

Nele McElvany, Charlotte Dignath, Michael Becker,  
Hanna Gaspard, Annika Ohle-Peters (Hrsg.)

# „Welche Kompetenzen soll die Schule von heute für die Gesellschaft von morgen vermitteln?“



DORTMUNDER

SYMPOSIUM

IFS

tu technische universität  
dortmund

WAXMANN

Dortmunder Symposium der  
Empirischen Bildungsforschung

herausgegeben von Nele McElvany

Band 8

Nele McElvany, Charlotte Dignath, Michael Becker,  
Hanna Gaspard, Annika Ohle-Peters (Hrsg.)

»Welche Kompetenzen  
soll die Schule von heute  
für die Gesellschaft von morgen  
vermitteln?«



Waxmann 2025  
Münster · New York

### **Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Dortmunder Symposium der Empirischen Bildungsforschung, Band 8  
ISSN 2366-6439

Print-ISBN 978-3-8188-0011-6  
E-Book-ISBN 978-3-8188-5011-1

Waxmann Verlag GmbH, 2025  
Steinfurter Straße 555, 48159 Münster

[www.waxmann.com](http://www.waxmann.com)  
[info@waxmann.com](mailto:info@waxmann.com)

Umschlaggestaltung: Inna Ponomareva, Münster  
Umschlagbild: © Jacob Lund, Adobe Stock  
Satz: satz&sonders, Dülmen

Dieses E-Book steht open access unter der Lizenz CC BY-NC-SA 4.0 (Namensnennung, Nicht kommerziell, Weitergabe unter denselben Bedingungen) zur Verfügung.



Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Zitate etc.) sind von der CC-Lizenz ausgenommen und für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen.

Die freie Verfügbarkeit der E-Book-Ausgabe dieser Publikation wurde ermöglicht durch das Bundesministerium für Bildung und Forschung, den Fachinformationsdienst Erziehungswissenschaft und Bildungsforschung und ein Netzwerk wissenschaftlicher Einrichtungen und Bibliotheken zur Förderung von Open Access in den Erziehungs- und Bildungswissenschaften.

Bibliothek der Pädagogischen Hochschule Freiburg

Bibliothek der Pädagogischen Hochschule Heidelberg

Bibliothek der Pädagogischen Hochschule Zürich

Bibliotheks- und Informationssystem der Carl von Ossietzky Universität Oldenburg

DIPF I Leibniz-Institut für Bildungsforschung und Bildungsinformation Frankfurt a. M.

Fernuniversität Hagen/Universitätsbibliothek

Freie Universität Berlin/Universitätsbibliothek

Gottfried Wilhelm Leibniz Bibliothek – Niedersächsische Landesbibliothek Hannover

Hochschulbibliothek der Pädagogischen Hochschule Karlsruhe

Hochschulbibliothek Hochschule für Technik, Wirtschaft, Kultur Leipzig

Hochschulbibliothek Hochschule Mittweida

Hochschulbibliothek Zittau/Görlitz

Humboldt-Universität zu Berlin/Universitätsbibliothek

Interkantonale Hochschule für Heilpädagogik (HfH) Zürich

IU Internationale Hochschule GmbH Erfurt

Justus-Liebig-Universität Gießen/Universitätsbibliothek

Landesbibliothek Oldenburg

Leuphana Universität Lüneburg

Pädagogische Hochschule Thurgau/Campus-Bibliothek

RPTU Kaiserslautern-Landau/Universitätsbibliothek

Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB)

Technische Informationsbibliothek (TIB) Hannover

Technische Universität Berlin/Universitätsbibliothek

Technische Universität Braunschweig

Technische Universität Chemnitz/Universitätsbibliothek

Universität der Bundeswehr München

Universität Mannheim/Universitätsbibliothek

Universitäts- und Landesbibliothek Darmstadt

Universitäts- und Landesbibliothek Münster

Universitäts- und Stadtbibliothek Köln – im Auftrag der Universität zu Köln

Universitätsbibliothek Augsburg

Universitätsbibliothek Bayreuth

Universitätsbibliothek Bochum

Universitätsbibliothek der LMU München  
Universitätsbibliothek Dortmund  
Universitätsbibliothek Duisburg-Essen  
Universitätsbibliothek Eichstätt-Ingolstadt  
Universitätsbibliothek Erlangen-Nürnberg  
Universitätsbibliothek Graz  
Universitätsbibliothek Greifswald  
Universitätsbibliothek Hildesheim  
Universitätsbibliothek Johann Christian Senckenberg/Frankfurt a. M.  
Universitätsbibliothek Kassel  
Universitätsbibliothek Leipzig  
Universitätsbibliothek Marburg  
Universitätsbibliothek Passau  
Universitätsbibliothek Marburg  
Universitätsbibliothek Passau  
Universitätsbibliothek Potsdam  
Universitätsbibliothek Regensburg  
Universitätsbibliothek Rostock  
Universitätsbibliothek Vechta  
Universitätsbibliothek Würzburg  
Universitätsbibliothek Wuppertal

# Inhalt

Vorwort .....	9
<i>Rebecca Lazarides</i>	
»Welche Kompetenzen soll die Schule von heute für die Gesellschaft von morgen vermitteln?«.	
Erziehungswissenschaftliche Perspektive .....	11
<i>Stefan Krauss &amp; Patrick Wiesner</i>	
Ist die Flipped-Classroom-Methode besser als traditioneller Unterricht?.	
Kritische Bemerkungen zu sieben Metaanalysen für den Schulbereich .....	27
<i>Cordula Artelt</i>	
Bildung in Zeiten von Digitalität und Künstlicher Intelligenz .....	41
<i>Tina Seidel, Michael Nickl, Daniel Sommerhoff, Stefan Ufer, Anika Radkowitzsch, Elisabeth Bauer, Sina Huber, Elias Codreanu, Ralf Schmidmaier &amp; Birgit J. Neuhaus</i>	
Förderung von Zukunftskompetenzen durch videobasierte Simulationen für die Lehrkräftebildung am Beispiel des Projekts Visit-Math .....	57
<i>Franziska Baier-Mosch &amp; Lukas Schulze-Vorberg</i>	
»Künstliche Intelligenz« im eigenen Unterrichtsfach vermitteln.	
»Will und kann ich das lernen?« .....	81
<i>Justine Stang-Rabrig</i>	
Geschlechtsspezifische Unterschiede im Wohlbefinden und die Rolle des wahrgenommenen Leistungsdrucks .....	91



## Vorwort

Vor dem Hintergrund einer sich wandelnden Gesellschaft und der zunehmenden Bedeutung von Digitalisierung stellt sich die Frage, welche Kompetenzen Kinder und Jugendliche in Zukunft benötigen werden und wie es dem heutigen Schulsystem gelingen kann, diese Kompetenzen erfolgreich zu vermitteln. Dieses für Bildungsforschung, -praxis und -administration gleichermaßen hochrelevante Thema kann aus verschiedenen wissenschaftlichen Disziplinen betrachtet werden. Der achte Band der Reihe »Dortmunder Symposium der Empirischen Bildungsforschung« widmet sich daher zum 50-jährigen Jubiläum des Instituts für Schulentwicklungsforschung (IFS) dem Thema »Welche Kompetenzen soll die Schule von heute für die Gesellschaft von morgen vermitteln?«. Neben theoretischen Beiträgen zur Beantwortung dieser Frage aus unterschiedlichen Disziplinen der Empirischen Bildungsforschung werden auch aktuelle Forschungsbefunde zu ausgewählten Zukunftskompetenzen vorgestellt.

Der zentralen Frage dieses Herausgeberbandes nähert sich Rebecca Lazarides in dem ersten Beitrag aus erziehungswissenschaftlicher Perspektive. Bezugnehmend auf klassische Bildungstheorien wird dargelegt, inwiefern in aktuellen internationalen Referenzrahmen zu Kompetenzen der Zukunft erziehungswissenschaftliche Bildungsziele, wie beispielsweise Reflexionsfähigkeit, Berücksichtigung finden und welche Kritikpunkte hinsichtlich der Zielsetzungen und der Vermittlung anzubringen sind. Konkret auf die unterrichtliche Förderung von Kompetenzen bezieht sich der zweite Beitrag von Stefan Krauss und Patrick Wiesner. Auf der Grundlage von Befunden aus Metaanalysen setzen sich die beiden Autoren mit der Wirksamkeit der »Flipped Classroom«-Methode auseinander, die in den letzten Jahren im schulischen Unterricht immer stärker an Bedeutung gewonnen hat. Insgesamt weisen die Ergebnisse auf eine Überlegenheit von Flipped Classroom gegenüber traditionellem Unterricht hin, wobei auch die Grenzen der Metaanalysen diskutiert werden.

Konkret auf ausgewählte Zukunftskompetenzen eingehend beleuchtet der Beitrag von Cordula Artelt die Chancen und Herausforderungen von Digitalität und Künstlicher Intelligenz für Bildung, die sowohl als Mittel zur Unterstützung schulischer Lehr-Lernprozesse als auch als Bildungsgegenstand diskutiert werden. Die Bedeutung eines reflektiven und mündigen Umgangs mit digitalen Medien und Künstlicher Intelligenz (KI) aufgreifend, wird die Erfassung digitaler Kompetenzen im Nationalen Bildungspanel vorgestellt. Ebenfalls aktuelle Forschung darstellend widmet sich der vierte Beitrag von Tina Seidel und Kolleg\*innen dem Projekt »Visit-Math«, das auf die Ausbildung angehender Lehrkräfte fokussiert, welche

eine entscheidende Rolle bei der Vermittlung von zukünftig relevanten Zukunftskompetenzen spielen. Die Debatten um die Verknüpfung von Theorie und Praxis in der Lehrkräfteausbildung aufgreifend werden Konzeption und Validierung einer Simulation zur Förderung der Diagnosekompetenz angehender Mathematiklehrkräfte vorgestellt. Im anschließenden fünften Beitrag von Franziska Baier-Mosch und Lukas Schulze-Vorberg werden bereits im Schuldienst tätige Lehrpersonen in den Fokus genommen und es wird der Frage nachgegangen, inwiefern das Erleben kognitiver Anstrengung und Belastung beim Lernen über KI-Inhalte mit individuellen Faktoren zusammenhängt. Eine Querschnittsstudie mit mehr als 100 Lehrkräften liefert erste Anregungen dafür, was bei der Gestaltung von Professionalisierungsmaßnahmen vor dem Hintergrund des Conceptual Change Ansatzes berücksichtigt werden sollte. Im abschließenden Beitrag von Justine Stang-Rabrig rückt mit dem subjektiven Wohlbefinden eine weitere wichtige Kompetenz in den Fokus. Analysen einer Befragung von mehr als 1.200 Schüler\*innen in Kalifornien zeigten Zusammenhänge zwischen wahrgenommenem Leistungsdruck und negativen Wohlbefindensaspekten; außerdem konnten geschlechtsspezifische Unterschiede in Bezug auf das Wohlbefinden identifiziert werden.

Welche Kompetenzen für Bildungserfolg, Gesellschaft und gesellschaftliche Teilhabe zukünftig relevant sind, ist eine komplexe und interdisziplinär relevante Fragestellung, der sich die Beiträge dieses Herausgeberbandes nähern. Dabei werden sowohl Lernende als auch Lehrkräfte in den Blick genommen und verschiedene, ausgewählte Zukunftskompetenzen und deren Vermittlung aus unterschiedlichen Disziplinen der Empirischen Bildungsforschung diskutiert.

Dortmund im April 2025

Nele McElvany, Charlotte Dignath, Michael Becker, Hanna Gaspard & Annika Ohle-Peters

*Rebecca Lazarides*

## **»Welche Kompetenzen soll die Schule von heute für die Gesellschaft von morgen vermitteln?«**

### Erziehungswissenschaftliche Perspektive

Gesellschaftlich besteht ein Mindestkonsens über die pädagogischen Zielsetzungen von Schule, zu denen unter anderem die Vermittlung von Wissen, Fertigkeiten und Fähigkeiten, die Befähigung zu kritischem Urteilen und die Erziehung zu Freiheit, Demokratie und Toleranz gehören (Blömeke et al., 2007; KMK, 1973). Angesichts multipler sozialer, ökonomischer, ökologischer und gesellschaftlicher Herausforderungen ist es dementsprechend eine zentrale Aufgabe von Schule und Gesellschaft, Schüler\*innen dazu zu befähigen, selbstständig und verantwortungsbewusst an Gesellschaften teilzuhaben und mit den damit verbundenen Problemstellungen konstruktiv umzugehen (OECD, 2018). Vor diesem Hintergrund stellt sich angesichts aktueller gesellschaftlicher Herausforderungen wie der zunehmenden Fülle an Informationen, die nicht mit empirisch überprüfbaren Fakten übereinstimmen (Deichmann, 2019) und demokratiefeindlicher Äußerungen innerhalb und außerhalb von Schulen (Bokelmann, 2022) die Frage, welche Kompetenzen schulische Bildungsprozesse Lernenden vermitteln sollten, um zukünftige Herausforderungen bewältigen zu können. Dieser Frage widmete sich das 50-jährige Institutsjubiläum des Instituts für Schulentwicklungsforschung (IFS), das im Kontext des 8. Dortmunder Symposiums der Empirischen Bildungsforschung zum Thema »Welche Kompetenzen soll die Schule von heute für die Gesellschaft von morgen vermitteln?« stattfand. Der vorliegende Beitrag nimmt eine erziehungswissenschaftliche Perspektive auf das Tagungsthema ein und gibt einen Überblick über verschiedene Vorstellungen von Bildung im 20. und 21. Jahrhundert. Auf das Konzept der »21st Century Skills« der OECD (Ananiadou & Claro, 2009) sowie den damit verbundenen Konsequenzen für Schule, Unterricht und Lehrkräfte bezugnehmend setzt sich der vorliegende Beitrag mit pädagogischen Zielsetzungen auseinander. Der Beitrag geht dabei von der Annahme aus, dass in hohem Ausmaß kritisch-reflektierende, kommunikationsfähige Lernende, denen grundlegende fachliche Kompetenzen fehlen, ein ebenso wenig wünschenswertes Ziel schulischer Bildung sind, wie Lernende mit hohen fachlichen Kompetenzen, die nicht fähig sind, kritisch zu reflektieren und gesellschaftlich Verantwortung zu übernehmen. Diese These ist angesichts eines signifikant negativen Trends in fachbezogenen Kompetenzen in Deutschland (Stanat et al., 2022) und im Hinblick

auf eine gewisse Zustimmung<sup>1</sup> in der Bevölkerung zu Aussagen, die eine Neigung zu elitenkritischem und antipluralem Populismus aufzeigen (Friedrich Ebert Stiftung, 2021), von hoher Relevanz. Im Beitrag wird vor diesem Hintergrund der Frage nachgegangen, welche pädagogischen Zielsetzungen in internationalen Referenzrahmen hervorgehoben werden und welche Herausforderungen mit diesen Vorschlägen verbunden sind. Dabei wird zunächst die historische Entwicklung vom erziehungswissenschaftlichen Bildungsbegriff zu aktuellen Referenzrahmen für Kompetenzen skizziert, um anschließend auf die internationalen Referenzrahmen, damit verbundene Herausforderungen und Kritik einzugehen.

## **Erziehungswissenschaftliche Perspektiven auf Bildung und Kompetenzen**

Im Folgenden werden erziehungswissenschaftliche, normativ begründete Konzeptualisierungen des Bildungsbegriffes vor ihrem jeweiligen historischen Kontext dargestellt, um aufzuzeigen, welche Diskussionen und Paradigmenwechsel die Konzeption aktueller Referenzrahmen für Kompetenzen prägen.

### **Vom klassischen Bildungsverständnis zu »Schlüsselproblemen« der Gesellschaft**

Pädagogische Diskurse zum Bildungsbegriff sind stark durch das neuhumanistische Verständnis von Bildung nach Humboldt (1767–1835) beeinflusst, aus dem unterschiedlichste Bildungskonzepte abgeleitet wurden (Seel & Hanke, 2015). Im Kontext der durch die Französische Revolution entstandenen bürgerlichen Gesellschaft und orientiert an Ideen des politischen Liberalismus formuliert Humboldt (1809/1963) die neuhumanistische Idee einer allgemeinen Menschenbildung als »die höchste und proportionierlichste Bildung seiner Kräfte zu einem Ganzen« (Humboldt, 1792/1980, S. 64). Humboldt beschreibt in seiner Theorie der Menschenbildung, dass Bildung dem Individuum zur Selbstentfaltung und Persönlichkeitsentwicklung dienen sollte und frei von äußeren Anforderungen durch gesellschaftliche oder religiöse Autoritäten oder wirtschaftliche Entwicklungen sein müsse (Koller, 2009). Das von Humboldt entworfene Konzept der allgemeinen Bildung sieht dementsprechend auch ab von der Berücksichtigung der Anforder-

---

1 »Knapp 20 % sind der Meinung, es würde zu viel Rücksicht auf Minderheiten genommen, ein knappes Viertel ist der Auffassung: ›Im nationalen Interesse können wir nicht allen die gleichen Rechten gewähren‹. Und rund 16 % unterstellen: ›Die regierenden Parteien betrügen das Volk‹. [...] 25 % der Befragten neigen zum Populismus, 13 % zum Rechtspopulismus. Legt man ein weiches Kriterium an, sind dies sogar noch mehr.« (Friedrich-Ebert-Stiftung 2021, S. 4)

rungen der Ausbildung – der allseitig gebildete Mensch soll selbstständig und aus eigenem Willen in Gesellschaft und Staat tätig werden können, nicht aber zum Beruf oder zur sozialen Position erzogen werden (Raithel et al., 2009). Humboldt (1809/1963) formulierte dabei Kernbereiche einer allgemeinen Bildung – diese umfassen die historischen, mathematischen, linguistischen und ästhetisch-expressiven Dimensionen der Bildung, die als schulische Lernbereiche verstanden werden können. Diese von Humboldt beschriebenen Kernbereiche stellen auch aktuell noch klassische Elemente aller Grundstrukturen allgemeiner Bildung dar (Tenorth, 2003). Gleichzeitig wurde Kritik am Bildungsverständnis nach Humboldt insbesondere im Hinblick auf die geringe Adaptierbarkeit auf die Komplexität moderner Gesellschaften, auf Anforderungen veränderter Arbeitswelten, aber auch in Bezug auf die limitierten institutionellen Zugänge bestimmter Gesellschaftsgruppen zu allgemeiner Bildung sowie hinsichtlich der ausschließlichen Ausrichtung auf das Individuum geübt (für einen Überblick siehe Koller, 2009).

Den letztgenannten Kritikpunkt greifen Horkheimer (1952/1985) und Adorno (1959/1972) vor dem Hintergrund der Geschehnisse des 2. Weltkrieges auf und schlagen einen Gegenpol zur individualisierenden Ausrichtung des Bildungsgedankens durch die explizite Berücksichtigung der Auseinandersetzung mit den gesellschaftlichen Verhältnissen im Bildungsbegriff vor. Horkheimer betont die Notwendigkeit der Neufassung des Bildungsbegriffes angesichts veränderter gesellschaftlicher Verhältnisse – der Kern eines modernen Bildungsverständnisses sei es, Manipulationen zu erkennen und ihnen mit Zivilcourage und innerer Unabhängigkeit entgegenzutreten. Horkheimer (1952/1985, S. 416) formuliert dementsprechend »Niemand ist gebildet, der nicht in Hingabe an seine eigene Sache ihren Zusammenhang mit dem Ganzen erkennt und der nicht dieselbe Freiheit von Schlagwörtern, Klischees und Vorurteilen, die man im akademischen Beruf in seiner Wissenschaft sich erwerben soll, gegen den Zeitgeist auch in öffentlichen Dingen tätig anwendet« und betont damit besonders die gesellschaftliche Dimension des Bildungsbegriffes.

Die Kritik an der geringen Übertragbarkeit des klassischen Bildungsbegriffes auf moderne Gesellschaften greift auch Klafki (1985, 2007) auf und schlägt in seiner »kritisch-konstruktiven« Didaktik ein Konzept von Allgemeinbildung vor, das zum Ziel hat, Menschen zur Selbstbestimmung, Mitbestimmung und Solidarität zu befähigen und dabei Anforderungen der aktuellen Gesellschaft zu adressieren und praxisbezogen vorzugehen (Gudjons & Traub, 2016). Das Allgemeinbildungskonzept von Klafki ist auf eine allseitige Bildung ausgerichtet, die anhand epochaltypischer Schlüsselprobleme den Menschen als mitverantwortlich für diese Probleme und ihre Lösungen betrachtet – Themen der Schlüsselprobleme sind beispielsweise Frieden, Umwelt, politische und gesellschaftliche Ungleichheiten und informationstechnologische Gefahren und Möglichkeiten (Gudjons & Traub, 2016; Raithel et al., 2009). Zudem betont Klafki in seiner »Theorie der kategorialen Bildung«

die Notwendigkeit einer »Bildung in allen Grunddimensionen menschlicher Interessen und Fähigkeiten« (Klafki, 1985, 2007, S. 54) sowie die Notwendigkeit des anforderungsbezogenen Wissenserwerbs.

Die hier dargestellten erziehungswissenschaftlichen Konzepte allgemeiner Bildung heben Zielgrößen wie allgemeine Bildung, Persönlichkeitsentwicklung und gesellschaftliche Verantwortungsübernahme mit wechselnder Intensität wiederholt als bedeutsam hervor (Bohl et al., 2014; Gudjons & Traub, 2016; Raithel et al., 2009). Gleichzeitig werden auch deutliche Bruchstellen in der erziehungswissenschaftlichen Diskussion um Bildungsziele deutlich, wie die Kritik am neuhumanistischen Bildungsideal des allseitig gebildeten mündigen Menschen durch Vertreter der Kritischen Theorie nach dem zweiten Weltkrieg (Adorno, 1959/1972; Horkheimer, 1952/1985). Einen weiteren zentralen Paradigmenwechsel stellt die Schwerpunktlegung auf messbare Kompetenzen nach den ernüchternden Ergebnissen internationaler Leistungsvergleichsstudien zur Sicherung fachlicher Mindeststandards im deutschen Bildungssystem Anfang der 2000er Jahre dar (Klieme et al., 2003; Köller, 2008; Stanat et al., 2002).

### **Bildungsstandards, fachbezogene und überfachliche Kompetenzen**

In den internationalen Schulleistungsvergleichsstudien der beginnenden 2000er Jahre wurden ausgeprägte Defizite des deutschen Schul- und Bildungssystems deutlich (für einen Überblick siehe Klieme et al., 2010). Die Perspektive der empirischen Bildungsforschung prägte in den folgenden Jahren stark die Diskussion um schulische Bildungsziele. Dabei wurde insbesondere die Normativität bildungstheoretischer Diskurse zu allgemeiner Bildung stark kritisiert und die Notwendigkeit einer zuverlässigen Generierung erfahrungswissenschaftlich abgesicherten Wissens betont, um Prozesse der Wissensaneignung, Erziehung und Bildung im schulischen Kontext besser zu verstehen und gegebenenfalls zielgerichtet zu verändern (Baumert et al., 2001; Prenzel, 2005). Hierbei kommen den Bildungsstandards eine bedeutsame Rolle zu, da sie konkrete pädagogische Ziele pädagogischer Arbeit im Sinne erwünschter Lernergebnisse von Schüler\*innen benennen, die sowohl eine fachliche Systematik, funktionale Anforderungen der Lebens- und Arbeitswelt und Lernvoraussetzungen und Entwicklungsbedürfnisse der Lernenden einbeziehen (Klieme, 2004a). Inwiefern Bildungsstandards erreicht werden, wird durch die Erfassung fachspezifischer Kompetenzen abgebildet. Nach Weinert (2001) werden Kompetenzen als »verfügbare oder erlernte kognitive Fähigkeiten und Fertigkeiten verstanden, die Individuen dazu befähigen, komplexe Probleme zu lösen sowie die individuellen motivationalen Orientierungen, die es ermöglichen, Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll zu nutzen« (zitiert nach Drechsel et al., 2009, S. 345). Dieser Kompetenzbegriff bezieht folglich sowohl kognitive als auch motivationale und volitionale Aspekte

ein und fokussiert sich auf einen handlungsbezogenen Umgang mit Inhalten und auf die Bewältigung von aufgabenspezifischen Anforderungen. Kompetenzen wurden zunächst fachspezifisch formuliert, wobei sich die fokussierten fachbezogenen Kompetenzen wesentlich von der klassischen Auflistung bildungsrelevanter Inhalte durch die Identifikation aufgabenspezifischer Anforderungen unterschieden, die auf bestimmten Kompetenzniveaus bewältigbar sein sollten (Klieme, 2004b). Überfachliche Kompetenzen, wie Argumentationsfähigkeit, Problemlösefähigkeit, Nutzung von Darstellungs- und Präsentationstechniken, fanden sich vor allem als Aspekte der jeweiligen Fachkompetenzen wieder.

Die hier dargestellte Entwicklung führte zu Kontroversen zwischen Bildungsforschung und Bildungstheorie. So wurde aus erziehungswissenschaftlicher Perspektive kritisiert, dass die Konzeption von Bildung als messbare Kompetenzen reduktionistisch und der angestrebte Erwerb überfachlicher und fachlicher Kompetenzen zu stark auf die Erfüllung von Anforderungen des Arbeits- und Ausbildungsmarkts ausgerichtet sei (z. B. Gruschka, 2006). Diese Kritik zurückweisend argumentiert Tenorth (2016), dass die in internationalen Leistungsvergleichsstudien gemessenen Kompetenzen auch zum bildungstheoretisch zentralen Thema der Grundbildung gehören. Weiterhin wurde in Bezug auf die internationale Schulleistungsvergleichsstudie PISA darauf verwiesen, dass diese sich auf die Erfassung grundlegender Basiskompetenzen konzentriere, um zur Teilhabe an modernen Gesellschaften zu befähigen (Baumert et al., 2001). Allerdings wurde diese Zielsetzung nur teilweise erreicht – so fand beispielsweise keine Nivellierung sozialer Ungleichheiten im Bildungssystem statt, obwohl auch dies Anfang der 2000er Jahre als Chance der Kompetenzmessung verstanden wurde (Blömeke, 2006). Auch fast zwanzig Jahre nach der Einführung der Bildungsstandards persistieren soziale Disparitäten in fachbezogenen Kompetenzen und haben sich seit dem Beginn der 2000er Jahre in einzelnen Kompetenzbereichen noch verschärft (McElvany et al., 2023). Gleichzeitig zeigen die Ergebnisse des nationalen Bildungsmonitorings (Stanat et al., 2022) und internationaler Schulleistungsvergleichsstudien (z. B. McElvany et al., 2023), dass es im deutschen Bildungssystem noch zu wenig gelingt, Schüler\*innen im Grundschulalter grundlegende Kompetenzen in den Bereichen Lesen und Zuhören, Orthografie und Mathematik zu vermitteln. Diese Befunde verdeutlichen, dass auch zukünftig die Vermittlung grundlegender fachlicher Kompetenzen eine wichtige Rolle in der Diskussion um notwendige Kompetenzen zur Bewältigung zukünftiger Herausforderungen einnehmen sollte.

Neben den fachbezogenen Kompetenzen wird seit den 1980er und 1990er Jahren als Reaktion auf die Expansion des Bildungswesens und die Ausdifferenzierung moderner Gesellschaften, Arbeitswelten und Wissenschaften auch verstärkt die Bedeutung inhaltsübergreifender Zielsetzungen betont (Klieme, 2004a; Klieme et al., 2003). Dieses Verständnis von Bildungszielen prägte auch die Konzeptionen zentraler von der OECD beschriebener Referenzrahmen für Kompetenzen (Ana-

niadou & Claro, 2009), die versuchen, eine systematische Antwort auf die Frage nach den aktuell und zukünftig zu vermittelnden Kompetenzen zur Verfügung zu stellen. Im Folgenden werden diese Referenzrahmen und ihre Bezüge zu erziehungswissenschaftlichen Konzepten pädagogischer Zielsetzungen vorgestellt.

## **Internationale Referenzrahmen für Kompetenzen**

Im Folgenden werden internationale Referenzrahmen für Kompetenzen skizziert. Exemplarisch bezieht sich das Kapitel auf das Konzept der »21st century skills« der OECD (Ananiadou & Claro, 2009) sowie auf den OECD Lernkompass 2030 (OECD, 2018). Anschließend wird auf Kritikpunkte an den Referenzrahmen aus Sicht der Erziehungswissenschaft und Bildungsforschung eingegangen. Schließlich wird thematisiert, welche Herausforderungen sich in der schulischen Praxis bei der Kompetenzvermittlung stellen.

### **21st century skills und OECD 2030 Learning Compass**

Angesichts komplexer ökonomischer, technologischer, gesellschaftlicher und sozialer Herausforderungen wurden in den letzten beiden Jahrzehnten verschiedene Referenzrahmen für Kompetenzen Lernender verfasst, die sich unter Bezugnahme auf Konzepte wie »21st century skills« (Ananiadou & Claro, 2009), »Future of Education and Skills« (OECD, 2018) und »life skills« (UNESCO, 2015) mit der Frage auseinandersetzen, wie neben fachlichen Kompetenzen auch relevante *überfachliche* Kompetenzen konzeptualisiert und erfasst werden können. Bereits 1997 entstand das OECD-Programm »Definition and Selection of Competencies (DeSeCo)«, das vor dem Hintergrund internationaler Programme wie PISA (OECD, 2005b) zum Ziel hatte, Kompetenzmessungen längerfristig auch in Bezug auf neue Kompetenzdimensionen zu ermöglichen. Dabei standen Kompetenzen im Vordergrund, die es Personen ermöglichen sollten, in verschiedenen Kontexten dazu in der Lage zu sein, ihr Wissen, ihre Fähigkeiten und psychosozialen Ressourcen selbstständig und verantwortlich zu nutzen (OECD, 2005a). Insbesondere drei zentrale Kompetenzkategorien wurden dafür als relevant erachtet: (1) (kulturelle) Werkzeuge/Tools interaktiv nutzen, (2) in heterogenen Gruppen interagieren, und (3) autonom handeln. Aus der DeSeCo-Konzeption entstand der Referenzrahmen »21st century skills and competencies« der OECD (Ananiadou & Claro, 2009), der fachbezogene Kompetenzen in Schlüssel domänen wie Sprachen, Mathematik oder Naturwissenschaften gemeinsam mit themenbezogenen Kompetenzen für »21st century themes« wie globales Bewusstsein oder politische Bildung in den Mittelpunkt rückt (Fadel, 2008). Der Erwerb der fachbezogenen Kompetenzen stellt die Grundlage für den Erwerb überfachlicher Kompetenzen dar – diese werden in »life and career skills« (z. B. Flexibilität, Produktivität, Verantwortungsübernahme, so-

ziale und kulturelle Fähigkeiten), »learning and innovation skills« (z. B. kritisches Denken, Problemlösen, Kommunizieren, Kollaborieren, Kreativität) sowie »information, media and technology skills« (z. B. Fähigkeiten im Umgang mit Informationen sowie in den Bereichen Kommunikation und Technologie [ICT]) unterteilt. Im Vergleich zu anderen internationalen Referenzrahmen betont der »21st century skills and competencies« Rahmen besonders die Bedeutung »neuer« kontextbezogener sowie affektiver und psychosozialer Kompetenzen (Dede, 2010). Eine vom Centre for Educational Research and Innovation der OECD durchgeführte Fragebogenstudie, an der 17 Länder und Regionen im Jahr 2009 teilnahmen, zeigte, dass die »21st century skills and competencies« (außer den ICT skills) vor allem als Querschnittsthemen in schulischen Curricula implementiert waren (Ananiadou & Claro, 2009).

