
Themenheft 65: Tagungsband des 1. AK Mediendidaktik der DGfE-Sektion Medienpädagogik (MEDIDA24). Herausgegeben von Maria Klar, Josef Buchner, Barbara Getto, Marco Kalz und Michael Kerres

Wie lernt es sich mit KI-basierten Conversational Agents?

Entwicklung eines Instruments zur Erfassung des Lernempfindens am Beispiel des spielbasierten ökonomischen Lernens

Jan-Martin Geiger¹  und Anna Wenzel² 

¹ Universität Münster

² Technische Universität Dortmund

Zusammenfassung

Im Rahmen dieses Beitrags wird ein Fragebogeninstrument entwickelt, das zur Erfassung des subjektiven Lernempfindens im Kontext des Lernens mit natürlichsprachlichen KI-basierten Conversational Agents dient. Zu diesem Zweck wurde ein Fragebogen zum subjektiven Lernempfinden unter Zusammenführung bestehender Skalen entwickelt und exemplarisch in zwei Lernszenarien erprobt. In diesen Szenarien, die die Durchführung zweier ökonomischer Serious Games zum Gegenstand haben, wurde jeweils ein auf die Veranstaltungsinhalte abgestimmter KI-basierter Conversational Agent eingesetzt, um Studierende in ihrem Lernprozess individuell zu unterstützen. Wir berichten eine erste explorative Faktorenanalyse, die fünf für das Lernen mit Conversational Agents relevante Faktoren mit 17 Items identifiziert. Der Beitrag erweitert auf diese Weise das diagnostische Instrumentarium rund um die individuelle Wahrnehmung KI-gestützter pädagogischer Unterstützungswerkzeuge aus Lernendenperspektive. Während hierdurch aus wissenschaftlicher Sicht insbesondere das methodische Spektrum vergrößert wird, werden auf lehrpraktischer Ebene Impulse für neue Möglichkeiten zur Lernbegleitung und Evaluation angeboten.

Learning with AI-Based Conversational Agents. Developing an Instrument to Measure Learning Experience in Game-Based Economic Learning

Abstract

Within the scope of this study, a questionnaire instrument is developed to assess subjective learning experiences in the context of learning with natural language AI-based conversational agents. To this end, a questionnaire on subjective learning was

constructed by synthesizing existing scales and subsequently validated in two learning scenarios. These scenarios involved the implementation of two economic serious games, during which an AI-based conversational agent, specifically tailored to the content of the sessions, was utilized to provide individualized support to students throughout their learning process. An initial exploratory factor analysis revealed five factors relevant to learning with conversational agents, represented by 17 items. This study contributes to the diagnostic repertoire for examining learners' perceptions of AI-supported pedagogical tools. From a scientific perspective, it enhances methodological approaches, while on a practical pedagogical level, it offers novel insights for advancing learning support and evaluation strategies.

1. Einführung

Seit generative Künstliche Intelligenz (KI) weit verbreitet und allgemein zugänglich ist, erfahren KI-basierte Conversational Agents (CA) zunehmend Einzug in schulische und akademische Lehr-Lernkontexte (Gimpel et al. 2023). KI-basierte CA eröffnen auch die Möglichkeit, bestehende Korpora wie Literatur, Studienbriefe oder Skripte einzubinden und zum inhaltlichen Gegenstand von Mensch-zu-Maschinen Interaktion mittels natürlichsprachlicher Fähigkeiten zu machen. In derartigen Szenarien übernehmen CA eine fundamentale Rolle der Stoffvermittlung und versprechen, lehrbezogene Zielsetzungen wie individuelle Förderung zu unterstützen und Lehrkräfte zu entlasten. Inwieweit diesen Bedarfen tatsächlich durch die Einbindung KI-basierter CA begegnet werden kann, ist aus didaktischer Perspektive bislang jedoch kaum untersucht.

So bestehen zentrale Erwartungen dahingehend, dass CA als Werkzeug fungieren können und eine Orientierung an individuellen Lernvoraussetzungen und damit verbundenen Bedarfen erlauben. Derartige Systeme sollen selbstreguliertes Lernen fördern und heterogene Lernbedingungen berücksichtigen, etwa indem sie sich in ihrem Interaktionsverhalten an sprachliche Fähigkeiten oder individuelle Kenntnisstände Lernender anpassen (Hobert und Meyer von Wolff 2019; Weber et al. 2021). Befürchtungen bestehen jedoch dahingehend, dass durch eine Verlagerung der Lehrendenrolle auf CA eine Verschärfung der Distanz zwischen Lehrenden und Lernenden provoziert wird und gerade diejenigen Lernenden benachteiligt werden, die einer kontinuierlicheren und intensiveren Betreuung bedürfen.

Wenngleich die Debatte über lehrbezogene Anwendungsszenarien generativer KI im Allgemeinen sowie hinsichtlich CA im Besonderen innerhalb des wissenschaftlichen wie auch des bildungspraktischen Diskurses noch geführt wird, werden entsprechende Werkzeuge in der Praxis längst durch Lehrende und Lernende genutzt (Weber et al. 2021). Aus didaktischer Perspektive besteht daher ein dringendes Desiderat darin, konzeptionelle wie empirische Ansätze zur Einbettung und Evaluation

entsprechender Systeme zu entwickeln. Auf diese Weise soll nicht nur die pädagogische Qualität von CA in der Lehre gewährleistet werden; vielmehr muss auch die Überprüfbarkeit der Wirkung entsprechender Kurse und Programme sichergestellt sein.

Vor diesem Hintergrund gilt unser Augenmerk dem Lernen mit CA verbunden mit der Frage, welche Lernmechanismen durch diese Technologie tangiert werden. Die wenigen bislang bestehenden Untersuchungen, die sich auf Bildungskontexte beziehen, deuten an, dass die mit CA verbundene Lernqualität auf selbstregulative Aspekte des Lernens, die wahrgenommene Handhabbarkeit solcher Werkzeuge oder die Interaktionsauthentizität abstellen kann (Han et al. 2023). Gleichwohl fehlt es nach unserem Kenntnisstand noch an Ansätzen, die weitere für die Lernqualität ebenfalls bedeutende Aspekte wie die individuelle Bedeutsamkeit des Lerngegenstands selbst oder die Sicht der Lernenden auf konkrete erzielte Lernfortschritte einbeziehen.

Es ist daher das Ziel dieses Beitrags, einen diagnostischen Ansatz für die Erfassung der Wirkung von CA auf das subjektive Lernempfinden zu entwickeln und darauf basierend ein Fragebogeninstrument für Forschungs- und Lehrzwecke zu erproben. Wir entwickeln hierzu einen CA zur Unterstützung von spielbasiertem Lernen im Studienfach Wirtschaft, mit dem Studierende durch Abfrage interagieren, um bspw. Fachwissen zu erfragen oder Handlungsalternativen zu eruieren. Dabei handelt es sich um ein Börsenplanspiel, in dem Lernende Entscheidungen am Finanzmarkt treffen, sowie ein Unternehmensplanspiel, in dessen Rahmen Studierende über Investitionen oder die Preisgestaltung entscheiden.

Während aus forschungsmethodischer Perspektive auf diese Weise ein Instrument entspringt, das künftige Wirkungsstudien hinsichtlich KI-basierter CA unterstützen kann, eröffnet der Beitrag aus lehrpraktischer Sicht Impulse, subjektive Lernprozesse auch bei Verlagerung in digitale Szenarien hinein und damit potenziell über räumliche Distanzen hinweg oder in grossen Gruppen zu evaluieren. Letztlich sollen hierdurch auch Möglichkeiten aufgezeigt werden, die bildungsstufenübergreifend immer prominenter werdenden KI-Systeme in Veranstaltungsevaluationen zielgerichteter einzubinden und auf diese Weise die Sicherung von Qualität im Bildungswesen zu unterstützen.

