-related digitalization. Type 2 in t1 spends most of the time on recommendations (85,5%), of which there is a clear focus on the tech use recommendations. In that sense, the other three types spend a comparable amount of time in the different AOIs, when aggregated for AOI type. For these interaction types, the difference is in the sequence (Figure 10), which indicates different degrees of switching between visuals and recommendations. Considering the outlier in Type 3, it seems viable that the large share of recommendations on content-related digitalization (29,96%) can be assigned to that outlier, reducing the overall capacity type 3 has on content-related digitalization. Type 2, in contrast, differs from type 3 and 4 through the recurrent fixations on the visual results of tech use, which accumulate to a similar percentage but are concentrated more heavily throughout the sequences. The fixations on the visuals for tech use in type 4, for example, mostly happen right at the beginning of the sequence. Type 2 also focuses 5% longer on the recommendations for tech use than the other types, but also about 5% less on the recommendations for content-related digitalization.

Overall, type 1 seems to focus more on the visuals and the scores, while the other three types are more focused on the recommendations. It is also the only type that fixates on the AOI information. The remaining three types vary in the order in which they fixate on the different AOIs and how often they refer back to the visual results and score. While type 2 and 4 each consider both subcategories, the high fixation percentage on content-related digitalization for type 3 seems to stem from an outlier, that was assigned to the type due to the remainder of their interaction. Thus, type 3 focuses on the overview first and then proceeds to only one subcategory's recommendations.

5.1.3 *Aggregated interaction types*

In addition to this separate sequence analysis, we now share the results of the aggregated analysis, where we only differentiated between AOI type and not between subcategories. Thus, all recommendations, visuals and scores of t0 and t1 were assigned to the same AOI label. The interaction types for this aggregated sequence analysis are presented in Figure 13. Similarly to t0, we chose a three-cluster-solution, with 12 sequences assigned to type 1, 17 to type 2 and 13 to type 3. In this instance, the types are more evenly distributed than in the detailed sequence analyses.

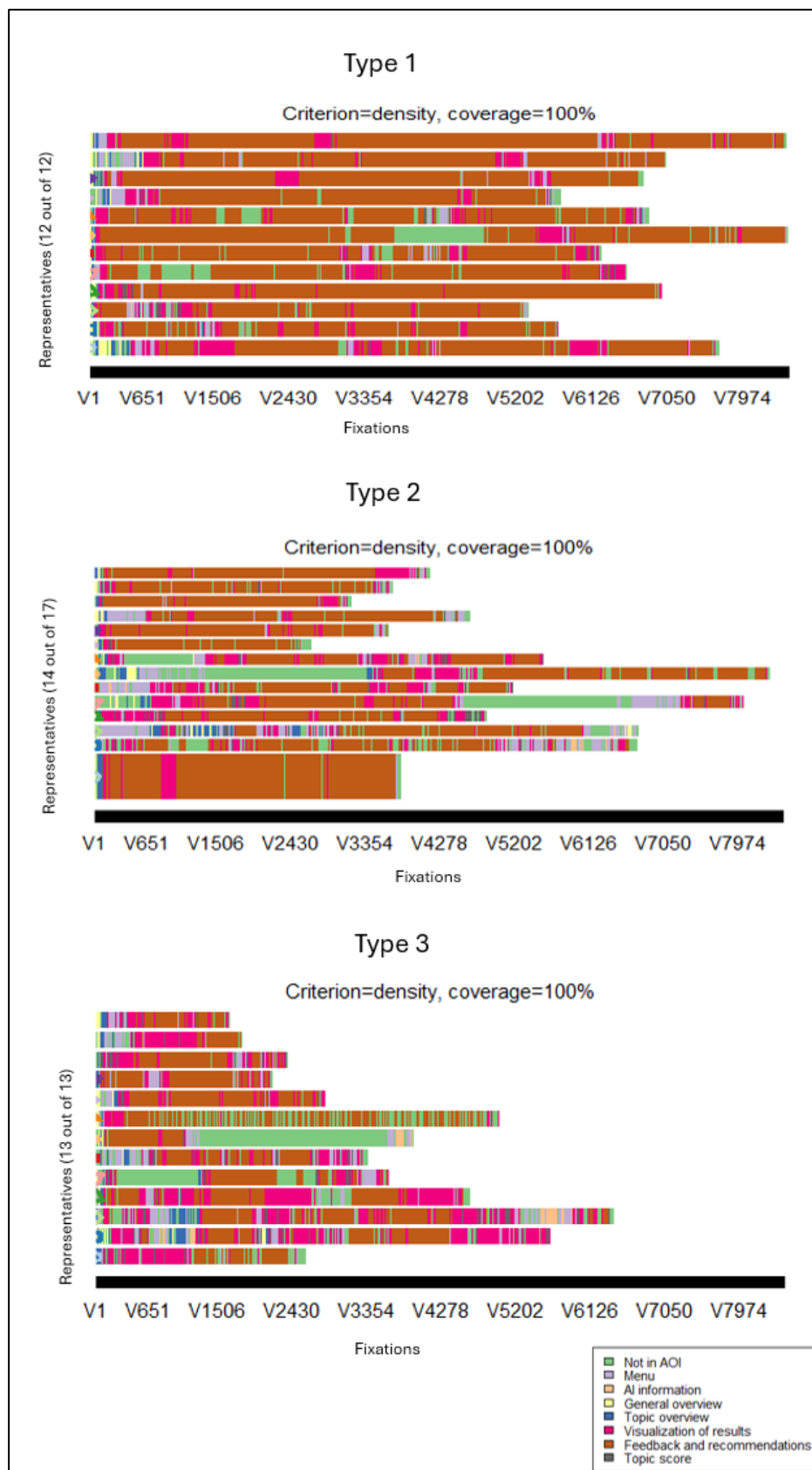


Fig. 13 Representative types for aggregated interaction types

The sequences that are assigned to type 1 start out with the general overview but move on to the visual results relatively quickly. Then, longer times are spent fixating on the recommendations. The visual results are fixated at

in-between the different recommendation AOIs, but the proportion of the visual results is relatively small compared to the recommendations. The sequences in this type also seem to refer to multiple feedback dimensions, as the sequence goes visual – recommendation – visual – recommendation. For type 2, this finding is more varied. While the representative sequences start out similarly to type 1 and the visual results are also considered, there are fewer switches between the visual results and the recommendations. Students of type 2 refer back to the menu, the overview or the topic results more frequently than type 1. The representative sequences of type 2 are also generally shorter and show longer periods in the visual results, compared to type 1. The representative sequences for type 3 show a similar, but even more detailed result: Here, the visualization of results is fixated on even longer and more frequently than in type 2. While students also fixate on the recommendations, there are larger chunks that indicate a focus on the visual results. Fixations on the scores and overviews are relatively short, except for the topic scores (formerly scores on SDG, SC, tech use and content-related digitalization), which are referred to in-between the fixations on the visual results.

The distribution of average total fixation time in percent for every type of the aggregated kind is shown in Figure 14. The times in milliseconds that are referred to below can be found in Appendix C.