Die Frage nach einem geeigneten Referenzrahmen für Kompetenzen im 21. Jahrhundert wurde im Jahr 2015 von der OECD durch die Initiierung des Projekts »Future of Education and Skills 2030« weiter aufgegriffen. Im Projekt wurde der Referenzrahmen »OECD Learning Framework 2030« entwickelt, der ähnlich wie das »21st century framework« auf der Idee des Zusammenwirkens verschiedener überfachlicher und fachbezogener Kompetenzdimensionen basiert. Der Referenzrahmen bezieht sich auf insgesamt sieben Kompetenzdimensionen, deren Förderung und Vermittlung Lernende dazu befähigen soll, individuelles und gesellschaftliches Wohlergehen zu gewährleisten, indem ökologische, soziale und ökonomische Herausforderungen bewältigt werden können (OECD, 2018). Im Zentrum des Referenzrahmens steht die Handlungs- und Gestaltungskompetenz von Lernenden (»Student Agency und Co-Agency«). Lernumgebungen sollen Schüler\*innen im Rahmen unterstützender Beziehungen dazu in die Lage versetzen, von ihrem Willen und ihrer Fähigkeit zur positiven Beeinflussung des eigenen Lebens und des Lebens anderer überzeugt zu sein. Gleichzeitig sollen sie dazu befähigt werden, zielgerichtet, reflektiert und verantwortungsbewusst zu handeln, um Veränderungen tatsächlich herbeizuführen (OECD, 2020). Neben dieser zentralen Kategorie werden die folgenden weiteren fünf Kompetenzbereiche beschrieben: (1) Transformationskompetenzen (Schaffung neuer Werte, Ausgleich von Spannungen und Dilemmata, Verantwortungsübernahme), (2) Lerngrundlagen (Schlüsselkompetenzen, Wissen, Haltungen und Werte), (3) Wissen (disziplinäres, interdisziplinäres, epistemisches und prozedurales Wissen), (4) Skills (kognitive, metakognitive, soziale, emotionale, praktische und physische Skills) sowie (5) Haltungen und Werte (Prinzipien und Überzeugungen). Seit dem Jahr 2019 wird das Projekt »Future of Education and Skills 2030« mit einer Schwerpunktsetzung auf Lehrkräftekompetenzen und Curriculaentwicklung weitergeführt.

## **Erziehungswissenschaftliche Kritik an den Referenzrahmen**

Inhaltliche Kritik an den Referenzrahmen wird insbesondere in internationalen Arbeiten formuliert (z. B. Hughson, 2022; Hughson & Wood, 2022; Mertanen & Brunila, 2022). Kritisiert wird insbesondere, dass die im »Learning Compass 2030« betonte Bedeutung des Erwerbs fachbezogener Kompetenzen letztendlich nicht einer Bildung im Sinne einer allgemeinen Persönlichkeitsbildung dient, sondern bei genauer Betrachtung Bildung und Wissenserwerb utilitaristisch auf ökonomische Zwecke wie die Qualifikation von zukünftigen Arbeitskräften für die ›knowledge economy‹ ausrichtet sei (Hughson & Wood, 2022). In aktuellen erziehungswissenschaftlichen Diskursen wird Bildung verstanden als der Prozess des Sich-Herausbildens einer Persönlichkeit, eines Selbst- und Wertbewusstseins und einer Individualität, die aber auch Prozesse der Wissensaneignung und der Entwicklung von Können einbezieht (Seel & Hanke, 2015; Klafki, 1998; Koller, 2009). Im Referenzrahmen steht eine Fokussierung auf Bildung zur Bewältigung zukünftiger sozialer, ökonomischer, ökologischer und gesellschaftlicher Herausforderungen im Vordergrund. In ihrer Analyse verschiedener OECD Stellungnahmen, Working Paper und Berichte arbeiten Mertanen and Brunila (2022) zudem heraus, dass die Referenzrahmen der OECD bestimmte positive und negative Zukunftsvisionen (›Utopias and Dystopias‹) in Bezug auf technologische, soziale, ökologische und ökonomische Neuerungen skizzieren, aus denen abgeleitet wird, dass Lernende im Rahmen stark individualisierter Lernprozesse insbesondere zu Flexibilität und Anpassungsfähigkeit befähigt werden müssen, um in den entworfenen Zukunftsvisionen handlungsfähig zu bleiben. Neben dem impliziten normativ-moralischen Gehalt der Referenzrahmen durch das Hinweisen auf gesellschaftliche Risiken, auf die nur mit bestimmten Skills reagiert werden kann, werden auch blinde Flecken kritisiert. Dazu gehört die kaum vorhandene Thematisierung von ethischen Aspekten im Umgang mit künstlicher Intelligenz, die geringe Problematisierung datenschutzrechtlicher Fragen angesichts eines wachsenden privatwirtschaftlichen Marktes für Bildungstechnologie im Kontext der geforderten individualisierten Lernprozesse und die offene Frage, wie in hoch individualisierten (technologiestützten) Lernumgebungen gemeinschaftliche demokratische Werte und bürgerliches Engagement vermittelt werden können (Mertanen & Brunila, 2022).

## **Wie soll Schule Kompetenzen zukünftig vermitteln?**

In internationalen Referenzrahmen der OECD wird die Vision eines hoch individualisierten, personalisierten Bildungssystems entworfen, in dem fachspezifisches Wissen eine zentrale Verständnisgrundlage für den Erwerb neuer Kompetenzen bildet, die als Querschnittsthemen in verschiedene Fächerbereiche Eingang finden sollen (OECD, 2020). Dabei wird der Learning Compass 2030 als erster Versuch

der Richtungsentwicklung einer umfassenden internationalen Curriculumsreform eingeordnet (Hughson & Wood, 2022). Vor diesem Hintergrund stellen sich verschiedene Fragen bezüglich der schulischen Vermittlung. Im Folgenden wird exemplarisch auf drei Herausforderungen eingegangen: die Messung neuer Kompetenzdimensionen, die Fokussierung auf individualisierte, technologiegestützte Lernsettings sowie die Vermittlung von gesellschaftlich relevanten Haltungen und Werten.

Die Messung der in den Referenzrahmen beschriebenen neuen Kompetenzdimensionen wurde als »One of the »hottest« topics in the educational measurement community« bezeichnet (Geisinger 2016, S. 245). Fragen zur Messung neuer Kompetenzdimensionen sind relevant, da sie die Grundlage zur Generierung erfahrungswissenschaftlich abgesicherten Wissens darstellt, das es ermöglicht, Prozesse der Kompetenzaneignung besser zu verstehen und gegebenenfalls zielgerichtet zu verändern (Baumert et al., 2001; Prenzel, 2005). Besonders die Multidimensionalität neuer Kompetenzdimensionen wie (kollaborativem) Problemlösen, Kreativität oder digitalen Kompetenzen, die starke Vernetzung der einzelnen Kompetenzen und der Fokus auf die Bewältigung authentischer, alltäglicher Problemstellungen in variierenden Kontexten erschweren dabei ihre Erfassung (Ercikan & Oliveri, 2016). Zwar lassen sich die häufig komplexen Konstrukte nachweislich reliabel messen, Fragen der Validität sind allerdings noch wenig geklärt (Geisinger, 2016).

Im Hinblick auf die Vermittlung neuer Kompetenzen wird aktuell in internationalen (OECD, 2020) und in nationalen bildungspolitischen Diskussionen (SMK, 2023) insbesondere eine noch stärkere Ausrichtung des Bildungssystems auf die Implementierung individualisierter Lernprozesse mittels neuer Lerntechnologien diskutiert. In diesem Kontext ist es notwendig, auch die Herausforderungen individualisierter Lernprozesse mit neuen Technologien in den Blick zu nehmen. Dazu gehören unter anderem die viel diskutierte Notwendigkeit der Vermittlung digitaler Kompetenzen an Lernende (KMK, 2016) und Lehrkräfte (Redecker, 2017) sowie die Tatsache, dass die Lernergebnisse individualisierter Lernprozesse mit neuen Technologien auch stark von den individuellen Lernvoraussetzungen der Schüler\*innen abhängen. So verweisen Ergebnisse einer Metaanalyse von Steenbergen-Hu and Cooper (2013) für das Fach Mathematik zwar auf schwache positive Effekte von Intelligenten Tutoriellen Systemen in Gruppen Lernender mit durchschnittlichem Leistungsniveau, jedoch auf negative schwache Effekte in Gruppen Lernender mit geringem Leistungsniveau.<sup>2</sup> Für andere offene und potentiell digitale Lern-

---

2 Eingeschlossen wurden  $n = 26$  Forschungsarbeiten, die zwischen 1997 und 2010 publiziert wurden, der Fokus lag auf Lernenden in den Bildungsbereichen vom Kindergarten bis zur zwölften Klasse der High School (»K-12 students«). Die Effekte waren signifikant in den unstandardisierten Modellen, in standardisierten Modellen zeigte sich der hier beschriebene signifikant negative Effekt in den fixed effects Modellen, nicht aber in den random effects Modellen.

settings wie Blended Learning Szenarien mangelt es an metaanalytischer Evidenz zur Bedeutung von Lernendenmerkmalen für die Effektivität der Lernsettings, da bestehende Metaanalysen nur die moderierende Rolle von Untersuchungsdesigns und Lernumgebungen untersuchen (Li & Wang, 2022). Allerdings ist es von hoher Relevanz, vor einer umfassenden Implementierung neuer Technologien über empirische Erkenntnisse über die möglicherweise unterschiedliche Effektivität individualisierter und digitaler Lernumgebungen in unterschiedlichen Lernendengruppen zu verfügen. Dies gilt besonders angesichts der Befunde nationaler (Stanat et al., 2022) und internationaler Schulleistungsvergleichsstudien (McElvany et al., 2023), die darauf verweisen, dass grundlegende fachliche Kompetenzen in den Bereichen Orthografie, Lesen, Zuhören und Mathematik bei einem erheblichen Anteil von Schüler\*innen im deutschen Bildungssystem nicht ausreichend vorhanden sind.

Schließlich stellt sich die Frage, wie neben den fachbezogenen Kompetenzen auch die in den Referenzrahmen beschriebenen zivilgesellschaftlichen Werte in individualisierten, technologiegestützten Lernumgebungen vermittelt werden können. Der Institution Schule kommt im Bereich der Vermittlung demokratischer Werte wie Freiheit, Gleichheit, Gerechtigkeit und Solidarität eine besondere Rolle zu (KMK, 1973). Diese Vermittlung findet in traditionellen schulischen Lernsettings durch Mitbestimmungsmöglichkeiten auf Ebene der Schulklasse und der Schule sowie durch Diskussionen und Kontroversen statt (Bokelmann, 2022). Bislang ungeklärt ist, wie zukünftige Lerngelegenheiten die Auseinandersetzung mit und Aneignung der in Referenzrahmen fokussierten Werte ermöglichen können.

## Ausblick

Ziel des vorliegenden Beitrages war es, aus erziehungswissenschaftlicher Perspektive der Frage nachzugehen, welche Kompetenzen in aktuellen internationalen Referenzrahmen berücksichtigt werden und welche Kritikpunkte mit den Konzeptionen der Referenzrahmen verbunden sind. Deutlich wurde, dass sowohl erziehungswissenschaftliche Bildungsziele als auch die internationalen Referenzrahmen die Notwendigkeit des Erwerbs fachbezogener Kenntnisse und Fertigkeiten als Grundlage für die Aneignung überfachlicher Fähigkeiten wie die zur kritischen Reflexion, zur Kommunikation und zur Verantwortungsübernahme betonen. Die in erziehungswissenschaftlichen Arbeiten hervorgehobene Relevanz des Erwerbs reflexiver Kompetenzen und die Forderung nach Bildung zum Zweck gesellschaftlicher Verantwortungsübernahme, die von Vertretern der Kritischen Theorie herausgestellt wurde, finden sich unter anderem in Konzepten wie der »student agency«, also der Befähigung zur Handlung und zur Gestaltung individueller und gesellschaftlicher Veränderungsprozesse, wieder. Auch die von Klafki (1998) beschriebene Vermittlung fachlicher Kenntnisse und Fertigkeiten unter

Bezugnahme auf gesellschaftliche Schlüsselprobleme findet gewisse Parallelen in den Referenzrahmen, wenn beispielsweise betont wird, dass die Vermittlung fachlicher und überfachlicher Kompetenzen der Bewältigung ungekannter sozialer und wirtschaftlicher Herausforderungen dient, um individuelles und gesellschaftliches Wohlergehen zu sichern (OECD, 2020). Gleichzeitig existieren klare Unterschiede – so legen die vorgestellten Referenzrahmen ihren Schwerpunkt nicht auf die normativ begründete Vermittlung von Inhalten, sondern orientieren sich an einem anwendungsbezogenen, funktionalen und operationalisierbaren Kompetenzbegriff. Zudem variiert der Zweck der Vermittlung grundlegender Kompetenzen zwischen Referenzrahmen und erziehungswissenschaftlichen Grundideen. Während in erziehungswissenschaftlichen Konzeptionen die Herausbildung einer sich selbst erkennenden und reflektierenden Persönlichkeit im Vordergrund steht, wird in aktuellen Referenzrahmen zwar ebenfalls hervorgehoben, dass Individuen gebildet werden müssen, um reflexions- und handlungsfähig zu sein, allerdings wird dabei eine zu eng an ökonomischen Zwecken orientierte Ausrichtung dieser pädagogischen Zielsetzung kritisiert (Hughson & Wood, 2022). Ohne hier vertiefen zu können, inwiefern diese Kritik berechtigt ist, stellt sich angesichts des in den Referenzrahmen formulierten Anspruches, international die Curriculumsentwicklung maßgeblich mitzuprägen, durchaus die Herausforderung, die eigenen Normative bei der Konzeption pädagogischer Zielsetzungen kritisch zu hinterfragen. Ungeklärt sind zudem Fragen der Vermittlung der in internationalen Referenzrahmen vorgeschlagenen komplexen überfachlichen Kompetenzdimensionen – konkret ist zu überlegen, wie beispielsweise gemeinschaftliche Werte in den vorgeschlagenen stark individualisierten und personalisierten Lernumgebungen sozialisiert werden können.

Insgesamt wird deutlich, dass die Frage, welche Kompetenzen für die Bewältigung zukünftiger Herausforderungen vermittelt werden sollten, zahlreiche Fragestellungen auf der Ebene der Konzeptualisierung, Messung und schulischen Vermittlung aufwirft, deren Diskussion in den nächsten Jahren zentral für Bildungspolitik, Bildungspraxis und Bildungsforschung sein wird.

## **Danksagung**

Ich möchte mich bei Prof. Dr. Dirk Richter (Universität Potsdam) für die konstruktive und kritische Rückmeldung zum Manuskript und die spannenden Gespräche über Bildungstheorien und Bildungsforschung bedanken, die für das Schreiben des Manuskripts (und auch sonst generell) sehr hilfreich waren.

## Literatur

- Adorno, T. W. (1959/1972). Theorie der Halbbildung. In A. Busch (Hrsg.), *Soziologie und moderne Gesellschaft: Verhandlungen des 14. Deutschen Soziologentages vom 20. bis 24. Mai 1959 in Berlin* (S. 169–191). Ferdinand Enke. VA: *Gesammelte Schriften* 8, Frankfurt a. M. (S. 93–121).
- Ananiadou, K., & Claro, M. (2009). *21st Century Skills and Competences for New Millennium Learners in OECD Countries*. OECD Education Working Papers No. 41. OECD. <https://dx.doi.org/10.1787/218525261154>
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (Hrsg.). (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Leske und Budrich.
- Blömeke, S. (2006). Literacy, Kompetenzen und Standards versus Bildung? Eine kritische Analyse der Veränderungen im deutschen Schulsystem (und in der deutschen Erziehungswissenschaft). In H.-J. Gamm, W. Keim, D. Kirchhöfer, G. Steffens, C. Uhlig & E. Weiß (Hrsg.), *Jahrbuch für Pädagogik* (S. 211–237). Peter Lang.
- Blömeke, S., Herzig, B., & Tulodziecki, G. (2007). *Gestaltung von Schule: Eine Einführung in Schultheorie und Schulentwicklung*. Julius Klinkhardt.
- Bohl, T., Wacker, A., & Harant, M. (2014). *Schulpädagogik und Schultheorie*. UTB.
- Bokelmann, O. (2022). *Demokratiepädagogik in Kooperation von Jugendhilfe und Schule: Eine explorativ-empirische Untersuchung der Aneignung demokratischer Kompetenzen*. Springer.
- Dede, C. (2010). Comparing frameworks for 21st century skills. In J. A. Bellanca (Hrsg.), *21st century skills: Rethinking how students learn* (Vol. 20, S. 51–76). Solution Tree.
- Deichmann, C. (2019). Neue Herausforderungen für die politische Bildung durch Populismus und Extremismus im »postfaktischen Zeitalter«. In C. Deichmann & M. May (Hrsg.), *Orientierungen politischer Bildung im »postfaktischen Zeitalter«* (S. 83–100). Springer.
- Drechsel, B., Prenzel, M., & Seidel, T. (2009). Nationale und internationale Schulleistungsstudien. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 353–380). Springer.
- Ercikan, K., & Oliveri, M. E. (2016). In search of validity evidence in support of the interpretation and use of assessments of complex constructs: Discussion of research on assessing 21st century skills. *Applied Measurement in Education*, 29(4), 310–318.
- Fadel, C. (2008). *21st century skills: How can you prepare students for the new global economy*. <https://www.oecd.org/site/educeri21st/40756908.pdf>
- Friedrich-Ebert-Stiftung. (2021). *Ergebnisszusammenfassung. Die geforderte Mitte. Rechts-extreme und demokratiegefährdende Einstellungen in Deutschland 2020/2021*. <https://www.fes.de/index.php?eID=dumpFile&t=f&f=68448&token=75a4869c95360ae597f072b724eab4c44ac29a22>
- Geisinger, K. F. (2016). 21st century skills: What are they and how do we assess them? *Applied Measurement in Education*, 29(4), 245–249.
- Gruschka, A. (2006). Bildungsstandards oder das Versprechen, Bildungstheorie in empirischer Bildungsforschung aufzuheben. *Pädagogische Korrespondenz*, 35, 5–22.

- Gudjons, H., & Traub, S. (2016). *Pädagogisches Grundwissen: Überblick – Kompendium – Studienbuch* (12. akt. Aufl.). UTB.
- Horkheimer, M. (1952/1985). Begriff der Bildung. In Schmid Noerr, G. (Hrsg.), *Gesammelte Schriften. Bd. 8.* (S. 409–419). Fischer.
- Hughson, T. A. (2022). The OECD and epistemic (de)colonisation: Globalising visions for knowledge in the Learning Compass 2030. *Compare: A Journal of Comparative International Education*, 2022, 1–17. <https://doi.org/https://doi.org/10.1080/03057925.2022.2078955>
- Hughson, T. A., & Wood, B. E. (2022). The OECD Learning Compass 2030 and the future of disciplinary learning: A Bernsteinian critique. *Journal of Education Policy*, 37(4), 634–654. <https://doi.org/https://doi.org/10.1080/02680939.2020.1865573>
- Humboldt, W. v. (1792/1980). Ideen zu einem Versuch, die Grenzen der Wirksamkeit des Staats zu bestimmen (Auszug). In A. Flitner & K. Giel (Hrsg.), *Werke in fünf Bänden, Bd. 1* (S. 64–69). Wissenschaftliche Buchgesellschaft.
- Humboldt, W. v. (1809/1963). Der Königsberger und der litauische Schulplan. Schriften zur Politik und zum Bildungswesen. In A. Flitner & K. Giel (Hrsg.), *Werke in fünf Bänden, Bd. 4* (S. 169–195). Wissenschaftliche Buchgesellschaft.
- Klafki, W. (1985). Konturen eines neuen Allgemeinbildungskonzepts. In W. Klafki, *Neue Studien zur Bildungstheorie und Didaktik: Beiträge zur kritisch-konstruktiven Didaktik* (S. 12–30). Beltz.
- Klafki, W. (1998). *Grundzüge kritisch-konstruktiver Erziehungswissenschaft*. Michio Ogawara. <https://archiv.ub.uni-marburg.de/sonst/1998/0003/k04.html>
- Klafki, W. (2007). *Neue Studien zur Bildungstheorie und Didaktik*. Beltz.
- Klieme, E. (2004a). Begründung, Implementation und Wirkung von Bildungsstandards: Aktuelle Diskussionslinien und empirische Befunde. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 50(5), 625–634.
- Klieme, E. (2004b). Was sind Kompetenzen und wie lassen sie sich messen? *Pädagogik*, 56, 10–13.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., Schneider, W., & Stanat, P. (Hrsg.). (2010). *PISA 2009: Bilanz nach einem Jahrzehnt*. Waxmann.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., & Tenorth, H.-E. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. BMBF.
- KMK. (1973). *Zur Stellung des Schülers in der Schule. Beschluss der Kultusministerkonferenz vom 25. 5. 1973*. [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/1973/1973\\_05\\_25\\_Stellung\\_Schueler.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1973/1973_05_25_Stellung_Schueler.pdf)
- KMK. (2016). *Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz. Beschluss der Kultusministerkonferenz vom 08. 12. 2016*. [https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2016/2016\\_12\\_08-Bildung-in-der-digitalen-Welt.pdf](https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2016/2016_12_08-Bildung-in-der-digitalen-Welt.pdf)
- Koller, H.-C. (2009). *Grundbegriffe, Theorien und Methoden der Erziehungswissenschaft: Eine Einführung* (4. Aufl.). W. Kohlhammer Verlag.
- Köller, O. (2008). Bildungsstandards-Verfahren und Kriterien bei der Entwicklung von Messinstrumenten. *Zeitschrift für Pädagogik*, 54(2), 163–173.

- Li, S., & Wang, W. (2022). Effect of blended learning on student performance in K-12 settings: A meta-analysis. *Journal of Computer Assisted Learning*, 38(5), 1254–1272. <http://dx.doi.org/10.1111/jcal.12696>
- McElvany, N., Lorenz, R., Frey, A., Goldhammer, F., Schilcher, A., & Stubbe, T. C. (2023). *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre*. Waxmann.
- Mertanen, K., & Brunila, K. (2022). Fragile utopias and dystopias? Governing the future(s) in the OECD youth education policies. *Globalisation, Societies, and Education*, 1–12. <https://doi.org/10.1080/14767724.2022.2121687>
- OECD. (2005a). *The definition and selection of key competencies. Executive summary*. OECD. <https://www.oecd.org/pisa/definition-selection-key-competencies-summary.pdf>
- OECD. (2005b). *PISA 2003 Technical Report*. OECD. [https://www.oecd-ilibrary.org/pisa-2003-technical-report\\_5lgw89r8jdvc.pdf?itemId=%2Fcontent%2Fpublication%2F9789264010543-en&mimeType=pdf](https://www.oecd-ilibrary.org/pisa-2003-technical-report_5lgw89r8jdvc.pdf?itemId=%2Fcontent%2Fpublication%2F9789264010543-en&mimeType=pdf)
- OECD. (2018). *The future of education and skills. Education 2030*. OECD position paper. [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf)
- OECD. (2020). *OECD Lernkompass 2030. OECD-Projekt Future of Education and Skills 2030 Rahmenkonzept des Lernens*. [https://www.oecd.org/education/2030-project/contact/OECD\\_Lernkompass\\_2030.pdf](https://www.oecd.org/education/2030-project/contact/OECD_Lernkompass_2030.pdf)
- Prenzel, M. (2005). Zur Situation der Empirischen Bildungsforschung. In H. Mandl & B. Kopp (Hrsg.), *Impulse für die Bildungsforschung: Stand und Perspektiven. Dokumentation eines Expertengesprächs*. Deutsche Forschungsgemeinschaft (S. 7–21). Akademie Verlag.
- Raithel, J., Dollinger, B., & Hörmann, G. (2009). *Einführung Pädagogik*. Springer.
- Redecker, C. (2017). *European framework for the digital competence of educators. DigComp-Edu*. (EUR, Scientific and technical research series, Bd. 28775). <https://dx.doi.org/10.2760/159770>
- Sächsisches Staatsministerium für Kultus (SMK), R. P.-u. Ö. (2023). Der Weg zum Bildungsland. *KLASSE. Das Magazin für Schulen in Sachsen*, 4, 4–7.
- Seel, N. M., & Hanke, U. (2015). *Erziehungswissenschaft: Lehrbuch für Bachelor-, Master- und Lehramtsstudierende*. Springer VS.
- Stanat, P., Artelt, C., Baumer, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Tillmann, K.-J., & Weiß, M. (2002). *PISA 2000: Die Studie im Überblick*. Max-Planck-Institut für Bildungsforschung.
- Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (Hrsg.). (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (Aufl. 4). Waxmann.
- Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K-12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970–987.

- Tenorth, H.-E. (2003). Bildungsziele, Bildungsstandards und Kompetenzmodelle – Kritik und Begründungsversuche. *RdJB Recht der Jugend und des Bildungswesens*, 51(2), 156–164.
- Tenorth, H.-E. (2016). Bildungstheorie und Bildungsforschung, Bildung und kulturelle Basiskompetenzen – ein Klärungsversuch, auch am Beispiel der PISA-Studien. *Zeitschrift für Erziehungswissenschaft*, 19(1), 45–71. <http://dx.doi.org/10.1007/s11618-016-0707-1>
- United Nations Educational Scientific and Cultural Organisation (UNESCO). (2015). *Education 2030: Incheon declaration and framework for action*. UNESCO. [https://uis.unesco.org/sites/default/files/documents/education-2030-incheon-framework-for-action-implementation-of-sdg4-2016-en\\_2.pdf](https://uis.unesco.org/sites/default/files/documents/education-2030-incheon-framework-for-action-implementation-of-sdg4-2016-en_2.pdf)



Stefan Krauss & Patrick Wiesner

## Ist die Flipped-Classroom-Methode besser als traditioneller Unterricht?

Kritische Bemerkungen zu sieben Metaanalysen für den Schulbereich<sup>1</sup>

### 1. Einleitender Überblick

Flipped-Classroom-Unterricht hat in den letzten Jahren auch aufgrund der COVID-19-Pandemie nochmal an Popularität gewonnen. Auf den ersten Blick könnte der Wissenserwerb zuhause durchaus eine verheißungsvolle Methode sein, um die Unterrichtszeit selbst systematischer zum Einüben beziehungsweise zur Vertiefung von Inhalten verwenden zu können (vgl. 2.1). Wie beinahe jedes schulische Konzept bietet aber natürlich auch der Flipped-Classroom-Unterricht sowohl potenzielle Chancen als auch mögliche Fallstricke (Cevikbas & Kaiser, 2023).

Im vorliegenden Beitrag geben wir einen kurzen Überblick zur Wirksamkeitsforschung bezüglich der Flipped-Classroom-Methode, wobei wir uns auf den schulischen Unterricht beschränken (vgl. 2.2). Hierzu stellen wir zu Beginn einige Metaanalysen vor, die allesamt einen positiven Effekt dieser Unterrichtsform gegenüber traditionellem Unterricht konstatieren (vgl. 2.3). Diese klare empirische Evidenz verunsichert zunächst, würde das doch bedeuten, dass der in Deutschland praktizierte Unterrichtsalltag ernsthaft überdacht werden müsste.

Anschließend werfen wir deshalb einen Blick auf die zugrundeliegenden Originalstudien (für den Primar- und den Sekundarbereich gibt es mittlerweile zahlreiche Einzelstudien). Dabei stellt sich interessanterweise heraus, dass die weit überwiegende Anzahl dieser individuellen Studien lediglich *eine* Klasse im Flipped-Classroom-Unterricht mit *einer* Klasse im traditionellen Unterricht verglichen hat (vgl. 2.4). Um reine Klasseneffekte (z. B. durch Mehrebenenmodellierung) ausschließen zu können, wäre jedoch die empirische Implementation einer substantiellen Anzahl unterschiedlicher Klassen pro Experimentalbedingung erforderlich. Die beiden einzigen Flipped-Classroom-Studien für den Schulbereich, die jeweils

---

<sup>1</sup> Die theoretischen Überlegungen des vorliegenden Beitrags stammen im Wesentlichen aus Wiesner, Krauss, Stegmüller und Binder (eingereicht). Dort werden Ergebnisse einer Flipped-Classroom-Studie mit  $N=950$  Schüler\*innen berichtet ( $N=12$  Klassen pro Experimentalbedingung; siehe Tabelle 2).

eine zweistellige Anzahl von Klassen pro Versuchsbedingung implementiert haben, haben jedoch *keine* positiven Effekte für die Flipped-Classroom-Methode gefunden.

In der empirischen Psychologie bzw. Bildungsforschung ist es keine Seltenheit, dass zahlreiche kleinere Studien wiederholt Effekte finden, die sich in größeren Studien nicht replizieren lassen (vgl. 2.5). Abschließend greifen wir deshalb einige generelle Kritiken an Meta-Studien auf, die eventuell auch alternative Erklärungen für die wiederholte Bestätigung positiver Effekte von Flipped-Classroom-Unterricht liefern könnten (vgl. 2.6). Die im vorliegenden Beitrag angestellten theoretischen Überlegungen sollen dessen Wirksamkeit nicht grundsätzlich widerlegen oder vor einem Einsatz der Flipped-Classroom-Methode warnen (siehe Fazit in Abschnitt 3). Das Ziel des Kapitels ist vielmehr eine Relativierung der scheinbar übermächtigen Evidenz der existierenden Metaanalysen in Bezug auf den Primar- und den Sekundarbereich.

## 2. Flipped Classroom

### 2.1 Eine kurze Begriffsklärung

Flipped-Classroom-Unterricht (dt. »umgedrehter« Unterricht) hat nicht nur aufgrund der Corona Pandemie in der Praxis an Bedeutung gewonnen, sondern rückte gerade in den letzten fünf Jahren zunehmend auch in den Fokus empirischer Forschung (Cevikbas & Kaiser, 2020, 2023; Hwang et al., 2019; Kapur et al., 2022). Fast alle Definitionen beinhalten die »invertierte« Reihenfolge des traditionellen Unterrichts als Mindestcharakterisierung. Im Gegensatz zum »traditionellen« Unterricht (d. h., der Wissenserwerb findet in der Schule statt und das Üben überwiegend zuhause) bedeutet Flipped Classroom, dass neues Wissen von den Schüler\*innen bereits vor der eigentlichen Unterrichtsstunde selbstständig erworben wird, um im Unterricht dann mehr Zeit für Reflexion, Vertiefung und begleitete Übung zu haben (Bishop & Verleger, 2013; Lage et al., 2000). Oftmals werden für das Lernen zuhause Erklärvideos zur Verfügung gestellt, gelegentlich in Kombination mit zusätzlichen kurzen Quizfragen zur Verständnisüberprüfung (z. B. Hew et al., 2021; Wagner et al., 2020). Ein effektiver Flipped-Classroom-Unterricht setzt jedenfalls voraus, dass a) einerseits den Lernenden Materialien zur Verfügung gestellt werden, die sich für eine eigenständige Erarbeitung des Stoffs eignen, und b) andererseits die Schüler\*innen diese Wissensaneignung dann zuhause auch tatsächlich erfolgreich absolvieren.

## 2.2 Einschränkung auf schulischen Unterricht

Flipped-Classroom-Unterricht ist für alle Fächer und alle Jahrgangsstufen denkbar, und auch Veranstaltungen an der Universität können prinzipiell im Flipped-Classroom-Format stattfinden (in der Tat stammt die Methode sogar ursprünglich aus der universitären Bildung). Es ist wichtig festzuhalten, dass wir uns im Folgenden ausschließlich auf den schulischen Bereich konzentrieren (d. h. Primar- und Sekundarstufe) und das Lernen an der Hochschule (d. h. den tertiären Bereich) unberücksichtigt lassen.