2. Conversational Agents in der Lehre: ein Kurzüberblick

Digitale Gesprächsagenten (engl. Conversational Agents, CA), verstanden als Computersysteme, die menschliche Kommunikation in natürlicher Sprache nachahmen sollen (Miner et al. 2016), finden bereits in vielen alltäglichen Bereichen Anwendung und übernehmen beratende Aufgaben, etwa im Kundenservice (Bavaresco et al. 2020), im Gesundheitswesen (Montenegro, da Costa, und da Rosa Righi 2019) oder

im Finanzsektor (Quah und Chua 2019). Die Auslagerung von beratenden Aufgaben in diese Systeme erfreut sich nicht nur aufgrund erwarteter Kosteneinsparungen grosser Beliebtheit, sondern auch wegen ihrer permanenten Verfügbarkeit und der damit verbundenen Orientierung an den Lebensumständen von Nutzenden.

Auch in Lehrkontexte werden nach ähnlichen Logiken operierende CA, die daher als pädagogische (Gesprächs-)Agenten (engl. Pedagogical Conversational Agents, PCA) bezeichnet werden (Weber et al. 2021), vermehrt implementiert. Entsprechende Prototypen finden insbesondere in schulischen (z. B. Wiegand 2024) oder akademischen (z. B. Ait Baha et al. 2024) Kontexten Erprobung und verdeutlichen die Durchdringungskraft dieser spezifischen Technologie auch im Bildungsbereich. In ihrer Rolle als Experten, Mentoren, Motivatoren oder Peers (Baylor und Kim 2005) erfüllen sie verschiedene pädagogische Funktionen: Sie demonstrieren Handlungs- oder Lösungsabläufe, fungieren durch die Abfrage von Wissensständen als Prüfende oder dienen den Lernenden als Informationsquelle, Coach (Clarebout et al. 2002) oder auch als Lernbegleiter, indem sie zum Lernen motivieren oder beim individuellen Zeitmanagement unterstützen (Schlimbach et al. 2024). Um Lernende bei der Überwindung von Barrieren oder Defiziten im Lernprozess zu unterstützen, greifen PCA auf vielfältige Kommunikationsformen wie Text, Sprache, Grafiken, Gesten oder kombinierte Ansätze zurück (Miner et al. 2016).

Seit Weizenbaums Pionierarbeit mit ELIZA im Jahr 1966, die oftmals als Geburtsstunde von CA bezeichnet wird, haben technologische Fortschritte in der KI und hierbei insbesondere im Bereich des «Natural Language Processing» und maschinellen Lernens zahlreiche Innovationen im Zusammenhang mit CA hervorgebracht und neue Wege für die Mensch-Computer-Interaktion eröffnet. Während in den 2000er-Jahren der Fokus der Entwicklung von CA auf der Verarbeitung sprachlicher Strukturen und der grammatikalisch korrekten Ausgabe lag (Reiter und Dale 1997), hat sich der Schwerpunkt in jüngster Zeit auf die Art und Weise verlagert, wie eine natürlichsprachliche und inhaltlich tiefgründige Interaktion gestaltet werden kann. Basierend auf den Absichten, den Entitäten und der kognitiven Information der Benutzereingaben können KI-gestützte CA semantische Repräsentationen erzeugen, diese Daten für zukünftige Referenzen speichern und passende Antworten generieren (Kulkarni et al. 2019). Die neueste Generation von CA imitiert damit realitätsnahe menschliche Interaktionen. Beschrieben wird die Bedeutung dieses Authentizitätsfaktors bspw. durch die CASA-Theorie (Computers as Social Actors, Reeves und Nass 1996), derzufolge Menschen auf digitale Agenten ähnlich reagieren wie auf reale Menschen. Das verwendete Sprachregister beeinflusst dabei Glaubwürdigkeit, Vertrauen und Überzeugungskraft und spielt somit eine vertrauensbildende Rolle in pädagogischen Kontexten, in denen die Förderung selbstregulierten Lernens sowie die Adaptionfähigkeit von KI-Anwendungen als zentrale Einsatzfaktoren gelten (Schöbel et al. 2023).

Das für Lehrkontexte innovative Potenzial schöpfen neueste PCA aus der Verbindung eben jenes natürlichsprachlichen KI-Modells mit einer konkreten Wissensbasis, bestehend z. B. aus Skripten, Lehrbüchern oder Tabellen, die durch Lehrende implementiert werden können und die den inhaltlichen Gegenstand der Konversation rahmen (Gimpel et al. 2023; König 2024; Wiegand 2024). Einerseits wird Lernen so eine natürlichsprachlich wirkende Interaktion mit durch Lehrende ausgewählten Inhalten ermöglicht, die eine Verengung auf einen bestimmten Lernstoff vornimmt und damit eine höhere fachliche Präzision erwarten lässt als KI-Systeme ohne individualisierte Wissensgrundlage. Andererseits senken neueste PCA die Einsatzbarrieren für Lehrkräfte erheblich, da diese kaum oder fast keine Programmierkenntnisse voraussetzen und vergleichsweise einfach in Webseiten und bestehende Lernplattformen integriert werden können. Gerade in Anbetracht dieser immer weiter sinkenden Einsatzbarrieren bedarf es umso mehr vor allem forschungsbasierter medienpädagogischer sowie (fach-)didaktischer Zugänge, die die Voraussetzungen und Wirkungen für eine lehrpraktische Einbindung solcher PCA in den Blick nehmen.

3. Bedarf diagnostischer Zugänge im Hinblick auf das subjektive Lernempfinden mit Conversational Agents

Hinsichtlich der Wirkung von PCA besteht eine vergleichsweise breite Studienlage insbesondere mit Blick auf Lernfortschritte, aktuelle Analysen deuten – auch unter Einbezug experimenteller Forschungsdesigns – eine positive Wirkung entsprechender Werkzeuge auf Kompetenzzuwächse oder eine gesteigerte Lernmotivation an (Ait Baha et al. 2024; Schroeder et al. 2025). In diesem Kontext weist die Meta-Analyse von Castro-Alonso, Wong, Adesope und Paas (2021) darauf hin, dass PCA-unterstütztes Lernen in multimedialen Umgebungen effektiver ist als Lernen ohne diese Unterstützung, allerdings wird eine vergleichsweise geringe Effektstärke beobachtet. Darüber hinaus scheinen PCA das Lernen durch intrapersonale Kommunikation auf individueller Ebene zu berühren, insbesondere in Bezug auf Motivation, Selbstregulation, Selbstwirksamkeit und Metakognition (S. Han et al. 2023). Auf zwischenmenschlicher Ebene scheinen PCA lernförderlich zu wirken, indem sie Feedback, Hinweise sowie Anregungen zum Lernprozess und zu den Lernergebnissen geben. Zudem können PCA Diskussionen erleichtern und die Zusammenarbeit unter Studierenden moderieren (Sikström et al. 2022).