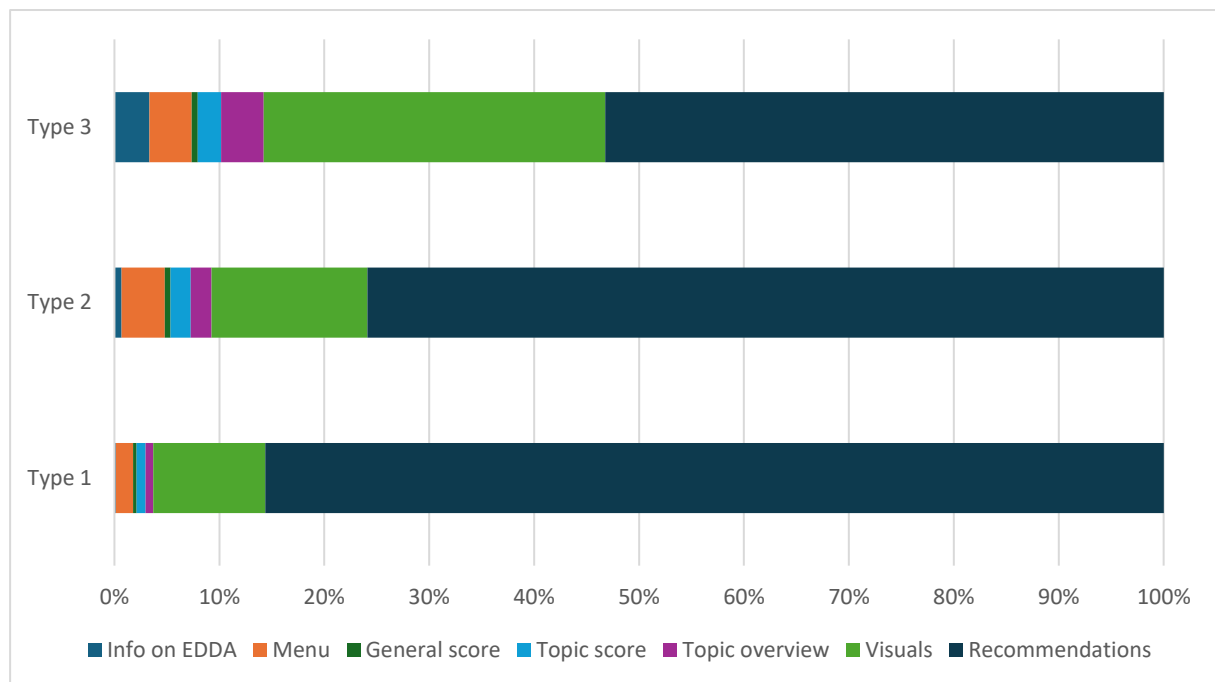


Fig. 14 Average total fixation duration in each AOI for the three aggregated types in percent

Type 1 is dominated by the fixation on the recommendations – students spend about 85% of their fixations there. The visuals are not fixated at as much, they only make up about 11%. The least relevant to type 1 seem to be the scores and overviews, as these only make up 1,92% of the fixations. Type 1 also does not fixate much on AI information. Type 2 spends less time in the recommendations (75,88%) and more time fixating the visual results (14,88%). This type also fixates much longer on the scores and overviews (4,44%) and the menu (4,13%). Type 3 fixates even less on the recommendations (53,22%) and even longer on the visuals (32,54%) and the scores and overviews (6,89%). This type also fixates the most on the AI information.

In summary, type 1 focuses the most on the recommendations, as also indicated by the representative sequences. Interestingly, when looking at the sequences in Figure 11, it looks like type 1 and type 2 both focus the most on the recommendations. However, the detailed analysis shows that type 1 indeed fixates longer on the recommendations whereas type 2 focuses more on the visuals. This aligns with the findings from the sequence analysis for the aggregated types (Figure 13), in which the students from type 1 consider the recommendations more frequently as compared to type 2, where these fixations are broken up more by the visuals, the scores and the menu. Lastly, for type 3, the sequence analysis in Figure 11 indicated a stronger focus on the visuals and scores, which proves to be true when looking at the percentages of the total fixation durations.

As a consequence of these analyses, we accept H1 that there are different interaction types in the sample. For each instance and for the aggregated analysis, there are distinct interaction types. In addition, the interactions became more nuanced throughout the semester, indicating increasing proficiency in the use of the technology.

5.2 Effects of the interaction types on the lesson plans

In this section, we present the effects of the interaction types on the lesson plans that were developed by the students throughout the semester. These were rated with a coding manual based on the categories introduced in Section 4.4.2. To identify whether or not the interaction types have an effect on the ratings of the lesson plans, we now present the results of the probit regression. We used the **aggregated types** from the previous section for these calculations. In the following, the regression results are presented by category (Section 4.4.2).

Sustainability

Sustainability is conceptualized on a content level through the *SDG* and on a competence level through *competences for sustainable change agents*. Table 2 provides an overview of both subcategories, the PPP values as well as R squared. Base indicates the development of the lesson plans, indicating that those who started off the semester on an already high rating did not show as much growth. These values are only provided for the sake of completeness but not interpreted. Type 2 and type 3 are the dummy variables that will be interpreted in this section.

Table 2. R-squared and standardized probit regression results for sustainability

Variable	PPP	R ²	Base	Type 2	Type 3
<i>Sustainable Development Goals</i>					
SDG8 – Decent work and economic growth	0.522	0.737	-0.722*	-0.418*	-0.071
SDG9 – Industry, innovation and infrastructure	0.541	0.359	-0.554*	-0.090	0.068
SDG12 – Sustainable consumption and production	0.538	0.634	-0.670*	0.117	0.163
SDG13 – Climate action	0.525	0.761	-0.535*	-0.365	0.167
SDG15 – Life on land	0.533	0.518	-0.127	0.038	-0.653*
<i>Sustainable Change Agent Competences</i>					
Content knowledge	0.490	0.871	-0.701*	-0.068	0.241
Network knowledge	0.529	0.810	-0.883*	-0.039	-0.036
Methodological competences	0.478	0.909	-0.925*	-0.024	-0.003
Social competences	0.482	0.897	-0.933*	0.043	0.063
Decision-making	0.505	0.869	-0.889	0.243**	0.220
Open and critical perspective	0.498	0.892	-0.883*	-0.139	0.081
Reflexivity	0.529	0.449	-0.433*	0.169	0.454*

Note: $N=32$, $*=p<0.05$; $**=p<0.10$. PPP cut off is at 0.15; no changes for SDG 1, 2, 3, 4, 5, 6, 7, 10, 11, 14, 16 and 17. Base = starting level.

Type 2 has a significant ($p<0.05$) negative effect on the implementation of *SDG 8* when compared to the effect of the reference type of the dummy variable. This indicates that the differences in the lesson plans between students who belong to type 1 and those who belong to type 2 are statistically significant, and that type 1 performs better than type 2. This also applies to the effect of type 3 on the implementation of *SDG 15*, where there is a significant ($p<0.05$) negative effect compared to the reference type.

Two dimensions of sustainability competences are influenced by the interaction type, which are *decision-making competence* and *reflexivity*. Compared to the reference type (type 1), type 2 has a significant ($p<0.1$) positive effect on *decision-making competence*, indicating that students who belong to this type receive statistically significant higher ratings than those of type 1. Similar is true for type 3 on the dimension of *reflexivity*, where there is also a significant ($p<0.05$) positive effect compared to the reference type.