Der Grund hierfür ist, dass das eigenständige Lernen zuhause ein gewisses Maß an Selbstregulation erfordert, wofür wiederum Motivation und Eigeninitiative der Lernenden notwendig sind. In dieser Hinsicht sind Schule und Universität nicht direkt vergleichbar. Schon die Wahl eines Studienfachs geschieht auf freiwilliger Basis, genauso die Aufrechterhaltung und der erfolgreiche Abschluss eines Studiums. Die Dozent\*innen von Vorlesungen und Seminaren sind wiederum nur bedingt für die individuellen Erfolge ihrer Studierenden (sowie deren »permanente Überwachung«) zuständig.

Für Schüler\*innen gibt es dagegen generelle Schulpflicht (zumindest bis zum Ende der 9. Klasse) und auch die zu lernenden Fächer können größtenteils nicht selbst gewählt werden. Weiterhin erarbeiten sich Lernende an der Schule neuen Stoff in der Regel nicht selbst. Lehrkräfte an der Schule wiederum haben Aufsichts- und Fürsorgepflichten und stehen für das Lernen der Schüler\*innen in einer größeren Verantwortung als Lehrende an der Universität (zumindest was die gängigen unterrichtlichen Qualitätskriterien *kognitive Aktivierung*, *konstruktive Unterstützung* oder *Klassenführung* betrifft; Kunter et al., 2011).

Aus diesen Gründen fokussieren wir im Folgenden den *schulischen* Unterricht. Zum Einsatz der Flipped-Classroom-Methode an der Hochschule gibt es ebenfalls bereits zahlreiche Einzel- sowie Meta-Studien (diesbezügliche empirische Evidenz findet sich z. B. in Güler et al., 2023; Jang & Kim, 2020; Chen et al., 2018).

## 2.3 Empirische Evidenz aus Meta-Analysen

Tabelle 1 stellt sieben Metaanalysen vor, die die Wirkung der Flipped-Classroom-Methode in der Schule im Vergleich zu traditionellem Unterricht untersucht haben (Cheng et al., 2019; Doğan et al., 2023; Låg & Sæle, 2019; Strelan et al., 2020; van Alten et al., 2019; Wagner et al., 2020; Zhu, 2021).

Alle Metaanalysen kommen zum Schluss, dass Flipped Classroom effektiver ist als traditioneller Unterricht (mit Effektgrößen von  $g = 0.21$  bis zu einem Ausreißer mit  $d = 1.89$ ). Es gibt mittlerweile sogar Meta-Studien über Metaanalysen (Hew et al., 2021; Kapur et al., 2022), die sich jedoch überwiegend mit dem Bereich der universitären Bildung befassen.

Tabelle 1: Sieben Metaanalysen zur Wirksamkeit der Flipped Classroom Methode (Effektstärken jeweils im Vergleich zu traditionellem Unterricht)

Metaanalyse Zeitschrift	Altersstufe (Anzahl berücksichtigter Einzelstudien)	Effektstärke (zugunsten von Flipped Classroom) Cohen's $d$ / Hedges $g^*$
<b>Doğan et al. (2023) **</b> <i>Education &amp; Information Technologies</i>	Sekundarstufe (11)	$d = 0.58$
	Primarstufe (4)	$d = 1.89$
<b>Zhu (2021)</b> <i>Educational Technology Research and Development</i>	Primar- und Sekundarstufe zusammen (27)	$d = 0.54$
<b>Wagner et al. (2020)</b> <i>Zeitschrift für Pädagogische Psychologie</i>	Sekundarstufe (25)	$d = 0.42$
<b>Strelan et al. (2020) **</b> <i>Educational Research Review</i>	Sekundarstufe (21)	$g = 0.64$
	Primarstufe (3)	$g = 0.47$
<b>Låg and Sæle (2019) **</b> <i>AERA Open</i>	Sekundarstufe (16)	$g = 0.45$
	Primarstufe (12)	$g = 0.44$
<b>van Alten et al. (2019) **</b> <i>Educational Research Review</i>	Sekundarstufe (11)	$g = 0.36$
<b>Cheng et al. (2019) **</b> <i>Education Tech Research</i>	Sekundarstufe (12)	$g = 0.21$

Bemerkung: \* Die Effektstärkemaße  $g$  und  $d$  unterscheiden sich nur durch eine leicht abweichende Berechnung der gepoolten Standardabweichung:

$|g$  bzw.  $d| = 0,2$ : kleiner Effekt,  $|g$  bzw.  $d| = 0,5$ : mittlerer Effekt,  $|g$  bzw.  $d| = 0,8$ : großer Effekt

\*\* In diesen Metaanalysen wurde auch der tertiäre Bereich untersucht (nicht in Tab. 1 berücksichtigt)

## 2.4 Empirische Evidenz aus Einzelstudien

Die Metaanalysen aus Tabelle 1 greifen insgesamt auf etwa 50 Einzelstudien zurück, die die Wirkung von Flipped Classroom im Vergleich zu traditionellem Unterricht empirisch für die Primar- beziehungsweise Sekundarstufe untersucht haben. Tabelle 2 stellt diese verfügbaren Einzelstudien, die in mindestens einer der sieben Metaanalysen aus Tabelle 1 enthalten waren, erstmals im kompletten Überblick dar (inklusive Dissertationen, lediglich unveröffentlichte Studien wurden ausgeschlossen). Die Anfangsbuchstaben der Autor\*innen (fett) in Tabelle 1 werden in der linken Spalte in Tabelle 2 verwendet, um anzugeben, welche Studie in welche Metaanalyse einbezogen wurde.

Tabelle 2: Übersicht über die Einzelstudien, die in die Metaanalysen von Tabelle 1 eingegangen sind (angeordnet nach der durchschnittlichen Anzahl an Klassen pro Experimentalbedingung); die ersten beiden Studien (Wiesner et al., eingereicht; Wagner & Urhahne, 2021) sind noch in keine Metaanalysen eingegangen

	Autoren (Jahr) Zeitschrift [berücksichtigt in Metaanalyse] <sup>a</sup>	insgesamt: # SuS, (# Klassen)	Fach (Jahrgangstufe) zusätzliche Info	Flipped Classroom # SuS, (# Klassen)	Kontroll- gruppe # SuS, (# Klassen)	Effektstärke (positiv: zugunsten von Flipped Classroom)	Quelle <sup>b</sup>
1	Wiesner et al. (eingereicht)	950, (37)	Mathematik (6. / 7.)	295, (12)	339, (13)	$d = -0.50^*$	(c)
			Mathematik (6. / 7.) + Unterstützung der Selbstregulation	316, (12)		$d = 0.25$	(c)
2	Wagner & Urhahne (2021) <i>Learning and Instruction</i>	848, (38)	Englisch (9.) + schülerzentriert	215, (10)	169, (8)	$d = -0.14$	(c)
			Englisch (9.) + lehrkraftzentriert	201, (9)	263, (11)	$d = -0.13$	(c)
3	Ramaglia (2015) <i>Dissertation</i> [L]	520, (24)	Mathematik (7.)	137, (6)	137, (6)	$g = 0.13$	(m)
			Mathematik (9.–12.)	129, (5)	49, (3)	$g = 1.09^*$	(m)
			Mathematik (9.–10.)	48, (2)	23, (2)	$g = -0.17$	(m)
4	Bell (2015) <i>Master Thesis</i> [L, Z]	196, (7)	Physik (n.v.)	n.v., (3)	n.v., (4)	$g = 0.07$	(m)
5	Lo et al. (2018) <i>Computers &amp; Education</i> [S, W]	382, (12)	Mathematik (10.–12.)	28, (1)	27, (1)	$d = 0.82^*$	(m)
			Physik (10.–12.)	119, (4)	125, (4)	$d = 0.61^*$	(m)
			Chinesisch (10.–12.)	12, (1)	12, (1)	$d = 0.95$	(m)
			Informatik (10.–12.)	11, (1)	11, (1)	$d = 0.02$	(m)
6	Ripley (2015) <i>Dissertation</i> [C, L]	322, (4 <sup>c</sup> )	Mathematik (6.)	55, (2 <sup>c</sup> )	267, (10 <sup>c</sup> )	$g = 0.01$	(m)
7	Dixon (2017) <i>Dissertation</i> [L, D]	123, (4 <sup>c</sup> )	Biologie (12.)	61, (2 <sup>c</sup> )	56, (2 <sup>c</sup> )	$d = 0.04$	(m)
8	Županec et al. (2018) <i>Journal of Baltic Science Education</i> [D]	112, (4 <sup>c</sup> )	Biologie (7.)	56, (2 <sup>c</sup> )	56, (2 <sup>c</sup> )	$d = 2.15^*$	(c)
9	Atwa et al. (2016) <i>Journal of Personalized Learning</i> [S, D]	109, (4)	Physik (11.)	53, (2)	56, (2)	$d = 0.50^*$	(c)
10	Baki Mohammed Diab, K. M. A. (2016) <i>Dissertation</i> [L]	89, (4)	Mathematik (11.)	39, (2)	40, (2)	$g = 0.97^*$	(m)
11	Bhagat et al. (2016) <i>Educational Technology &amp; Society</i> [L, S, W, Z]	82, (4 <sup>c</sup> )	Mathematik (9.)	41, (2 <sup>c</sup> )	41, (2 <sup>c</sup> )	$g = 0.52^*$	(m)
12	Ceylan & Kesici (2017) <i>Journal of Human Sciences</i> [L]	72, (4 <sup>c</sup> )	Informatik (6.)	37, (2 <sup>c</sup> )	35, (2 <sup>c</sup> )	$g = 0.92^*$	(m)
13	Duffy (2016) <i>Dissertation</i> [D, Z]	87, (4)	Biologie (8.)	44, (2)	43, (2)	$d = 0.28$	(m)
14	Leo & Puzio (2016) <i>Journal of Science Education and Technology</i> [D, L, S]	69, (4)	Biologie (9.)	40, (2)	29, (2)	$g = 0.30$	(m)
15	Salimi & Yousefzadeh (2015) <i>Advances in Language and Literary Studies</i> [C, S, W, Z]	250, (10)	Englisch (7.)	25, (1)	25, (1)	$g = 1.63^*$	(m)
			Arabisch (7.)	25, (1)	25, (1)	$g = 0.51$	(m)
			Mathematik (7.)	25, (1)	25, (1)	$g = 1.83^*$	(m)
			NWT (7.)	25, (1)	25, (1)	$g = 1.57^*$	(m)
			Geographie (7.)	25, (1)	25, (1)	$d = 1.22^*$	(m)

16	Jong (2017) <i>Educational Technology &amp; Society</i> [W, Z]	213, (6)	Wirtschaft (11.) high academic	36, (1)	36, (1)	$g = 0.10$	(m)
			Wirtschaft (11.) medium academic	36, (1)	36, (1)	$g = 0.92^*$	(m)
			Wirtschaft (11.) bottom academic	35, (1)	34, (1)	$g = 0.62^*$	(m)
17	Sergis et al. (2018) <i>Computers in Human Behavior</i> [S, Z]	128, (6)	Informatik (8.)	23, (1)	23, (1)	$g = 0.92^*$	(m)
			Mathematik (10.)	20, (1)	20, (1)		
			Geisteswissen- schaft (8.)	20, (1)	20, (1)		
18	Charles-Ogan & Williams (2015) <i>British Journal of Education</i> [S, Z]	100, (2)	Mathematik (n.v.)	55, (1)	45, (1)	$g = 1.69^*$	(m)
19	Tsai et al. (2015) <i>International Journal of Information and Communication Technology Education</i> [C, L]	96, (3)	Informatik (6.)	50, (2)	46, (1)	$g = 0.41^*$	(m)
20	Chao et al. (2015) <i>Engineering Education</i> [A, L, S, W, Z]	91, (2)	Ingenieurwissen- schaft (11.)	46, (1)	45, (1)	$d = 0.78^*$	(m)
21	Esperanza et al. (2016) <i>Lecture Notes in Computer Science</i> [W]	91, (2)	Mathematik (10.–11.)	45, (1)	46, (1)	$d = 0.59^*$	(o)
22	Pengfei & Mingxuan (2015) <i>Proceedings 2015 International Conference of Educational Innovation through Technology (EITT)</i> [L]	90, (2)	Mathematik (5.)	45, (1)	45, (1)	$g = 0.26^*$	(m)
23	Mohanty & Parida (2016) <i>Creative Education</i> [D, S, Z]	90, (2)	Naturwissenschaft & Geschichte (8.)	45, (1)	45, (1)	$d = 1.09^*$	(c)
24	Sezer (2017) <i>Journal of Educational Computing Research</i> [D, W, Z]	68, (2)	Science (6.)	35, (1)	33, (1)	$g = 0.92^*$	(m)
25	Clark (2015) <i>Journal of Educators Online</i> [A, C, L, S, W, Z]	42, (2)	Mathematik (9.)	21, (1)	21, (1)	$g = 0.03$	(m)
26	Denprapat & Chuaychoowong (2016) <i>Journal of Social Academic</i> [L]	67, (2)	Englisch (7.)	n.v., (1)	n.v., (1)	$g = 0.76^*$	(m)
27	Olakanmi (2017) <i>Journal of Science Education and Technology</i> [D, L, W, Z]	66, (2)	Chemie (7.)	33, (1)	33, (1)	$g = 1.16^*$	(m)
28	Wiginton (2013) <i>Dissertation</i> [A, L, Z]	64, (3)	Mathematik (9.)	24, (1)	22, (1)	$g = 1.69^*$	(m)
29	Afrilyasanti et al. (2016) <i>International Journal of English Language and Linguistics Research</i> [S]	62, (2)	Englisch (10.)	30, (1)	32, (1)	$d = 2.77^*$	(c)
30	Schultz et al. (2014) <i>Journal of Chemical Education</i> [A, C, L, S, W]	61, (2)	Chemie (10.–12.)	29, (1)	32, (1)	$g = 0.96^*$	(m)
31	Saunders (2014) <i>Dissertation</i> [A, C, L, Z]	58, (2)	Mathematik (11.)	30, (1)	28, (1)	$g = -0.06$	(m)
32	Špilka & Popper (2014) <i>ICERI 2014 Proceedings</i> [A, L]	56, (2)	Mathematik (8.)	29, (1)	27, (1)	$g = 0.74^*$	(m)
33	Howell (2013) <i>Dissertation</i> [A, D, W]	56, (2)	Physik (9.)	31, (1)	25, (1)	$g = 0.38$	(m)

34	Kazu & Demirkol (2014) <i>The Turkish Online Journal of Educational Technology</i> [Z]	54, (2)	Biologie (12.)	27, (1)	27, (1)	$g = 0.57^*$	(m)
35	Kirvan et al. (2015) <i>Computers in the Schools</i> [A, C, L, S, W, Z]	54, (2)	Physik (7./ 8.)	25, (1)	29, (1)	$g = 0.28$	(m)
36	Tugun (2018) <i>TEM Journal</i> [S]	52, (2)	Informatik (9.)	28, (1)	24, (1)	$d = 0.52^*$	(c)
37	Johnson & Renner (2012) <i>Dissertation</i> [L]	52, (2)	Informatik (9.-12.)	26, (1)	26, (1)	$g = -0.15$	(m)
38	Aidinopoulou & Sampson (2017) <i>Educational Technology &amp; Society</i> [L, W, Z]	49, (2)	Geschichte (5.)	26, (1)	23, (1)	$g = 0.21$	(m)
39	DeSantis et al. (2015) <i>Journal of Interactive Learning Research</i> [A, L, S, Z]	47, (2)	Mathematik (9.)	26, (1)	21, (1)	$g = -0.12$	(m)
40	Kostaris et al. (2017) <i>Educational Technology &amp; Society</i> [L, S, W, Z]	46, (2)	Informatik (8.)	23, (1)	23, (1)	$g = 1.00^*$	(c)
41	Al-Harbi & Alshumaimeri (2016) <i>English Language Teaching</i> [L, S, W, Z]	43, (2)	Englisch (n.v.)	20, (1)	23, (1)	$g = 0.33$	(m)
42	Katsa M. (2016) <i>International Association for Development of the Information Society</i> [L, W, Z]	40, (2)	Mathematik (10.)	20, (1)	20, (1)	$g = 0.90^*$	(m)
43	Abdelrahman et al. (2017) <i>Online Journal of Educational Technology</i> [S, W]	28, (2)	Englisch (6.)	14, (1)	14, (1)	$d = 1.24^*$	(m)
44	Casem (2016) <i>European Journal of STEM Education</i> [S, W]	24, (2)	Mathematik (12.)	n.v., (1)	n.v., (1)	$d = 0.83$	(c)
45	Gayathri & Vijayarani (2018) <i>International Journal of Research and Analytical Reviews</i> [D]	n.v., (n.v.)	Chemie (11.)	n.v., (n.v.)	n.v., (n.v.)	$d = 1.61^*$	(c)
46	Malto et al. (2018) <i>KnE Social Science</i> [D]	80, (4°)	Biologie (10.)	40, (2°)	40, (2°)	n.v.*	
47	Grazia & Hall (2017) <i>Journal of Computers in Maths and Science Teaching</i> [S]	78, (4°)	Mathematik (9.-12.)	39, (2°)	39, (2°)	n.v.	
48	Loza et al. (2014) <i>Proceedings of the Latin American Conference</i> [W]	72, (4°)	Mathematik (9.)	37, (2°)	35, (2°)	n.v.*	
49	Chen (2016) <i>Journal of Educational Technology Systems</i> [S]	64, (2)	Gesundheitswissenschaft (9.)	33, (1)	31, (1)	n.v.*	
50	Yang (2017) <i>Journal of Information Technology Education</i> [W]	57, (2)	Englisch (8.)	31, (1)	27, (1)	n.v.	
51	Gross (2014) <i>Dissertation</i> [C]	30, (2)	Englisch (12.)	16, (1)	18, (1)	n.v.	
a	A = van Alten et al. [9 eingegangene Studien] C = Cheng et al. [7 eingegangene Studien] D = Doğan et al. [11 eingegangene Studien] L = Låg and Sele [27 eingegangene Studien]			S = Strelan et al. [21 eingegangene Studien] W = Wagner et al. [20 eingegangene Studien] Z = Zhu [21 eingegangene Studien]			
b	(c) = Effektstärke berechnet (da sie weder aus der Originalquelle noch aus den Metaanalysen entnommen werden konnte)			(m) = Effektstärke von einer Metaanalyse entnommen (o) = Effektstärke von der Originalpublikation entnommen			
-	Anzahl der Klassen von den Autoren geschätzt (1-35 SuS entsprechen einer Klasse)						
*	Effektstärke: * signifikant ( $p < 0.05$ )						

In Tabelle 2 haben wir die Flipped-Classroom-Studien als primäres Kriterium nach der *durchschnittlichen Anzahl der Klassen pro Versuchsbedingung* angeordnet (inklusive der Kontrollbedingung mit regulärem Unterricht) und als zweites Kriterium nach der Gesamtanzahl der untersuchten Schüler\*innen. Man beachte, dass letztere sich immer bereits in Klassen befinden und sich auch für Interventionsstudien in der Regel nicht ohne Weiteres für eine Randomisierung aus dem Klassenverband lösen lassen. Da deshalb meistens lediglich die Zuweisung von Klassen zu Versuchsbedingungen zufällig erfolgen kann, macht oftmals erst die Implementation mehrerer Klassen eine empirische Studie im Schulkontext zu einem randomisierten Experiment (aber natürlich ist man auch bei Klassen meist auf Gelegenheitsstichproben angewiesen). Da auch die statistische Mehrebenenmodellierung, die potenzielle Klasseneffekte abschwächen kann, nur möglich ist, wenn mehrere Klassen pro Bedingung implementiert werden, haben wir die durchschnittliche Anzahl der Klassen pro Experimentalbedingung als Hauptkriterium für die Beurteilung der »Größe« einer Flipped-Classroom-Studie gewählt. Wenn diese Anzahl der Klassen in den Metaanalysen nicht explizit angegeben war, haben wir diese den Originalstudien entnommen.

Aus Tabelle 2, die insgesamt 51 Artikel enthält, wird deutlich, dass etwa 40 Studien (Nr. 11–41) nur *eine* Klasse pro Versuchsbedingung untersuchten. Weitere vier Studien (Nr. 6–9) haben *zwei* Klassen für jede Bedingung implementiert und lediglich fünf Studien (Nr. 1–5) haben eine größere Anzahl von Klassen pro Experimentalbedingung untersucht. Die bisherige Evidenz ist also zu einem großen Teil quasi-experimentell. Diese Einschränkung wurde auch in den meisten Metaanalysen erwähnt (z. B. Cheng et al., 2019), wenn auch oftmals nicht in dieser Deutlichkeit.

Bislang scheint es demnach lediglich zwei »größere« Flipped-Classroom-Studien für den Schulbereich zu geben (Wiesner et al., eingereicht; Wagner & Urhahne, 2021). Hierzu ist festzuhalten, dass diese beiden Studien noch zu neu waren, um in die Metaanalysen aus Tabelle 1 mit einzugehen.

In Wagner und Urhahne (2021) wurden durchschnittlich zehn Klassen pro Bedingung untersucht (im Englischunterricht). In dieser Studie war jedoch nicht der Flipped Classroom die effektivste Methode, sondern das gemeinsame Anschauen des Erklärvideos in der Klasse, gefolgt von schüler\*innenzentriertem Unterricht. Da die Lernmaterialien in allen Bedingungen grundsätzlich identisch waren, hat die Studie von Wagner und Urhahne (2021) eine hohe *interne Validität*, aber gleichzeitig gibt es eine Einschränkung hinsichtlich der *ökologischen Validität*, da Erklärvideos in der Regel nicht im Unterricht gezeigt werden. Interessanterweise liefert diese Studie auch Hinweise auf die entscheidende Rolle, die selbstregulative Fähigkeiten speziell in Bezug auf die Wissensaneignung zuhause und somit für die Wirksamkeit der Flipped-Classroom-Methode im Allgemeinen haben könnten. Von den fünf unterschiedlichen von Wagner und Urhahne (2021) identifi-

zierten »Lerntypen« profitieren »Risikoschüler\*innen«, »durchschnittliche Schüler\*innen« und »ausgezeichnete Schüler\*innen« tendenziell weniger vom Flipped Classroom, während sich diese Unterrichtsform laut dieser großangelegten Studie im Besonderen für »selbstständige« und »selbstbewusste« Schüler\*innen eignet.

In Wiesner et al. (eingereicht) wurden zwölf Klassen pro Bedingung untersucht (im Mathematikunterricht). Neben einer »normalen« Flipped-Classroom-Bedingung, in der die Schüler\*innen sich zuhause stochastische Inhalte per Erklärvideos aneignen sollten, gab es eine weitere Flipped Classroom Bedingung, in der die *Selbstregulation* der etwa 10–11-jährigen Kinder explizit unterstützt wurde (vgl. »Unterstützung der Selbstregulation« in Tab. 2). Dies geschah durch eine zusätzliche »mathematikfreie« Unterrichtsstunde, in der den Klassen Strategien zum Umgang mit Lernvideos erläutert und mit ihnen anhand eines nicht-mathematischen Erklärvideos eingeübt wurden (und zwar jeweils für die Phasen »vor dem Video«, »während des Anschauens«, und »nach dem Video«; siehe Wiesner et al., eingereicht). Interessanterweise war der Lernzuwachs im »normalen« Flipped-Classroom-Unterricht geringer als in der Kontrollgruppe mit traditionellem Unterricht. Mit der zusätzlichen Unterstützung durch explizit auf Erklärvideos bezogene Lernstrategien konnte ein kleiner positiver (jedoch nicht-signifikanter) Effekt im Hinblick auf den Lernzuwachs im Vergleich zur Kontrollgruppe und ein signifikanter positiver Effekt gegenüber dem »normalen« Flipped Classroom (d. h. ohne zusätzliche »Strategiestunde«) erzielt werden.

Insgesamt sind die beiden größeren (noch nicht in Metaanalysen berücksichtigten) Studien also weit davon entfernt, die in den vielen kleineren Einzelstudien gefundenen positiven Effekte zugunsten der Flipped-Classroom-Methode zu bestätigen. Das wirft natürlich die Frage auf, ob (und ggfs. welche) alternativen Erklärungen es für die wiederholten positiven Resultate in den kleineren Studien geben könnte.

## 2.5 Kritische Bemerkungen zu Metastudien im Allgemeinen

In der experimentellen Psychologie ist es nicht ungewöhnlich, dass Effekte, die in kleineren Studien gefunden wurden, in großen und gut kontrollierten Experimenten nicht repliziert werden können. Abgesehen von den bekannten Problemen, die im Rahmen der Replikationskrise wiederholt angeführt wurden (z. B., dass in der Regel nur signifikante Ergebnisse veröffentlicht werden und Studien auch nicht präregistriert werden müssen; vgl. Maxwell et al., 2015; Wiggins & Christopherson, 2019), gibt es gerade in jüngster Zeit vermehrt auch Arbeiten, die insbesondere Metaanalysen kritisch in den Blick nehmen (z. B. Bartoš et al., 2023; Kvarven et al., 2020; Sotola, 2022).

Sotola (2022) beispielsweise untersuchte im *Psychology Bulletin* veröffentlichte Metaanalysen und zeigte, dass das Risiko, dass eine solche Metaanalyse einen posi-

tiven Effekt bestätigt, obwohl in Wirklichkeit ein Nulleffekt vorliegt, immer dann besonders hoch ist, wenn die Anzahl der Versuchspersonen in den berücksichtigten Studien sehr klein ist. Sotola mutmaßt darüber hinaus, dass es für jede in einer Metaanalyse aufgenommene Studie eine weitere Studie geben muss, die nicht veröffentlicht wurde. Der Grund für die Nichtveröffentlichung könnte sein, dass diese geschätzten 50 Prozent aller *tatsächlich durchgeführten* Studien keinen oder einen negativen Effekt gefunden haben.

Während Kvarven et al. (2020) konkret davon ausgehen, dass in Metaanalysen berichtete Effektgrößen um den Faktor drei überschätzt sein könnten, gehen Bartoš et al. (2023) sogar von einer noch größeren Überschätzung aus. Die allgemeine Schlussfolgerung aller drei Beiträge lautet, dass a) die berichteten Effektstärken in der Regel wahrscheinlich überschätzt werden, und dass b) Metaanalysen nicht per se unzuverlässig sind, aber mit Vorsicht interpretiert werden sollten. Metaanalysen erlauben zwar die Synthese einer größeren Anzahl ähnlicher Studien und ermöglichen somit Generalisierungen, sind aber natürlich immer nur so gut wie die eingehenden Einzelstudien (was treffend aus dem Titel des Beitrags von Sotola (2022) deutlich wird: »Garbage in, garbage out«).

## 2.6 Alternative Erklärungsmöglichkeiten für Flipped-Classroom-Metaanalysen

Es gibt keine spezifischen Besonderheiten, die die Metaanalysen in Tabelle 1 von denen unterscheiden, die von Bartoš et al. (2023), Kvarven et al. (2020) und Sotola (2022) im Allgemeinen untersucht wurden. In Ermangelung gegenteiliger Argumente vermuten wir deshalb, dass die in diesen drei Arbeiten angestellten theoretischen Analysen auch auf die Meta-Ergebnisse in Tabelle 1 zutreffen könnten. Die Hypothese, dass diese Effektstärken ebenfalls überschätzt sein dürften, steht im Einklang mit der Übersichtsarbeit von Kapur et al. (2022), die ebenfalls nahelegt, dass speziell die Effektgrößen in Flipped-Classroom-Metaanalysen aufgrund der oftmals geringen Stichprobengröße der eingeschlossenen Studien und eines Publikationsbias überschätzt sein könnten (siehe auch Hew et al., 2021). Allerdings fokussieren sich Kapur et al. (2022) und Hew et al. (2021) auf das universitäre Lernen (90 % der z. B. von Kapur et al. (2022) berücksichtigten Studien stammen aus dem Hochschulbereich).

## 3. Fazit

Zunächst haben wir Metaanalysen zur Flipped-Classroom-Methode für den Primar- und Sekundarbereich vorgestellt, die scheinbar ein eindeutiges Bild zugunsten dieser Methode zeichnen (Tab. 1). Anschließend haben wir die individuellen Studien, die diesen Metaanalysen zugrunde liegen, einzeln und der Größe nach

aufgeführt. Dabei kann man feststellen, dass ein Großteil dieser Studien nur ein oder zwei Klassen pro Versuchsbedingung implementiert hat (vgl. Tab. 2).

Demgegenüber stehen mittlerweile zwei größere Flipped-Classroom-Studien mit durchschnittlich 10 bzw. 12 Klassen pro Versuchsbedingung (Wagner et al., 2021; Wiesner et al., eingereicht). Diese beiden Studien, die noch zu neu für die Metaanalysen in Tabelle 1 waren, haben auf der Basis von Mehrebenenmodellierungen *keine* generellen positiven Effekte für die Flipped-Classroom-Methode gefunden.

Anschließend wurde auf der Suche nach einer Aufklärung dieses Widerspruchs allgemein-psychologische Literatur zu Rate gezogen, die sich generell kritisch mit Metaanalysen auseinandersetzt (Bartoš et al., 2023; Kvarven et al., 2020; Sotola, 2022). In diesen Arbeiten wird ausgeführt, dass Effekte in Metaanalysen vor allem dann überschätzt sein dürften, wenn die eingehenden Studien auf kleinen Versuchspersonenanzahlen beruhen (was im Besonderen für Tabelle 1 zutrifft).

Wir können im Rahmen des vorliegenden Beitrags keine Aussage darüber machen, ob die Flipped-Classroom-Methode für das Lernen an der Universität geeignet ist (wozu es sogar noch deutlich mehr Metaanalysen als zum schulischen Unterricht gibt; vgl. z. B. Kapur et al., 2022; bzw. Hew et al., 2021), sondern haben universitäre Lehrformate bewusst vom allgemeinbildenden Unterricht an Schulen abgegrenzt. Da für die Organisation eines Hochschulstudiums ohnehin ein gewisses Maß an Selbstregulation erforderlich ist, könnte es durchaus plausibel sein, anzunehmen, dass ein »Flipped Hörsaal« vorteilhaft sein könnte.

Mit Blick auf Schüler\*innen kann man dagegen berechtigterweise die Frage stellen, ob es ausreicht, sie zwar mit Erklärvideos zu versorgen, dann damit aber zuhause »alleine zu lassen«. Gerade jüngere Kinder sind vermutlich auf Interaktionen mit der Lehrkraft angewiesen, z. B. wenn sie etwas nicht verstanden haben und Rückfragen stellen möchten. Lehrkräfte können – im Gegensatz zu Erklärvideos – adaptiv auf solche Fragen eingehen, sie können Gestik und Mimik von Kindern interpretieren und sich bei Lern- und Verständnisschwierigkeiten aus ihrem fachdidaktischen Repertoire bedienen.

Interessanterweise unterstützt die Übersicht über Flipped-Classroom-Metaanalysen von Kapur et al. (2022) indirekt die Behauptung einer größeren Relevanz der Unterstützung der Selbstregulation der Lernenden, je jünger diese sind. Kapur et al. (2022) zeigen nämlich, dass auf Hochschulebene die durchschnittliche Effektgröße zugunsten der Flipped-Classroom-Methode (ohne spezifische Unterstützung der Selbstregulation und im Vergleich zum traditionellen Unterricht) am höchsten ist ( $g = 0.93$ ), während der Effekt geringer wird, je jünger die Lernenden werden (high school:  $g = 0.63$ , elementary school:  $g = 0.40$ ). Auch in der Studie von Wiesner et al. (eingereicht) mit Schüler\*innen der sechsten bzw. siebten Jahrgangsstufe, war die Unterstützung der Selbstregulation (in Form von spezifisch auf Erklärvideos zugeschnittenen Lernstrategien) das entscheidende Element. Bei

alleiniger Bereitstellung von Erklärvideos war der Lernzuwachs im Flipped-Classroom-Unterricht unter allen drei implementierten Bedingungen am geringsten.