Trotz der scheinbar vielfältigen Potenziale von PCA zur Unterstützung des Lernens bestehen nach wie vor blinde Flecken hinsichtlich der Bedingungen, unter denen diese Werkzeuge ihre intendierte Wirkung entfalten. Unklar ist, inwieweit Lernende durch solche Werkzeuge tatsächlich ihr eigenes Lerngeschehen autonom gestalten können und inwieweit dies wiederum individuelle Lernprozesse unterstützt. Auch ist denkbar, dass Lernende unterschiedliche Affinitäten und

Überzeugungen hinsichtlich PCA haben, die wiederum die tatsächliche und aktive Nutzung für Lernzwecke beeinflussen. Wir fokussieren daher das subjektive Lernempfinden, das in Anlehnung an Chavez Rojas, Faure Ninoles und Barril Madrid (2023, 259) die diskursive Rekonstruktion von Lernsituationen durch Lernende und die damit verbundene Zuschreibung von Bedeutung umfasst. Üblicherweise findet das subjektive Lernempfinden Eingang in unterschiedliche Konzeptionen von Lehrvaluationen (z. B. Landes und Ziegler 2015) und liefert Informationen dahingehend, wie ein Lernarrangement auf Teilnehmende wirkt. Anerkennung findet hierdurch insbesondere die Individualität von Lernprozessen, in denen es Lernenden unterschiedlich leicht oder schwer fällt, sich auf Problemlagen einzulassen oder neues Wissen zu erarbeiten.

Mit den Studien von Hew et al. (2021) und Han et al. (2023) liegen bereits erste empirische Ansätze zur Erfassung des subjektiven Lernempfindens mit PCA vor. Diese Ansätze greifen insbesondere auf das Technologie-Akzeptanz-Modell zurück (TAM, Davis et al. 1989), welches subjektive Überzeugungen von Lehrenden sowie Lernenden im Hinblick auf die Wirkung von Technologien in pädagogischen Kontexten betrachtet. Zentrale Elemente in diesen Studien sind einerseits die *wahrgenommene Benutzerfreundlichkeit* (Perceived Ease of Use, PEU) und der *wahrgenommene Nutzen* (Perceived Usefulness, PU) andererseits. Demzufolge umfasst die PEU die Überzeugung, dass ein Werkzeug einfach und mühelos zu bedienen ist, während die PU dasjenige Ausmass anzeigt, in dem das Werkzeug die Leistung oder Produktivität verbessert (Davis et al. 1989). Tatsächlich beziehen sich die bestehenden Studien mit «massive open online courses» (MOOCs, J. Han et al. 2023) und «flipped classrooms» (Hew et al. 2021) auf Lehrformate, in denen insbesondere die Überbrückung physischer Distanz adressiert wird und für Lernende entsprechend soziale Lernaspekte bei der Nutzung von PCA relevant scheinen. Es ist daher zu vermuten, dass je nach Lernszenario und damit verbundenen Bedürfnissen Lernender womöglich auch weitere, bisher noch nicht betrachtete subjektive Lernaspekte wichtig sein können. So werden weitere kognitive (z. B. subjektiv wahrgenommene Lernfortschritte) und emotionale Aspekte des Lernempfindens als bedeutend angesehen, etwa wenn Lernende sich konzeptionell vielschichtigem und für sie zugleich bedeutendem Lernstoff zuwenden (Pekrun et al. 2017; Vogl et al. 2019). Es bedarf daher einer Erweiterung diagnostischer Zugänge, die den Einfluss von PCA auf die Lernqualität betrachten und Rückschlüsse dahingehend zulassen, wie Lernende sich auch tatsächlich durch die Nutzung entsprechender Werkzeuge unterstützt sehen.

Das mit einer solchen diagnostischen Möglichkeit einhergehende Einsatzspektrum für die Lehre kann ganz unterschiedliche Schwerpunkte einnehmen, wobei die nachfolgende Aufzählung sicherlich weder abschliessend noch überschneidungsfrei ist: Aus *didaktischer* Sicht eröffnet sich die Möglichkeit zur Gewinnung diagnostischer Informationen über die Auseinandersetzung Lernender mit dem

Lerngegenstand und potenziell auftretenden Beschwerlichkeiten, etwa im Hinblick auf die Nützlichkeit eines PCA zur Ermittlung eines konkreten fachlichen Phänomens. Lehrende könnten auf diese Weise Schlüsse dahingehend ziehen, wie hemmend oder förderlich sich PCA in bestimmten Lernabschnitten erweisen, und bedarfsorientiert intervenieren. Aus *evidenzorientierter* Perspektive bieten auf das Lerngeschehen abstellende diagnostische Zugänge eine Option, die Effektivität von PCA zu betrachten. Denkbar ist dies sowohl formativ, also den Lernverlauf über eine längere Unterrichts- oder Veranstaltungsreihe betreffend, als auch summativ, also im Nachgang an ein bestimmtes Lernarrangement oder eine bestimmte Veranstaltungsreihe (Maier 2010). Auch bestünde so eine Möglichkeit für eine integrative Betrachtung von formativer und summativer Lernbeobachtung, die ein besseres Verständnis dahingehend erlaubt, warum es zu potenziellen Abweichungen zwischen erwarteten und beobachteten Lernfortschritten kommen kann. Eine Begleitung des subjektiven Lernempfindens mit Blick auf CA kann ferner einen *evaluatorischen* Zweck verfolgen, der insbesondere auf die Sicherstellung der Qualität des mit einem Lernarrangement jeweils verfolgten pädagogischen Anspruchs zielt. Auf diese Weise werden wichtige Erkenntnisse hinsichtlich der didaktischen Konfiguration von Lernarrangements generiert, die eine wichtige Feedbackfunktion für Lehrende, Lernende sowie Programmverantwortliche erfüllen kann.

Es ist daher das Ziel der vorliegenden Studie, die bestehenden Instrumente zur Erfassung des subjektiven Lernempfindens mit PCA vor dem Hintergrund weiterer Lernszenarien zu erweitern, was nachfolgend am Beispiel spielbasierten Lernens im Studienfach Wirtschaft erfolgt.

4. Spielbasiertes Lernszenario und Fragebogenkonstruktion

Spielbasiertes Lernen erfreut sich innerhalb des Bildungsbereichs grosser Beliebtheit und wird u. a. in solchen Lernszenarien als vielversprechend gesehen, die eine Personalisierung des Lerngeschehens fokussieren, Feedback für Lernende generieren und komplexe fachliche Phänomene veranschaulichen (Plass et al. 2020). Speziell in der ökonomischen Hochschullehre hat sich das spielbasierte Lernen als ein zentrales Lehrformat etabliert (Arndt 2020; Liening 2019; Weyland 2016), das maßgeblich durch das Zusammenwirken von digitalen Plattformen, Social-Media-Apps und digitalen Serious Games gestaltet wird (Chen et al. 2021; Hochmuth und Kirchner 2019/20). Im Einklang mit dem «Integrated Design Framework of Game-based and Playful Learning» (Plass et al. 2015) wird der individuelle Lernprozess im (digitalen) spielbasierten Lernen durch kognitive, affektive, motivationale sowie soziokulturelle Aspekte und deren Beziehungen beeinflusst. Das Potenzial (digitalen) spielbasierten Lernens wird durch mehrere interdisziplinäre Metaanalysen gestützt, die dessen positive Wirkung auf kognitive und affektive Lernziele hervorheben (Barz

et al. 2024; Clark et al. 2016; Wouters et al. 2013). Speziell in der Ökonomischen Bildung belegen Studien positive Effekte auf den Erwerb ökonomischen Wissens, die Förderung kognitiver und sozialer Kompetenzen sowie eine gesteigerte unternehmerische Handlungsintention (Faisal et al. 2022; Platz 2022), wenngleich Studien Verbesserungspotenziale etablierter Serious Games hinsichtlich der Gestaltung ansprechenderer Spielelemente, der Bereitstellung von formativem Feedback und der Förderung von Reflexionsmöglichkeiten aufzeigen (Bellotti et al. 2014; Fox et al. 2018). Entsprechend anschlussfähig scheinen spielbasierte Lernumgebungen für PCA und damit erwarteten pädagogischen Mehrwerten.