Action orientation, process orientation and problem orientation

These categories represent didactic criteria that are important dimensions for lesson plans in vocational business education. Table 3 shows the regression results for these categories. Interaction type 2 has a significant positive effect ($p<0.1$) on *reflect*, the last phase of the complete action cycle, where students are supposed to reflect on their choices in the learning process. This finding indicates that students who belong to type 2 differ from the reference type in a statistically significant way regarding this phase. The opposite applies for the results on *context*, which reflects the learning place for which the situation and methods in the lesson plan are authentic for. The significant ($p<0.05$) negative result indicates that students with interaction type 2 integrate this subcategory significantly worse than those in type 1. Both type 2 and type 3 have a significant ($p<0.05$) positive effect on *process orientation*, indicating that both types perform significantly better than interaction type 1. Lastly, type 3 has a significant ($p<0.1$) negative effect on the category *information streams*, indicating that type 1 performs significantly better than type 3.

Table 3. R-squared and probit regression results for action orientation, process orientation and problem orientation

Variable	PPP	R ²	Base	Type 2	Type 3
<i>Action orientation</i>					
Mental planning	0.503	0.647	-0.880*	-0.131	-0.262
Actual execution	0.455	0.863	-0.876*	0.085	0.094
Kind of action	0.500	0.727	-0.858*	-0.205	-0.227
Inform	0.439	0.863	-0.895*	0.028	0.065
Plan	0.501	0.756	-0.805*	-0.086	0.191
Decide	0.440	0.707	-0.648*	0.034	0.281
Act	0.493	0.762	-0.865*	-0.081	-0.173
Check	0.438	0.845	-0.885*	0.041	0.051
Reflect	0.535	0.592	-0.666*	0.231**	-0.085
Context	0.488	0.621	0.155*	-0.402*	0.236
<i>Process orientation</i>					
Process orientation	0.504	0.701	-0.261	0.667*	0.615*
Value streams	0.509	0.682	-0.787*	-0.068	-0.044

Financial streams	0.539	0.210	-0.308**	-0.153	-0.013
Information streams	0.444	0.761	-0.727	0.116	-0.271**
<i>Problem orientation</i>					
Problem orientation	0.470	0.869	-0.902*	-0.116	-0.048

Note: $N=32$, $*=p<0.05$; $**=p<0.10$. PPP cut off is at 0.15; Base = starting level.

Motivation

A high degree of self-determination is supposed to increase the chance for qualitative motivation. There are no significant differences between the interaction types regarding the implementation of the dimensions *competence* and *relatedness*. The dimension of autonomy was further specified into three subcategories, of which only one is significantly impacted by an interaction type. The dimension *contains problems* means that the tasks contain a challenge that students have to structure by themselves, and students have to solve it autonomously. The results for this category are presented in Table 4.

Table 4. R-squared and probit regression results for motivation

Variable	PPP	R ²	Base	Type 2	Type 3
Contains problems	0.441	0.692	-0.687*	-0.188	-0.556*
Multiple solution strategies	0.499	0.688	-0.756*	-0.152	-0.044
Multiple solutions possible	0.522	0.377	-0.498*	-0.172	-0.076
Competence	0.510	0.616	-0.654*	-0.217	-0.038
Relatedness	0.509	0.795	-0.868*	0.046	0.043

Note: $N=32$, $*=p<0.05$; $**=p<0.10$. PPP cut off is at 0.15; Base = starting level.

Type 3 has a significant ($p<0.05$) effect on this subcategory, which means that type 3 includes this category worse than reference type 1 in a statistically significant way. Thus, type 1 promotes the implementation of this subcategory better than type 3.

Digitalization

In vocational business education, digitalization needs to be considered on two levels: on the methodological level that describes the way technology is used and on a content level that describes the developments from the vocation. The results for both dimensions are presented in Table 5.

Table 5. R-squared and probit regression results for digitalization

Variable	PPP	R ²	Base	Type 2	Type 3
<i>Methodological digitalization</i>					
Technology use	0.473	0.931	-0.944*	0.061	0.072
Engagement	0.486	0.709	-0.783*	-0.105	-0.425*
<i>Content-related digitalization</i>					
General technologies	0.513	0.486	-0.507*	-0.279**	-0.570*
Verbalization	0.509	0.553	-0.490*	-0.185	-0.669*
Topic	0.506	0.688	0.020*	-0.264*	-0.773*

Note: $N=32$, $*=p<0.05$; $**=p<0.10$. PPP cut off is at 0.15; Base = starting level.

Regarding methodological digitalization, there is a significant ($p<0.05$) negative effect for type 3 on the subcategory *engagement* (i.e., that students engage actively with technology rather than consuming content). This shows

that the reference type integrates this category better than type 3. On the other level of digitalization, both type 2 and type 3 show a significant ($p < 0.1$ and 0.05 , respectively) negative effect on *general technologies*, indicating that the reference type implements these technologies better. In addition, there is another significant ($p < 0.05$) negative effect of type 3 on *verbalization* (i.e., verbalizing information from technologies), thus, type 1 also integrates this dimension better. The same applies to the significant ($p < 0.05$) effects of type 2 and type 3 on *topic* (i.e., using technology to structure the learning process or derive learning needs), which students who belong to type 1 implement significantly better.

Basic dimensions

For the last category, the basic dimensions, the regression shows two significant results (Table 6).

Table 6. R-squared and probit regression results for the basic dimensions

Variable	PPP	R ²	Base	Type 2	Type 3
Lesson and content structure	0.472	0.914	-0.949*	0.013	-0.230**
Learner activation	0.486	0.872	-0.933*	0.011	-0.174
Cross-linked learning	0.529	0.645	-0.669*	-0.265**	-0.105

Note: $N=32$, $*=p < 0.05$; $**=p < 0.10$. PPP cut off is at 0.15; Base = starting level.

Type 3 has a significant ($p < 0.1$) negative effect on *lesson structure* (i.e., following a clear structure), thus, the reference type 1 integrates this dimension significantly better than type 3. In addition, type 2 also has a significant ($p < 0.1$) negative effect on the subcategory *cross-linked learning*, which describes the connection of the topic to other relevant aspects of business or management. This result indicates that type 1 implements this category significantly better than type 2.

Regarding H2, which stated that students who engage more with EDDA receive higher ratings in their lesson plans, we can hesitantly accept the hypothesis. In Section 5.1 we concluded that both types 1 and 2 engage more deeply with the recommendations. The results show that type 1 clearly shows the best results, but type 2 performs similarly well, as everywhere where type 3 is worse than type 1, type 2 is at least as good as type 1, as there is no significant positive or negative result. This could point at a connection between the intensity of the interaction with the recommendations and the number of significant effects.

As most significant results indicate that either students who follow type 1 or type 2 achieve the highest ratings, we can to some extent accept the hypothesis. However, there are still many categories where no type outperformed the others, indicating no effect of the unique interaction type.