Solange jedoch Versuche scheitern, scheinbar positive Effekte zugunsten der Flipped-Classroom-Methode auch in kontrollierten »large-scale« Designs mit mehreren Klassen pro Versuchsbedingung nachzuweisen, stellt die scheinbar erdrückende Evidenz aus Tabelle 1 für die Institution Schule noch keinen Anlass zur Besorgnis dar. Insgesamt gewinnt man hierbei eher den Eindruck, dass die Forschung zu Flipped Classroom gerade erst am Anfang steht. In der Tat ist ein weiteres Ergebnis aus Tabelle 2, dass die diesbezügliche Wirksamkeitsforschung erst vor etwa 10 Jahren begonnen hat: Alle 51 Einzelstudien stammen aus dem Zeitraum von 2012–2021 (und die sieben Metaanalysen aus Tab. 1 sind konsequenterweise alle aus den letzten 6 Jahren). Wir stimmen mit Wagner et al. (2020, S. 14) jedenfalls überein, dass »in particular, more randomized controlled trials with larger sample sizes and objective quantitative measures are needed.«

Insbesondere ist noch offen, für welche Fächer und für welche Jahrgangsstufen, aber auch spezifischer für welche Unterrichtsthemen pro Fach, sich Flipped-Classroom-Unterricht im Besonderen anbieten könnte. Gerade in Bezug auf die Unterrichtsfächer kommen Metaanalysen derzeit noch zu unterschiedlichen Schlussfolgerungen. Während Wagner et al. (2020) auf Basis metaanalytischer Verfahren größere Effekte für MINT-Fächer finden, konstatieren Strelan et al. (2020) im Gegenteil gerade gesellschaftswissenschaftlichen Fächern ein größeres Potential für den Flipped-Classroom-Unterricht. Diesbezüglich kann mit Spannung auch auf die Ergebnisse der Studie FALKE-d (Frei et al., 2020) gewartet werden, von der Wiesner et al. (eingereicht) eine Teilstudie war und in der in heterogenen Unterrichtsfächern (Mathematik, Chemie, Musik, Deutsch, Sachunterricht in der Grundschule) auf Grundlage eines parallelen Designs und mit jeweils großen Stichproben ebenfalls die drei Bedingungen (»normaler« Flipped Classroom vs. Flipped Classroom mit zusätzlicher Unterstützung des selbstregulierten Lernens vs. Kontrollgruppe mit traditionellem Unterricht) implementiert wurden.

Weiterhin werden in Flipped-Classroom-Studien die Schüler\*innen oftmals mit von den Forschenden aufwändig konzipierten Materialien für die Wissensaneignung zu Hause versorgt. Weitestgehend noch offen ist diesbezüglich, welche Effekte ökologisch valider Flipped-Classroom-Unterricht hat, d. h., wenn die Lehrkräfte im Unterrichtsalltag selbständig Materialien entwerfen oder Erklärvideos produzieren müssen.

## Literatur

- Bartoš, F., Maier, M., Shanks, D. R., Stanley, T. D., Sladekova, M., & Wagenmakers, E.-J. (2023). Meta-analyses in psychology often overestimate evidence for and size of effects. *Royal Society Open Science*, *10*(7), 230224. <https://doi.org/10.1098/rsos.230224>
- Bishop, J., & Verleger, M. (2013). *The Flipped Classroom: A survey of the research*. 2013 ASEE Annual Conference & Exposition, Atlanta, Georgia, United States.
- Cevikbas, M., & Kaiser, G. (2020). Flipped classroom as a reform-oriented approach to teaching mathematics. *ZDM*, *52*(7), 1291–1305.
- Cevikbas, M., & Kaiser, G. (2023). Can flipped classroom pedagogy offer promising perspectives for mathematics education on pandemic-related issues? A systematic literature review. *ZDM*, *55*(1), 177–191.
- Chen, K. S., Monrouxe, L., Lu, Y. H., Jenq, C. C., Chang, Y. J., Chang, Y. C., & Chai, P. Y. C. (2018). Academic outcomes of flipped classroom learning: A meta-analysis. *Medical education*, *52*(9), 910–924.
- Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019). Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis. *Educational Technology Research and Development*, *67*(4), 793–824.
- Dogan, Y., Batdi, V., & Yasar, M. D. (2023). Effectiveness of flipped classroom practices in teaching of science: A mixed research synthesis research. *Science & Technological Education*, *41*(1), 393–421.
- Frei, M., Asen-Molz, K., Hilbert, S., Schilcher, A., & Krauss, S. (2020). *Die Wirksamkeit von Erklärvideos im Rahmen der Methode Flipped Classroom*. Waxmann.
- Güler, M., Kokoç, M., & Önder Bütüner, S. (2023). Does a flipped classroom model work in mathematics education? A meta-analysis. *Education and Information Technologies*, *28*(1), 57–79.
- Hew, K. F., Bai, S., Dawson, P., & Lo, C. K. (2021). Meta-analyses of flipped classroom studies: A review of methodology. *Educational Research Review*, *33*, 100393. <https://doi.org/10.1016/j.edurev.2021.100393>
- Hwang, G.-J., Yin, C., & Chu, H.-C. (2019). The era of flipped learning: promoting active learning and higher order thinking with innovative flipped learning strategies and supporting systems. *Interactive Learning Environments*, *27*(8), 991–994.
- Jang, H. Y., & Kim, H. J. (2020). A meta-analysis of the cognitive, affective, and interpersonal outcomes of flipped classrooms in higher education. *Education Sciences*, *10*(4), 115.
- Kapur, M., Hattie, J., Grossman, I., & Sinha, T. (2022). Fail, flip, fix, and feed – Rethinking flipped learning: A review of meta-analyses and a subsequent meta-analysis. *Frontiers in Education*, *7*, Article 956416. <https://doi.org/10.3389/educ.2022.956416>
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Waxmann.
- Kvarven, A., Strömland, E., & Johannesson, M. (2020). Comparing meta-analyses and pre-registered multiple-laboratory replication projects. *Nature Human Behaviour*, *4*(4), 423–434.

- Låg, T., & Sæle, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-Analysis. *AERA Open*, 5(3), 233285841987048.
- Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, 31(1), 30–43.
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does »failure to replicate« really mean? *American Psychologist*, 70(6), 487.
- Sotola, L. K. (2022). Garbage In, Garbage Out? Evaluating the evidentiary value of published meta-analyses using z-curve analysis. *Collabra: Psychology*, 8(1), Article 32571.
- Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review*, 30, 100314.
- van Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review*, 28, 100281.
- Wagner, M., Gegenfurtner, A., & Urhahne, D. (2020). Effectiveness of the flipped classroom on student achievement in secondary education: A meta-Analysis. *Zeitschrift Für Pädagogische Psychologie*, 35(1), 11–31.
- Wagner, M., & Urhahne, D. (2021). Disentangling the effects of flipped classroom instruction in EFL secondary education: When is it effective and for whom? *Learning and Instruction*, 75, 101490.
- Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, 39(4), 202.
- Wiesner, P., Krauss, S., Stegmüller, N., & Binder, K. (eingereicht). Is Flipped Classroom really superior? Questioning the flip in K-12 teaching.
- Zhu, G. (2021). Is flipping effective? A meta-analysis of the effect of flipped instruction on K-12 students' academic achievement. *Educational Technology Research and Development*, 69(2), 733–761.

*Cordula Artelt*

## **Bildung in Zeiten von Digitalität und Künstlicher Intelligenz**

### **1. Einführung**

Bildung in Zeiten von Künstlicher Intelligenz (KI) und Digitalität ist ein weites Feld. Nicht nur der Einsatz digitaler Medien im und für den Unterricht und die richtige Balance zwischen klassischen Unterrichtsformen und digital unterstützten und begleiteten Formen ist auszuloten, auch die jeweiligen Tools und Plattformen gilt es für die Zwecke des Fachunterrichts weiterzuentwickeln und in die Breite zu bringen. Eine besondere Herausforderung ist dabei der »intelligente« Einsatz von KI, aber auch von neuen Formen der Veranschaulichung und des Erlebens, wie sie über immersive Verfahren, den Einsatz von Virtual oder Augmented Reality möglich sind. Und die Frage, wann welche Formen entwicklungsadäquat, lernförderlich und in das Curriculum integrierbar sind. Im Nationalen Bildungsbericht 2020 (Autor:innengruppe Bildungsberichterstattung, 2020, Abb. H1-2) haben wir die Funktionen digitaler Medien und Werkzeuge in Anlehnung an Diethelm (2018) u. a. als Hilfsmittel für Bildungsprozesse (z. B. Vokabelapps, Geometriesoftware) und als Gegenstand von Bildungsprozessen (z. B. Wissen über typische Anwendungen und Funktionen digitaler Medien) unterschieden. Darüber, welche Auswirkungen digitale Medien (und KI) auf Lernen, Einstellungen und Haltungen von Lernenden haben und welche Rezeptionsanforderungen sie an Lernende stellen, d. h. warum sie als Gegenstand von Bildungsprozessen zu begreifen sind, wird im zweiten Teil dieses Beitrags genauer betrachtet und in dem sich daraus ergebenden Bildungsauftrag beleuchtet. Digitale Medien haben aber auch Funktionen als Organisationsmittel und unterstützen das pädagogische Personal durch Lernmanagementsysteme und digitale Prüfungsformate bei der Organisation von Bildungsprozessen. Oder sie fungieren als Lehr-Lern-Werkzeug (z. B. zur Gestaltung und Modifikation von digitalen Lerninhalten wie Makros und Skripten). Insbesondere multimediale, interaktive und adaptive digitale Werkzeuge können die Qualität individueller und kollaborativer Lehr-Lernprozesse steigern. Digitale Werkzeuge ermöglichen dabei, durch den zeit- und ortsunabhängigen Zugang zu umfangreichen Bildungsressourcen für eine breite Nutzer\*innengruppe, ein großes Potential für die Bildung. Gleichzeitig sind die Anforderungen an die intelligente Nutzung dieser Funktionen hoch, und Lehrkräfte müssen in der ersten und zweiten Phase der Ausbildung, aber auch in der Weiterbildung, hierauf vorbereitet werden.

Der Beitrag gliedert sich in drei Teile und ein Fazit. Zunächst stehen jene Funktionen digitaler Medien im Mittelpunkt, bei denen sie als Hilfsmittel zum Lern- und Kompetenzerwerb oder zur Organisation und Orchestrierung des Lehrens und Lernens (noch nicht) verwendet werden. Im Abschnitt »Digitale Medien zur Unterstützung von Lern- und Lehrprozessen« werden Arbeiten und Entwicklungen sowie Bedarfe in diesem Bereich dargestellt. Anschließend geht es um »Digitale Medien und KI als Gegenstand von Bildung«, also um die Anforderungen und die wichtigen Kompetenzen in diesem Bereich, wie um die Notwendigkeit, diese mit als schulischen Auftrag zu verstehen. Im letzten Teil »ICT und digitale Kompetenzen im Nationalen Bildungspanel« werden aktuelle Arbeiten im Kontext des Nationalen Bildungspanels (NEPS) dargestellt. Hier wird in der 2022 neu gestarteten Kohorte von Schüler\*innen der Sekundarstufe I ergänzend zur ICT Literacy gemessen, wie es um die Ausprägung jener digitalen Kompetenzen bestellt ist, die einen verantwortungsvollen Umgang mit der Reflexion und Praktiken im Umgang mit digitalen Medien und KI umfassen. Im abschließenden Fazit werden Desiderata für die schulische Bildung im Bereich Digitaler Souveränität benannt.

## **2. Digitale Medien zur Unterstützung von Lern- und Lehrprozessen**

Durch die Corona-Pandemie wurde für alle Beteiligten deutlich, dass Deutschland in Bezug auf die intelligente Nutzung digitaler Medien im Bereich des Lehrens und Lernens vor einer großen Entwicklungsaufgabe steht (z. B. Nationale Akademie der Wissenschaften, 2020; Wolter et al., 2020). Anders als in anderen Ländern waren Lösungen für alternative digitale Lehr-Lern-Szenarien und Angebote, die in Zeiten der Kontaktbeschränkungen notwendig waren, nicht vorhanden und/oder in der Breite nicht umsetzbar. Aber auch unabhängig von der Frage, ob Unterricht auch digital durchgeführt werden muss oder soll, ist die Nutzung von lernförderlichen und in die didaktische Gestaltung des Fachunterrichts eingebetteten digitalen Medien im Unterrichtskontext ein wichtiges Forschungsdesiderat und eine Entwicklungsaufgabe. Die ICILS-Studie zeigt für Deutschland auch in diesem Bereich Defizite: digitale Medien werden vergleichsweise selten in Lehr- und Lernsettings eingebunden (Beblavý et al., 2019; Eickelmann et al., 2019), was teilweise auf eine unzureichende Ausstattung und/oder fehlenden technischen Support zurückgeführt wurde. Zwar war in der Pandemie ein deutlicher Anstieg bzgl. Ausstattung und Nutzungshäufigkeit digitaler Medien und Tools zu erkennen (Mußmann et al., 2021), der durch die Coronakrise bedingte Digitalisierungsschub hat Deutschland – so Huber et al. (2020) – jedoch international nicht weit nach vorne gebracht. Auch entfalten digitale Medien allein nicht automatisch eine lernförderliche Wirkung. Vielmehr muss eine an sie angepasste Lernkultur der Schlüssel für Verbesserungen sein (Cress et al., 2018). Durch Digitalisierung ergeben sich Veränderungen in fachspezifischen Arbeitsweisen, die auch in den fachlichen Bildungs-

zielen berücksichtigt werden müssen. Eine zentrale Aufgabe für das pädagogische Personal besteht dabei darin, digitale und analoge Unterrichtsangebote zu einem Gesamtkonzept zu verbinden und zu orchestrieren. Didaktiken und damit auch die Fachdidaktiken in der Lehrkräfteausbildung müssen diese neuen Angebote ebenso integrieren wie entsprechende Angebote für die Weiterbildung entwickelt und implementiert werden müssen.

Vor dem Hintergrund der Befundlage legt die neue BMBF-Förderinitiative einen Schwerpunkt auf die Stärkung digitaler Transformation im Bildungsbereich und die Professionalisierung des Lehrens und Lernens mit Medien. So sind in Bezug auf die Frage, wie digitale Medien und Tools – neuerdings auch KI – den Kompetenzerwerb im fachgebundenen Unterricht stärken können, im Kontext des BMBF-geförderten Kompetenzverbunds *lernen:digital* (2023) zahlreiche Initiativen und Vorhaben entstanden, die darauf abzielen, den »Dialog zwischen Wissenschaft und Praxis für die digitale Transformation von Schule und Lehrkräftebildung« zu stärken. In drei der vier Kompetenzzentren, die sich mit den Bereichen MINT, Sprachen/Gesellschaft/Wirtschaft, Musik/Kunst/Sport befassen, steht die digitale Transformation in fachgebundenen Kontexten im Vordergrund. Im Mittelpunkt stehen Fort- und Weiterbildungsangebote sowie Materialien für digital gestützten Unterricht in den jeweiligen Fächern, die forschungsbasiert entwickelt, erprobt und implementiert werden sollen. Bei Fragen nach Digitalen Medien und KI als Mittel zum Kompetenzerwerb geht es primär darum, welche Medien, Tools und Methoden fachgebundenes Lehren und Lernen unterstützen, wie sie didaktisch eingebunden werden könnten und sollten und wie sie entwickelt werden können, ohne zu starke Abhängigkeiten von großen Technologiekonzernen zu erzeugen. Und es geht auch darum, welches Wissen, welche technischen/didaktischen und digitalen Kompetenzen und Voraussetzungen dafür auf Seiten der Lehrenden notwendig sind und wo diese vermittelt werden.

Diese Entwicklungen sind wichtig, und erhebliche Anstrengungen sind notwendig, um die Rahmenbedingungen dafür zu schaffen, dass das schulische Lehren und Lernen die positiven/lernförderlichen Wirkungen des Einsatzes digitaler Medien und KI für den Unterricht nutzt. Gleichwohl rückt nachfolgend ein anderer Aspekt in den Mittelpunkt der Betrachtung. Es geht es um die Notwendigkeit des Einsatzes von Digitalen Medien und KI als *Gegenstand des Kompetenzerwerbs* und der sich vollziehenden Veränderung und sich daraus ergebenden Implikationen von Bildung durch Digitalität.

### 3. Digitale Medien und KI als Gegenstand von Bildung

Digitale Medien und Werkzeuge selber müssen als Gegenstand von Bildung betrachtet werden. Als Lehr-Lern-Gegenstand bezieht sich dies einerseits auf Fähigkeiten der Handhabung von Informations- und Kommunikationstechnologien (ICT Literacy), andererseits auch auf das Wissen darüber, wie Systeme funktionieren und welche Qualität Produkte haben. In diesem Sinne bedeutet die Berücksichtigung von digitalen Medien und KI als Gegenstand von Bildung, dass die Grundlagen dafür geschaffen werden, dass diese produktiv aufgenommen und mitgestaltet werden können – etwa durch die Nutzung digitaler Medien zur Gestaltung und Modifikation von digitalen Lerninhalten wie Makros und Skripten als Lerninhalt, also als Werkzeuge. Obwohl digitale Medien und Technologien in der Freizeit häufig genutzt werden (z.B. Medienpädagogischer Forschungsverbund Südwest, 2017), hat die international vergleichende ICILS-Studie aus den Jahren 2013 und 2018 für den Bereich der Schüler\*innenkompetenzen im Bereich der Computer- and Information-Literacy keine guten Leistungen attestiert (Bos et al., 2014; Eickelmann et al., 2019).

Neben diesen auch produktorientierten Aspekten, die die Förderung von ICT-Kompetenzen nahelegen, müssen die Gesellschaft und insbesondere das gesellschaftlich getragene Bildungssystem ein Konzept für das Bild des Menschen entwickeln, der Digitalisierung nicht nur produktiv aufnimmt und mitgestaltet, sondern auch kritisch verarbeitet. Der Aktionsrat Bildung (Blossfeld et al., 2018) fasst dieses unter dem Begriff *digitale Souveränität*: »Bildung in einer digital vernetzten Welt erweitert heute Medienkompetenz um den Begriff der digitalen Souveränität. Neben dem Erlernen der Kulturtechniken Lesen, Schreiben und Rechnen ist der souveräne Umgang mit digitalen Medien die Voraussetzung für eine systematische Verankerung der Medienbildung im Handeln jedes Einzelnen« (Blossfeld et al., 2018, S. 17). Neben der technischen Souveränität, die sowohl als passiv-rezeptive als auch als aktiv-gestaltende Kompetenz aufgefasst wird und die hier nicht weiter differenziert werden soll, führen die Mitglieder des Aktionsrats den Begriff der ethisch-reflexiven Souveränität ein, der über die Facetten Wissen, Kompetenz und Bildung gekennzeichnet wird. Geschichte, Theorie und Implikationen von (digitaler) Information und Kommunikation stellen die Wissensfacetten dar, als Kompetenz wird beispielhaft das Hinterfragen von Quellen, Erkennen »falscher« Tatsachen, die Unterscheidung von Fakten und Meinungen und die verantwortungsvolle Formulierung und Platzierung eigener Beiträge aufgeführt. Aus der weiteren Perspektive von Bildung wird Souveränität als individuell und gesellschaftlich verantwortungsvoller Umgang mit der Rezeption, Erstellung und Umsetzung von digitaler Information und Kommunikation verstanden (Blossfeld et al., 2018, S. 18). Auch in den über den Bildungsbereich hinausgehenden Diskursen ist der Begriff der Digitalen Souveränität, definiert als »die Summe aller

Fähigkeiten und Möglichkeiten von Individuen und Institutionen, ihre Rolle(n) in der digitalen Welt selbstständig, selbstbestimmt und sicher ausüben zu können« (Goldacker, 2017, S. 3) mittlerweile etabliert (vgl. Friedrichsen & Bisa, 2016; Peucker, 2020; Rohleder, 2019; Koziol & Vogel, 2020; Ernst, 2020; Frederking, 2022; Müller et al., 2020).

Medien und KI selber stellen somit einen eigenen wichtigen Gegenstandsreich dar, dem insbesondere im formalen Bildungskontext/im schulischen Curriculum bisher jedoch zu wenig Beachtung geschenkt wird. Der kompetente Umgang mit Daten und digitalen Informationen sowie ein proaktives Verständnis der Potenziale, aber auch der Gefahren ihrer Nutzung sind in unserer digital geprägten Welt wichtige Schlüsselkompetenzen und ein grundlegender Baustein der digitalen Transformation unserer Gesellschaft. Digitale und datenbezogene Kompetenzen sind u. a. Voraussetzung für die Entfaltung eigener Handlungsspielräume, für bürgerschaftliches Engagement und den mündigen Umgang mit eigenen und fremden Daten.

Welche Kompetenzen benötigen Kinder und Jugendliche (und Lehrkräfte) beim Umgang mit digitalen Medien und KI? Anders als Schulbücher und konventionelle Unterrichtsmedien stellen die durch Algorithmen gesteuerten Medien neue Anforderungen und gehen mit anderen Interaktionsformen einher. Die Nutzung sozialer Medien ist unter Kindern und Jugendlichen heute enorm weit verbreitet (Medienpädagogischer Forschungsverbund Südwest, 2017) und YouTube und Google sind schon lange, ChatGPT und weitere generische KIs erst seit kürzerem vertraute Partner bei der Suche nach Antworten/Lösungen oder Informationen. Die Veränderungen durch KI und die immer weniger mit einfachen Modellen zu verstehenden Interaktionen mit Medien gehen jedoch auch mit einem Auftrag an die Schule einher. Mündigkeit im Umgang und digitale Souveränität beinhalten auch, dass Kinder und Jugendliche über die Potenziale genauso wie über die Gefahren im Bilde sind. Sowohl die private als auch die berufliche Nutzung digitaler Medien bergen Chancen und Risiken in sich, die es differenziert zu betrachten gilt. In den Fokus rücken damit Kompetenzen, die zur Bewältigung der mit der Digitalisierung einhergehenden, gesellschaftlichen Herausforderungen nötig sind (Brandhofer et al., 2019; Lachner et al., 2020). Die Nutzung von digitalen Technologien, KIs und Social-Media-Plattformen erfordert eine kritisch-reflexive Auseinandersetzung mit den damit einhergehenden Anforderungen und Risiken. Es gilt entsprechende Kenntnisse u. a. zu »Fake News«, Privatsphäre und Datensicherheit zu vermitteln, um dem gesellschaftlichen Anspruch an den Bildungssektor gerecht zu werden, »souveräne, mündige Bürger und Bürgerinnen in einer digitalen Gesellschaft« (Seufert et al., 2018, S. 179) hervorzubringen. Digitale Kompetenzen umfassen somit nicht nur medienspezifische, sondern auch nicht-medienspezifische Fähigkeiten und bestimmte Sozial- und Selbstkompetenzen, die im Umgang mit Digitalität wichtig sind (Brandhofer et al., 2019; Petko et al., 2018). Die Arbeiten

der Arbeitsgruppe um Olga Zlatkin-Troitschanskaia machen deutlich, dass auch bei Berufsanfänger\*innen in Jura, Medizin und im Lehramt zum Teil erhebliche Defizite im kritischen Umgang mit online-Informationen bestehen (z. B. Nagel et al., 2022).

Auch der von der Europäischen Kommission im Jahr 2020 verabschiedete aktualisierte *Aktionsplan für Digitale Bildung (DEAP)*, der Strategien für die Laufzeit 2021 bis 2027 konkretisiert, betont diesen Aspekt (European Commission, 2020). Die 14 konkreten Maßnahmen für die schulformübergreifende Transformation des Bildungswesens aller Mitgliedstaaten fokussieren verschiedene Handlungsebenen, wie die Entwicklung von Richtlinien für Lehrpersonen und Kompetenzmodellen oder Kooperations- sowie Reflexionsmöglichkeiten durch entsprechende Tools (z. B. European Digital Education Hub). Digitale Kompetenzen werden verstanden als »Kenntnisse, Fertigkeiten und Einstellungen [, um] in einer zunehmend durch digitale Technologien geprägten Welt zu leben, zu arbeiten, zu lernen und sich zu entwickeln« (Europäische Kommission, 2020, S. 2). Der Ausbau der digitalen Fähigkeiten und Kompetenzen für die digitale Transformation umfasst sowohl grundlegende digitale Kompetenzen (digital literacy; diese umfassen das Erkennen von Fehlinformationen, Informatik-Bildung und gute Kenntnisse und ein Verständnis von datenintensiven Technologien, wie z.B. künstlicher Intelligenz) als auch fortgeschrittene digitale Kompetenzen »which produce more digital specialists [...]« (European Commission, 2020, Absatz 4).

#### 4. KI, Bildung und Kulturtechniken

Die Veröffentlichung und breite Verfügbarkeit des KI-basierten Sprachmodells ChatGPT im Herbst 2022 (und seither einer Reihe ähnlicher Anwendungen) markierte eine Zeitenwende, auch und insbesondere für die Bildung (Artelt, 2023). Das erste allein durch eine KI generierte Buch erschien 2019 (Beta Writer, 2019) zum Thema Lithium-Ionen-Batterien. Zwei Jahre später kam, ebenfalls im Springer Verlag, ein Literaturüberblick mit dem Titel »Climate, Planetary and Evolutionary Sciences« heraus, das Guido Visconti (2021) in Zusammenarbeit mit einer KI geschrieben hat. Der Ausdruck »Zusammenarbeit« passt nicht wirklich, da er Absichten suggeriert, die Menschen haben, nicht jedoch Sprachmodelle. Gleichwohl weist die Arbeit im Wechselspiel mit und unter Nutzung des von der KI generierten Wissens und Formulierungen Elemente einer dialogischen Zusammenarbeit auf. So geschieht es seit der Veröffentlichung von ChatGPT in zahlreichen Anwendungen, bei denen im Wechselspiel mit generativen Sprachmodellen Texte erzeugt werden. Dies kann sehr unterschiedliche Formen annehmen, von einer fast 1:1-Übernahme des vorgeschlagenen Textes bis hin zu sehr differenzierten Interaktionsschritten und einer dialogartigen Auseinandersetzung mit der KI. Eliza Schüler (2023) berichtet in einem FAZ-Artikel von mehr als 250 auf Amazon im

Selbstverlag veröffentlichten Titeln, die KI als (Ko-)Autor\*in angeben. Unabhängig von den massiven Veränderungen im Selbstverständnis von und der Nachfrage nach den Produkten von Schriftsteller\*innen und Drehbuchautor\*innen ist der Prozess des Schreibens mit Unterstützung eines generativen Sprachmodells ein genuin anderer. Auch die Anforderungen an einen mündigen Umgang mit digitalen und durch Algorithmen gesteuerten Informationstechnologien und sozialen Medien verändern sich und sind deshalb hoch, weil psychische Verhaltensmuster der ggf. so wichtigen kritischen Reflexion zuwiderlaufen (Artelt, 2023). So hat der Psychologe Daniel Kahneman gezeigt, dass wir bei Entscheidungen, von denen wir im Umgang mit digitalen Medien täglich sehr viele treffen (etwa: Bleibe ich auf dieser Webseite? Klicke ich auf diesen Link? Erkenne ich die Wirkung von Filterblasen?) oft nicht rational vorgehen. Das Hinterfragen, Durchdenken und das Abwägen von Alternativen sind kognitiv aufwendige Vorgehensweisen, die im täglichen Handlungsvollzug eher selten zur Anwendung kommen. Oft und aus guten Gründen funktioniert die menschliche Informationsverarbeitung intuitiv und bedient sich dabei überlernter Schemata und Handlungsmuster. Werden diese auf digitale Interaktionen übertragen, besteht ggf. nicht einmal die Einsicht, dass es einer resilient-souveränen Haltung im Umgang mit digitalen Manipulationen, Desinformationen und Fake News bedarf.

Für Jugendliche, die Ergebnisse von schulischen Aufgabenstellungen quasi auf Knopfdruck geliefert bekommen, stellt sich die motivational bedeutsame Frage danach, was sie vor dem Hintergrund der Leistungen generativer Sprachmodelle überhaupt noch selber tun oder erlernen müssen. Sind klassische Kulturtechniken, die so eng mit Tätigkeiten der Rezeption von und der intellektuellen Auseinandersetzung mit Texten, Sprache und Argumenten verbunden sind, in Zeiten von automatisch generierten Texten überhaupt noch relevante Zieldimensionen von Bildung und wertgeschätzte Kulturtechniken? Warum Übersetzungen üben, wenn DeepL besser übersetzt? Warum Romane lesen und zusammenfassen, wenn ChatGPT diese Aufgabe passgenau übernimmt? Warum das Schreiben von Aufsätzen lernen, wenn die KI dies sowieso besser kann? Die Nutzung von KI übt eine ungeheure Faszination aus und geht gleichzeitig auch mit einer intellektuellen Kränkung einher. Die Angebote sind attraktiv und niederschwellig. Wissen, Antworten und Lösungen können auf Knopfdruck generiert werden, was jedoch mit einer Abwertung individueller Wissensbestände und Kompetenzen einhergeht, bei gleichzeitiger Steigerung der Anforderung an die eigene Urteilskraft und einer sinkenden Motivation zur Aneignung und Pflege von Kulturtechniken. Die Art und Weise, wie diese Techniken schulisch vermittelt werden, ist vor diesem Hintergrund neu zu betrachten. Paradoxerweise ist jedoch die Beherrschung der Kulturtechniken des Lesens, Schreibens und Argumentierens und das kritische Reflektieren dessen, was die KI anbietet, eine Grundvoraussetzung für den souveränen Umgang mit ihr. Sprache ist ein Machtinstrument und ein Bewusstsein für diese Technologie und

darüber, welcher Umgang mit dieser Technologie im Bildungsbereich sinnvoll ist, ist dringend geboten. Ein unreflektierter/naiver Optimismus ist hierbei genauso wenig das Gebot der Stunde wie das Verbannen der Technologie aus dem Schulgebäude und das Beharren auf der Nutzung ausschließlich tradierter Formen des Lesens und Schreibens.

Texte, Erklärungen, Programmcodes, die so gut wie zu jedem Thema und in Duktus und Tenor beliebig auswählbarer Art und Weise verfasst bzw. generiert werden können, sind Ausdruck einer disruptiven kulturellen Veränderung, die auch den Bildungsbereich betrifft. Die neuen Technologien verändern letztendlich auch das Verständnis dessen, was Wissen ist, wie es entsteht und was es ausmacht. Etablierte Prozesse und Kulturtechniken werden durch sie in Frage gestellt und neu geformt und gestaltet. Auch die Ideen davon, was Lernen und Bildung ausmacht, sind dadurch einem unaufhaltsamen Wandel unterworfen. Dabei betrifft die Omnipräsenz digitaler Medien und KI-basierter Anwendungen mittlerweile fast alle Lebensbereiche. Digitale Transformation ist so gesehen technische Revolution und medienkultureller Paradigmenwechsel zugleich. Kultur, Gesellschaft und Ökonomie verändern sich nachhaltig (Stalder, 2016). Gesellschaftliche Regelmäßigkeiten, Strukturen und Muster bilden dabei, so Nassehi (2019), das Material, aus dem die Digitalisierung u. a. ihr ökonomisches und politisches Kontroll- und Steuerungspotential schöpft. Es ist leicht vorstellbar, dass die Menge an Quellen, die z. B. mit bestimmten Täuschungsabsichten oder (politischer) Einflussnahme generiert werden, in großen Mengen verfügbar werden und nicht leicht von jenen Quellen unterschieden werden können, für die diese Merkmale nicht gelten. Besonders dann, wenn durch KI erzeugte Texte auf bestimmte Personen optimiert werden, wenn also KI Texte generiert, die dem Duktus und der Form nach auch von einer existierenden Person hätten stammen können oder gar von einem Avatar der Person mit passender Stimme und Intonation vorgebracht werden, werden die Grenzen zwischen Fiktion und Realität fließend und für die Rezipient\*innen schwer zu durchschauen. Darüber, was eine diese Entwicklungen berücksichtigende »agile« digitale Kompetenz ausmacht und wie sich eine kritisch-reflexive Haltung gegenüber den konsumierten Medien und digitalen Artefakten ausdrücken und manifestieren kann, wird mit neuen Entwicklungen immer neu justiert und durch Bildungsprozesse begleitet werden müssen. Es sollte jedoch klar sein, dass das kritische Hinterfragen in mehr und mehr Situationen von Bedeutung sein wird. Die Unsicherheit wird nicht weniger, sondern mehr, genauso wie die Anforderung an das Individuum, sich mit der Qualität und Belastbarkeit der rezipierten Information auseinanderzusetzen.