Für die Herleitung eines Instruments zur Erfassung des subjektiven Lernempfindens wurden daher zwei auf textgenerativer KI (dem GPT-Modell von OpenAI, vgl. Tabelle 1) basierende PCA zur individuellen Unterstützung im spielbasierten Lernen im Rahmen von zwei rundenbasierten Serious Games zur finanziellen Bildung sowie zur Unternehmensführung in einem wirtschaftswissenschaftlichen Studiengang eingesetzt.

Während Studierende im Serious Game zur Finanzbildung («CCC Shareholder», Liening und Mittelstädt 2011; Sender 2017) in der Rolle als Investmentmanager die Aufgabe haben, in einem ganztägigen Seminar das Startkapital von 200.000 Euro durch Aktienhandel zu erhöhen und die Erwartungen der Investmentgesellschaft zu erfüllen, gilt es im Serious Game zur Unternehmensführung («Start-Up Competition», Liening et al. 2013), in Kleingruppen von vier bis fünf Studierenden ein sich im Wachstum befindendes Unternehmen durch eine Reihe wöchentlicher Entscheidungen in einem Zeitraum von fünf Wochen erfolgreich am Markt zu etablieren und sich gegen konkurrierende Gruppen zu behaupten. Die wichtigsten Design-Charakteristika der PCA, basierend auf der Taxonomie nach Weber et al. (2021), können Tabelle 1 entnommen werden.

Dimension		Umsetzung des PCA im Kontext des Serious Game zur finanziellen Bildung	Umsetzung des PCA im Kontext des Serious Game zur Unternehmensführung
Technologie	TE1: PCA Design	Lernende chatten mit einem freundlichen Planspiel-Bot in Messenger-Form	
	TE2: Programmierung	Nutzung des LLM GPT 3.5 zum Zeitpunkt der Datenerhebung (Januar bis März 2024)	Nutzung des LLM GPT 4.0/4.0 mini (hybrid) zum Zeitpunkt der Datenerhebung (Oktober bis November 2024)
	TE3: Interaktionsdesign (Input)	Ermöglicht individuelle Dialoge zwischen Lernenden und dem PCA in Messenger-Form	
	TE4: Interaktionsdesign (Output)	Antworten in kurzen, prägnanten Sätzen unter Nutzung passender Emojis	Wöchentliches elaboriertes Feedback zum Spielerfolg, offener Dialog in kurzen, prägnanten Sätzen unter Nutzung passender Emojis
	TE5: PCA Embedding	Zugänglich über einen zur Verfügung gestellten Link/QR-Code	Einbettung in die Planspielumgebung auf der Lernplattform Moodle sowie webbasiert über einen Link
Aufgaben/Zielgruppe	TA1: Rolle des PCA	Bietet Hilfestellung bei der Orientierung auf der Benutzeroberfläche, unterstützt die Lernenden bei der Auswahl passender Order-Typen und gibt ihnen Anreize zur Reflexion	Bietet Feedback zum Spielerfolg, indem produktions- und absatzrelevante Kennzahlen und deren Zusammenhang der Unternehmen analysiert werden (z. B. Rentabilität, Produktionsmenge, Absatz) und Hilfestellung zu inhaltlichen und organisatorischen Fragen im Planspiel
	TA2: Erwarteter Nutzen	Konzeptuelles Wissen über die verschiedenen Ordertypen, prozedurales Wissen über den Kauf und Verkauf von Aktien, metakognitives Wissen über das eigene Anlageverhalten und Verhaltensanomalien	Konzeptuelles Wissen über betriebswirtschaftlich relevante Kennzahlen, prozedurales Wissen über das Durchführen von Wirtschaftlichkeitsrechnungen sowie Strategieanpassung auf Grundlage von Unternehmenserfolg/ gesamtwirtschaftlichen Veränderungen, metakognitives Wissen über eigenes Managementverhalten, Entscheidungsfindung in Gruppen
	TA3: Zielgruppe	Teilnehmerinnen des Börsenplanspiels «CCC Shareholder» (Bachelor- und Masterstudierende)	Teilnehmende des Unternehmensplanspiels «Start-Up Competition» (Bachelor- und Masterstudierende)
Struktur	S1: Anwendungsbereich	Finanzielle Bildung	Unternehmensführung/ Entrepreneurship
	S2: Facetten des Lernprozesses	Informationsorganisation und -integration: Wahl der Ordertypen vor dem Hintergrund individueller Anlagestrategien, Reflexion des Kauf- und Verkaufsverhaltens	Informationsorganisation und -integration: Entscheidungsfindung innerhalb der Gruppen Reflexion: Wöchentliches Feedback sowie reflexives Feedback im offenen Dialog

Tab. 1: Design der eingesetzten pädagogischen *Conversational Agents*.

Zur Identifikation von Aspekten des Lernens mit dem CA wird eine explorative Faktorenanalyse durchgeführt, für die zunächst basierend auf bestehenden Instrumenten ein Itempool generiert wurde. In dieses Verfahren gehen solche Instrumente ein, die auf das subjektive Lernempfinden rekurrieren und sich auf das Lernen mit natürlichsprachlichen PCA beziehen lassen. Dieser Pool umfasst zunächst die «scale to measure students' learning experiences with knowledge-based chatbots in MOOCs» (S. Han et al. 2023) sowie die «Chatbot Usability Scale» (Borsci et al. 2022; Borsci et al. 2023), wobei aus letzterem – sich auf Bots im Allgemeinen beziehenden Konstrukt – lediglich Items verwendet wurden, die noch nicht in den Skalen von S. Han et al. (2023) enthalten sind. Zusätzlich zu diesen beiden Skalen wurden weitere Konstrukte berücksichtigt, die die Bedeutsamkeit des Lernstoffs und wahrgenommene Lernfortschritte fokussieren. In diesem Kontext greifen wir auf das Vorgehen in der Studie von Sender (2017) zurück, in der kognitive, motivationale und emotionale Konstrukte zum subjektiven Lernempfinden in ökonomischen Lernprozessen adaptiert werden. Diesem Vorgehen folgend wurden dem «Achievement Emotions Questionnaire» (Pekrun et al. 2011), dem «Motivated Strategies for Learning Questionnaire» (Pintrich et al. 1991) sowie dem «Therapie-Prozessbogen» (Schiepek et al. 2012) entsprechende Items zum subjektiven Lernempfinden entnommen und auf das Lernen mit PCA übertragen (Tabelle 2).