6 Discussion

This study investigated how students differ in their interactions with technology as a feedback medium for lesson planning by the example of the analytical AI platform EDDA. It was examined how these interaction types change throughout the semester and whether the interaction types impact the implementation of didactical criteria and planning dimensions into lesson plans. To achieve this, we employed eye tracking technology that recorded the interaction with the AI platform. We calculated a sequence analysis and a probit regression to analyze the data.

Our results indicate that there are indeed distinct interaction types that can be differentiated in both individual settings and in an aggregated setting. At the beginning of the semester in t0 with focus on sustainability, we chose

the three-cluster solution. Each cluster is characterized by a difference in interaction duration, frequency and intensity. Type 1 shows signs of being feedback-forward, meaning that they refer mostly to the written recommendations for improving their lesson plans. Type 2 can be interpreted as a somewhat balanced type, as they consider the visuals and then choose the feedback that they want to engage with more closely (sustainability competences, in this case). In contrast, type 3 spends the most time on scores and visual results, indicating that they are more of a results-forward type, who is less interested in concrete details for improvement. Besides this typology, there is another possible interpretation of the types in t0, which refers to the preferences in feedback presentation. Type 1 focuses mostly on the written feedback and type 2 poses a stark contrast with mostly visual representations. Type 3 is balanced between written and visual feedback. The issue with type 2 is that while the visuals show the results and allow them to subsume where improvements are needed, they do not include specific suggestions or elaborations. Thus, there are two directions for interpretation. First, the general preference of *feedback intensity*, to either receive input on possible improvements (type 1), to just see what was done well or not well (type 3) or using the visual results to derive where improvements are needed and then make an informed decision on what feedback to read (type 2). Second, the general preference for *feedback representation*, e.g. written (type 1), mixed (type 2) or visual (type 3).

Regarding the end of the semester, t1 with focus on digitalization, we identified four distinct interaction types. In this instance, type 1 shows a preference for scores and visual results (results-forward type), while type 2 and type 4 show a similar interaction strategy that shows a preference for the recommendations (feedback-forward type). Type 3 decides which feedback to read based on the results (balanced type). Similar to t0, the types can also be characterized by their preference for a specific feedback representation; type 1 focuses more on the visual feedback, type 2 and 4 focus more on the written text and type 3 focuses on both. Overall, the types for t0 and t1 show similarities in strategies regarding feedback intensity and feedback representation. However, at the second instance at the end of the semester, the balanced type from t0 is split up into two variations (type 2 and 4) of which one refers back to the visual results more than the other.

Through the aggregation of the fixations on the areas of interest of t0 and t1, we could detect a more general typology and validate if our conclusions from the individual types can be applied to a general typology. In the aggregated typology, three types were detected. Type 1 focuses more consequently on the written recommendations, while type 2 breaks up these fixations by checking back with the visual results. Similar to the topic-specific level, type 1 can be considered to be a feedback-forward type in this context, as they focus more intensively on the recommendations. Type 2 can be seen as a more balanced type, as students with this type check the visuals more frequently. In contrast to the topic-specific instances, the differences between the feedback-forward and the balanced type are smaller. In the general typology, type 3 shows a higher preference for the visuals, scores and overviews, indicating that this type is results-forward. The preferences regarding feedback representation can also be applied to the aggregated typology, however, it is not as distinct. Type 1 and 2 show signs of being slightly more focused on written representations, however type 2 refers back to the visual results slightly more often. Type 3 can be vaguely assigned to be more focused on the visual representations. While with the topic-specific types, there were clear distinctions regarding feedback representation, this does not hold true for the types on an aggregated level.

Regarding the question whether specific interaction types produce better lesson plans, we found that no interaction type outperforms the others on all planning dimensions. While interaction type 2 outperforms interaction type 1 in

three categories and interaction type 3 outperforms type 1 in two categories, type 1 performs best in eleven categories. Generally, interaction types 1 and 2 seem to produce the best outcomes. These are also the types that are more feedback-forward or balanced and seem to prefer written recommendations over the visual representations. As a consequence, students with a more balanced or feedback-forward interaction type tend to integrate more subcategories better than students with a rather results-forward type.

The findings of this study support the conclusion by C. W. Mayer et al. (2023) that higher performing participants focus their attention more on relevant areas for the task. In addition, in their meta-analysis on digital feedback, van der Kleij et al. (2015) found that elaborate feedback shows larger effects than corrective feedback. This aligns with our findings, as students who focus more on the recommendations also integrate the analyzed categories better in their own lesson plans. Our findings expand the current state of research as we add insights into vocational business education. We agree with Hahn and Klein (2022) that there is a link between visual attention and levels of performance, however in the context of this study, this link might also be between the preferred feedback representation and the performance. Strohmaier et al. (2020) indicate that eye tracking studies are needed which focus on the interaction itself. We contribute to this gap with our typology for AI feedback integration on lesson plans in teacher education. In addition, we expand the current research on the interaction with learning materials and learning platforms (Beach & McConnel, 2019) by establishing an effect of the interaction types between the integration of didactic criteria into lesson plans. This allows conclusions on the development of expertise as there are differences in the lesson plans between students who focus on their results and those who focus on elaborate feedback.

In summary, this study expands the research on the interaction and effects of analytical AI platforms on learning, as it adds valuable insights into the typology that produces the best results, which is an interaction strategy that focuses on the recommendations in terms of feedback-forward.

Limitations

The present study is not without limitations. Perhaps the biggest limitation is the sample size. According to Beach and McConnel (2019) eye tracking studies in educational science can often only be interpreted as exploratory, as the contexts usually are highly unique, and the sample sizes do not allow for representative interpretations. It would have been desirable to include more students in this study, however, eye tracking studies remain highly experimental for teacher education. This becomes clear when looking at the sample sizes of other relevant studies, such as N=35 in Brückner and Zlatkin-Troitschanskaia (2024a), N=43 in Winter et al. (2021) or N=28 in Wyss et al. (2021). Of course, there are also studies with significantly larger sample sizes, but our sample size of N=32 is well within range for other related studies.

Another limitation lies in the development of the lesson plans as students worked in groups. Thus, the independence of observation was violated. Technically, we could have applied a multilevel model for the probit regression. However, most groups only sent one person to the eye tracking study, thus, there was no eye tracking data for the group-level. This limits the informative value of our interpretations and conclusions, as the group dynamic could have had effects on the integration of the didactical criteria.

In addition, we cannot make any statements on the development of expertise, as we did not test the students before and after each eye tracking instance on what they retained from interaction. We only documented their expertise development through the lesson plans. We also did not interview the students or provide surveys on the

interactions. This would have been helpful to gain insight into their reasons behind interacting with certain areas of interest more than others and if they preferred written or visual feedback. Lastly, we did not gather data on the user experience or the usability of the platform. This information would have also helped with understanding the motivation behind the interaction. Further research activities could therefore address pre-post designs on AI use and in-depth interviews with users on their learning experience through AI feedback.

7 Conclusion

In this study, we showed that there exist distinct interaction types with the use of an analytical AI platform and thus the way they perceive feedback impact the lesson planning process in teacher education. Different interaction types focus on different building blocks of feedback and thus need different support in the implementation process. Those students with a feedback-forward interaction type focus on the recommended changes, while students with a results-forward type only focus on where they performed well or badly. Students with a balanced interaction type focus on both what needs to be improved and how it can be improved, indicating a more holistic understanding of feedback than the results-forward type. It therefore seems to be of great importance for individual student support that the AI platform offers different feedback formats and representations.