In seinem Buch »Home Digitales« ruft Rafael Capurro (2017) zu einer reflektierten Auseinandersetzung mit der (bzw. einer resilient-souveränen Haltung zur) digitalen Transformation auf und betont ihre Bedeutung für die individuelle und gesellschaftliche Entwicklung. Ähnlich wie Gerd Gigerenzer (2021), der die un-

reflektierte Haltung gegenüber der vielfältigen digitalen Einflussnahme und die Geschäftsmodelle von Tech-Unternehmen anprangert, geht es auch Capurro um digitale Aufklärung. Gleichwohl beschreibt er die Veränderung durch Digitalisierung als einen unaufhaltsamen Prozess, der alle Sphären umfasst und tradierte Vorstellungen und Konzepte wie Identität und Privatsphäre in Frage zu stellen vermag. Gefahren sehen Capurro wie Gigerenzer darin, Entscheidungen zum Einsatz von Technologien unüberlegt zu treffen. Transparenz, Respekt, Verantwortung und Reflexion sind in der Bewertung von und im Umgang mit kommerziell agierenden Unternehmen und auch Machthaber\*innen, die die Einflussnahme gezielt betreiben, zentrale Kategorien. Mündige Teilhabe ist die Voraussetzung, um der Volatilität, Unsicherheit, Komplexität und Ambiguität (sog. »VUKA Welt«) der digitalen Welt aktiv entgegenzuwirken.

## 5. ICT und Digitale Kompetenzen im Nationalen Bildungspanel

Eine zentrale bildungsbezogene Facette der Digitalisierung bezieht sich auf die Ausbildung der Fähigkeit, sich mit medial vermittelten Inhalten auseinanderzusetzen. Menschliche Qualitätskontrolle und die Fähigkeit, die Informationen, die auf Webseiten dargeboten werden, durch besondere Techniken zu prüfen, gewinnen vermehrt an Bedeutung, laufen sie doch Tendenzen zuwider, Urteilsprozesse auf Basis von einfachen Heuristiken zu vollziehen. Warum soll ich einer gut aufgemachten und seriös wirkenden Webseite nicht glauben? Wieso den umständlichen Weg gehen, die Webseite selber zu recherchieren und die dort enthaltenen Informationen über andere Quellen zu validieren (vgl. Gigerenzer, 2021)? Viel einfacher und attraktiver ist es doch, sich in den über Algorithmen gesteuerten Vorschlägen für Webseiten, Filme oder Posts weiter zu bewegen und damit eher mehr von Gleichem zu hören. Digitale Kompetenzen gehen auch über das hinaus, was klassischerweise als ICT-Kompetenz (Information and Communication Technology) gemessen wird. Die vom Educational Testing Service (ETS) zusammen mit dem International ICT Literacy Panel erarbeitete Rahmenkonzeption zu ICT-Literacy (ETS, 2002) stand auch Pate bei der Konzeption von ICT-Literacy im Nationalen Bildungspanel (Senkbeil et al., 2013; zu NEPS: Artelt & Sixt, 2023) und auch bei der International Computer and Information Literacy Study (ICILS) der IEA, (für Deutschland: Eickelmann et al., 2019). Grundlegend sind dabei die Prozesskomponenten, die technologische und kognitive Aspekte des Konstruktes ICT Literacy vereinen, hier mit Beispielen aus dem NEPS-Framework zu ICT ergänzt (Senkbeil et al., 2013): *Access*: Kenntnis grundlegender Operationen zum Abrufen von Informationen (z. B. Eingabe eines Suchbegriffs in einen Internetbrowser, Öffnen und Speichern eines Dokuments); *Create*: Fähigkeit, Dokumente und Dateien zu erstellen und zu bearbeiten (z. B. Einrichten von Tabellen, Erstellen von Formeln); *Manage*: Fähigkeit, Informationen innerhalb eines Programms zu finden (z. B. Abruf von

Informationen aus Tabellen, Verarbeitung der von einer Suchmaschine zurückgegebenen Treffer); *Evaluate*: die Fähigkeit, Informationen zu bewerten und sie als Grundlage für fundierte Entscheidungen zu nutzen (z. B. Bewertung der Glaubwürdigkeit der abgerufenen Informationen). Im NEPS werden diese auf unterschiedliche Softwareanwendungen (die benötigt werden, um Informationen zu lokalisieren, zu prozessieren, zu präsentieren und zu kommunizieren) bezogen und jeweils über Items abgedeckt (s. a. Senkbeil et al., 2013). Auch ICILS 2018 verwendet die Prozesskomponenten bei der Spezifikation der Teilbereiche. So bezieht sich der Teilbereich »Informationen sammeln und organisieren« auf die o. g. *access*, *evaluate* und *manage* Komponente und »Informationen erzeugen« auf die o. g. *create* Komponente. Zudem geht es im Teilbereich »Über Wissen zur Nutzung von Computern verfügen«, um die Kenntnis und das Verstehen (und die Nutzung) der Grundlagen und grundlegenden Konventionen der Computernutzung und im Teilbereich »Digitale Kommunikation« steht das Austauschen und die verantwortungsvolle und sichere Nutzung von Informationen im Vordergrund (vgl. Eickelmann et al., 2019).

Durch die fortschreitende Digitalisierung der Gesellschaft und die daraus resultierenden Herausforderungen und Chancen für Lehr- und Lernprozesse werden bei den Erhebungen im Nationalen Bildungspanel (NEPS) zukünftig ergänzend zu ICT auch digitale Kompetenzen erhoben. In der neuen Startkohorte, die im Jahr 2022 mit Fünftklässler\*innen gestartet ist, wird der eigens entwickelte neue Test zur Messung »Digitaler Kompetenzen« erstmals eingesetzt. Die Rahmenkonzeption (s. a. Tural & Schwaß, 2023) enthält die Hauptfacetten: *Kommunizieren und Interagieren*, *Datensouveränität*, *Informationen bewerten* und *Absichten und Strategien erkennen*. Beim *Kommunizieren und Interagieren* geht es um soziale Auswirkungen der Interaktion und Kommunikation im Internet, also Phänomene wie Hashtags und Hashtag-Debatten, Fake-Profile, Cybermobbing sowie Shitstorms und Candystorms. *Datensouveränität* als weitere Facette bezieht sich a) auf den Umgang mit eigenen Daten, hierbei behandeln wir Privatsphäre-Einstellungen, Fragen der Datenerfassung, Möglichkeiten der Reduzierung von Datenspuren und Fragen zum Teilen von persönlichen Informationen. Zudem b) auf den Umgang mit Daten und geistigem Eigentum anderer, wobei das Teilen und Schützen der Daten anderer Personen und das Wissen über Urheberrechte und Persönlichkeitsrechte abgebildet werden. Die Teilfacette *Informationen bewerten*, die einen bereits in der ICT-Konzeption enthaltenen Aspekt weiter ausbaut, umfasst die Themen der Quellenbewertung und der Glaubwürdigkeit von Online-Quellen sowie Informationsfilter, die als Algorithmen für Werbung, Suchmaschinen und andere Websites fungieren. In der Teilfacette *Absichten und Strategien erkennen* geht es schließlich um das Erkennen kommerzieller Absichten (Werbung unterschiedlicher Subtilität auf Websites und in sozialen Medien) auf der einen, und um Meinungsbildung in sozialen Medien, also das Wissen über Phänomene der Manipulation von und Einflussnahme auf Meinungen, der Einschränkung der Meinungs-

vielfalt über Filterblasen und Echokammern sowie Fake News und Bots, auf der anderen Seite. Die Konzeption des Tests, als Ergänzung zum ICT-Test vorgenommen, nimmt die ethisch-reflexive Facette digitale Souveränität (Blossfeld et al., 2018) stärker in den Blick. Inwieweit es sich hierbei tatsächlich um komplementäre Facetten einer zugrundeliegenden digitalen Souveränität handelt, die sich im Laufe der Schulzeit differenziell, aber in Abhängigkeit voneinander entwickeln und unterschiedliche Prädiktoren und Korrelate aufweisen, werden die Daten des NEPS (Startkohorte 8) in den nächsten Jahren zeigen. Dadurch, dass auch die Kontextfragebögen des NEPS vermehrt Informationen zu digitalen Medien und Unterrichtspraktiken und auch Haltungen und Praktiken im Elternhaus und in der Peergroup aufweisen, können Entstehungsbedingungen und Korrelate beleuchtet werden.

## 6. Fazit

Die Phänomenologie des Lernens verändert sich rasant. Schon jetzt nutzen Schüler\*innen und Studierende mehr oder weniger eigenständig – aber auch sozial selektiv – digitale Medien und Künstliche Intelligenz zum Lernen, zur Erledigung von Aufträgen, zum Verfassen von Arbeiten und für zahlreiche andere Anwendungen. Bildung in Zeiten von Digitalität bedeutet, dass die Möglichkeiten der Nutzung von digitalen Medien und KI für fachliche und überfachliche Zwecke, die Thematisierung der Qualität und Nutzungsmöglichkeiten ihrer Produkte, genauso wie die Reflexion inhärenter *biases* und »Nebeneffekte« aktiv im Unterricht aufgegriffen und in den schulischen Bildungsauftrag und das Curriculum aufgenommen werden. Dies bezieht sich auch auf die Entwicklung einer digitalitätsbezogenen reflexiven Haltung mit Blick auf die kulturell-gesellschaftlichen Implikationen der digitalen Transformation. Bildungsziele – genauso wie der Fokus, die Inhalte und die Bewertungsschemata von Hausaufgaben und Prüfungen – müssen sich als Folge digitaler Transformation verändern und um Kompetenzen für eine kritisch-reflektierte Nutzung digitaler Angebote erweitert werden (Leibniz-Forschungsnetzwerk Bildungspotentiale, 2020).

Eine mündige Auseinandersetzung mit KI-basierten Produkten und den gefilterten Angeboten der digitalen Welt, die darauf bezogene Argumentation, Beurteilung und das Hinterfragen der Lösungen sind wichtige Kompetenzen, die moderne Bildung ausmachen werden. Das Bildungssystem wird keine andere Wahl haben, als sich neu mit Inhalten und Formen des Lernens und Lehrens auseinanderzusetzen. Ob und in welchem Maße dies fachbezogen und aufbauend auf fachbezogenen Epistemologien, oder aber (auch) quer zu den Fächern geschehen sollte, ist nicht zuletzt eine empirische Frage. In dem Maße, in dem die digitale Transformation auch fachliche Gegenstände und damit auch Arbeitsweisen und ggf. auch Epistemologie der Fächer verändert, ist die reflexive Perspektive auf jeden Fall auch fachlich zu integrieren.

Bildung in einer durch Digitalität und Künstliche Intelligenz geprägten Welt muss dabei kontinuierlich reflektiert und weiterentwickelt werden. Das gilt nicht nur für den Unterricht und die pädagogische Arbeit in Schulen, sondern auch für die Aus-, Fort- und Weiterbildung von Lehrkräften. Stakeholder und Akteur\*innen auf unterschiedlichen Ebenen sollten dabei auch Experimentierräume schaffen, die es erlauben, den Anforderungen an eine sich im Wandel befindliche Welt und eine Kultur der Digitalität (Stalder, 2016) dynamisch(er) und direkter zu begegnen. Diese Experimentierräume können und sollten Forschungsergebnisse aufgreifen und durch Forschung begleitet werden. Reziprozität und Austausch zwischen Praxis und Forschung ist dabei auch bei den aus der Forschung oder der Administration stammenden Implementierungsvorhaben zentral, um vielfältig Optionen der Integration von digitalen Medien und insbesondere künstlicher Intelligenz in institutionelle Formen des Lehrens und Lernens zu sammeln, zu evaluieren und in die Aus- und Weiterbildung von Lehrkräften einfließen zu lassen. Aber auch die Rahmenbedingungen des Lehrens und Lernens bedürfen der Anpassung, genauso wie Bildungsadministration und -politik Orientierungsrahmen und Optionen eröffnen müssen, unter denen der Einsatz möglich und wünschenswert ist. Die Vermittlung digitaler Souveränität kann dabei nicht früh genug beginnen.

## Literatur

- Artelt, C. (2023). Digitale Souveränität – Ein Bildungsauftrag! *Akademie Aktuell* 79(1), 41.
- Artelt, C., & Sixt, M. (2023). The National Educational Panel Study (NEPS) – framework, design, and research potential. *Zeitschrift für Erziehungswissenschaft*, 26, 277–298. <https://doi.org/10.1007/s11618-023-01156-w>
- Autor:innengruppe Bildungsberichterstattung (2020). *Bildung in Deutschland 2020. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung in einer digitalisierten Welt*. wvb media. <https://doi.org/10.3278/6001820gw>
- Beblavý, M., Baiocco, S., Kilhoffer, Z., Akgüç, M., & Jacquot, M. (2019). *Index of readiness for digital lifelong learning: Changing how europeans upgrade their skills. CEPS Final Report*. Centre for European Policy Studies in partnership with Grow with Google. <https://www.ceps.eu/ceps-publications/index-of-readinessfor-digital-lifelong-learning/>
- Beta Writer. (2019). Lithium-Ion batteries. A machine-generated summary of current research. Springer Cham. <https://doi.org/10.1007/978-3-030-16800-1>
- Blossfeld, H. P., Bos, W., Daniel, H. D., Hannover, B., Köller, O., Lenzen, D., & Wößmann, L. (2018). *Digitale Souveränität und Bildung: Gutachten des Aktionsrats Bildung*. Waxmann.
- Bos, W., Eickelmann, B., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M. Schulz-Zander, R., & Wendt, H. (2014). *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich*. Waxmann.
- Brandhofer, G., Baumgartner, P., Ebner, M., Köberer, N., Trültzsch-Wijnen, C., & Wiesner, C. (2019). Bildung im Zeitalter der Digitalisierung. In S. Breit, F. Eder, K. Krainer, C.

- Schreiner, A. Seel, & C. Spiel (Hrsg.), *Nationaler Bildungsbericht Österreich 2018, Band 2: Fokussierte Analysen und Zukunftsperspektiven für das Bildungswesen* (S. 307–362). Leykam.
- Capurro, R. (2017). *Homo Digitalis*. Springer VS.
- Cress, U., Diethelm, I., Eickelmann, B., Köller, O., Nickolaus, R., Pant, H. A., & Reiss, K. (2018). *Schule in der digitalen Transformation – Perspektiven der Bildungswissenschaften*. acatech DISKUSSION.
- Diethelm, I. (2018). *Digitalisierung in Schule, Ausbildung und Hochschule – Strukturierungshilfen, Bildungsziele und Handlungsempfehlungen für das Feld »Digitale Bildung«*. Stellungnahme Öffentliches Fachgespräch »Digitalisierung in Schule, Ausbildung und Hochschule«. Ausschussdrucksache 19(18)37g. Ausschuss für Bildung, Forschung und Technikfolgenabschätzung. [https://www.bundestag.de/resource/blob/573972/37590b970d6d530bcce7825efe39160e/Diethelm\\_Stellungnahme\\_37g-data.pdf](https://www.bundestag.de/resource/blob/573972/37590b970d6d530bcce7825efe39160e/Diethelm_Stellungnahme_37g-data.pdf)
- Eickelmann, B., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., & Vahrenhold, J. (Hrsg.). (2019). *ICILS 2018 #Deutschland – Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking*. Waxmann.
- Ernst, C. (2020). *Der Grundsatz digitaler Souveränität: Eine Untersuchung zur Zulässigkeit des Einbindens privater IT-Dienstleister in die Aufgabenwahrnehmung der öffentlichen Verwaltung*. Schriften zum Öffentlichen Recht, Band 1426. Duncker & Humblot. <https://doi.org/10.3790/978-3-428-55931-2>
- ETS [Educational Testing Service]. (2002). *Digital transformation. A framework for ICT literacy. A report of the International ICT Literacy Panel*. ETS.
- European Commission. (2020). *Digital Education Action Plan 2021–2027. Resetting education and training for the digital age*. [https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan\\_en](https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan_en)
- Europäische Kommission. (2020). *Aktionsplan für digitale Bildung 2021–2027. Resetting education and training for the digital age*. <https://education.ec.europa.eu/de/focus-topics/digital-education/action-plan>
- Frederking, V. (2022). *Digitale Textsouveränität. Funktional-anwendungsorientierte und personal-reflexive Bildungsherausforderungen in der digitalen Weltgesellschaft im 21. Jahrhundert: Eine Theorieskizze* [Version 3, Januar 2022]. <https://www.deutschdidaktik.phil.fau.de/files/2021/09/digitale-textsouveraenitaet.pdf>
- Friedrichsen, M., & Bisa, P.-J. (Hrsg.). (2016). *Digitale Souveränität. Vertrauen in der Netzwerkesellschaft*. Springer VS.
- Gigerenzer, G. (2021). *Klick – Wie wir in einer digitalen Welt die Kontrolle behalten und die richtigen Entscheidungen treffen*. C. Bertelsmann.
- Goldacker, G. (2017). *Digitale Souveränität*. Kompetenzzentrum Öffentliche IT. Fraunhofer-Institut für Offene Kommunikationssysteme. <https://www.oeffentliche-it.de/documents/10181/14412/Digitale+Souver%C3%A4nit%C3%A4t>
- Huber, S. G., Günther, P. S., Schneider, N., Helm, C., Schwander, M., Schneider, J., & Pruitt, J. (2020). *COVID-19 – aktuelle Herausforderungen in Schule und Bildung. Erste Befunde des Schul-Barometers in Deutschland, Österreich und der Schweiz*. Waxmann.

- Kompetenzverbund lernen:digital. (2023). *Wissenschaft und Praxis im Dialog*. <https://lernen.digital/>
- Koziol, K., & Vogel, N. (2020). *Bildung und Medienkompetenz: Wege zur digitalen Souveränität*. kopaed.
- Lachner, A., Scheiter, K., & Stürmer, K. (2020). Digitalisierung und Lernen mit digitalen Medien als Gegenstand der Lehrerinnen- und Lehrerbildung. In C. Cramer, J. König, M. Rothland, & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 67–75). Klinkhardt.
- Leibniz-Forschungsnetzwerk Bildungspotenziale. (2020). *Bildung in der digitalen Welt: Potenziale und Herausforderungen. Positionspapier des Leibniz-Forschungsnetzwerks Bildungspotenziale*. [https://www.leibniz-bildung.de/wp-content/uploads/2020/10/LERN-Positionspapier\\_Digitale-Bildung-1.pdf](https://www.leibniz-bildung.de/wp-content/uploads/2020/10/LERN-Positionspapier_Digitale-Bildung-1.pdf)
- Medienpädagogischer Forschungsverbund Südwest. (2017). *KIM-Studie 2016. Basisstudie zum Medienumgang 6- bis 13-Jähriger in Deutschland*. [https://www.mpfs.de/fileadmin/files/Studien/KIM/2016/KIM\\_2016\\_Web-PDF.pdf](https://www.mpfs.de/fileadmin/files/Studien/KIM/2016/KIM_2016_Web-PDF.pdf)
- Müller, J., Thumel, M., Potzel, K., & Kammerl, R. (2020). Digital sovereignty of adolescents. *MedienJournal*, 44(1), 30–40. <https://dx.doi.org/10.24989/medienjournal.v43i1.1926>.
- Mußmann, F., Hardwig, T., Riethmüller, M., & Klötzer, S. (2021). *Digitalisierung im Schulsystem 2021. Arbeitszeit, Arbeitsbedingungen, Rahmenbedingungen und Perspektiven von Lehrkräften in Deutschland; Ergebnisbericht*. Kooperationsstelle Hochschulen und Gewerkschaften der Georg-August-Universität Göttingen. <https://doi.org/10.3249/ugoe-publ-10>
- Nagel, M.-T., Zlatkin-Troitschanskaia, O., & Fischer, J. (2022). Validation of newly developed tasks for the assessment of generic Critical Online Reasoning (COR) of university students and graduates. *Frontiers in Education*, 7. <https://doi.org/10.3389/educ.2022.914857>
- Nassehi, A. (2019). *Muster. Theorie der digitalen Gesellschaft* (2. Aufl.). C.H. Beck.
- Nationale Akademie der Wissenschaften Leopoldina. (2020). *Coronavirus-Pandemie: Für ein krisenresistentes Bildungssystem. Ad-hoc-Stellungnahme der Nationalen Akademie der Wissenschaften Leopoldina*. Leopoldina. [https://www.leopoldina.org/uploads/tx\\_leopublication/2020\\_08\\_05\\_Leopoldina\\_Stellungnahme\\_Coronavirus\\_Bildung.pdf](https://www.leopoldina.org/uploads/tx_leopublication/2020_08_05_Leopoldina_Stellungnahme_Coronavirus_Bildung.pdf)
- Petko, D., Döbeli Honegger, B., & Prasse, D. (2018). Digitale Transformation in Bildung und Schule: Facetten, Entwicklungslinien und Herausforderungen für die Lehrerinnen- und Lehrerbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 36(2), 157–174.
- Peuker, E. (2020). *Verfassungswandel durch Digitalisierung*. Mohr Siebeck.
- Rohleder, B. (2019). *Digitale Souveränität: Positionsbestimmung und erste Handlungsempfehlungen für Deutschland und Europa*. BITKOM. <https://bitkom.org/si-tes/default/files/file/import/BITKOM-Position-Digitale-Souveraenitaet.pdf>.
- Schüler, E. (2023, 28. Februar). Kann ChatGPT den Autor ersetzen? Kommentar. *Frankfurter Allgemeine Zeitung*. <https://www.faz.net/aktuell/feuilleton/kuenstliche-intelligenz-chatgpt-jetzt-auch-buchautor-18707948.html>

- Senkbeil, M., Ihme, J. M., & Wittwer, J. (2013). The Test of Technological and Information Literacy (TILT) in the National Educational Panel Study: Development, empirical testing, and evidence for validity. *Journal for Educational Research Online*, 5, 139–161.
- Seufert, S., Guggemos, J., & Tarantini, J. (2018). Digitale Transformation in Schulen – Kompetenzanforderungen an Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung: Zeitschrift zu Theorie und Praxis der Aus- und Weiterbildung von Lehrerinnen und Lehrern*, 36(2), 175–193.
- Stalder, F. (2016). *Kultur der Digitalität*. Suhrkamp.
- Tural, S., & Schwaß, M. (2023, February 28–March 3). *Digital natives = digital experts? Designing a novel instrument to measure digital competence in NEPS-SC8 secondary school students* [Posterpräsentation]. 10. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Universität Duisburg-Essen, Deutschland.
- Visconti, G. (2021). Climate, planetary and evolutionary sciences. A machine-generated literature overview. Springer. <https://doi.org/10.1007/978-3-030-74713-8>
- Wolter, I., Nusser, L., Attig, M., & Fackler, S. (2020). *Corona-bedingte Schulschließungen – ... und nun funktioniert alles digital? Wie Eltern mit Kindern in der 8. Klasse die Zeit der Schulschließungen in Deutschland erlebt haben* (NEPS Corona & Bildung 1). Leibniz-Institut für Bildungsverläufe. <https://doi.org/10.5157/NEPS:Bericht:Corona:01:1.0>



*Tina Seidel, Michael Nickl, Daniel Sommerhoff, Stefan Ufer,  
Anika Radkowitzsch, Elisabeth Bauer, Sina Huber, Elias Codreanu,  
Ralf Schmidmaier & Birgit J. Neuhaus*

## **Förderung von Zukunftskompetenzen durch videobasierte Simulationen für die Lehrkräftebildung am Beispiel des Projekts Visit-Math**

### **1. Einleitung**

Die gegenwärtigen gesellschaftlichen Entwicklungen sind gekennzeichnet von Volatilität, Unsicherheit, Komplexität und Ambiguität (Barber, 1992). Um junge Heranwachsende auf diese Entwicklungen vorzubereiten, stellen viele Länder derzeit ihre Bildungssysteme um und verankern verstärkt Zukunftskompetenzen in ihren Lehrplänen und Curricula (OECD, 2019). Zu diesen Zukunftskompetenzen zählen unter anderem sogenannte Life und Career Skills, Learning and Innovation Skills, sowie Information, Media and Technology Skills (P21 Partnership for 21st Century Skills, 2019). Diese sollen in den Kernfächern und anhand ausgewählter gesellschaftlich relevanter Lernkontexte systematisch in allen Bildungsphasen vermittelt werden. Zu den Lern- und Innovationsfähigkeiten als zentralen Bereich von Zukunftskompetenzen zählen die sogenannten vier »K«: Kommunikation, Kollaboration, Kritisches Denken und Kreativität (Aktionsrat Bildung, 2022). Bei diesen Lern- und Innovationsfähigkeiten handelt es sich um Bereiche, die bereits in den meisten Lehrplänen als übergeordnete Bildungsziele genannt werden. Allerdings erfahren sie noch nicht den Umfang und die Tiefe in der Realisierung im Unterricht, wie es in Anbetracht der gesellschaftlichen Entwicklungen und der notwendigen Anpassung unserer Bildungssysteme notwendig wäre (Aktionsrat Bildung, 2022).

Betrachtet man den empirischen Forschungsstand dazu, wie wirksam Unterrichtsansätze zur Förderung von Zukunftskompetenzen im Bereich der Lern- und Innovationsfähigkeiten sind, ist die Befundlage in positiver Weise relativ eindeutig: Werden im Unterricht verstärkt Instruktionsansätze wie kooperatives Lernen, soziales und emotionales Lernen, forschendes Lernen, selbstreguliertes Lernen und Lernen mit digitalen Tools und Medien umgesetzt, erreichen diese Instruktionsansätze im Vergleich zu herkömmlichem Unterricht positive Lernergebnisse, und zwar in motivational-affektiver, metakognitiver und kognitiver Hinsicht (Aktionsrat Bildung, 2022; Knogler et al., 2022; Knogler et al., 2019). Es spricht also viel

dafür, unseren Unterricht tiefgreifend zu verändern und diese Instruktionsansätze zur täglichen Praxis werden zu lassen.

Für diese notwendigen Veränderungen spielen Lehrpersonen die entscheidende Rolle (Cochran-Smith & Zeichner, 2005): Sie sind es, die diese Instruktionsansätze kennen, verstehen und umsetzen können müssen, damit Schüler\*innen verstärkt Zukunftskompetenzen erwerben können. Die Lehrpersonen sind dabei in der Rolle als Lehrende und als Lernende. Als Lernende müssen sie daher selbst in ihrer Ausbildung diese Arten von Lernumgebungen erfahren, um selbst entsprechende Zukunftskompetenzen zu entwickeln. Die bisherige, durchaus heterogene, Praxis der Vorlesungs- und Seminargestaltung, beispielsweise in der universitären Lehrkräftebildung, dürfte weit entfernt sein von den Lernumgebungen, wie sie auf Basis des empirischen Forschungsstands in positiver Weise Lern- und Innovationsfähigkeiten unterstützen sollten (Bauer et al., 2012; Bauer & Möller, 2023; Kunter et al., 2017; Möller et al., 2023). Von daher ist es dringend erforderlich, die universitäre Lehrkräftebildung weiter zu reformieren und innovative Instruktionsansätze zu realisieren (Darling-Hammond, 2010; Grossman, 2008).

In der universitären Lehrkräftebildung besteht die Herausforderung, komplexe Ideen und Begriffe der Unterrichtsfächer, der Fachdidaktik und der Bildungswissenschaften sowohl evidenzbasiert als auch praxisorientiert zu vermitteln (Seidel et al., 2021). Zu einer Evidenzbasierung zählt die Berücksichtigung des empirischen Forschungsstands und die Vermittlung zentraler Erkenntnisse der Bildungswissenschaften an zukünftige Lehrer\*innen (Bauer & Prenzel, 2012). Dazu erfahren vor allem solche Befunde besondere Berücksichtigung, die einen hohen Anwendungsbezug für die Unterrichtspraxis aufweisen (Gröschner et al., 2015). Gleichzeitig orientiert sich eine moderne Lehrkräftebildung an den professionellen Anforderungen des Lehrberufs und nutzt dabei in der Vermittlung relevanter Inhalte konkrete Anwendungskontexte zu den Standards in der Lehrkräftebildung, beispielsweise in den Bereichen Erziehen, Diagnostizieren, Unterrichten und Innovieren (Kulturministerkonferenz, 2014; Terhart, 2000).

Bei der Fülle an möglichen Praxisbezügen besteht eine Herausforderung für die universitäre Lehrkräftebildung darin, Entscheidungen bezüglich der Auswahl der Praxisbeispiele zu treffen (Seidel et al., 2021). Eine Möglichkeit besteht darin, sich dabei auf sogenannte »Core Teaching Practices«, also Kernpraktiken im Lehrberuf, zu konzentrieren (Grossmann, 2018; McDonald et al., 2013). Für die unterschiedlichen Unterrichtsfächer und für die generischen pädagogischen Praktiken werden hierzu derzeit in verschiedenen Konsortien Festlegungen für solche Kernpraktiken getroffen. Beispiele für generische Praktiken stellen die Orchestrierung von klassenöffentlichen Diskussionen, die Modellierung von Kleingruppenarbeiten, oder die Bereitstellung instruktionaler Erklärungen dar. Fachspezifische Kernpraktiken können sich beispielsweise auf die Identifikation von Gelegenheiten für forschendes Lernen im naturwissenschaftlichen Unterricht, auf den Umgang mit der Vertrau-

enswürdigkeit von Quellen im Geschichtsunterricht oder die instruktionale Unterstützung bei Interpretationen literarischer Texte beziehen (Grossmann, 2018).

Außerdem sind Rahmenmodelle hilfreich, die Vermittlungsaspekte im professionellen Lernen von Lehrpersonen abzubilden versuchen. Das Rahmenmodell zum praxisorientierten Lernen von Lehrpersonen stellt einen solchen Ansatz dar (Grossman et al., 2009; Grossman & McDonald, 2008). Hier wird ein Dreischritt in der instruktionalen Aufarbeitung und Vermittlung komplexer Konzepte mit Praxisbezug vorgeschlagen. Zunächst wird ein komplexes Konzept in zentrale Einheiten strukturiert und damit eine Dekomposition erreicht. Für Konzepte wie beispielsweise der »kognitiven Aktivierung« würde dies bedeuten, diese in wesentliche Unterbegriffe und Komponenten des Konzepts zu zerlegen. Auf Basis dieser Dekomposition werden in einem zweiten Schritt dann sogenannte Praxisrepräsentationen bestimmt. Dies bedeutet, dass für die entsprechenden Unterbegriffe und Teilkomponenten besonders einschlägige Kernpraktiken ausgewählt werden. Auf Basis dieser Auswahl erfolgen dann in der instruktionalen Vermittlung dieser so strukturierten Konzepte Zusammenstellungen sogenannter Praxisapproximationen, die von stärker vereinfachten Approximationen mit hoher Salienz der zu vermittelnden Komponenten (z. B. Simulationen und mediale Tools) bis hin zu einer möglichst authentischen Annäherung an die reale Praxis (z. B. Unterrichtspraktika) reichen (Chernikova et al., 2024; Fischer et al., 2022).