Instrument	Quellen	Dimensionen (Anzahl verwendeter Items)	Gesamtanzahl Items
Achievement Emotions Questionnaire	Pekrun et al. 2011	Spaß (2)	2
Bot usability Scale	Borsci et al. (2022), Borsci et al. (2023)	Privatsphäre (1), Wahrgenommene Qualität der Chatbot-Funktion (2), wahrgenommene Qualität der Konversation und der bereitgestellten Informationen (2)	5
Motivated Strategies for Learning Questionnaire	Pintrich et al. 1991	Motivation (6)	6
Scale to measure students' learning experiences with knowledge-based chatbots in MOOCs	S. Han et al. (2023)	Soziale Präsenz (6), Kognitive Präsenz (5), Förderung selbstregulierten Lernens (4), Nutzungsintention (3), Benutzerfreundlichkeit (4)	22
Therapie-Prozessbogen	Haken und Schiepek 2006; Schiepek et al. 2012	Therapeutische Fortschritte (3) Perspektivenerweiterung (4)	7

Tab. 2: Generierung der Fragebogenitems mit Ursprungsstrukturen.

Um im Zuge der Fragebogenbearbeitung ein möglichst eindeutiges Begriffsverständnis sicherzustellen, wurde in der Formulierung der Items anstelle des Begriffs «Conversational Agent» die im Alltagssprachgebrauch gebräuchliche Bezeichnung «Chatbot» verwendet. Auf diese Weise wurde ein Fragebogen mit 42 Items generiert, den die Teilnehmenden im Nachgang an die Nutzung im Zuge der spielbasierten Lernumgebung bearbeiten konnten.

Es werden die für eine explorative Faktorenanalyse gängigen Kriterien ($KMO > 0,6$; Bartlett-Test auf Sphärizität mit Signifikanzniveau von 5%) herangezogen (Döring und Bortz 2016; Zwick und Velicer 1986), die Faktoren werden via Hauptkomponentenanalyse und Varimaxrotation extrahiert. Die Bildung der Faktoren erfolgt im ersten Schritt unter Berücksichtigung des Eigenwert-Kriteriums (Eigenwert > 1 , Zwick und Velicer 1986). Berücksichtigung finden zunächst diejenigen Faktoren, die mindestens vier Items mit Ladungen $> 0,6$ aufweisen (Döring und Bortz 2016). Für die Skalenbildung werden diejenigen Items herangezogen, die eine Faktorenladung $\geq 0,4$ aufweisen, nicht in ähnlicher Höhe auf mehrere Faktoren laden (Döring und Bortz 2016) und mindestens zu einer akzeptablen Skalenkonsistenz (Cronbachs $\alpha > 0,7$, Saad et al. 1999; Streiner et al. 2015) beitragen.

5. Ergebnisse

In den beiden Pilotstudien, in deren Rahmen Studierende des Fachs Wirtschaftswissenschaften an den beschriebenen Spielen teilnahmen, wurde der Fragebogen zur Evaluierung des Lernempfindens von 183 Personen vollständig bearbeitet. 60,8% der Teilnehmenden sind weiblich, das Alter beträgt durchschnittlich $23,10 \pm 2,55$ Jahre. 51,6% der Studierenden befinden sich im Bachelorstudium (48,4% im Master).

Zur Identifikation der dem Lernen mit dem natürlichsprachlichen Chatbot zugrundeliegenden Aspekte wurde eine explorative Faktorenanalyse mittels Hauptkomponentenanalyse und Varimax-Rotation durchgeführt. Sowohl das Kaiser-Meyer-Olkin Kriterium ($KMO = ,953$) als auch der Bartlett-Test auf Sphärizität ($c^2 = 8.315,751$, $p < ,001$) deuten die Eignung der zugrundeliegenden Stichprobe für eine explorative Faktorenanalyse an. Hierbei wurden fünf Faktoren identifiziert, die 71,2% der Varianz aufklären. Unter Berücksichtigung insbesondere der originären Itemquellen werden dabei folgende Faktorenbezeichnungen verwendet: (1) Lernfortschritt (60,4%), (2) Interaktionsqualität (8,0%), (3) Bedeutsamkeit des Gelernten (4,0%), (4) Nutzungsintention (2,9%) und (5) Selbstregulation (2,7%).

Faktor/ Item n=183				Faktorladungen				
		AM	SD	(1)	(2)	(3)	(4)	(5)
Faktor (1) Lernfortschritt (5 Items, $\alpha = .957$, mittlere Item-Trennschärfe = .879)								
1	Durch den Chatbot sind mir Zusammenhänge klar geworden, die ich bisher nicht gesehen habe.	2,81	1,8	.78				
2	Durch den Chatbot haben sich für mich neue Perspektiven ergeben.	2,72	1,81	.78				
3	Die Inhalte, mit denen ich mich bei der Nutzung des Chatbots beschäftigt habe, sind neu und ungewöhnlich.	2,66	1,78	.83				
4	Durch den Chatbot wird es mir immer besser möglich, den Lernstoff zu verstehen.	2,95	1,85	.79				
5	Durch den Chatbot bin ich meinen Lernzielen nähergekommen.	2,84	1,81	.75				
Faktor (2) Interaktionsqualität (3 Items, $\alpha = .941$, mittlere Item-Trennschärfe = .878)								
6	Mit dem Chatbot macht es Spass, Antworten auf Fragen zu finden.	3,08	1,88		.66			
7	Ich mag die Unterhaltung mit dem Chatbot.	2,92	1,78		.65			
8	Der Chatbot gab klare Anweisungen.	2,96	1,87		.65			
Faktor (3) Bedeutsamkeit des Gelernten (3 Items, $\alpha = .900$, mittlere Item-Trennschärfe = .805)								
9	Für mich ist es wichtig, dass ich den Lernstoff so gut wie möglich verstehe.	4,17	1,77			.86		
10	Für mich ist es sehr wichtig, hier etwas zu lernen.	4,13	1,72			.89		
11	Ich bin sicher, dass ich die Herausforderungen dieses Kurses meistern kann.	3,78	1,77			.66		
Faktor (4) Nutzungsintention (3 Items, $\alpha = .906$, mittlere Item-Trennschärfe = .813)								
12	Ich würde eher einen Chatbot nutzen, anstatt eine E-Mail an die Kursleitung zu schreiben	3,53	2,05				.79	
13	Ich würde eher einen Chatbot nutzen, anstatt eine Forumsdiskussion zu führen.	3,49	2,06				.78	
14	Ich würde eher einen Chatbot nutzen, anstatt Literatur zu dem Thema zu lesen.	3,37	2,08				.81	
Faktor (5) Selbstregulation (3 Items, $\alpha = .904$, mittlere Item-Trennschärfe = .810)								
15	Ich war mir meiner Konzentration bei der Nutzung des Chatbots bewusst.	3,40	1,85					.60
16	Ich habe meine Eingabestrategien bewertet, nachdem ich meine Fragen an den Chatbot gestellt hatte.	3,16	2,01					.72
17	Ich habe mein vorhandenes Wissen während der Chatbot-Nutzung überprüft.	3,10	1,91					.69

Tab. 3: Fragebogenitems mit Faktorladungen (AM=arithmetisches Mittel, SD=Standardabweichung, Ladungen $\geq 0,6$) sind angegeben. Skala von 0=«trifft gar nicht zu» bis 6=«trifft voll zu».

Während die Faktoren (4) Nutzungsintention und (5) Selbstregulation nahezu unverändert die entsprechenden Skalen von S. Han et al. (2023) wiedergeben, stellen sich die weiteren Faktoren in neuer Konstellation für das subjektive Lernempfinden mit PCA dar. Im Rahmen der Faktorenanalyse wurden zahlreiche Items ausgeschlossen, wenn sich diese als uneindeutig erwiesen oder zu hohe Kreuzladungen mit anderen Faktoren bestanden. Zu Überschneidungen kommt es insbesondere zwischen den Faktoren (1) Lernfortschritt und (2) Interaktionsqualität, wie etwa das Item «Mir hat es Spass gemacht, mithilfe des Chatbots neues Wissen zu erlangen», sodass die entsprechenden Items nicht in die finalen Skalen eingegangen sind. Aufgrund von Kreuzladungen mussten auch Aussagen zur Bedienbarkeit des Chatbots, bspw. das Item «Ich glaube, dass der Chatbot zuverlässige Informationen gibt», ausgeschlossen werden, da dies ähnlich hoch sowohl auf den Lernfortschritt, die Interaktionsqualität sowie die Bedeutsamkeit des Gelernten lädt.