These results directly impact those who wish to develop similar software or would like to personalize existing AI platforms. Assessing feedback types can help instructors who wish to use electronic or AI-based feedback in their classes to understand their learners' needs. In this sense, instructors can integrate regular feedback practices that help students to develop a holistic understanding that meets empirically recognized feedback standards. Instructors can also integrate opportunities for reflection that help students to move towards a more balanced approach. On a more practical level, this could also mean to highlight relevant areas by using a spotlight technique, that strategically shows beneficial interaction sequences with a highlight around the respective area (Winter et al., 2021). Additionally, the software should motivate students to read all relevant areas of the feedback. This could be realized by providing students with a visual feedback preference with condensed feedback that focuses on the most urgent aspects.

References

- Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, 21(1), 93–113. <https://doi.org/10.1146/annurev.so.21.080195.000521>
- Beach, P., & McConnel, J. (2019). Eye tracking methodology for studying teacher learning: a review of the research. *International Journal of Research & Method in Education*, 42(5), 485–501. <https://doi.org/10.1080/1743727X.2018.1496415>
- Beach, P., & Willows, D. (2014). Investigating teachers’ exploration of a professional development website: An innovative approach to understanding the factors that motivate teachers to use Internet-based resources / Investigation de l’exploration par les enseignants d’un site Web. *Canadian Journal of Learning and Technology / La Revue Canadienne De L’apprentissage Et De La Technologie*, 40(3). <https://doi.org/10.21432/T2RP47>
- Berding, F., Slopinski, A., & Frerichs, R. (2020). Auszubildende als zukünftige Change Agents for Sustainable Innovations. *Betriebswirtschaftliche Forschung Und Praxis*, 72(3), 313–337.
- Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A., & Bethel, E. C. (2009). A Meta-Analysis of Three Types of Interaction Treatments in Distance Education. *Review of Educational Research*, 79(3), 1243–1289. <https://doi.org/10.3102/0034654309333844>
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, 18(1), 32. <https://doi.org/10.2307/1176008>
- Brückner, S., & Zlatkin-Troitschanskaia, O. (2024a). Effects of teacher students’ study progress on their gaze behavior while solving of an economics knowledge test. *Empirical Research in Vocational Education and Training*, 16(1). <https://doi.org/10.1186/s40461-024-00172-2>
- Brückner, S., & Zlatkin-Troitschanskaia, O. (2024b). The role of confidence in the gaze bias effect among economics trainee teachers — Results from a digital assessment of economic content knowledge. *Empirical Research in Vocational Education and Training*, 16(1), 1–20. <https://doi.org/10.1186/s40461-024-00156-2>
- Cain, M. K., & Zhang, Z. (2019). Fit for a Bayesian: An Evaluation of PPP and DIC for Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(1), 39–50. <https://doi.org/10.1080/10705511.2018.1490648>
- Canela, M. Á., Alegre, I., & Ibarra, A. (2019). Dummy Variables. In M. Á. Canela, I. Alegre, & A. Ibarra (Eds.), *Springer eBook Collection. Quantitative Methods for Management: A Practical Approach* (1st ed. 2019, pp. 57–63). Springer. https://doi.org/10.1007/978-3-030-17554-2_6
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*(11), 227–268.
- Depaoli, S., Winter, S. D., & Liu, H. (2024). Under-Fitting and Over-Fitting: The Performance of Bayesian Model Selection and Fit Indices in SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, 31(4), 604–625. <https://doi.org/10.1080/10705511.2023.2280952>
- Dörner, D. (1979). *Problemlösen als Informationsverarbeitung* (2. Aufl.). Kohlhammer.
- Dunn, M. J., Alexander, R. G., Amiebenomo, O. M., Arblaster, G., Atan, D., Erichsen, J. T., Ettinger, U., Giardini, M. E., Gilchrist, I. D., Hamilton, R., Hessels, R. S., Hodgins, S., Hooge, I. T. C., Jackson, B. S., Lee, H., Macknik, S. L., Martinez-Conde, S., Mcilreavy, L., Muratori, L. M., . . . Sprenger, A. (2024).

- Minimal reporting guideline for research involving eye tracking (2023 edition). *Behavior Research Methods*, 56(5), 4351–4357. <https://doi.org/10.3758/s13428-023-02187-1>
- Ebner, A., & Sageder, M. (2023). Künstliche Intelligenz im Kreditrisikomanagement - Akzeptanz und Herausforderungen bei der Einführung in österreichischen Banken. In T. A. Herberger (Ed.), *Digitale Transformation und Nachhaltigkeit in der Finanzwelt* (pp. 27–48). Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783748939245-27>
- Eklund, R., Sellberg, C., & Osvalder, A.-L. (2020). From Tacit Knowledge to Visual Expertise: Eye-Tracking Support in Maritime Education and Training. In S. Nazir, T. Ahram, & W. Karwowski (Eds.), *Advances in Intelligent Systems and Computing: Vol. 1211, Advances in Human Factors in Training, Education, and Learning Sciences: Proceedings of the AHFE 2020 Virtual Conference on Human Factors in Training, Education, and Learning Sciences, July 16-20, 2020, USA* (1st ed. 2020, pp. 269–275). Springer International Publishing; Imprint Springer. https://doi.org/10.1007/978-3-030-50896-8_39
- Emhardt, S. N., Jarodzka, H., Brand-Gruwel, S., Drumm, C., Niehorster, D. C., & van Gog, T. (2022). What is my teacher talking about? Effects of displaying the teacher's gaze and mouse cursor cues in video lectures on students' learning. *Journal of Cognitive Psychology*, 34(7), 846–864. <https://doi.org/10.1080/20445911.2022.2080831>
- Faiella, F., Mannese, E., Savarese, G., Plutino, A., & Lombardi, M. G. (2019). Eye-tracking glasses for improving teacher education: the e-Teach project. *Research on Education and Media*, 11(1), 85–92. <https://doi.org/10.2478/rem-2019-0012>
- Frenzke–Shim, A., Kohl–Dietrich, D., Standl, B., Langner, M., Maedche, A., & Kerber, U. (2024). The Evolving Role of Generative AI in Text Literacies: Exploring the Potential of Cognition-Adaptive Assistance. In *2024 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1–3). IEEE. <https://doi.org/10.1109/EDUCON60312.2024.10578629>
- Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37. <https://doi.org/10.18637/jss.v040.i04>
- Gegenfurtner, A., Gruber, H., Holzberger, D., Keskin, Ö., Lehtinen, E., Seidel, T., Stürmer, K., & Säljö, R. (2023). Towards a cognitive theory of visual expertise. In C. Damşa, A. Rajala, G. Ritella, & J. Brouwer (Eds.), *Re-theorising Learning and Research Methods in Learning Research* (pp. 142–158). Routledge. <https://doi.org/10.4324/9781003205838-10>
- Hacker, W., & Sachse, P. (2014). *Allgemeine Arbeitspsychologie: Psychische Regulation von Tätigkeiten*. Hogrefe.
- Hahn, L., & Klein, P. (2022). Eye tracking in physics education research: A systematic literature review. *Physical Review Physics Education Research*, 18(1). <https://doi.org/10.1103/PhysRevPhysEducRes.18.013102>
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van Weijer, J. de. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- Hox, J., & van de Schoot, R. (2013). Robust methods for multilevel models. In M. A. Scott, J. S. Simonoff, & B. D. Marx (Eds.), *The SAGE handbook of multilevel modeling* (pp. 387–402). SAGE.
- Jacobsen, L. J., & Weber, K. E. (2024). *Chatgpt & Co.: Feedback as a Use Case for Large Language Models in Higher Education [Preprint]*. <https://doi.org/10.13140/RG.2.2.17993.35685>
- Jarodzka, H., Holmqvist, K., & Gruber, H. (2017). Eye tracking in Educational Science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research*, 10(1). <https://doi.org/10.16910/jemr.10.1.3>