Simulationen, wie sie in vielen anderen Professionen für die Ausbildung eingesetzt werden (beispielsweise in der Medizin, in den Ingenieurwissenschaften oder im Pilotentraining), haben bislang in der Lehrkräftebildung noch wenig Nutzung erfahren (Chernikova et al., 2020). Sie bieten allerdings eine sehr effektive Lernumgebung, in der dekomponierte komplexe Konzepte vermittelt und Praxisannäherungen mit einer Balance aus Authentizität und Komplexitätsreduktion realisiert werden können (Fischer et al., 2022; Heitzmann et al., 2019).

## **2. Simulationen für die Lehrkräftebildung am Beispiel des Projekts Visit-Math**

### **2.1 Zielstellungen im Projekt**

Im Projekt Visit-Math, das von der Deutschen Forschungsgemeinschaft als Teilprojekt der Forschungsgruppe Cosima (FOR 2385) von 2017 bis 2024 gefördert wurde, erfolgte die Realisierung einer Simulation für die universitäre Lehrkräftebildung mit dem Schwerpunkt auf zentralen und in der Regel schwer zu vermittelnden Konzepten in der Mathematikdidaktik. Ein Konzept, das hier eine besondere Relevanz und Herausforderung in der Vermittlung aufweist, sind Diagnosekompetenzen angehende Lehrpersonen (Artelt & Gräsel, 2009; Südkamp et al., 2012).

Mit der Entwicklung einer Simulation bestand das Ziel, Lehramtsstudierende mit Fach Mathematik beim Erwerb zentraler Diagnosekompetenzen zu unterstützen (Codreanu, 2021a). Wichtig für Diagnosekompetenzen (Artelt & Gräsel, 2009) war das korrekte Wahrnehmen und wissensbasierte Schlussfolgern von Schüler\*innenaktivitäten bei Beweisaufgaben in der frühen Sekundarstufe (7./8. Jahrgangsstufe) mit Bezug zu dem zugrunde liegenden mathematischen Basiswissen, mathematischen Methodenwissen bezüglich des Vorgehens bei Beweisaufgaben und den wichtigen Problemlösefähigkeiten zur Bearbeitung der Beweisaufgaben (Codreanu et al., 2022; Codreanu et al., 2021b).

Lehramtsstudierende sollten über die Simulation die Möglichkeit erhalten, anhand ausgewählter Praxisrepräsentationen wichtige Teilkomponenten in der Diagnose dieser Argumentations- und Beweiskompetenzen bei Schüler\*innen anzuwenden. In der Realisierung der Simulation wurden für die Lernaufgabe der Diagnose von Argumentations- und Beweiskompetenzen bei Lernenden zentrale mathematikdidaktische Konzepte dekomponiert, geeignete Praxisrepräsentationen definiert und diese in Form von Praxisannäherungen in der Simulation realisiert. Um in diesem Prozess eine valide Lernumgebung zu generieren, erfolgte die Entwicklung der Simulation unter der besonderen Zielstellung, eine möglichst optimierte Balance im Erleben von Authentizität in der Bearbeitung der Simulation und der erfolgreichen Bewältigung kognitiver Anforderungen bei den zu bearbeitenden Aufgabenstellungen in der Simulation zu gewährleisten (Codreanu et al., 2020).

Darüber hinaus wurden im Projekt Visit-Math adaptiv orientierte Optimierungen der Simulation realisiert, in dem a) die kognitiven und motivational-affektiven Lernvoraussetzungen und Lernprozesse während der Bearbeitung der Simulation genauer untersucht wurden, um eine Basis für eine differential abgestimmte Anpassung der Simulation für Nutzer\*innen mit unterschiedlichen Lernvoraussetzungen zu schaffen (Plass & Pawar, 2020). Darüber hinaus wurden b) verschiedene instruktionale Unterstützungsmaßnahmen empirisch untersucht, beispielsweise in Form von Prompts für eine erfolgreiche Bewältigung der Aufgabenstellungen in der Simulation (Belland et al., 2011).

## 2.2 Zentrale Ergebnisse zur Entwicklung und Validierung der Simulation

Für die Entwicklung und Validierung der Simulation Visit-Math waren zwei Fragen zentral (Codreanu, 2021a): Inwiefern repräsentieren geskriptete Videosequenzen mit diagnostischen Situationen, in denen Interaktionen einer Lehrkraft mit Schüler\*innen zum Thema mathematisches Argumentieren dargestellt sind, eine authentische Unterrichtspraxis? Inwiefern gibt eine Dekomposition der in der Simulation dargestellten Konzepte einen validen Einblick in die Messung diagnostischer Kompetenzen bei angehenden Lehrpersonen?

### 2.2.1 Konzeptualisierung und Realisierung der geskripteten Videosequenzen

Um in möglichst optimaler Weise eine Balance zwischen der Vereinfachung komplexer Konzepte wie Diagnosekompetenzen und der Realisierung einer authentischen Praxisrepräsentation zu erreichen, wurden für die Simulation Videosequenzen von Interaktionen einer Lehrkraft mit Schüler\*innen auf Basis eines gezielt entwickelten und vorgefertigten Skripts erstellt bzw. nachgedreht (im Englischen häufig als staged video bezeichnet) (Dieker et al., 2009; Piwowar et al., 2018; Seidel et al., 2022). Dazu interagierten vier schauspielernde Schüler\*innen der achten Jahrgangsstufe mit einer schauspielernden Lehrkraft. Jede bzw. jeder der vier simulierten Lernenden repräsentierte dabei ein typisches Profil von den in der mathematikdidaktischen Forschung häufig festgestellten Lernvoraussetzungen beim mathematischen Argumentieren und Beweisen bei geometrischen Beweisaufgaben (Ufer et al., 2008). Grundlage waren dabei unterschiedliche Voraussetzungen der Lernenden in den Bereichen a) mathematisches Basiswissen, b) methodisches Wissen zum Beweisen und Argumentieren und c) Problemlösestrategien.



<b>Mathematisches Basiswissen</b>							
Begriffsinhalt	★	★ ★	★ ★ ★ ★	★ ★ ★ ★			
Begriffsumfang	★	★ ★	★ ★ ★ ★	★ ★ ★ ★			
Begriffnetz	★	★ ★	⊙	★ ★ ★ ★			
<b>Methodenwissen zum Beweisen</b>							
Beweisschema	★	★	★ ★ ★	★ ★ ★ ★			
Beweisstruktur	★	★ ★	★ ★ ★	★ ★ ★			
Beweiskette	★	★	★ ★	★ ★ ★			
<b>Problemlösestrategien</b>							
Heuristische Strategien	★ ★ ★	★ ★ ★ ★	★ ★ ★ ★	★ ★ ★ ★			
Metakognitive Strategien	★	★ ★	★ ★ ★	★ ★ ★ ★			
<div style="display: flex; justify-content: space-between; align-items: center; padding: 5px;"> <span>⊙ nicht beobachtbar</span> <span>★ überwiegend nicht vorhanden</span> <span>★ ★ unzureichend vorhanden</span> <span>★ ★ ★ überwiegend vorhanden</span> <span>★ ★ ★ ★ vorhanden</span> </div>							

Abbildung 1: Vier simulierte Lernende mit unterschiedlichen Lernvoraussetzungen beim mathematischen Beweisen und Argumentieren als Grundlage für die Erstellung geskripteter Videosequenzen für die Simulation Visit-Math (Codreanu, 2021a, S. 38)

Für eine möglichst hohe Vergleichbarkeit zwischen den vier simulierten Lernenden arbeiteten diese in den Videosequenzen an der gleichen Beweisaufgabe in der Geometrie. Auf Basis dieser vergleichbaren Aufgabensituation sollten in der Simulation die Unterschiede in den Voraussetzungen der zu diagnostizierenden Schüler\*innen salient werden.

Weiterhin wurden für jeden simulierten Lernenden deren Aufgabenbearbeitungen in je insgesamt acht Videoclips aufgeteilt. Jeder Videoclip hatte eine Dauer von ca. 1 Minute und über den Zeitverlauf der acht Videos wurden die jeweiligen Argumentations- und Beweiskompetenzen der simulierten Lernenden zunehmend salient. Die Rolle der schauspielenden Lehrkraft blieb dabei über die Clips hinweg konstant: jeder Clip startete mit der Bitte an die jeweiligen Lernenden, den bisherigen Bearbeitungsverlauf der Aufgabe zu erklären.

Die Skriptrealisierung und Produktion der Videosequenzen erfolgte nach den bisher publizierten Empfehlungen zur Realisierung geskripteter Videosequenzen für die Lehrkräftebildung (Dieker et al., 2009; Piwowar et al., 2018; Seidel et al., 2022). Die final erstellten Videosequenzen wurden von vier externen Expert\*innen im Bereich Mathematikdidaktik geprüft und von drei weiteren Expert\*innen aus dem Forschungsteam entlang der dargestellten Kategorien (siehe Abbildung 1) einem Rating unterzogen. Zudem validierten drei weitere Studierende die Simulation in Bezug auf Authentizität und Verständlichkeit der Aufgabenstellungen. Die Expert\*innenratings entsprachen im Wesentlichen den im Produktionsprozess intendierten Ausprägungen für die einzelnen schauspielenden Lernenden und dienten als Grundlage für spätere Kodierungen und Messungen der in der Simulation erreichten Diagnoseakkuratheit (Übereinstimmung mit Expert\*innenratings für die vier zu diagnostizierenden Lernenden).

### **2.2.2 Integration der Videosequenzen in das Design der Simulationsumgebung**

Die Simulation Visit-Math besteht in der Grundstruktur aus vier Teilen (Co-dreanu, 2021a). Sie startet mit einer Einführung und Kontextualisierung: Nutzer\*innen der Simulation werden aufgefordert, die Rolle einer Assistenz für eine Lehrperson einzunehmen mit dem Auftrag, vier Lernende genauer zu beobachten und deren mathematische Argumentations- und Beweiskompetenzen zu diagnostizieren. In der Rahmenhandlung der Simulation möchte die Lehrperson diesen vier Lernenden im weiteren Verlauf des Unterrichts gezielte Unterstützungen (z. B. durch zusätzliche Aufgabenstellungen) anbieten. Im zweiten Teil der Simulation werden die Nutzer\*innen technisch und inhaltlich mit den Anforderungen vertraut gemacht. Sie erhalten auch zur Aktivierung ihres mathematikdidaktischen Wissens eine Kurzzusammenfassung zu den Themen mathematisches Basiswissen, Methodenwissen und Problemlösestrategien (vgl. Abbildung 2). Im dritten Teil der

Simulation erfolgt die Realisierung des diagnostischen Prozesses. Hier setzen die Nutzer\*innen die Aufgabe um, die vier Lernenden bezüglich ihrer Argumentations- und Beweiskompetenzen zu diagnostizieren. Der Diagnoseprozess erfolgt in mehreren Zyklen. Jeder Zyklus startet mit der Beobachtung eines ersten Videoclips, in dem die Lehrperson mit einem der vier Lernenden interagiert und sich nach dem Bearbeitungsstand erkundigt. Die verbalen und non-verbalen Erklärungen (z. B. Deuten auf Zeichnungen) des jeweiligen Lernenden dienen als Grundlage für den diagnostischen Prozess. Im ersten Zyklus beobachten die Nutzer\*innen die jeweils ersten Videoclips der vier Lernenden nacheinander. Während der Beobachtung werden die Nutzer\*innen aufgefordert, sich zu ihrem diagnostischen Prozess in einem Freitextfenster schriftliche Notizen zu machen, die sie für die finale Erstellung eines schriftlichen Diagnoseberichts für die Lehrperson nutzen können. Am Ende des ersten Zyklus entscheiden die Proband\*innen, für welche der vier Lernenden sie weitere Videoclips beobachten möchten. Je nach Entscheidung stehen ihnen im nächsten Zyklus die nächsten Videoclips der ausgewählten Lernenden zur Verfügung. Um den Zeitrahmen für die Bearbeitung zu strukturieren und auch eine authentische Begrenztheit der zeitlichen Ressourcen einer Lehrperson zu simulieren, werden die Nutzer\*innen bereits bei der Einführung instruiert, nur so viele Videos wie nötig anzusehen. Zudem werden sie informiert, dass ihnen ein Maximum von insgesamt 20 nutzbaren Videoclips zur Verfügung steht. Wie sie diese auf die Videoclips der vier Lernenden verteilen, bleibt ihnen offen gestellt. Sobald die Nutzer\*innen der Meinung sind, dass sie nun bereit sind für die finalen Diagnosen, gehen sie über in den vierten Teil der Simulation.

Im vierten Teil erfolgt abschließend die Erstellung eines schriftlichen Diagnoseberichts für jeden der vier Lernenden. Dabei können die Nutzer\*innen auf ihre während des Diagnoseprozesses erstellten Notizen zurückgreifen und diese als Textbausteine nutzen. Nach Erstellung der Diagnoseberichte werden die Nutzer\*innen gebeten, für verschiedene Subaspekte der drei Dimensionen Basiswissen, Methodenwissen und Problemlösestrategien auf einer vierstufigen Likert-Skala Ausprägungen für jeden der vier Lernenden einzuschätzen. Diese Einschätzungen werden abgeglichen mit den Expert\*innenratings (Treffer, kein Treffer) und ein Übereinstimmungsmaß als Diagnoseakkuratheit berechnet.

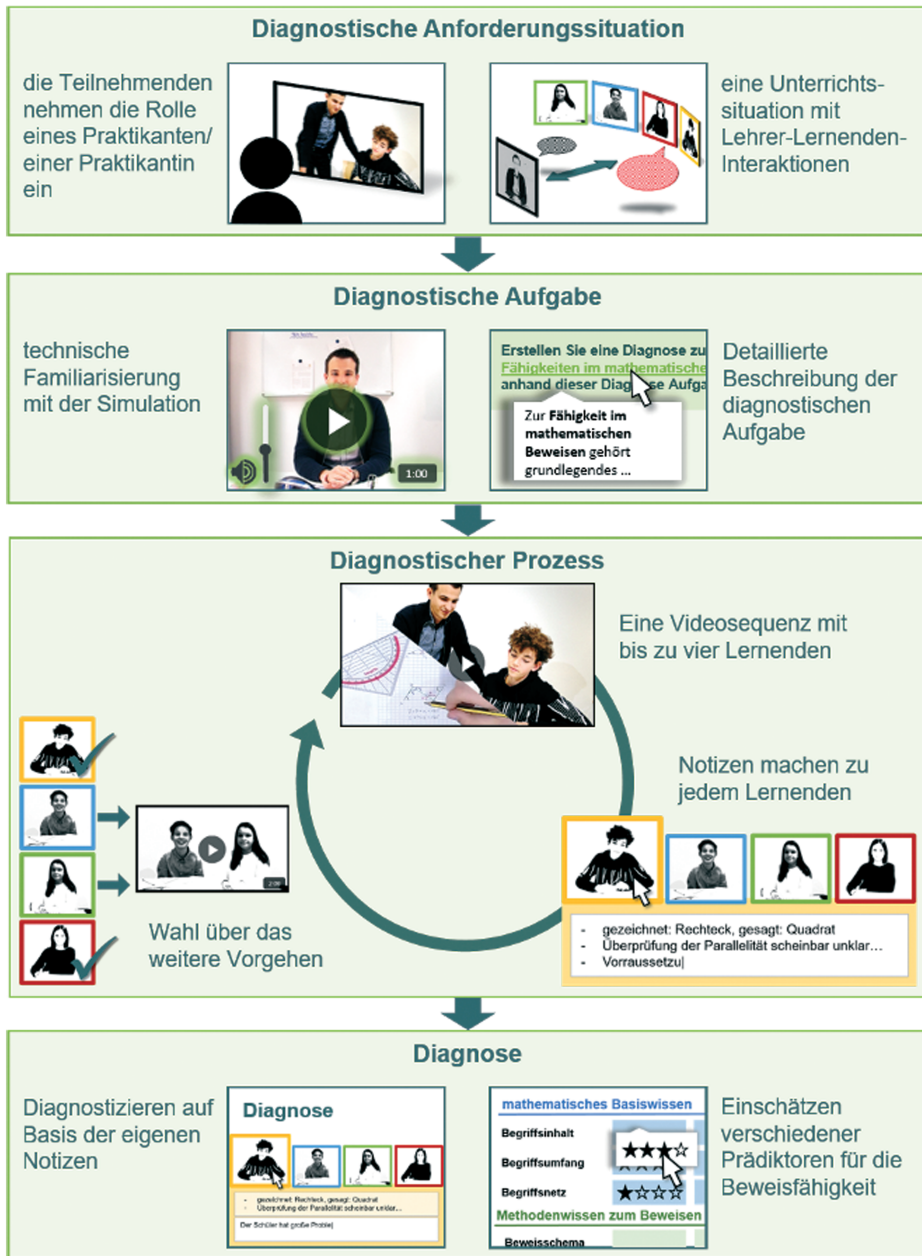


Abbildung 2: Struktur und Ablauf der Simulation Visit-Math (eigene Darstellung).

### 2.2.3 Empirische Untersuchungen zur Validierung der Simulation

Zur Validierung erfolgte die Durchführung mehrerer empirischer Studien. Als erstes erfolgte eine Pilotierungsstudie mit Befragung von vier ausgewählten Expert\*innen der Mathematikdidaktik und der videobasierten Lehrkräfteforschung zur Realisierung der geskripteten Videosequenzen und deren Implementation im Design der Simulation. Darüber hinaus wurden die anvisierten Nutzer\*innen in einer ersten Studie bestehend aus  $N=28$  Lehramtsstudierenden in der universitären Lehrkräftebildung, zu ihren Wahrnehmungen und Erfahrungen mit der Simulation befragt. Kriterien für die Realisierung einer gelungenen authentischen Praxisrepräsentation bildeten die wahrgenommene Authentizität der Simulation (Schubert et al., 2001), die erlebte Involviertheit (Vorderer et al., 2004) und der persönliche Wert (Wigfield & Eccles, 2000) der simulierten Aufgabenstellung als Indikator für Motivation (Codreanu et al., 2020). Diese wurden mittels Fragebogenskalen im Likert-Format erfasst, die von den Nutzer\*innen im Anschluss an die Simulationsbearbeitung eingeschätzt wurden. Als Kriterien für eine valide Realisierung der kognitiven Anforderungen erfolgten Auswertungen des Diagnoseprozesses und der Diagnoseakkuratheit für die vier simulierten Lernenden auf Basis der Vergleiche zwischen intendierten und realisierten Messungen (Codreanu et al., 2020; Codreanu et al., 2021b). Dazu wurde unter anderem geprüft, ob die Nutzer\*innen die übergeordneten Rangreihen der vier Lernenden bezüglich deren Argumentations- und Beweiskompetenzen korrekt einschätzen konnten, sowie eine Analyse der diagnostischen Fähigkeiten für jeden der vier Lernenden entlang der fokussierten mathematikdidaktischen Dimensionen durchgeführt (Basiswissen, Methodenwissen, Problemlösestrategien).

Die empirischen Ergebnisse verweisen insgesamt auf eine recht positive Einschätzung der Simulation durch die Nutzer\*innen (zusammenfassend aus Codreanu, 2021a; Codreanu et al., 2020; Codreanu, 2021b). Die Simulation sowie die Videosequenzen werden als authentische Praxisrepräsentationen wahrgenommen. Darüber hinaus berichten die Nutzer\*innen über eine hohe kognitive Involviertheit und eine positive Motivation durch Erkennen eines hohen persönlichen Werts der Nützlichkeit der Simulation für die Ausbildung als Lehrperson. Bezüglich der kognitiven Anforderungen verdeutlichen die empirischen Analysen ebenfalls eine gelungene Realisierung der Simulation. Lehramtsstudierende können nach anfänglichen Kursen in der Fachdidaktik im Wesentlichen die Rangreihen der vier simulierten Lernenden bezüglich deren mathematischer Argumentations- und Beweiskompetenzen korrekt bestimmen. Relativ zu den intendierten Schwierigkeiten bei den drei fokussierten Dimensionen und den simulierten Unterschieden zwischen den vier Lernenden zeigt sich eine hohe Trefferquote. Dimensionen und Ausprägungen bei den vier Lernenden, die aufgrund theoretischer Erwartun-

gen schwieriger zu diagnostizieren waren, zeigten sich auch empirisch schwieriger zu diagnostizieren. Leichter zu diagnostizierende Dimensionen und Unterschiede zwischen Lernenden gingen auch empirisch mit einer höheren Diagnoseakkuratesse und entsprechenden Diagnoseprozessen einher.

## **2.3 Zentrale Ergebnisse zur instruktionalen Optimierung der Simulation**

Die bisherigen Ausführungen zu Visit-Math beziehen sich auf das Design und die Struktur einer ersten Grundversion der Simulation (im Folgenden als Originalversion bezeichnet). In weiteren Studien wurden im Projekt Visit-Math Fragen der instruktionalen Optimierungen geklärt. Dazu erfolgten in einem ersten Schritt Analysen der Lernvoraussetzungen der Nutzer\*innen bezüglich ihres fachdidaktischen Wissens, ihrer Studienmotivation, ihres Selbstwirksamkeitserlebens während der Bearbeitung der Simulation sowie deren Lernaktivitäten bei der Bearbeitung der Aufgabenstellungen während der Simulation. Außerdem wurden in einem zweiten Schritt verschiedene instruktionale Unterstützungsmaßnahmen, wie sie sich in anderen Feldern der Instruktionspsychologie bewährt haben, auf die Simulation für die Lehrkräftebildung übertragen und deren Wirksamkeit weiter empirisch geprüft (Derry et al., 2014).

### **2.3.1 Lernvoraussetzungen und Lernaktivitäten während der Simulation als Grundlage für instruktionale Optimierungen**

Als Grundlage für weitere adaptiv an die Lernenden angepasste instruktionale Unterstützungen für die Simulation erfolgten im Projekt eine Reihe von Analysen zu den Lernvoraussetzungen und zur Rekonstruktion der Lernaktivitäten während der Simulationsbearbeitung.

Mit Blick auf die Lernvoraussetzungen wurden in einer Studie mit  $N=150$  Lehramtsstudierenden mit Unterrichtsfach Mathematik latente Lernprofilanalysen durchgeführt (Nickl et al., 2022). Als Kriterien wurden die professionelle Wissensbasis in Form des fachdidaktischen und fachlichen Wissens, das Interesse bezüglich der in der Simulation repräsentierten Gegenstandsbereiche (Diagnose von Lernenden, Geometrie, Mathematikdidaktik) (Rotgans & Schmidt, 2014), die Selbstwirksamkeit bezüglich der in der Simulation geforderten Diagnosekompetenzen (Kunter et al., 2002), sowie selbst eingeschätzte Selbstregulationsfähigkeiten (Kunter et al., 2016) herangezogen.

Die latente Profilanalyse zeigte, dass sich die Datenstruktur der Proband\*innen empirisch am besten durch drei Profile beschreiben ließ (Abbildung 3). Eine Gruppe von Studierenden zeichnete sich durch eine überdurchschnittlich hohe professionelle Wissensbasis (fachdidaktisches und fachliches Wissen) und durchschnittlich ausgeprägte motivational-affektive Lernvoraussetzungen (Interesse,

Selbstwirksamkeit, Selbstregulation) aus (40 % der Proband\*innen). Die zweite Gruppe hatte eine durchschnittlich ausgeprägte Wissensbasis, aber vergleichsweise hohe positive motivational-affektive Voraussetzungen (motiviertes Profil, 25 %). Eine dritte Gruppe zeichnete sich durch ein über alle Lernvoraussetzungen hinweg unterdurchschnittlich ausgeprägtes Profil aus (potentiell gefährdetes Profil, 35 %).

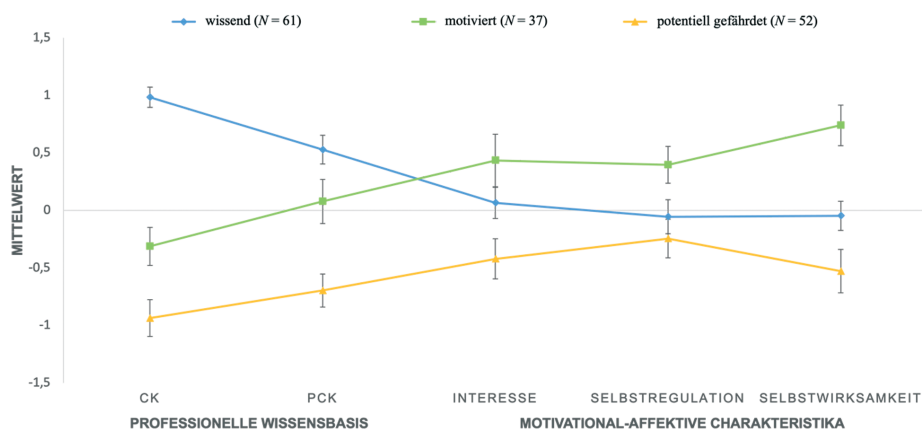


Abbildung 3: Lernvoraussetzungen für die Bearbeitung der Simulation Visit-Math (Nickl et al., 2022, S. 13).

Zur weiteren Validierung wurde untersucht, welche Lernaktivitäten sich für die Proband\*innengruppen mit unterschiedlichen Voraussetzungen in der Bearbeitung der Simulation nachzeichnen ließen (Nickl et al., 2022). Dabei zeigte sich, dass das potentiell gefährdete Profil im Vergleich zu den anderen beiden Profilen die geringste Erfolgserwartung aufwies und sich am wenigsten intensiv mit der Bearbeitung der Simulation beschäftigte (z. B. über die Anzahl der genutzten und beobachteten Videoclips). In der Tendenz erlebten Proband\*innen mit einem wissenden Profil am stärksten einen positiven persönlichen Wert (als Indikator für Motivation) in der Bearbeitung der Simulation und berichteten über eine relativ hohe lernbezogene kognitive Belastung. Das motivierte Profil zeichnete sich im Vergleich insbesondere durch besonders hohe Erfolgserwartungen während der Simulationsbearbeitung aus. Alle drei Gruppen erlebten die Simulation in ähnlich positiver Weise als authentische und involvierende Lernumgebung.

In einer zweiten Studie wurden die schriftlichen Notizen während der Simulationsbearbeitung zur Rekonstruktion der Lernaktivitäten einer vertiefenden Analyse unterzogen (Radkowitz et al., 2023). Kriterium war hierbei, ob die Proband\*innen wichtige mathematikdidaktische Aspekte erkannt haben (Noticing) und wie sie ihre Beobachtungen auf Basis ihres bislang erworbenen professionellen Wissens einordnen (Reasoning). Diese Kriterien sind wichtige Elemente des

Professional Vision Modells für Lehrpersonen (Seidel & Stürmer, 2014) und dessen Anwendung auf Simulationen für die Lehrkräftebildung (Kramer et al., 2021; Kramer et al., 2020). Beim Reasoning wurde differenziert zwischen der Fähigkeit, professionelles Wissen zur Beschreibung und Erklärung der beobachteten Situation heranzuziehen, sowie der Fähigkeit, auf Basis der Beobachtungen zu Vorschlägen für alternative Handlungsweisen zu gelangen (Entscheiden). Um dies zu messen, erfolgte die Segmentierung der schriftlichen Notizen während der Simulationsbearbeitung von  $N = 210$  Studierenden in Sinneinheiten (in der Regel einzelne Sätze oder Satzsegmente). Für jede Sinneinheit kodierten zwei unabhängige Personen auf Basis eines in einer vorherigen Studie entwickelten Kodiermanuals (Codreanu et al., 2021b) entlang der Kategorien des Beschreibens, Erklärens und Entscheidens. Die Kodierenden erzielten dabei zufriedenstellende Übereinstimmungen (Durchschnitt Cohen's Kappa = .72).

In einem nächsten Schritt wurden latente Profilanalysen (Bakk & Vermunt, 2016) für die kodierten Daten angewandt. Es zeigte sich, dass diese am besten durch vier Profile beschrieben werden konnten (Abbildung 4). Proband\*innen im ersten Profil (68 %) zeichneten sich durch unterdurchschnittliche Werte in allen drei Dimensionen (Beschreiben, Erklären, Entscheiden) aus. Die hohen Anteile von Lehramtsstudierenden in diesem Profil sind durchaus typisch und Ausdruck einer noch wenig entwickelten Professional Vision bei Lehranfänger\*innen (Farrell et al., 2022; Stürmer et al., 2016). Im zweiten Profil wurden verhältnismäßig ausgeprägt detaillierte Beschreibungen und Erklärungen der beobachteten Situationen genannt, mit vergleichsweise geringen Häufigkeiten in der Kategorie Entscheiden (8 %). Das dritte Profil weist dagegen eher durchschnittliche Werte beim Beschreiben und Erklären auf, dafür aber besonders ausgeprägte Angaben zu alternativen Handlungsoptionen (Entscheiden) (8 %). Die schriftlichen Notizen der Proband\*innen mit dem vierten Profil (16 %) beinhalteten Aspekte aller drei kodierten Dimensionen in verhältnismäßig ausgeglichener Weise (balanciertes Profil).

Die weitere Validierung der Profile verdeutlichte, dass Lehramtsstudierende mit dem zweiten Profil (vergleichsweise viele Beschreibungen und Erklärungen) besonders viele Videosequenzen genutzt und beobachtet haben und im Vergleich die höchste Diagnoseakkuratheit erreichen (Abbildung 5). Eine ähnlich hohe Diagnoseakkuratheit erreichten aber auch Studierende mit einem balancierten Profil. Dagegen hatten Proband\*innen mit dem ersten, eher unterentwickelten Professional Vision Profil die geringsten Häufigkeiten in der Anzahl beobachteter Videoclips und erreichten im Vergleich eine geringere Diagnoseakkuratheit. Bei den beiden erstgenannten Profilen mit verhältnismäßig hoher Diagnoseakkuratheit (balanciert, beschreibend/erklärend) ist interessant, dass die Studierenden im balancierten Profil das höchste fachliche Wissen aufwiesen. Im Verhältnis dazu war das fachliche Wissen der Studierenden mit dem zweiten Profil am niedrigsten ausgeprägt.