Die Reliabilität der extrahierten Skalen mit den verbliebenen Items ist gemessen an Cronbachs α jeweils sehr gut, auch die Item-Trennschärfen weisen auf eine hohe Skalengüte hin.

6. Zusammenfassender Ausblick

Ausgehend von einem Bedarf an lerndiagnostischen Zugängen, die den Einsatz von generativer KI in der Lehre angemessen begleiten, entwickelt der vorliegende Beitrag einen Ansatz zur Erfassung des subjektiven Lerngeschehens. Gerade in Anbetracht der rasanten Durchdringung der Lehrpraxis durch eine Vielzahl von Werkzeugen, die Lehrenden einen niedrigschwelligen Einsatz von CA versprechen, bearbeitet dieses Vorgehen ein wichtiges Anliegen. Hierzu wurde anhand der Einbindung von CA in eine spielbasierte ökonomische Lernumgebung exemplifiziert, wie ein Fragebogen zur Begleitung des Lernens mit auf generativer KI basierenden Gesprächsagenten entwickelt werden kann.

Das im Rahmen dieses Beitrags identifizierte Faktorenmodell deutet neue Aspekte an, die für das Lernen mit PCA relevant sein können und im didaktischen Gesamtgefüge zu beachten sind. So weisen die hier gewonnenen Daten darauf hin, dass der mithilfe von PCA erzielte wahrgenommene Lernfortschritt einen grossen Anteil (ca. 60% der Varianzaufklärung) des subjektiven Lernempfindens ausmacht, gefolgt von der Interaktionsqualität (8,0%) und der Bedeutsamkeit des Gelernten (4,0%). Spannend ist dieser Befund insbesondere deshalb, weil bestehende Skalen zum Einsatz von PCA in MOOCs die soziale Präsenz als zentralen Faktor identifizieren (S. Han et al. 2023). In diesem Zusammenhang ist ebenfalls zu erwähnen, dass die in der vorliegenden Studie extrahierten Faktoren «Bedeutsamkeit des Gelernten», «Nutzungsintention» und «Selbstregulation» vergleichsweise geringe Varianzanteile des subjektiven Lernempfindens aufklären. Womöglich hat auch dies mit

dem spezifischen Lernszenario zu tun, das – im Vergleich zu MOOCs – weniger auf die Überbrückung physischer Distanz ausgelegt ist, sondern mehr auf das Durchdringen und Reflektieren disziplinärer (hier: ökonomischer) Phänomene. Ferner vertiefen unsere Befunde zur Bedeutung des subjektiven Lernfortschritts bereits bestehende Skalen zur PU, die sich bisher eher auf die selbstregulative Wirkung von PCA (S. Han et al. 2023; Hew et al. 2021) konzentrieren. Aus diesen Beobachtungen lässt sich das bestehende Plädoyer weiter unterfüttern, die Wirkung von generativer KI nicht rein technikbezogen, sondern im jeweiligen didaktischen Gesamtkontext zu betrachten (z. B. Klar und Schleiss 2024).

Limitierend ist zu erwähnen, dass das hier vorgestellte Modell mittels einer vergleichsweise kleinen Stichprobe identifiziert wurde und – in Verbindung mit diesem Umstand – eine konfirmatorische Überprüfung der Modellstruktur noch aussteht. Gerade zur Schärfung von Skalen sowie zur Identifizierung weiterer Faktoren, die potenziell relevant für das Lernempfinden in Bezug auf PCA sind, bedarf es daher weiterer Untersuchungen mit einem grösseren Stichprobenumfang. Auch im Hinblick auf die Validität sind weitere Untersuchungen nötig, um die Güte des entwickelten Instruments zu verbessern. So müsste etwa der Zusammenhang des subjektiven Lernempfindens im Hinblick auf externe Kriterien wie bspw. objektivierbare Lernfortschritte (Kriteriumsvalidität) sowie das Verhältnis zu bestehenden Konstrukten des Lernempfindens (konvergente und diskriminante Validität) empirisch untersucht werden (Döring und Bortz 2016).

Ferner scheint es gerade auch für eine Generalisierbarkeit des Instruments notwendig, entsprechende Studien auch in weiteren Lehr-Lernszenarien anzuberaumen. Dies betrifft sowohl spielbasierte Szenarien in weiteren Fachdisziplinen als auch weitere Lehrformate jenseits spielbasierten Lernens wie bspw. Projektarbeiten oder lehrendenzentriertem Unterricht. Letztlich bedarf es auch der Erprobung mit weiteren Zielgruppen ausserhalb der akademischen Lehre, bspw. innerhalb des schulischen Kontexts. Auf diese Weise könnten weitere Erkenntnisse hinsichtlich der Generalisierbarkeit des in dieser Studie entwickelten Instruments zur Begleitung subjektiven Lernempfindens in Bezug auf PCA gewonnen werden.

Nichtsdestotrotz wird mit der vorliegenden Studie aus Perspektive der Lehrpraxis ein erstes Instrument angeboten, das das subjektive Lernempfinden in Bezug auf PCA in den Vordergrund stellt und damit die für Lehrende wie Programmverantwortliche wichtige Diagnostikfunktion adressiert. Das Instrument kann aufgrund seiner Itemanzahl (17) vergleichsweise aufwandsarm sowohl im Längsschnitt zur Gewinnung von Lernprozessdaten und damit für formative Diagnostikzwecke eingesetzt werden als auch für summative Zwecke im Nachgang an eine Unterrichts- oder Veranstaltungsreihe. Aus Perspektive der Forschung erweitert das Modell das Methodenrepertoire dahingehend, dass es subjektive Sichtweisen Lernender in das Lernen mit PCA einbezieht. Dies wiederum ist ein wichtiger Bestandteil, das Verständnis

über die Wirkung dieses spezifischen digitalen Werkzeugs auf Mechanismen des Lernens zu explorieren. Wir hoffen, damit einen wichtigen medienpädagogischen Beitrag zu leisten angesichts der hohen Geschwindigkeit, mit der generative KI sehr unterschiedliche Etappen des Lehrens und Lernens durchdringt.