- Jarodzka, H., Skuballa, I., & Gruber, H. (2021). Eye-Tracking in Educational Practice: Investigating Visual Perception Underlying Teaching and Learning in the Classroom. *Educational Psychology Review*, 33(1), 1–10. <https://doi.org/10.1007/s10648-020-09565-7>
- Kok, E. M., Jarodzka, H., Sibbald, M., & van Gog, T. (2023). Did You Get That? Predicting Learners' Comprehension of a Video Lecture from Visualizations of Their Gaze Data. *Cognitive Science*, 47(2), e13247. <https://doi.org/10.1111/cogs.13247>
- Krämer, H., Jordanski, G., & Goertz, L. (2017). *Medien anwenden und produzieren. Entwicklung von Medienkompetenz in der Berufsausbildung*. (Wissenschaftliche Diskussionspapiere No. 181).
- Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, 10, 90–115. <https://doi.org/10.1016/j.edurev.2013.10.001>
- Langner, A., Graulich, N., & Nied, M. (2022). Eye-Tracking as a Promising Tool in Pre-Service Teacher Education—A New Approach to Promote Skills for Digital Multimedia Design. *Journal of Chemical Education*, 99(4), 1651–1659. <https://doi.org/10.1021/acs.jchemed.1c01122>
- Li, S., Duffy, M. C., Lajoie, S. P., Zheng, J., & Lachapelle, K. (2023). Using eye tracking to examine expert-novice differences during simulated surgical training: A case study. *Computers in Human Behavior*, 144, 107720. <https://doi.org/10.1016/j.chb.2023.107720>
- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023). Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2), 100017. <https://doi.org/10.1016/j.metrad.2023.100017>
- Mayer, C. W., Rausch, A., & Seifried, J. (2023). Analysing domain-specific problem-solving processes within authentic computer-based learning and training environments by using eye-tracking: a scoping review. *Empirical Research in Vocational Education and Training*, 15(1). <https://doi.org/10.1186/s40461-023-00140-2>
- Mayer, R. E. (2024). The Past, Present, and Future of the Cognitive Theory of Multimedia Learning. *Educational Psychology Review*, 36(1). <https://doi.org/10.1007/s10648-023-09842-1>
- McIntyre, N. A., Jarodzka, H., & Klassen, R. M. (2019). Capturing teacher priorities: Using real-world eye-tracking to investigate expert teacher priorities across two cultures. *Learning and Instruction*, 60, 215–224. <https://doi.org/10.1016/j.learninstruc.2017.12.003>
- O'Meara, P., Munro, G., Williams, B., Cooper, S., Bogossian, F., Ross, L., Sparkes, L., Browning, M., & McClounan, M. (2015). Developing situation awareness amongst nursing and paramedicine students utilizing eye tracking technology and video debriefing techniques: A proof of concept paper. *International Emergency Nursing*, 23(2), 94–99. <https://doi.org/10.1016/j.ienj.2014.11.001>
- Pargmann, J., Berding, F., Rebmann, K., & Riebenbauer, E. (in review). How feedback from an analytical AI platform helps lesson planning in teacher education: A longitudinal intervention study.
- Pargmann, J., Leube, A., Berding, F., Riebenbauer, E., Rebmann, K., Slopinski, A., & Gillert, M. (2024). Die elektronisch-didaktische Assistenz (EDDA) zur Unterrichtsplanung auf Basis künstlicher Intelligenz: Funktionsweise, Anwendungsbereiche und Forschungsperspektiven. *Empirische Pädagogik*, 38(1), 118–146. <https://doi.org/10.62350/WLQO3003>

- Pargmann, J., Riebenbauer, E., Leube, A., Berding, F., Slopinski, A., & Rebmann, K. (2024). Der Einsatz von künstlicher Intelligenz zur Unterrichtsplanung in der beruflichen Bildung – exemplarische Möglichkeiten mit der elektronisch-didaktischen Assistenz EDDA. *Lernen&lehren*, 39(153), 22–28.
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking.
- Rausch, A., Deutscher, V., Seifried, J., Brandt, S., & Winther, E. (2021). Die web-basierte Bürosimulation LUCA – Funktionen, Einsatzmöglichkeiten und Forschungsausblick. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 117(3), 372–394. <https://doi.org/10.25162/zbw-2021-0017>
- Riebenbauer, E. (2021). *Kompetenzentwicklung im Masterstudium Wirtschaftspädagogik. Längsschnittstudie zur Unterrichtsplanung im Rechnungswesen*. wbv.
- Riebenbauer, E., Rebmann, K., Berding, F., Leube, A., Pargmann, J., Slopinski, A., & Gillert, M. (2024). Zum Planungshandeln von Studierenden der Wirtschaftspädagogik - Teil 1: Orientierung bei der Planung mit Unterrichtsentwürfen. *Zeitschrift Für Berufs- Und Wirtschaftspädagogik*, 4.
- Riebenbauer, E., Rebmann, K., Berding, F., Leube, A., Pargmann, J., Slopinski, A., & Gillert, M. (2025). Zum Planungshandeln von Studierenden der Wirtschaftspädagogik – Teil 2: Empirische Befunde zur Analyse von Unterrichtsentwürfen. *Zeitschrift Für Berufs- Und Wirtschaftspädagogik*, 1.
- Schweinberger, M., Watzka, B., & Girwidz, R. (2023). Eye tracking as feedback tool in physics teacher education. *Frontiers in Education*, 8, Article 1140272. <https://doi.org/10.3389/educ.2023.1140272>
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Strohmaier, A. R., MacKay, K. J., Obersteiner, A., & Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics*, 104(2), 147–200. <https://doi.org/10.1007/s10649-020-09948-1>
- Stürmer, K., Seidel, T., Müller, K., Häusler, J., & S. Cortina, K. (2017). What is in the eye of preservice teachers while instructing? An eye-tracking study about attention processes in different teaching situations. *Zeitschrift Für Erziehungswissenschaft*, 20(S1), 75–92. <https://doi.org/10.1007/s11618-017-0731-9>
- Stütz, S., Berding, F., Reincke, S., & Scheper, L. (2022). Characteristics of learning tasks in accounting textbooks: An AI assisted analysis. *Empirical Research in Vocational Education and Training*, 14(1), 1–33. <https://doi.org/10.1186/s40461-022-00138-2>
- Suits, D. B. (1957). Use of Dummy Variables in Regression Equations. *Journal of the American Statistical Association*, 52(280), 548–551. <https://doi.org/10.1080/01621459.1957.10501412>
- Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What Forty Years of Research Says About the Impact of Technology on Learning. *Review of Educational Research*, 81(1), 4–28. <https://doi.org/10.3102/0034654310393361>
- Tunga, Y., & Cagiltay, K. (2023). Looking through the model's eye: A systematic review of eye movement modeling example studies. *Education and Information Technologies*, 28(8), 9607–9633. <https://doi.org/10.1007/s10639-022-11569-5>
- United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development*. <https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>