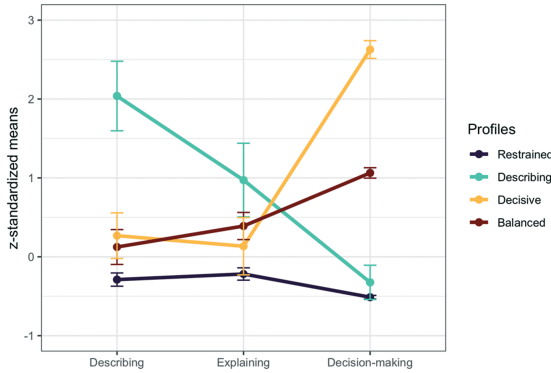


Abbildung 4: Profile von Lernaktivitäten während der Simulationsbearbeitung mit Bezug zu Professional Vision (Radkowsch et al., 2023, S. 10)

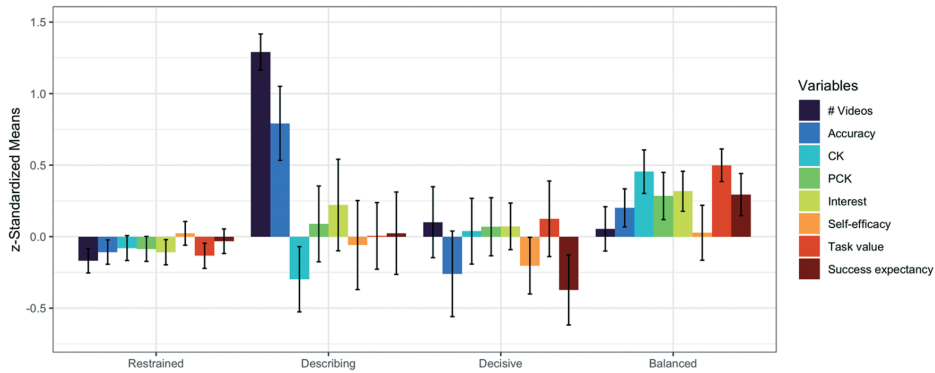


Abbildung 5: Profile für Lernaktivitäten während der Simulation mit Bezug zu weiteren Lernendenmerkmalen (Radkowsch et al., 2023, S. 11)

Tentativ könnte man vermuten und weiter untersuchen, welche Diagnosestrategien Studierende mit unterschiedlichen Lernvoraussetzungen in Simulationen einsetzen. Beispielsweise könnte es sein, dass Studierende mit dem zweiten Profil durch hohe Lernaktivitäten beim Beschreiben und Erklären und damit Sammeln von Informationen ihr im Vergleich niedriger ausgeprägtes fachliches Wissen kompensieren konnten, mit dem Ergebnis einer vergleichbaren Diagnoseakurtheit wie bei Studierenden mit verhältnismäßig hohem fachlichem Vorwissen. Keine wesentlichen Differenzen zwischen den Profilen gab es bei fachdidaktischem Wissen, Interesse, Selbstwirksamkeit und den Erfolgserwartungen.

### 2.3.2 Optimierung der Simulation durch adaptive instruktionale Unterstützung

Neben der Anwendung personenzentrierter Ansätze für Profilanalysen der Lernvoraussetzungen und Lernaktivitäten bei der Bearbeitung der Simulation wurden im Projekt Visit-Math verschiedene evidenzbasierte Unterstützungsmaßnahmen aus der Instruktionsforschung auf den Kontext von Simulationen in der Lehrkräftebildung übertragen und deren Wirksamkeit geprüft. Insbesondere wurde die Wirksamkeit verschiedener Prompts als zusätzliche Unterstützung während der Simulationsbearbeitung untersucht.

Mit Blick auf die unterschiedlichen Voraussetzungen von Lehramtsstudierenden bezüglich ihres fachdidaktischen Wissens gab es begründete Annahmen, dass Prompts die Anwendung und Nutzung dieses Wissens während der Simulationsbearbeitung positiv beeinflussen. Für Studierende mit wenig ausgeprägtem fachdidaktischen Vorwissen wurde angenommen, dass konzeptuelle Prompts, die gezielt einzelne Wissensdimensionen ansprechen (z. B. das für einen Videoclip zu beobachtende Basiswissen, Methodenwissen oder Problemlösestrategie), eine unterstützende Wirkung mit Blick auf die zu erreichende Diagnoseakkuratheit haben. Dagegen sollten verknüpfende Prompts insbesondere wirksam sein für Studierende mit fortgeschrittener fachdidaktischer Wissensbasis. Diese Prompts unterstützen den gezielten Abgleich von mindestens zwei der drei fokussierten Dimensionen in der Simulation Visit-Math (z. B. Beobachtung von dem vorhandenen Basiswissen in Kombination mit Methodenwissen oder Problemlösestrategien).

Die Wirksamkeit dieser beiden Prompt-Varianten wurde in einer experimentellen Studie mit  $N=59$  Lehramtsstudierenden mit Unterrichtsfach Mathematik untersucht (Sommerhoff et al., 2023). Die Proband\*innen wurden zufällig drei Bedingungen (Kontrollgruppe, Konzeptuelle Prompt Intervention, Verknüpfende Prompt Intervention) zugeteilt. Für diese Studie diagnostizierten die Proband\*innen zu drei Messzeitpunkten (Vortest, Intervention und Nachtest) je zwei Lernende aus der Originalversion mit den vier simulierten Lernenden. Für die beiden Lernenden konnten im Vor- und Nachtest bis zu acht Videoclips genutzt werden. Das obere Limit für die Anzahl maximal nutzbarer Videoclips wurde für diese Studie im Vor- und Nachtest auf 10 begrenzt (Originalversion: 20 für vier Lernende). Als abhängige Variablen dienten die Diagnoseakkuratheit während der Intervention und die erzielte Akkuratheit im Vor- und Nachtest. Die Ergebnisse dieser Interventionsstudie zeigten einen Vorteil der konzeptuellen Prompts zur gezielten Wissensaktivierung einzelner Dimensionen mit Auswirkungen auf die erzielte Diagnoseakkuratheit. Studierende in der Bedingung mit zusätzlich bereitgestellten konzeptuellen Prompts zur Aktivierung einzelner fachdidaktischer Wissensdimensionen erreichten eine höhere Akkuratheit sowohl während der In-

tervention als auch in der Entwicklung von Vor- auf Nachtest. Die verknüpfenden Prompts erwiesen sich im Vergleich zur anderen Prompt-Bedingung und der Kontrollgruppe nicht als zusätzlich unterstützend. Weitere explorative Analysen, insbesondere unter Berücksichtigung des fachdidaktischen Wissens, verdeutlichten, dass vor allem Studierende mit vergleichsweise hoch ausgeprägtem fachdidaktischem Wissen von den verknüpfenden Prompts profitierten. Die Ergebnisse deuten damit darauf hin, dass der Einsatz unterschiedlicher Prompts in Simulationen für die Lehrkräftebildung in Abhängigkeit des bereits erworbenen fachdidaktischen Wissens erfolgen sollte. Studierende mit geringem Vorwissen scheinen eher von gezielten konzeptuellen Aktivierungen einzelner Wissens Elemente zu profitieren, während Studierende mit fortgeschrittenerem Vorwissen Vorteile durch gezielte Unterstützungen in der Verknüpfung mehrerer Wissens Elemente erreichen.

Eine zweite Studie zur Prüfung der Wirksamkeit instruktorischer Maßnahmen beschäftigte sich insbesondere mit der Unterstützung kognitiver und motivationaler Voraussetzungen für die Simulationsbearbeitung (Nickl et al., 2023). Gerade in der Lehramtsausbildung verfügen Studierende häufig über eine gering ausgeprägte Motivation in Bezug auf die Nützlichkeit der zu erwerbenden professionellen Wissensbasis und wünschen sich weniger Vermittlung konzeptuellen Wissens und eine stärkere Praxisbetonung im Studium (Alles et al., 2019). Natürlich bieten Simulationen für die Lehrkräftebildung bereits eine Lernumgebung, die einen stärkeren Praxisbezug durch die Integration der Videosequenzen zu Interaktionen von Lehrkräften und Schüler\*innen beim mathematischen Beweisen und Argumentieren herstellt. Allerdings deuteten die Profilanalysen auch darauf hin, dass es eine Gruppe von Studierenden gibt, die sich auch im Bereich der wahrgenommenen persönlichen Nützlichkeit (Wertkomponente der Motivation) der Simulationsbearbeitung schwertut. Vor diesem Hintergrund wurde mit Bezug zu den Themen der Simulation eine sogenannte Utility-Value Intervention entwickelt und wichtige bisher als wirksam erwiesene Komponenten solcher Interventionen aus der Motivationsforschung für diesen Simulationskontext angewandt (Hulleman & Harackiewicz, 2021). Diese motivational unterstützende Instruktionsmaßnahme wurde kombiniert mit der erfolgreichen Integration konzeptueller Prompts. Die Fragestellung, ob und in welchen Kombinationen dieser beiden Interventionsmaßnahmen wirksame Ergebnisse mit Blick auf die erreichten Diagnosekompetenzen im Vergleich zu einer Kontrollgruppe erzielt werden, war die Grundlage für eine weitere experimentell angelegte Studie mit  $N=108$  Lehramtsstudierenden. Die Studierenden wurden randomisiert vier verschiedenen Bedingungen zugeordnet (Kontrollgruppe, Prompt Intervention, Utility-Value Intervention, beide Interventionen). Eine Woche vor der Intervention bearbeiteten die Studierenden die Simulation als Vortest in der gleichen Version wie in der ersten Intervention beschrieben (Original mit zwei zu diagnostizierenden Lernenden). Eine Woche später erfolgte die Intervention mit anschließender Durchführung des Nachtests.

Die Ergebnisse dieser Studie zeigten erneut positive Wirkungen der konzeptuellen Prompt-Bedingung auf die Entwicklung der Diagnoseakkuratheit. Die Utility-Value Intervention zeigte für die Entwicklung der Diagnoseakkuratheit sowohl als alleinstehende Bedingung als auch in der kombinierten Variante keine erhöhte Wirksamkeit. Vermittelt über die Förderung der Motivation konnten folglich keine Zuwächse in den gezeigten Diagnosekompetenzen erreicht werden. Besonders interessant war, dass in der kombinierten Variante beider Interventionsmaßnahmen die niedrigsten Lernentwicklungen von Vor- zu Nachtest zu verzeichnen waren. Zusätzlich durchgeführte explorative Analysen deuteten dann darauf hin, dass die Erfolgserwartung der Studierenden eine moderierende Rolle spielte. Studierende mit niedriger Erfolgserwartung profitierten am stärksten von den konzeptuellen Prompts, während Studierende mit hoher Erfolgserwartung von der Utility-Value Intervention profitierten.

Auch mit der zweiten Studie konnte die Wirksamkeit der konzeptuellen Prompts zur Wissensaktivierung nachgewiesen und damit Grundlagen für Optimierungen der Simulation geschaffen werden. Für die motivationale Förderung deuten die Ergebnisse sehr tentativ darauf hin, dass Utility-Value Interventionen für ausgewählte Studierendengruppen in Betracht gezogen werden können (z. B. solche mit hohen Erfolgserwartungen, die durch die Intervention einen zusätzlichen Boost erhalten).

### **3. Zusammenfassung und Ausblick**

Mit dem Projekt Visit-Math erfolgte eine evidenzbasierte Entwicklung, Validierung und Optimierung einer videobasierten Simulation für die Lehrkräftebildung. Insbesondere die universitäre Lehrkräftebildung sollte von dieser Art von Simulation profitieren. Es wurde die Möglichkeit genutzt, empirische Evidenz aus der Mathematikdidaktik zu typischen Problemlagen bei Schüler\*innen beim Beweisen und Argumentieren für die Gestaltung und Realisierung von simulierten Lernen anzuwenden. Mit der Simulation sollten authentische und dennoch komplexitätsreduzierte Lerngelegenheiten zu relevanten Kernpraktiken des Unterrichtens generiert werden (Fischer et al., 2022; Grossman et al., 2018; Seidel et al., 2021). Zudem folgten das Design und die Struktur der Simulation wichtigen evidenzbasierten Kriterien der Instruktionsforschung durch sorgfältige Pilotierung und Validierung der anvisierten Lernprozesse und Lernentwicklungen im Bereich der Diagnosekompetenzen angehender Lehrpersonen.

Zusammenfassend zeigen die empirischen Studien des sechsjährigen Projekts, dass Lehramtsstudierende vom praxisorientierten und evidenzbasierten Lernen in der Simulation profitierten. Die Simulation war unabhängig von den Lernvoraussetzungen der Studierenden für alle motivierend und kognitiv aktivierend. Damit scheint das Ziel, eine gute Relation zwischen authentischer Praxisapproximation

und Dekomposition komplexer Lerninhalte aus der Mathematikdidaktik erreicht worden zu sein. Die Integration von Videosequenzen in die Simulation hat sich dabei als durchaus erfolgreich erwiesen, vor allem auch mit Blick auf die noch relativ selten praktizierte Realisierung von geskripteten Videosequenzen mit schauspielenden Schüler\*innen sowie einer Lehrperson. Allerdings ist hier auch zu berücksichtigen, dass der Aufwand für diese Art von Videoproduktionen relativ hoch ist, insbesondere auch mit Blick auf die Einbindung realer Lernender aus dem Schulkontext. Die Schüler\*innen, die sich bereit erklärt haben, an der Videoproduktion mitzuwirken, haben dies freiwillig und in ihrer Freizeit getan. Für weitere Forschungsarbeiten ist es deshalb interessant, weiter zu erkunden, inwieweit auch über digitale Avatare entsprechend authentisch wirkende Praxisrepräsentationen produziert werden können. Insbesondere aktuelle KI-Entwicklungen im Bereich der Text-to-Speech oder Text-to-Video Integration bieten für die Gestaltung von Simulationen für die universitäre Bildung interessante Perspektiven. Diese müssen aber systematisch erforscht und mit Blick auf entwickelte Qualitätskriterien für Videoproduktionen weiter geprüft werden (Dieker et al., 2009; Piwowar et al., 2018; Seidel et al., 2022).

Neben der Tatsache, dass bereits die Grundversion der Simulation Visit-Math vielen Kriterien der Instruktionspsychologie in positiver Weise entsprach, hat das Projekt auch zur Weiterentwicklung dieses Forschungsfelds in Fragen der Optimierungen beigetragen. In der Anwendung von Prompts zur Wissensaktivierung und -verknüpfung, insbesondere für das besonders relevante fachdidaktische Wissen angehender Lehrpersonen, hat sich die Wirksamkeit wissensaktivierender Prompts auch für diesen Anwendungskontext erwiesen. Allerdings sind, wie auch in anderen Feldern der Instruktionspsychologie, die Lernvoraussetzungen der Nutzer\*innen einer Simulation zu berücksichtigen und die unterstützenden Prompts in Abhängigkeit der individuellen Voraussetzungen einzusetzen (Plass & Pawar, 2020). Weitere Forschungsstudien, auch in diesem Projekt, zielen derzeit darauf ab, den Lernprozess bei der Simulationsbearbeitung mikroadaptiv zu unterstützen (Nickl et al., 2024). Dies setzt fortlaufende Auswertungen der Lernaktivitäten im Simulationsprozess voraus. Auch hier erfolgen derzeit eine Reihe explorativer Prüfungen der Anwendung KI-basierter Verfahren, beispielsweise im Bereich des Machine Learning, Deep Learning und des Natural Language Processing.

In diesem Kapitel wurden die zentralen Zielstellungen und erzielten Ergebnisse des Projekts Visit-Math zusammenfassend dargestellt. Dabei ist zu berücksichtigen, dass das Projekt eingebettet war in die von der Deutschen Forschungsgemeinschaft geförderte Forschungsgruppe Cosima (FOR 2385). In dieser Forschungsgruppe sind eine Reihe von Simulationen sowohl für die Lehrkräftebildung als auch die Ausbildung von Mediziner\*innen entstanden, die eine besondere Grundlage für weiterführende Forschungsarbeiten bieten. Einen besonderen Vorteil bie-

tet das gemeinsam realisierte Rahmenmodell, in denen eine Vielzahl von Studien Fragen des Einflusses von Lernvoraussetzungen auf Lernprozesse und -ergebnisse, aber auch die Wirksamkeit ähnlich gestalteter instruktorischer Maßnahmen (z. B. im Bereich der Prompts und des Scaffolding) untersuchten (Heitzmann et al., 2019). Diese Ergebnisse werden derzeit zusammenfassend für die Forschungsgruppe integriert und meta-analytischen Untersuchungen zur Wirksamkeit unterzogen.

## Literatur

- Aktionsrat Bildung (Hrsg.). (2022). *Bildung und Resilienz*. Waxmann.
- Alles, M., Apel, J., Seidel, T., & Sturmer, K. (2019). How candidate teachers experience coherence in university education and teacher induction: The influence of perceived professional preparation at university and support during teacher induction. *Vocations and Learning*, 12(1), 87–112. <https://doi.org/10.1007/s12186-018-9211-5>
- Artelt, C., & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 23(34), 157–160. <https://doi.org/10.1024/1010-0652.23.34.157>
- Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(1), 20–31. <https://doi.org/10.1080/10705511.2014.955104>
- Barber, H. (1992). Developing strategic leadership: The US army war college experience. *Journal of Management Development*, 11(6), 4–12.
- Bauer, J., Diercks, U., Rösler, L., Möller, J., & Prenzel, M. (2012). Lehramtsstudium in Deutschland: Wie groß ist die strukturelle Vielfalt? *Unterrichtswissenschaft*, 40, 101–120.
- Bauer, J., & Möller, J. (2023). Lernumwelten Lehramtsstudierender. Zu Studienstrukturen und -inhalten in der PaLea-Stichprobe. In T. Kauper, A. Bernholt, J. Möller, & O. Köller (Hrsg.), *PaLea: Professionelle Kompetenzen und Studienstrukturen im Lehramtsstudium* (S. 89–108). Waxmann.
- Bauer, J., & Prenzel, M. (2012). European teacher training reforms. *Science*, 336(6089), 1642–1643. <https://doi.org/10.1126/science.1218387>
- Belland, B. R., Glazewski, K. D., & Richardson, J. C. (2011). Problem-based learning and argumentation: Testing a scaffolding framework to support middle school students' creation of evidence-based arguments. *Instructional Science*, 39, 667–694. <https://doi.org/10.1007/s11251-010-9148-z>
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-based learning in higher education: A meta-analysis. *Review of Educational Research*, 90(4), 499–541, Article 0034654320933544. <https://doi.org/10.3102/0034654320933544>
- Chernikova, O., Holzberger, D., Heitzmann, N., Stadler, M., Seidel, T., & Fischer, F. (2024). Where salience goes beyond authenticity: A meta-analysis on simulation-based learning in higher education. *Zeitschrift für Pädagogische Psychologie*, 38 (1–2), 15–25. <https://doi.org/10.1024/1010-0652/a000357>

- Cochran-Smith, M., & Zeichner, K. M. (Hrsg.). (2005). *Studying teacher education: The report of the AERA Panel on research and teacher education*. Lawrence Erlbaum Associates.
- Codreanu, E. (2021a). *Approximating practice* [Dissertation, Technische Universität München]. mediaTUM. <http://mediatum.ub.tum.de.eaccess.ub.tum.de/?id=1609556>
- Codreanu, E., Huber, S., Reinhold, S., Sommerhoff, D., Neuhaus, B. J., Schmidmaier, R., Ufer, S., & Seidel, T. (2022). Diagnosing mathematical argumentation skills: A video-based simulation for pre-service teachers. In F. Fischer & A. Opitz (Hrsg.), *Learning to diagnose with simulations: Examples from teacher education and medical education* (S. 33–47). Springer International Publishing. [https://doi.org/10.1007/978-3-030-89147-3\\_4](https://doi.org/10.1007/978-3-030-89147-3_4)
- Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education*, 95, 103146. <https://doi.org/10.1016/j.tate.2020.103146>
- Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2021b). Exploring the process of preservice teachers' diagnostic activities in a video-based simulation [Original Research]. *Frontiers in Education*, 6(133). <https://doi.org/10.3389/feduc.2021.626666>
- Darling-Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education*, 61(1–2), 35–47. <https://doi.org/10.1177/0022487109348024>
- Derry, S. J., Sherin, M. G., & Sherin, B. L. (2014). Multimedia learning with video. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of multi-media learning* (S. 785–812). Cambridge University Press.
- Dieker, L. A., Lane, H. B., Allsopp, D. H., O'Brien, C., Butler, T. W., Kyger, M., Lovin, L., & Fenty, N. S. (2009). Evaluating video models of evidence-based instructional practices to enhance teacher learning. *Teacher Education and Special Education*, 32(2), 180–196. <https://doi.org/10.1177/0888406409334202>
- Farrell, M., Martin, M., Renkl, A., Rieß, W., Könings, K. D., van Merriënboer, J. J. G., & Seidel, T. (2022). An epistemic network approach to teacher students' professional vision in tutoring video analysis [Original Research]. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.805422>
- Fischer, F., Bauer, E., Seidel, T., Schmidmaier, R., Radkowsitch, A., Neuhaus, B. J., Hofer, S. I., Sommerhoff, D., Ufer, S., Kuhn, J., Küchemann, S., Sailer, M., Koenen, J., Gartmeier, M., Berberat, P., Frenzel, A., Heitzmann, N., Holzberger, D., Pfeffer, J., Lewalter, D., Niklas, F., Schmidt-Hertha, B., Gollwitzer, M., Vorholzer, A., Chernikova, O., Schons, C., Pickal, A. J., Bannert, M., Michaeli, T., Stadler, M., & Fischer, M. R. (2022). Representational scaffolding in digital simulations – learning professional practices in higher education. *Information and Learning Sciences*, 123(11/12), 645–665. <https://doi.org/10.1108/ILS-06-2022-0076>
- Gröschner, A., Müller, K., Bauer, J., Seidel, T., Prenzel, M., Kauper, T., & Möller, J. (2015). Praxisphasen in der Lehrerbildung – Eine Strukturanalyse am Beispiel des gymnasialen Lehramtsstudiums in Deutschland. *Zeitschrift für Erziehungswissenschaft*, 18(4), 639–665. <https://doi.org/10.1007/s11618-015-0636-4>

- Grossman, P. (2008). Responding to our critics: From crisis to opportunity in research on teacher education. *Journal of Teacher Education*, 59(1), 10–23. <https://doi.org/10.1177/0022487107310748>
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9), 2055–2100.
- Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, 45(1), 184–205. <https://doi.org/10.3102/0002831207312906>
- Grossman, P., Schneider Kavanagh, S., & Pupik Dean, C. G. (2018). The turn towards practice-based teacher education: Introduction to the work of the Core Practice Consortium. In P. Grossmann (Hrsg.), *Teaching core practices in teacher education* (S. 1–14). Harvard Educational Press.
- Grossmann, P. (Hrsg.). (2018). *Teaching core practices in teacher education*. Harvard Educational Press.
- Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M., Ufer, S., Schmidmaier, R., Neuhaus, B., Siebeck, M., Stürmer, K., Obersteiner, A., Reiss, K., Girwidz, R., & Fischer, F. (2019). Facilitating diagnostic competences in simulations: A conceptual framework and a research agenda for medical and teacher education. *Frontline Learning Research*, 7(4), 1–24. <https://doi.org/10.14786/flr.v7i4.384>
- Hulleman, C. S., & Harackiewicz, J. (2021). The utility-value intervention. In G. M. Walton & A. J. Crum (Hrsg.), *Handbook of wise interventions: How social psychology can help people change* (S. 100–125). The Guilford Press.
- Knogler, M., Hetmanek, A., & Seidel, T. (2022). Determining an evidence base for particular fields of educational practice: A systematic review of meta-analyses on effective mathematics and science teaching [Systematic Review]. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.873995>
- Knogler, M., Seidel, T., Wiesbeck, A., & Mok, S. Y. (2019). Das Clearing House Unterricht. Ein Angebot zur adressatengerechten Bereitstellung wissenschaftlicher Evidenz für die Lehrer\*innenbildung. In B. Priebe, C. Mattiesson, & K. Sommer (Hrsg.), *Dialogische Verbindungslinien zwischen Wissenschaft und Schule. Theoretische Grundlagen, praxisbezogene Anwendungsaspekte und zielgruppenorientiertes Publizieren* (S. 36–49). Verlag Julius Klinkhardt.
- Kramer, M., Förtsch, C., Seidel, T., & Neuhaus, B. J. (2021). Comparing two constructs for describing and analyzing teachers' diagnostic processes. *Studies in Educational Evaluation*, 68, 100973. <https://doi.org/https://doi.org/10.1016/j.stueduc.2020.100973>
- Kramer, M., Förtsch, C., Stürmer, J., Förtsch, S., Seidel, T., & Neuhaus, B. J. (2020). Measuring biology teachers' professional vision: Development and validation of a video-based assessment tool [Article]. *Cogent Education*, 7(1), Article 1823155. <https://doi.org/10.1080/2331186X.2020.1823155>
- Kulturministerkonferenz. (2014). *Standards für die Lehrerbildung: Bildungswissenschaften*. KMK.

- Kunter, M., Baumert, J., Leutner, D., Terhart, E., Seidel, T., Dicke, T., Holzberger, D., Kunina-Habenicht, O., Linninger, C., Lohse-Bossenz, H., Schulze-Stocker, F., & Stürmer, K. (2016). *Dokumentation der Erhebungsinstrumente der Projektphasen des BilWiss-Forschungsprogramms von 2009 bis 2016: Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung (BilWiss); die Bedeutung des bildungswissenschaftlichen Hochschulwissens für den Berufseinstieg von Lehrkräften (BilWiss-Beruf ) [Scale manual of BilWiss research program from 2009 to 2016]*. Goethe-Universität.
- Kunter, M., Kunina-Habenicht, O., Baumert, J., Dicke, T., Holzberger, D., Lohse-Bossenz, H., Leutner, D., Schulze-Stocker, F., & Terhart, E. (2017). Bildungswissenschaftliches Wissen und professionelle Kompetenz in der Lehramtsausbildung. In C. Gräsel & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals: Interdisziplinäre Betrachtungen, Befunde und Perspektiven* (S. 37–54). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-07274-2\\_3](https://doi.org/10.1007/978-3-658-07274-2_3)
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente [PISA 2000: scale manual]. Materialien aus der Bildungsforschung: Nr. 72. Max-Planck-Inst. für Bildungsforschung*. <http://hdl.handle.net/hdl:11858/00-001M-0000-0023-9987-C>
- McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education*, 64(5), 378–386. <https://doi.org/10.1177/0022487113493807>
- Möller, J., Bauer, J., & Zimmermann, F. (2023). Das Lehramtsstudium. Angebot, Nutzung, Lernergebnisse. In T. Kauper, A. Bernholt, J. Möller, & O. Köller (Hrsg.), *PaLea: Professionelle Kompetenzen und Studienstrukturen im Lehramtsstudium* (S. 7–34). Waxmann.
- Nickl, M., Huber, S. A., Sommerhoff, D., Codreanu, E., Ufer, S., & Seidel, T. (2022). Video-based simulations in teacher education: The role of learner characteristics as capacities for positive learning experiences and high performance. *International Journal of Educational Technology in Higher Education*, 19(1), 45. <https://doi.org/10.1186/s41239-022-00351-9>
- Nickl, M., Sommerhoff, D., Böheim, R., Ufer, S., & Seidel, T. (2023). Fostering pre-service teachers' assessment skills in a video simulation. *Zeitschrift für Pädagogische Psychologie*, 38(1–2), 27–34. <https://doi.org/10.1024/1010-0652/a000362>
- Nickl, M., Sommerhoff, D., Radkowsch, A., Huber, S. A., Bauer, E., Ufer, S., Plass, J. L., & Seidel, T. (2024). Effects of real-time adaptivity of scaffolding: Supporting pre-service mathematics teachers' assessment skills in simulations. *Learning and Instruction*, 94, 101994. <https://doi.org/https://doi.org/10.1016/j.learninstruc.2024.101994>
- OECD. (2019). OECD Future of Education and Skills 2030. OECD Learning Compass 2030 – A series of concept notes.
- P21 Partnership for 21st Century Skills. (2019). *Framework for 21st century learning*. [http://static.battelleforkids.org/documents/p21/P21\\_Framework\\_Brief.pdf](http://static.battelleforkids.org/documents/p21/P21_Framework_Brief.pdf)
- Piowar, V., Barth, V. L., Ophardt, D., & Thiel, F. (2018). Evidence-based scripted videos on handling student misbehavior: The development and evaluation of video cases for

- teacher education. *Professional Development in Education*, 44(3), 369–384. <https://doi.org/10.1080/19415257.2017.1316299>
- Plass, J. L., & Pawar, S. (2020). Toward a taxonomy of adaptivity for learning. *Journal of Research on Technology in Education*, 52(3), 275–300. <https://doi.org/10.1080/15391523.2020.1719943>
- Radkowsch, A., Sommerhoff, D., Nickl, M., Codreanu, E., Ufer, S., & Seidel, T. (2023). Exploring the diagnostic process of pre-service teachers using a simulation – A latent profile approach. *Teaching and Teacher Education*, 130, 104172. <https://doi.org/10.1016/j.tate.2023.104172>
- Rotgans, J. I., & Schmidt, H. G. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction*, 32, 37–50. <https://doi.org/https://doi.org/10.1016/j.learninstruc.2014.01.002>
- Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, 10(3), 266–281. <https://doi.org/10.1162/105474601300343603>
- Seidel, T., Farrell, M., Martin, M., Rieß, W., & Renkl, A. (2022). Developing scripted video cases for teacher education: Creating evidence-based practice representations using mock ups [Original Research]. *Frontiers in Education*, 7. <https://doi.org/10.3389/educ.2022.965498>
- Seidel, T., Renkl, A., & Rieß, W. (2021). Basisdimensionen für Unterrichtsqualität im Fachkontext konkretisieren: Die Rolle von Unterrichtsartefakten und Bestimmung von Standardsituationen. *Unterrichtswissenschaft*, 49(2), 293–301. <https://doi.org/10.1007/s42010-021-00108-9>
- Seidel, T., & Stürmer, K. (2014). Modeling the structure of professional vision in pre-service teachers. *American Educational Research Journal*. <https://doi.org/10.3102/0002831214531321>
- Sommerhoff, D., Codreanu, E., Nickl, M., Ufer, S., & Seidel, T. (2023). Pre-service teachers' learning of diagnostic skills in a video-based simulation: Effects of conceptual vs. interconnecting prompts on judgment accuracy and the diagnostic process. *Learning and Instruction*, 83, 101689. <https://doi.org/https://doi.org/10.1016/j.learninstruc.2022.101689>
- Stürmer, K., Seidel, T., & Holzberger, D. (2016). Intra-individual differences in developing professional vision: Preservice teachers' changes in the course of an innovative teacher education program. *Instructional Science*, 44(3), 293–309. <https://doi.org/10.1007/s11251-016-9373-1>
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762. <https://doi.org/10.1037/a0027627>
- Terhart, E. (2000). *Perspektiven der Lehrerbildung in Deutschland. Abschlußbericht der von der Kultusministerkonferenz eingesetzten Kommission*. Beltz.
- Ufer, S., Heinze, A., & Reiss, K. (2008). Individual predictors of geometrical proof competence. *PME 32 and PME-NA XXX*, 1(4), 361–368.
- Vorderer, P., Wirth, W., Gouveia, F., Biocca, F., Saari, T., Jäncke, L., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Laarni, J., Ravaja, N., Sacau, A., Baumgartner,

- T., & Jäncke, P. (2004). MEC spatial presence questionnaire (MEC-SPQ): Short documentation and instructions for application. *Report to the European Community, Project Presence: MEC (IST-2001-37661)*.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. <https://doi.org/https://doi.org/10.1006/ceps.1999.1015>



*Franziska Baier-Mosch & Lukas Schulze-Vorberg*

## »Künstliche Intelligenz« im eigenen Unterrichtsfach vermitteln

»Will und kann ich das lernen?«

### 1. Theoretischer Hintergrund

Anwendungen der Künstlichen Intelligenz (KI) durchdringen zunehmend Beruf und Alltag (Diethelm, 2021). Aus diesem Grund erscheint es äußerst relevant, Schüler\*innen auf den zukünftigen Umgang mit KI angemessen vorzubereiten. Eine Integration von Inhalten zum Thema KI in Lehrpläne wird zunehmend gefordert (Jaschke et al., 2023). Lehrkräfte haben allerdings wenig Erfahrung mit der Vermittlung von KI-Inhalten allgemein und speziell in ihrem eigenen Fachunterricht (Ciu & Chai, 2020; Lindern & Romeike, 2019). Vor allem ältere Lehrkräfte fühlen sich im Umgang mit digitalen Medien weniger kompetent (Hämäläinen et al., 2021). Was das Geschlecht angeht, sind Frauen in informatischen Studienfächern nach wie vor unterrepräsentiert (Pinkard et al., 2017) und Männer/männliche Lehrkräfte haben positivere Einstellungen gegenüber digitalen Technologien (Cai et al., 2017; Marpa, 2021). Möchte man Lehrkräfte also darauf vorbereiten, Inhalte zum Thema KI in ihrem Unterricht an Schüler\*innen zu vermitteln, könnte die Berücksichtigung dieser soziodemographischen Merkmale bei der Gestaltung von Maßnahmen zur Professionalisierung eine Rolle spielen. Alter und Geschlecht von Lehrkräften könnten sich auf die Annahme und Nutzung solcher Lerngelegenheiten auswirken.