Literatur

- Ait Baha, Tarek, Mohamed El Hajji, Youssef Es-Saady, und Hammou Fadili. 2024. «The impact of educational chatbot on student learning experience». *Education and Information Technologies*, 29(8), 10153-10176. <https://www.doi.org/10.1007/s10639-023-12166-w>.
- Arndt, Holger. 2020. *Ökonomische Bildung*. Erlangen: FAU University Press.
- Barz, Nathalie, Manuela Benick, Laura Dörrenbächer-Ulrich, und Franziska Perels. 2024. «The effect of digital game-based learning interventions on cognitive, metacognitive, and affective-motivational learning outcomes in school: A meta-analysis». *Review of Educational Research*, 94(2), 193–227. <https://doi.org/10.3102/00346543231167795>.
- Bavaresco, Rodrigo, Diórgenes Silveira, Eduardo Reis, Jorge Barbosa, Rodrigo Righi, Cristiano Costa, ... Mariangela Vanzin. 2020. «Conversational agents in business: A systematic literature review and future research directions». *Computer Science Review*, 36, 100239. <https://doi.org/10.1016/j.cosrev.2020.100239>.
- Baylor, Amy L., und Yanghee Kim. 2005. «Simulating instructional roles through pedagogical agents». *International Journal of Artificial Intelligence in Education*, 15(2), 95–115. <https://doi.org/10.1007/s40593-015-0055-y>.
- Bellotti, Francesco, Riccardo Berta, Alessandro de Gloria, Elisa Lavagnino, Alessandra Antonaci, Francesca Dagnino, ... Igor S. Mayer. 2014. «Serious games and the development of an entrepreneurial mindset in higher education engineering students». *Entertainment Computing*, 5(4), 357–66. <https://doi.org/10.1016/j.entcom.2014.07.003>.
- Borsci, Simone, Alessio Malizia, Martin Schmettow, Frank van der Velde, Gunay Tariverdiyeva, Divyaa Balaji, und Alan Chamberlain. 2022. «The chatbot usability scale: the design and pilot of a usability scale for interaction with AI-based conversational agents». *Personal and Ubiquitous Computing*, 26, 95–119. <https://doi.org/10.1007/s00779-021-01582-9>.
- Borsci, Simone, Martin Schmettow, Alessio Malizia, Alan Chamberlain, und Frank van der Velde. 2023. «A confirmatory factorial analysis of the Chatbot Usability Scale: a multilanguage validation». *Personal and Ubiquitous Computing*, 27(2), 317–30. <https://www.doi.org/10.1007/s00779-022-01690-0>.
- Castro-Alonso, Juan C., Rachel M. Wong, Olusola O. Adesope, und Fred Paas. 2021. «Effectiveness of multimedia pedagogical agents predicted by diverse theories: A meta-analysis». *Educational Psychology Review*, 33, 989–1015. <https://doi.org/10.1007/s10648-020-09587-1>.

- Chavez Rojas, Jorge, Faure Ninoles, Jaime und Barril Madrid, und Juan Pablo. 2023. «The construction of teachers' professional identity: An analysis of subjective learning experiences». *European Journal of Teacher Education*, 46(2), 256–73. <https://doi.org/10.1080/02619768.2021.1905627>.
- Chen, Li, Dirk Ifenthaler, und Jane Yau. 2021. «Online and blended entrepreneurship education: a systematic review of applied educational technologies». *Entrepreneurship Education*, 4(2), 191–232. <https://www.doi.org/10.1007/s41959-021-00047-7>.
- Clarebout, Geraldine, Jan Elen, W Lewis Johnson, und Erin Shaw. 2002. «Animated pedagogical agents: An opportunity to be grasped?» *Journal of Educational multimedia and hypermedia*, 11(3), 267–86.
- Clark, Douglas B., Emily E. Tanner-Smith, und Stephen S. Killingsworth. 2016. «Digital games, design, and learning: A systematic review and meta-analysis». *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/00346543155820>.
- Davis, Fred D., Richard P. Bagozzi, und Paul R. Warshaw. 1989. «User Acceptance of Computer Technology: A Comparison of Two Theoretical Models». *J Manag Sci*, 35(8), 982–1003.
- Döring, Nicola, und Jürgen Bortz. 2016. *Forschungsmethoden und Evaluation*. Wiesbaden: Springer. <https://doi.org/10.1007/978-3-642-41089-5>.
- Faisal, Nadia, Mehmood Chadhar, Anitra Goriss-Hunter, und Andrew Stranieri. 2022. «Business simulation games in higher education: A systematic review of empirical research». *Human Behavior and Emerging Technologies*, 2022(1), 1578791. <https://doi.org/10.1155/2022/1578791>.
- Fox, Joe, Luke Pittaway, und Ikenna Uzuegbunam. 2018. «Simulations in Entrepreneurship Education: Serious Games and Learning Through Play». *Entrepreneurship Education and Pedagogy*, 1(1), 61–89. <https://www.doi.org/10.1177/2515127417737285>.
- Gimpel, Henner, Kristina Hall, Stefan Decker, Torsten Eymann, Luis Lämmermann, Alexander Mädche, ... Steffen Vandirk. 2023. *Unlocking the Power of Generative AI Models and Systems such as GPT-4 and ChatGPT for Higher Education. A Guide for Students and Lecturers*.
- Haken, Hermann, und Günter Schiepek. 2006. *Synergetik in der Psychologie. Selbstorganisation verstehen und gestalten*. Göttingen: Hogrefe.
- Han, Jining, Geping Liu, und Yuxin Gao. 2023. «Learners in the Metaverse: A systematic review on the use of roblox in learning». *Education Sciences*, 13(3), 296. <https://doi.org/10.3390/educsci13030296>.
- Han, Songhee, Xiaofen Hamilton, Ying Cai, Peixia Shao, und Min Liu. 2023. «Knowledge-based chatbots: a scale measuring students' learning experiences in massive open online courses». *Educational technology research and development*, 71(6), 2431–56. <https://www.doi.org/10.1007/s11423-023-10280-7>.
- Hew, Khe Foon, Weijiao Huang, Jiahui Du, und Chengyuan Jia. 2021. *Using Chatbots in Flipped Learning Online Sessions: Perceived Usefulness and Ease of Use*. Cham: Springer Nature.
- Hoibert, Sebastian, und Raphael Meyer von Wolff. 2019. Say hello to your new automated tutor – a structured literature review on pedagogical conversational agents. Conference Paper.

- Hochmuth, Jörg und Kirchner, Vera. 2019/20. Einstellungen von Wirtschaft unterrichtenden Lehrpersonen zu Fortbildungen in diesem Feld: Ergebnisse eines Online-Surveys im Rahmen des ELF-Projekts.
- Klar, Maria, und Johannes Schleiss. 2024. «Künstliche Intelligenz im Kontext von Kompetenzen, Prüfungen und Lehr-Lern-Methoden: Alte und neue Gestaltungsfragen». *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 58: 41–57. <https://doi.org/10.21240/mpaed/58/2024.03.24.X>.
- König, Alexander. 2024. «KI-Tools für Schule und Unterricht». In *Praxisratgeber: Künstliche Intelligenz als Unterrichtsassistent. Wie KI-Tools das Lehrerleben erleichtern*, herausgegeben von Alexander König und Julia Mosbach. Hannover: Friedrich.
- Kulkarni, Pradnya, Ameya Mahabaleshwarkar, Mrunalini Kulkarni, Nachiket Sirsikar, und Kunal Gadgil. 2019. «Conversational AI: An overview of methodologies, applications & future scope». *Paper presented at the 2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*. <https://doi.org/10.1109/ICCU-BEA47591.2019.9129347>.
- Landes, Tom und Matthias Ziegler. 2015. *Ein Praxisbeispiel zur Konstruktion eines Lehrevaluationsinstruments: Berliner Lehrevaluationsinventar für Vorlesungen (BLEI-VL)* (Vol. 61): Göttingen: Hogrefe.
- Liening, Andreas. 2019. *Ökonomische Bildung: Grundlagen und neue synergetische Ansätze* (2. Aufl.). Wiesbaden: Springer Gabler. <https://doi.org/10.1007/978-3-658-24731-7>.
- Liening, Andreas, und Ewald Mittelstädt. 2011. «Börsen-Planspiele – Leuchttürme oder Irrlichter der Finanziellen Allgemeinbildung?» In *Finanzielle Bildung in der Schule. Mündige Verbraucher durch Konsumentenbildung*, herausgegeben von Thomas Retzmann, 99–114. Schwalbach i. Ts.: Wochenschau.
- Liening, Andreas, Guido Strunk, und Ewald Mittelstädt. 2013. «Management Learning in Times of Crisis: An Experimental Study based on Synergetics». *Nonlinear dynamics, psychology, and life sciences*, 17(4), 517–41.
- Maier, Uwe. 2010. «Formative Assessment – Ein erfolgversprechendes Konzept zur Reform von Unterricht und Leistungsmessung?» *Zeitschrift für Erziehungswissenschaft* 13 (2), 293–308. <https://doi.org/10.1007/s11618-010-0124-9>.
- Miner, Adam S., Arnold Milstein, Stephen Schueller, Roshini Hegde, Christina Mangurian, und Eleni Linos. 2016. «Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health». *JAMA internal medicine*, 176(5), 619–25. <https://doi.org/10.1001/jamainternmed.2016.0400>.
- Montenegro, Joao Luis Zeni, Cristiano André da Costa, und Rodrigo da Rosa Righi. 2019. «Survey of conversational agents in health». *Expert Systems with Applications* 129, 56–67. <https://doi.org/10.1016/j.eswa.2019.03.054>.
- Pekrun, Reinhard, Thomas Goetz, Anne C. Frenzel, Petra Barchfeld, und Raymond P. Perry. 2011. «Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ)». *Contemporary educational psychology* 36 (1), 36–48. <https://doi.org/10.1016/j.cedpsych.2010.10.002>.