- van den Berg, G., & Du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness in Teacher Education. *Education Sciences*, 13(10), 998. <https://doi.org/10.3390/educsci13100998>
- van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes. *Review of Educational Research*, 85(4), 475–511. <https://doi.org/10.3102/0034654314564881>
- van Gog, T., & Jarodzka, H. (2013). Eye Tracking as a Tool to Study and Enhance Cognitive and Metacognitive Processes in Computer-Based Learning Environments. In R. Azevedo & V. Aleven (Eds.), *Springer International Handbooks of Education. International Handbook of Metacognition and Learning Technologies* (Vol. 28, pp. 143–156). Springer New York. https://doi.org/10.1007/978-1-4419-5546-3_10
- Vijaya, Sharma, S., & Batra, N. (2019). Comparative Study of Single Linkage, Complete Linkage, and Ward Method of Agglomerative Clustering. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (pp. 568–573). IEEE. <https://doi.org/10.1109/COMITCon.2019.8862232>
- Vijayan, K., Mork, O., & Hansen, I. (2018). Eye Tracker as a Tool for Engineering Education. *Universal Journal of Educational Research*, 6(11), 2647–2655. <https://doi.org/10.13189/ujer.2018.061130>
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Winter, M., Pryss, R., Probst, T., & Reichert, M. (2021). Applying Eye Movement Modeling Examples to Guide Novices' Attention in the Comprehension of Process Models. *Brain Sciences*, 11(1), 72. <https://doi.org/10.3390/brainsci11010072>
- Wyss, C., Rosenberger, K., & Bühner, W. (2021). Student Teachers' and Teacher Educators' Professional Vision: Findings from an Eye Tracking Study. *Educational Psychology Review*, 33(1), 91–107. <https://doi.org/10.1007/s10648-020-09535-z>

Appendix A. Average total fixation duration in each AOI for the three types in ms and percent for t0

AOI	Type 1	Type 2	Type 3
Info on EDDA/AI info	/ (0.00%)	3,548.66 (1.02%)	17,332.52 (4.65%)
Menu	15,176.39 (3.50%)	29,947.51 (8.61%)	33,307.17 (8.94%)
General score	3,798.18 (0.88%)	3,497.84 (1.01%)	5,328.32 (1.43%)
Sustainability score	4,538.42 (1.05%)	6,584.68 (1.89%)	9,957.90 (2.67%)
SDG score	1,457.18 (0.34%)	2,307.43 (0.66%)	6,087.86 (1.63%)
SDG visuals	22,793.07 (5.26%)	17,481.36 (5.03%)	30,066.76 (8.07%)
SDG recommendations	177,380.55 (40.95%)	25,443.59 (7.32%)	55,159.64 (14.81%)
SC score	1,835.5 (0.42%)	5,351.17 (1.54%)	11,547.91 (3.10%)
SC visuals	40,875.33 (9.44%)	45,904.07 (13.20%)	96,012.58 (25.78%)
SC recommendations	165,354.99 (38.17%)	207,616.91 (59.71%)	107,561.12 (28.89%)
Total	433,209.70 (100.00%)	347,683.21 (100.00%)	372,361.77 (100.00%)

Note: Percentages are calculated by dividing the average total fixation duration in each AOI by the total for each type.

Appendix B. Average total fixation duration in each AOI for the four types in ms and percent for t1

AOI	Type 1	Type 2	Type 3	Type 4
Info on EDDA	8,825.71 (5.54%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Menu	18,321.70 (11.49%)	19,804.87 (4.07%)	7,830.62 (2.45%)	11,330.65 (3.77%)
General score	2,022.58 (1.27%)	3,695.63 (0.76%)	2,465.71 (0.77%)	889.75 (0.30%)
Digi score	3,348.18 (2.10%)	3,739.70 (0.77%)	2,371.00 (0.74%)	2,402.84 (0.80%)
Tech use score	3,813.59 (2.39%)	1,893.04 (0.39%)	2,411.96 (0.75%)	865.98 (0.29%)
Tech use visuals	24,396.17 (15.30%)	28,676.56 (5.90%)	19,827.85 (6.21%)	17,909.82 (5.95%)
Tech use recommen- dations	54,134.65 (33.95%)	300,578.61 (61.82%)	181,314.43 (56.75%)	170,274.25 (56.62%)
CD score	2,850.27 (1.79%)	467.30 (0.10%)	709.17 (0.22%)	923.23 (0.31%)
CD visuals	14,254.24 (8.94%)	12,212.44 (2.51%)	6,851.26 (2.14%)	11,868.63 (3.95%)
CD recommendations	27,465.36 (17.23%)	115,151.70 (23.68%)	95,701.91 (29.96%)	84,291.10 (28.03%)
Total	159,432.44 (100.00%)	486,219.85 (100.00%)	319,483.92 (100.00%)	300,756.25 (100.00%)

Note: Percentages are calculated by dividing the average total fixation duration in each AOI by the total for each type.

Appendix C. Average total fixation duration in each AOI for the three aggregated types in ms and percent

AOI	Type 1	Type 2	Type 3
Info on EDDA	1,017.76 (0.11%)	3,977.06 (0.67%)	15,453.19 (3.33%)
Menu	14,675.95 (1.65%)	24,456.61 (4.13%)	18,646.29 (4.02%)
General score	3,147.90 (0.35%)	3,192.30 (0.54%)	2,733.02 (0.59%)
Topic score	7,446.17 (0.84%)	11,414.33 (1.93%)	10,251.78 (2.21%)
Topic overview	6,440.89 (0.73%)	11,627.01 (1.97%)	18,954.72 (4.09%)
Visuals	94,690.37 (10.68%)	88,064.67 (14.88%)	150,937.67 (32.54%)
Recommendations	759,605.44 (85.64%)	448,960.93 (75.88%)	246,828.97 (53.22%)
Total	887,024.49 (100.00%)	591,692.91 (100.00%)	463,805.63 (100.00%)

Note: Percentages are calculated by dividing the average total fixation duration in each AOI by the total for each type.

F Zusammenfassung

Die vorliegende Dissertation untersucht die Wirkung und Nutzung einer KI-Plattform zur Unterstützung der Unterrichtsplanung in der wirtschaftsberuflichen Lehrkräftebildung und fokussiert aktuelle Themen wie Digitalisierung und Nachhaltigkeit. Speziell der Rechnungswesen- und Controllingunterricht ist dabei ein relevanter Bezugsbereich, da er essenziell für das wirtschaftliche Denken ist und zugleich in allen wirtschaftsberuflichen Curricula stattfindet. Vor dem Hintergrund der dynamischen Veränderungen in der Berufswelt müssen Lehrkräfte darauf vorbereitet werden, aktuelle Herausforderungen im Unterricht zu berücksichtigen, um die berufliche Handlungskompetenz der Auszubildenden zu stärken und diese langfristig handlungsfähig zu machen. Dies geschieht unter den komplexen Bedingungen der Unterrichtsplanung.

Im Zentrum steht die Frage, wie eine KI-Plattform in der wirtschaftsberuflichen Lehrkräftebildung gestaltet und genutzt werden kann, um Studierende bei ihrer Unterrichtsplanung zu aktuellen Themen zu unterstützen. Dabei werden vier Teilfragestellungen bearbeitet. Zunächst wird mittels zweier systematischer Literaturreviews herausgestellt, wie der Stand der Forschung zu Digitalisierung und Nachhaltigkeit im Rechnungswesen und Controlling aussieht, wie diese Inhalte in den Unterricht integriert werden und welche Kompetenzen befördert werden sollen. Daraufhin erfolgt die theoretisch-konzeptionelle Entwicklung von Gestaltungsmöglichkeiten für eine analytische KI-Plattform als Assistenz bei der Unterrichtsplanung, um den empirischen Teil der Arbeit vorzubereiten. Die Plattform, genannt EDDA, wurde in einem gemeinschaftlichen Vorhaben der Universitäten Hamburg, Oldenburg und Graz (Österreich) entwickelt. EDDA bietet sowohl schriftliches als auch visuelles Expertenfeedback. Die semesterbegleitende Interventionsstudie im Experimental-Kontrollgruppendesign trägt zur Beantwortung der Teilfrage nach der Wirksamkeit der Plattform bei und erlaubt Vergleiche zu menschlichem Feedback. Die letzte Teilfrage zur Interaktion mit der Plattform wird mittels einer Eye Tracking-Studie über zwei Erhebungszeitpunkte beantwortet.

Die empirischen Ergebnisse zeigen, dass EDDA signifikante Verbesserungen in der Qualität der Unterrichtspläne bewirken kann. Die Nutzung der Plattform ermöglicht es den Studierenden darüber hinaus, ihre Pläne in einigen Bereichen besser an die Bedürfnisse einer digitalisierten und nachhaltigen Wirtschaftswelt anzupassen. Bei der Interaktion mit EDDA konnten verschiedene Nutzungstypen identifiziert werden, die auch in der Nutzungsintensität voneinander abweichen. Diese Ergebnisse verstärken die Bedeutung von KI-

gestütztem Feedback als Werkzeug zur Professionalisierung von (angehenden) Lehrkräften. Verschiedene Präferenzen für schriftliches oder visuelles Feedback zeigen die Notwendigkeit hochschuldidaktische Lehr-Lernprozesse individualisiert zu gestalten. Studierende, die sich stärker auf die Gestaltungsempfehlungen für ihre Unterrichtsplanungen fokussieren, entwickeln in vielen analysierten Bereichen signifikant bessere Unterrichtsentwürfe. Jene, die sich eher auf die Darstellung ihrer Ergebnisse als auf Veränderungspotenziale fokussieren, schneiden schlechter ab. Diese Erkenntnis regt dazu an, die Kollaboration zwischen KI und Lehrenden weiter auszubauen, um optimal auf die vielfältigen Bedürfnisse der Lernenden eingehen zu können.

Zusammenfassend zeigt die vorliegende Arbeit, dass KI-gestütztes Feedback ein effektives Mittel zur Professionalisierung in der Lehrkräftebildung darstellt. Die gewonnenen Erkenntnisse tragen dazu bei, lernförderliche KI-Anwendungen zu gestalten und geben ferner Aufschluss darüber, unter welchen Gesichtspunkten die Integration von Digitalisierung und Nachhaltigkeit gelingen kann. Die Arbeit erweitert das Verständnis dafür, wie KI-Technologien als pädagogische Werkzeuge genutzt werden können, um qualitative Verbesserungen in Lehr- und Lernprozessen zu erreichen.

G Summary

This dissertation examines the impact and use of an AI platform to support lesson planning in vocational business teacher education, focusing on current topics such as digitalization and sustainability. Especially accounting and controlling lessons are of interest here, as this area plays a crucial role in economic thinking and is embedded into all vocational business education curricula. Against the backdrop of dynamic changes in the professional world, teachers must be prepared to address current instructional challenges, aiming to strengthen the vocational action competence of students and ensure their long-term ability to act. All of these developments have to be recognized under the already complex conditions of lesson planning.

The main question asks how an AI platform can be designed and utilized in vocational business teacher education to support students in their lesson planning on current topics. Four sub-questions are addressed. First, two systematic literature reviews highlight the state of research on digitalization and sustainability in accounting and controlling, examine how these contents are integrated into teaching, and identify the competences to be promoted. Following this, the theoretical and conceptual design possibilities for an analytical AI platform as an assistant in lesson planning are developed to inform the following empirical study. The platform EDDA was developed through a collaborative effort by the universities of Hamburg, Oldenburg, and Graz (Austria). EDDA provides both written and visual expert feedback. A semester-long intervention study using an experimental-control group design contributes to the sub-question on the platform's effectiveness and allows comparisons with human feedback. The final sub-question regarding students' interaction with the platform is addressed using an eye-tracking study across two data collection points.

Empirical results demonstrate that EDDA can significantly improve the quality of lesson plans in some of the analyzed areas. In the eye-tracking study, various interaction types were identified, differing in usage intensity. These results underscore the importance of AI-supported feedback as a tool for the professionalization of (prospective) teachers. Different preferences for written or visual feedback highlight the necessity of individualized instructional design in higher education teaching and learning processes. Students who focus more on the design recommendations for their lesson plans develop significantly better lesson drafts in many analyzed areas. Those focusing more on showcasing their results rather than on potential changes perform worse. This insight encourages further

development of collaboration between AI and educators to better meet the diverse needs of learners.

In summary, this work demonstrates that AI-supported feedback is an effective means of professionalization in teacher education. The findings contribute to designing learning-enhancing AI applications and provide insights into the successful integration of digitalization and sustainability. The work expands the understanding of how AI technologies can be used as pedagogical tools to achieve qualitative improvements in teaching and learning processes.

Eidesstattliche Versicherung

Ich, Julia Pargmann, versichere an Eides statt, dass ich die Dissertation mit dem Titel: „Einsatzpotenziale von künstlicher Intelligenz in der Lehrkräftebildung – Wirkung und Nutzung einer KI-basierten Assistenz für die Unterrichtsplanung im Kontext von Digitalisierung und Nachhaltigkeit“ selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Oldenburg, 08.05.2025

Ort, Datum

Unterschrift Doktorand:in

Erklärung

Darüber hinaus erkläre ich, dass ich keine kommerzielle Promotionsberatung in Anspruch genommen habe und die Arbeit nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt worden ist.

Oldenburg, 08.05.2025

Ort, Datum

Unterschrift Doktorand:in