- Pekrun, Reinhard, Elisabeth Vogl, Krista R. Muis, und Gale M. Sinatra. 2017. «Measuring emotions during epistemic activities: the Epistemically-Related Emotion Scales». *Cognition and Emotion*, 31(6), 1268–76. <https://doi.org/10.1080/02699931.2016.1204989>.
- Pintrich, Paul R., David A. F. Smith, Teresa Garcia, und Wilbert J. McKeachie. 1991. *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor: National Center for Research to Improve Postsecondary Teaching and Learning.
- Plass, Jan L, Bruce D. Homer, und Charles K. Kinzer. 2015. «Foundations of game-based learning». *Educational psychologist* 50 (4): 258–83. <http://doi.org/10.1080/00461520.2015.1122533>.
- Plass, Jan L., Bruce D. Homer, Richard E. Mayer, und Charles K. Kinzer. 2020. *Theoretical Foundations of Game-Based and Playful Learning*: Mit Press.
- Platz, Liane. 2022. «Learning with serious games in economics education a systematic review of the effectiveness of game-based learning in upper secondary and higher education». *International journal of educational research*, 115, 102031. <https://doi.org/10.1016/j.ijer.2022.102031>.
- Quah, Jon TS, und YW Chua. 2019. «Chatbot assisted marketing in financial service industry». *Paper presented at the Services Computing–SCC 2019: 16th International Conference, Held as Part of the Services Conference Federation, SCF 2019, San Diego, CA, USA, June 25–30, 2019, Proceedings* 16. https://doi.org/10.1007/978-3-030-23554-3_8.
- Reeves, Byron, und Clifford Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places*. Center for the Study of Language and Information, Cambridge University Press.
- Reiter, Ehud, und Robert Dale,. 1997. «Building applied natural language generation systems». *Natural Language Engineering*, 3(1), 57–87. <https://doi.org/10.1017/S1351324997001502>.
- Saad, Syed, Gary W. Carter, Mark Rothenberg, und Enid Israelson. 1999. *Testing and Assessment: An Employer's Guide to Good Practices*.
- Schiepek, Günter, Wolfgang Aichhorn, und Guido Strunk. 2012. «Der Therapie-Prozessbogen (TPB) – Faktorenstruktur und psychometrische Daten». *Zeitschrift für Psychosomatische Medizin und Psychotherapie* 58 (3): 257–66. <https://doi.org/10.13109/zptm.2012.58.3.257>.
- Schlimbach, Ricarda, Bijan Khosrawi-Rad, Tim C. Lange, Timo Strohmann, und Susanne Robra-Bissantz. 2024. «Design Knowledge for Virtual Learning Companions from a Value-centered Perspective». *Communications of the association for Information Systems* 54 (1): 8. <https://doi.org/10.17705/1CAIS.05411>.
- Schöbel, Sofia, Anuschka Schmitt, Dennis Benner, Mohammed Saqr, Andreas Janson, und Jan Marco Leimeister. 2023. «Charting the Evolution and Future of Conversational Agents: A Research Agenda Along Five Waves and New Frontiers». *Information Systems Frontiers*, 1–26.
- Schroeder, Noah L., Robert O. Davis, und Eunbyul Yang. 2025. «Designing and learning with pedagogical agents: An umbrella review». *Journal of Educational Computing Research* 62 (8): 1907–36. <https://doi.org/10.1177/07356331241288476>.

- Sender, Till. 2017. *Wirtschaftsdidaktische Lerndiagnostik und Komplexität. Lokalisierung liminaler Unsicherheitsphasen im Hinblick auf Schwellenübergänge*. Wiesbaden: Springer Gabler. <https://doi.org/10.1007/978-3-658-18947-1>.
- Sikström, Pieta, Chiara Valentini, Anu Sivunen, und Tommi Kärkkäinen. 2022. «How pedagogical agents communicate with students: A two-phase systematic review». *Computers & Education*, 188, 104564. <https://doi.org/10.1016/j.compedu.2022.104564>.
- Streiner, David L., Geoffrey R. Norman, und John Cairney. 2015. *Health measurement scales: a practical guide to their development and use*. Oxford: Oxford University Press. <https://doi.org/10.1093/med/9780199685219.001.0001>.
- Vogl, Elisabeth, Reinhard Pekrun, Kou Murayama, Kristina Loderer, und Sandra Schubert. 2019. «Surprise, Curiosity, and Confusion Promote Knowledge Exploration: Evidence for Robust Effects of Epistemic Emotions». *Frontiers in Psychology*, 10 (2474). <https://doi.org/10.3389/fpsyg.2019.02474>.
- Weber, Florian, Thiemo Wambsganss, Dominic Rüttimann, und Matthias Söllner. 2021. «Pedagogical agents for interactive learning: A taxonomy of conversational agents in education». *Paper presented at the Forty-Second International Conference on Information Systems*. Austin, Texas.
- Weyland, Michael. 2016. *Experimentelles Lernen und Ökonomische Bildung. Ein Beitrag zur fachdidaktischen Entwicklungsforschung*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-14611-5>.
- Wiegand, Dorothee. 2024. «Klett stellt neuen KI-Chatbot für Lehrer vor». Retrieved 04.09.2024, from <https://www.heise.de/news/Klett-stellt-neuen-KI-Chatbot-fuer-Lehrer-vor-9854872.html>.
- Wouters, Pieter, Christof van Nimwegen, Herre van Oostendorp, und Erik D. van Der Spek. 2013. «A meta-analysis of the cognitive and motivational effects of serious games». *Journal of Educational Psychology*, 105 (2), 249–65. <https://doi.org/10.1037/a0031311>.
- Zwack, William R., und Wayne F. Velicer. 1986. «Comparison of Five Rules for Determining the Number of Components to Retain». *Psychological Bulletin*, 99 (3), 432–42. <https://doi.org/10.1037/0033-2909.99.3.432>.