Neben soziodemographischen Merkmalen gehen Modelle zur Entwicklung der professionellen Kompetenz von Lehrkräften (z. B. Kunter et al., 2011) und Modelle zum *conceptual change* bei Lehrkräften (z. B. Gregoire, 2003) davon aus, dass kognitive (z. B. Vorwissen) und affektiv-motivationale (z. B. Selbstwirksamkeitserwartung) Voraussetzungen der Lehrkräfte eine entscheidende Rolle für die Verarbeitung und Nutzung von Lerngelegenheiten zur Professionalisierung spielen. Genauer gesagt legen solche Modelle nahe, dass Professionalisierungsmaßnahmen an die unterschiedlichen motivationalen und kognitiven Voraussetzungen von Lehrkräften angepasst werden sollten (Kunter et al., 2011).

Besonders relevant erscheint in diesem Zusammenhang das Conceptual-Change-Modell von Gregoire (2003), das als Ausgangspunkt die Implementation von Reformen durch Lehrkräfte hat. Das Modell geht davon aus, dass das

Vorhandensein beziehungsweise die Abwesenheit von kognitiven und affektiv-motivationalen Ressourcen von Lehrkräften beeinflusst, wie Lehrkräfte Lerngelegenheiten zu innovativen Unterrichtsthemen wie KI wahrnehmen und entsprechend verarbeiten und nutzen (Gregoire, 2003). Persönliche Voraussetzungen der Lehrkräfte wie Selbstwirksamkeitserwartungen und Vorwissen in Hinblick auf ein neues Thema spielen dabei eine besondere Rolle (Gregoire, 2003). Fehlen diese persönlichen Ressourcen und mangelt es an Zeit und Unterstützung (z. B. durch Kolleg\*innen; Gregoire, 2003) kann dies zu emotionaler Erschöpfung führen, die wiederum das Gefühl eines Mangels an emotionalen Ressourcen darstellt (Klusmann et al., 2008).

Insgesamt beeinflussen die beschriebenen Ressourcen, inwiefern Lehrkräfte eine Lerngelegenheit als Herausforderung und/oder Bedrohung wahrnehmen (Gregoire, 2003). Dabei sind vor allem eine hohe Selbstwirksamkeitserwartung und das Vorhandensein von Vorwissen mit der Wahrnehmung von Herausforderung verbunden. Herausforderung wird dabei als eher angenehm erlebt und führt zu einer Annäherung an die Lerngelegenheit und einer tiefen Auseinandersetzung mit ihren Inhalten (systematische Verarbeitung; Gregoire, 2003). Das Gefühl der Bedrohung ist hingegen mit Ängsten assoziiert und führt zu einer Vermeidung der Lerngelegenheit und einer wenig systematischen Verarbeitung ihrer Inhalte (Broughton et al., 2013; Gregoire, 2003). Systematische Verarbeitung meint dabei die tiefe Verarbeitung der Inhalte, die mit hoher investierter kognitiver Anstrengung einhergeht (Gregoire, 2003). Ängste und mangelnde Ressourcen (z. B. mangelndes Vorwissen, emotionale Erschöpfung) führen aber auch dazu, dass nicht nur weniger *kognitive Anstrengung* investiert wird (Feldon et al., 2019), sondern auch dazu, dass generell weniger kognitive Ressourcen zur Verarbeitung der Inhalte zur Verfügung stehen (Feuerhahn et al., 2013; Scheiter et al., 2009). Die Lerngelegenheit stellt somit eine höhere *kognitive Belastung* für diese Lehrkräfte dar und die kognitive Verarbeitung der Inhalte fällt schwerer (Krell, 2017; Orru & Longo, 2019).

Einzelne empirische Studien zu den angenommenen Zusammenhängen des Conceptual-Change-Modells von Gregoire (2003) bei Lehrkräften wurden zum Beispiel zu Reformthemen wie Forschendes Lernen (Ebert & Crippen, 2010) oder die Implementation eines kompetenzbasierten Mathematik-Curriculums (Boesen et al., 2014) durchgeführt. Insgesamt können die angenommenen Zusammenhänge des Modells allerdings nur teilweise bestätigt werden oder wurden nicht umfassend untersucht (Peltier, 2022; Wu & Yeziarski, 2023). In der vorliegenden Arbeit sollen die beschriebenen Annahmen auf das Thema KI übertragen und für das Lernen von Lehrkräften in diesem Bereich geprüft werden. Ziel ist es, die relative Bedeutsamkeit der in der Literatur postulierten soziodemographischen und kognitiv-affektiven Merkmale von Lehrkräften zur Vorhersage der kognitiven Anstrengung bei und kognitiven Belastung durch Professionalisierungsmaßnahmen

zu prüfen. Daraus können wichtige Implikationen zur differenzierten Gestaltung solcher Lerngelegenheiten für Lehrkräfte abgeleitet werden.

## 2. Fragestellung

Die vorliegende Studie untersucht die relative Vorhersagekraft von Alter und Geschlecht, der Selbstwirksamkeitserwartung, der emotionalen Erschöpfung, des Erlebens als Herausforderung und des Erlebens als Bedrohung einer Lerngelegenheit für die investierte *kognitive Anstrengung* und die erlebte *kognitive Belastung* beim Lernen über KI-Inhalte von Lehrkräften.

## 3. Methode

### 3.1 Stichprobe

Bei der Stichprobe handelt es sich um eine Gelegenheitsstichprobe von 113 Personen, die als Lehrkraft an hessischen Schulen zum Zeitpunkt der Erhebung tätig waren. 32 der Lehrkräfte befanden sich noch in der Ausbildung. Da es sich um eine Gelegenheitsstichprobe handelt, ist die Verallgemeinerung der Ergebnisse eingeschränkt. Die Rekrutierung der Lehrkräfte erfolgte über eine E-Mail-Einladung. Von den teilnehmenden Lehrkräften waren 66 % weiblich und 34 % männlich. Die Lehrkräfte waren im Mittel 39 Jahre alt ( $SD = 12$ , min. = 21, max. = 69). 56 % der Lehrkräfte unterrichteten an einem Gymnasium, 15,6 % an beruflichen Schulen, 11,9 % an Haupt- und Realschulen, 9,2 % an Grundschulen und 7,3 % an Förderschulen. Was die unterrichteten Fächer angeht, unterrichteten 50,4 % mindestens eine Sprache (inkl. Deutsch), 46 % unterrichteten mindestens ein naturwissenschaftliches Fach (inkl. Mathematik und Informatik), 42,5 % mindestens ein gesellschaftswissenschaftliches Fach und 14,2 % mindestens ein künstlerisches Fach.

### 3.2 Design & Durchführung

Bei der Studie handelt es sich um eine Querschnittsstudie, die eine Lerngelegenheit (einen zu lesenden Text) beinhaltete. Die Datenerhebung erfolgte mittels einer Online-Befragung in der Umfragesoftware *EFS Survey* (Unipark, 2022). Die Durchführung dauerte circa 30 Minuten. Die Lehrkräfte beantworteten zunächst Fragen zu ihrer Soziodemographie sowie ihrem Vorwissen und ihrer Erfahrung mit dem Thema Künstliche Intelligenz (KI). Es folgten Fragen zu ihren kognitiven und motivational-emotionalen Voraussetzungen (Selbstwirksamkeitserwartung, Emotionale Erschöpfung) und ihren Emotionen/Wahrnehmungen (Herausforderung, Bedrohung) hinsichtlich der bevorstehenden Lerngelegenheit (Text) zum Thema KI im eigenen Fachunterricht. Anschließend hatten die Lehrkräfte Zeit, einen Text

darüber zu lesen, wie das Thema KI, verknüpft mit Fachinhalten, im schulischen Unterricht vermittelt werden kann. Der Text erläuterte in einzelnen Schritten, wie die Schüler\*innen erlernen können, mit der Programmiersprache Scratch (Resnick et al., 2009) einen Chatbot zu programmieren, der zu einem spezifischen Unterrichtsthema Fragen beantwortet. Im Deutschunterricht könnten sich die Schüler\*innen zum Beispiel Fragen und Antworten zu einer Lektüre überlegen, die sie gelesen haben und einen Chatbot darauf trainieren, diese Fragen zu beantworten. Der Text umfasste zweieinhalb Seiten und endete mit weiterführenden Links zur Thematik. Entwickelt wurde der Text basierend auf Materialien von Junge Tüftler\*innen (2020). Nach dem Lesen des Textes beantworteten die Lehrkräfte Fragen zu ihrer erlebten kognitiven Belastung und kognitiven Anstrengung hinsichtlich des Textes (Krell, 2017).

### 3.3 Instrumente

Alle relevanten Prädiktoren und Kriterien wurden im Selbstbericht der Lehrkräfte erfasst. Das Alter wurde anhand einer Frage mit offenem Antwortformat erfasst (»Bitte geben Sie Ihr Alter in Jahren an«), das Geschlecht auf einer Nominalskala mit drei Antwortmöglichkeiten (weiblich, männlich, divers). Die anderen Variablen wurden anhand von Items erfasst, die die Lehrkräfte jeweils auf einer Skala von 1 – stimme überhaupt nicht zu bis 6 – stimme voll und ganz zu beurteilen.

Die Erfassung der *Selbstwirksamkeitserwartung* der Lehrkräfte hinsichtlich ihres Umgangs mit technologischen Anforderungen erfolgte anhand von sieben Items basierend auf Schmidt et al. (2009). Ein Beispielitem ist: »Mir fällt es leicht, den Umgang und Einsatz von digitalen Technologien zu erlernen«. Die Skala zeigte eine sehr gute interne Konsistenz:  $\alpha = .92$ . Der empirische Mittelwert der Skala betrug  $M = 4,31$  mit einer Standardabweichung von  $SD = 0,93$ .

Die *Emotionale Erschöpfung* der Lehrkräfte wurde anhand von 4 Items erfasst, basierend auf Hackman und Oldham (1975) und Merz (1986). Ein Beispielitem ist: »Ich fühle mich von der Arbeit insgesamt überlastet«. Die Skala zeigte eine gute interne Konsistenz:  $\alpha = .88$ . Der empirische Mittelwert der Skala lag bei  $M = 3,13$  mit einer Standardabweichung von  $SD = 1,19$ .

Die Bewertung der Lerngelegenheit (Text) als *Herausforderung* wurde anhand von 4 Items basierend auf Decker et al. (2017) und Kunter et al. (2016) erfasst. Ein Beispielitem ist: »Der Text über die Integration des Themas KI in meinen Fachunterricht ist eine spannende Herausforderung für mich«. Die Skala zeigte eine sehr gute interne Konsistenz:  $\alpha = .92$ . Der empirische Mittelwert der Skala lag bei  $M = 4,06$  mit einer Standardabweichung von  $SD = 1,25$ .

Die Bewertung der Lerngelegenheit (Text) als *Bedrohung* wurde anhand von 3 Items basierend auf Decker et al. (2017) und Kunter et al. (2016) erfasst. Ein Beispielitem ist: »Ich mache mir Sorgen, ob ich die Inhalte des Textes zum Thema

KI-Integration in meinen Unterricht verstehen werde.«. Die Skala zeigte eine gute interne Konsistenz:  $a = .87$ . Der empirische Mittelwert der Skala lag bei  $M = 2,67$  mit einer Standardabweichung von  $SD = 1,15$ .

Die Lehrkräfte beurteilten ihre *kognitive Belastung* und ihre *kognitive Anstrengung* in Hinblick auf den gelesenen Text jeweils anhand von 4 Items basierend auf Krell (2017). Ein Beispielitem für die kognitive Belastung ist: »Der Inhalt des Texts war kompliziert«. Ein Beispielitem für die kognitive Anstrengung ist: »Beim Lesen des Textes habe ich mich wenig bemüht«. Die Skala für die kognitive Belastung zeigte mit  $a = .90$  eine gute interne Konsistenz sowie die Skala für die kognitive Anstrengung mit  $a = .80$ . Der empirische Mittelwert der Skala kognitive Belastung lag bei  $M = 2,52$  mit einer Standardabweichung von  $SD = 0,90$ . Der empirische Mittelwert der Skala kognitive Anstrengung lag bei  $M = 3,87$  mit einer Standardabweichung von  $SD = 1,01$ .

### 3.4 Datenauswertung

Zunächst wurden bivariate Korrelationen aller Untersuchungsvariablen berechnet, um die einfachen Zusammenhänge der Variablen untereinander zu betrachten. Zur Untersuchung der Forschungsfragen wurden multiple Regressionsmodelle gerechnet, getrennt für die beiden Kriteriumsvariablen kognitive Belastung und kognitive Anstrengung. Trotz der Gelegenheitsstichprobe wurde eine inferenzstatische Auswertung vorgenommen und das Signifikanzniveau auf  $.05$  festgesetzt.

## 4. Ergebnisse

Die bivariaten Korrelationen zwischen den Untersuchungsvariablen lagen überwiegend im geringen ( $r = .10$ ) bis mittleren Bereich ( $r = .30$ ; Cohen, 1988). Die höchsten Korrelationen bestanden zwischen kognitiver Belastung und dem Erleben der Lerngelegenheit als Bedrohung,  $r = .45$ ,  $p < .001$  sowie zwischen der Selbstwirksamkeitserwartung und dem Erleben der Lerngelegenheit als Bedrohung,  $r = -.46$ ,  $p < .001$ .

Zunächst werden die Ergebnisse für die kognitive Belastung berichtet. Als Prädiktoren wurden Alter, Geschlecht, Selbstwirksamkeitserwartung, Emotionale Erschöpfung, das Erleben als Herausforderung und das Erleben als Bedrohung in das Regressionsmodell aufgenommen. Die erklärte Varianz lag bei  $R^2 = .34$ ,  $F(6, 100) = 8,45$ ,  $p < .001$ . Statistisch signifikante Prädiktoren für die mentale Belastung waren das Alter,  $b = .22$ ,  $t(100) = 2,44$ ,  $p = .02$ , die Wahrnehmung des Textes als Herausforderung,  $b = -.19$ ,  $t(100) = -2,15$ ,  $p = .03$  und die Wahrnehmung des Textes als Bedrohung,  $b = .39$ ,  $t(100) = 4,13$ ,  $p < .001$ . Dabei empfanden Lehrkräfte mit höherem Alter eine höhere kognitive Belastung durch den Text. Lehrkräfte, die die Beschäftigung mit dem Thema KI als Herausforderung wahrnahmen, emp-

fanden eine geringere Belastung und Lehrkräfte, die die Beschäftigung mit dem Thema KI als Bedrohung wahrnahmen, eine höhere Belastung.

Die Ergebnisse zur kognitiven Anstrengung beim Lesen des Textes fielen anders aus. Die erklärte Varianz des Modells lag bei  $R^2 = .11$ ,  $F(6, 101) = 2,02$ ,  $p = .07$ . Nur das Geschlecht der Lehrkräfte war ein signifikanter Prädiktor für die kognitive Anstrengung,  $b = -.21$ ,  $t(101) = -2,17$ ,  $p = .03$ . Männliche Lehrkräfte gaben an, sich mehr angestrengt zu haben als weibliche Lehrkräfte. Keiner der anderen Prädiktoren, das heißt, Alter, Selbstwirksamkeitserwartung, Emotionale Erschöpfung, das Erleben als Herausforderung und das Erleben als Bedrohung, hatte einen signifikanten Effekt auf die kognitive Anstrengung.

## 5. Diskussion

### 5.1 Interpretation

Die Ergebnisse zeigen, dass das emotionale Erleben der Lehrkräfte in Hinblick auf eine bevorstehende Lerngelegenheit mit der Verarbeitung und Nutzung dieser Lerngelegenheit zusammenhängt. Das Conceptual-Change-Modell von Gregoire (2003) postuliert, dass die emotionale Bewertung einer Lerngelegenheit als Bedrohung/Herausforderung die investierte *kognitive Anstrengung* bei der Auseinandersetzung mit Reform-Inhalten beeinflusst. In der aktuellen Untersuchung zeigt sich dieser Effekt allerdings vor allem in Hinblick auf die *kognitive Belastung* durch die Lerngelegenheit. Lehrkräfte, die die Lerngelegenheit zum Thema KI im Unterricht als Herausforderung wahrnahmen, empfanden den Text als weniger komplex und anspruchsvoll. Lehrkräfte, die sich davon bedroht fühlten, empfanden den Text hingegen als schwieriger. Letzterer Befund ist in Einklang mit Studien, die zeigen, dass negative Emotionen wie Ängste die kognitive Verarbeitung erschweren, beziehungsweise mit einer höheren wahrgenommenen Schwierigkeit assoziiert sind (Chen & Chang, 2009). Andere kognitive oder affektiv-motivationale Ressourcen waren in der vorliegenden Untersuchung nicht von relativer Bedeutsamkeit.

Neben dem Erleben als Bedrohung und Herausforderung spielten soziodemographische Aspekte der Lehrkräfte eine Rolle. So empfanden ältere Lehrkräfte den Text zum Thema KI als kognitiv fordernder, während sich weibliche Lehrkräfte weniger kognitiv anstrengten bei der Verarbeitung des Textes. Diese Befunde könnten damit erklärt werden, dass ältere Lehrkräfte berichten, in ihrer Ausbildung weniger gut auf den Einsatz von digitalen Technologien allgemein vorbereitet worden zu sein (Hämäläinen et al., 2021). Das dadurch geringere vorhandene Vorwissen könnte erklären, warum älteren Lehrkräften das Verständnis eines Textes zum Thema KI im Unterricht schwerer fällt. Was das Geschlecht angeht, zeigen Studien teilweise, dass weibliche Lehrkräfte und Frauen/Mädchen allgemein negativere Einstellungen und geringeres Interesse an digitalen Technologien und in-

formatischen Inhalten haben (Happe et al., 2021; Lin et al., 2021; Marpa, 2021). Dies könnte die geringere investierte kognitive Anstrengung in der vorliegenden Untersuchung erklären.

## 5.2 Limitationen

In der vorliegenden Arbeit wurden keine Mediationsmodelle gerechnet, die noch besser die angenommenen Zusammenhänge im Modell (zum Beispiel Selbstwirksamkeitserwartung wirkt sich positiv auf die Wahrnehmung als Herausforderung aus, die sich wiederum positiv auf die systematische Verarbeitung auswirkt) widerspiegeln würden. Dieser Umstand ist der eher kleinen Stichprobe und dem Pilotcharakter der Studie geschuldet. Auch wurde das Vorwissen der Lehrkräfte noch nicht als weitere zentrale Variable in die Analysen mit einbezogen. Der Fokus der Arbeit lag zudem auf der kognitiven Verarbeitung der Lerngelegenheit und nicht auf dem *conceptual change*, also der Untersuchung der Veränderung der Überzeugungen der Lehrkräfte zum Thema KI durch die Lerngelegenheit. Obwohl sich das herangezogene theoretische Modell (Gregoire, 2003) auf solch einen *conceptual change* bezieht, war es Ziel der Untersuchung, zunächst zu zeigen, wie sich mehrere im Modell postulierte Voraussetzungen der Lehrkräfte auf die kognitive Verarbeitung von Inhalten zu einem Reformthema auswirken. Umfassende Untersuchungen zu diesen Modellaspekten fehlen. Sie liefern aber wichtige Implikationen für die Anpassung von Lerngelegenheiten an die Voraussetzungen von Lehrkräften. Um genau diese Implikationen geht es im letzten Abschnitt.

## 5.3 Implikationen

Der Befund, dass ältere Lehrkräfte und sich bedrohtühlende Lehrkräfte eine höhere kognitive Belastung bei der Verarbeitung der Lerngelegenheit erleben, hat wichtige praktische Implikationen für die Gestaltung von Professionalisierungsmaßnahmen. Für die genannten Gruppen von Lehrkräften sollten Lerngelegenheiten zu KI entsprechend angepasst und darauf geachtet werden, dass die kognitive Belastung für diese Lehrkräfte nicht zu hoch ist. Eine hohe kognitive Belastung (mehr benötigte kognitive Ressourcen) geht mit schlechterer Performanz beim Lernen einher (Chen & Chang, 2009; Krell, 2017). Genauere und kleinschrittigere Erklärungen und die Zusammenarbeit mit jüngeren und in diesem Bereich bereits kompetenteren Lehrkräften (Christodoulou & Angeli, 2022; Clarke et al., 2019) könnten Möglichkeiten sein, um die kognitiven Anforderungen für ältere Lehrkräfte und solche mit negativen Emotionen zu reduzieren. Die erlebte kognitive Belastung könnte außerdem bereits im Vorfeld durch das Ansetzen an den Lehrkraftvoraussetzungen adressiert werden. Im Vorfeld zu Lerngelegenheiten könnten Ängste der Lehrkräfte zum Beispiel durch das Initiieren von Strategien zur

Emotionsregulation, wie die Neubewertung der Situation (Reappraisal), reduziert werden (Hofmann et al., 2009; Morgan & Atkin, 2016).

Auch für weibliche Lehrkräfte sollten Professionalisierungsmaßnahmen entsprechend angepasst werden. Um die investierte Anstrengung von weiblichen Lehrkräften bei der Auseinandersetzung mit KI-Inhalten zu erhöhen, könnte in den Lerngelegenheiten explizit mit weiblichen Rollenmodellen gearbeitet werden und zunächst stärker der Problemlöse- und Design-Aspekt von KI betont werden anstelle von Programmieraspekten (Happe et al., 2021).

## Literatur

- Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2014). Developing mathematical competence: From the intended to the enacted curriculum. *The Journal of Mathematical Behavior*, 33, 72–87.
- Broughton, S. H., Sinatra, G. M., & Nussbaum, E. M. (2013). »Pluto has been a planet my whole life!« Emotions, attitudes, and conceptual change in elementary students' learning about Pluto's reclassification. *Research in Science Education*, 43, 529–550.
- Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, 105, 1–13.
- Chen, I. J., & Chang, C. C. (2009). Cognitive load theory: An Empirical study of anxiety and task performance in language learning. *Electronic Journal of Research in Educational Psychology*, 7(2), 729–746.
- Chiu, T. K., & Chai, C. S. (2020). Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability*, 12(14), 5568.
- Christodoulou, A., & Angeli, C. (2022). Adaptive learning techniques for a personalized educational software in developing teachers' technological pedagogical content knowledge. *Frontiers in Education*, 7, 789397. <https://doi.org/10.3389/educ.2022.789397>
- Clarke, A. J., Burgess, A., van Diggele, C., & Mellis, C. (2019). The role of reverse mentoring in medical education: Current insights. *Advances in Medical Education and Practice*, 10, 693–701. <https://doi.org/10.2147/amep.s179303>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Erlbaum Associates.
- Decker, A., Kunter, M., Seiz, J., & Wilde, A. (2017). Professionelle Entwicklungsverläufe verstehen – Kinder mit Bildungsrisiken wirksam fördern. In M. Hartmann, M. Hasselhorn, & A. Gold (Hrsg.), *Entwicklungsverläufe verstehen – Kinder mit Bildungsrisiken wirksam fördern: Forschungsergebnisse des Frankfurter IDeA-Zentrums Forschungsergebnisse des Frankfurter IDeA-Zentrums* (S. 455–472). Kohlhammer.
- Diethelm, I. (2021). Informatische Bildung für alle Lehrkräfte – Position des GI-Arbeitskreises » Lehrkräftebildung «. *INFOS 2021–19. GI-Fachtagung Informatik und Schule*.
- Ebert, E. K., & Crippen, K. J. (2010). Applying a cognitive-affective model of conceptual change to professional development. *Journal of Science Teacher Education*, 21, 371–388.
- Feldon, D. F., Callan, G., Juth, S., & Jeong, S. (2019). Cognitive load as motivational cost. *Educational Psychology Review*, 31, 319–337.

- Feuerhahn, N., Stamov-Roßnagel, C., Wolfram, M., Bellingrath, S., & Kudielka, B. M. (2013). Emotional exhaustion and cognitive performance in apparently healthy teachers: A longitudinal multi-source study. *Stress and Health, 29*(4), 297–306.
- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review, 15*, 147–179.
- Hackman, J. R., & Oldham, G. R. (1975). Development of the job diagnostic survey. *Journal of Applied Psychology, 60*, 159–170.
- Hämäläinen, R., Nissinen, K., Mannonen, J., Lämsä, J., Leino, K., & Taajamo, M. (2021). Understanding teaching professionals' digital competence: What do PIAAC and TALIS reveal about technology-related skills, attitudes, and knowledge? *Computers in Human Behavior, 117*, 106672.
- Happe, L., Buhnova, B., Koziolk, A., & Wagner, I. (2021). Effective measures to foster girls' interest in secondary computer science education: A literature review. *Education and Information Technologies, 26*, 2811–2829.
- Hofmann, S. G., Heering, S., Sawyer, A. T., & Asnaani, A. (2009). How to handle anxiety: The effects of reappraisal, acceptance, and suppression strategies on anxious arousal. *Behaviour Research and Therapy, 47*(5), 389–394. <https://doi.org/10.1016/j.brat.2009.02.010>
- Jaschke, S., Klusch, M., Krupka, D., Losch, D., Michaeli, T., Opel, S., Schmid, U., Schwarz, R., Seegerer, S., & Stechert, P. (2023). *Künstliche Intelligenz in der Bildung*. Positionspapier der Gesellschaft für Informatik e. V. (GI).
- Junge Tüftler\*innen. (2020). *KI-Campus-Original. Schule macht KI*. <https://ki-campus.org/courses/kischule2020>
- Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Engagement and emotional exhaustion in teachers: Does the school context make a difference? *Applied Psychology, 57*, 127–151.
- Krell, M. (2017). Evaluating an instrument to measure mental load and mental effort considering different sources of validity evidence. *Cogent Education, 4*(1), 1280256.
- Kunter, M., Decker, A.-T., & Seiz, J. (2016). *Proliefs-Online-Studie. Dokumentation der Erhebungsinstrumente [Unveröffentlichtes Dokument]*. Goethe-Universität.
- Kunter, M., Kleickmann, T., Klusmann, U., & Richter, D. (2011). Die Entwicklung professioneller Kompetenz von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften – Ergebnisse des Forschungsprogramms COACTIV* (S. 56–68). Waxmann.
- Lin, P. Y., Chai, C. S., Jong, M. S. Y., Dai, Y., Guo, Y., & Qin, J. (2021). Modeling the structural relationship among primary students' motivation to learn artificial intelligence. *Computers and Education: Artificial Intelligence, 2*, 100006.
- Lindner, A., Romeike, R., Jasute, E., & Pozdniakov, S. (2019, November 18–20). *Teachers' perspectives on artificial intelligence*. 12th International conference on informatics in schools: Situation, evaluation and perspectives (ISSEP), Larnaca, Cyprus.

- Marpa, E. P. (2021). Technology in the teaching of mathematics: An analysis of teachers' attitudes during the COVID-19 Pandemic. *International Journal on Studies in Education (IJonSE)*, 3(2), 92–102.
- Merz, J. (1986). SAF: Fragebogen zur Messung von dispositioneller Selbstaufmerksamkeit. *Diagnostica*, 32, 142–152.
- Morgan, J., & Atkin, L. (2016). Expelling stress for primary school teachers: Self-affirmation increases positive emotions in teaching and emotion reappraisal. *International Journal of Environmental Research and Public Health*, 13(5), 500. <https://doi.org/10.3390/ijerph13050500>
- Orru, G., & Longo, L. (2019). The evolution of cognitive load theory and the measurement of its intrinsic, extraneous and germane loads: A review. In L. Longo & M. C. Leva (Hrsg.), *Human Mental Workload: Models and Applications: Second International Symposium, H-WORKLOAD 2018, Amsterdam, The Netherlands, September 20–21, 2018, Revised Selected Papers* (S. 23–48). Springer International Publishing.
- Peltier, T. (2022). *Helping preservice teachers understand dyslexia: A study on conceptual change and engagement with three text conditions* [Dissertation]. OU – Dissertations.
- Pinkard, N., Erete, S., Martin, C. K., & McKinney de Royston, M. (2017). Digital youth divas: Exploring narrative-driven curriculum to spark middle school girls' interest in computational activities. *Journal of the Learning Sciences*, 26(3), 477–516.
- Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM*, 52(11), 60–67. <https://doi.org/10.1145/1592761.1592779>
- Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction*, 19(5), 387–401.
- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological Pedagogical Content Knowledge (TPACK). *Journal of Research on Technology in Education*, 42(2), 123–149. <https://doi.org/10.1080/15391523.2009.10782544>
- Unipark. (2022). *EFS Survey* (Version EFS Fall 2022). Tivian XI GmbH.
- Wu, M. Y. M., & Yeziarski, E. J. (2023). Secondary chemistry teacher learning: Precursors for and mechanisms of pedagogical conceptual change. *Chemistry Education Research and Practice*, 24(1), 245–262.

*Justine Stang-Rabrig*

## **Geschlechtsspezifische Unterschiede im Wohlbefinden und die Rolle des wahrgenommenen Leistungsdrucks**

### **1. Einführung**

Das Wohlbefinden von Lernenden ist neben dem Kompetenzerwerb ein weiteres, wichtiges Bildungsziel (z. B. OECD, 2017; van Petegem et al., 2007) und von hoher Bedeutsamkeit. Es ist ein Wert an sich und zudem beispielsweise positiv mit Lernaktivitäten und Lernerfolg assoziiert (z. B. Stang-Rabrig et al., 2023b; Yang et al., 2019). Aufgrund der Wichtigkeit ist es zentral, Variablen zu identifizieren, die mit dem Wohlbefinden in einem positiven oder negativen Zusammenhang stehen. Theoretischen Modellen wie dem ökosystemischen Ansatz folgend nehmen Faktoren unterschiedlicher Systemebenen einen Einfluss auf Individuen, deren Merkmale und Entwicklung (Bronfenbrenner & Morris, 2006). Entsprechend können für das Wohlbefinden außer Individualfaktoren wie beispielsweise dem Geschlecht auch Umgebungsfaktoren wie der wahrgenommene Leistungsdruck, der beispielsweise in leistungsstarken Schulkontexten stärker ausgeprägt sein kann, von Relevanz sein (z. B. Hascher et al., 2018; Kleinkorres et al., 2023a). In Bezug auf das Geschlecht als einen Einflussfaktor des Wohlbefindens ist die empirische Befundlage in Teilen heterogen (z. B. Batz & Tay, 2018). Hinsichtlich des wahrgenommenen Leistungsdrucks zeigten sich geschlechtsbezogene Differenzen zuungunsten von Mädchen (z. B. Löfstedt et al., 2020) sowie, dass dieser insgesamt negativ für das Wohlbefinden zu sein scheint (z. B. Kleinkorres et al., 2023a). Die Zusammenhänge können dabei durch die Selbstbestimmungstheorie erklärt werden (z. B. Ryan & Deci, 2020).

Jugendliche Lernende in einer leistungsstarken Umgebung stellen eine besondere Gruppe dar, die als »at-risk group« bezeichnet werden (z. B. Luthar et al., 2020) und zentrale entwicklungspsychologische Herausforderungen wie zum Beispiel das Erreichen emotionaler Unabhängigkeit oder die Vorbereitung auf den Schulabschluss meistern müssen (z. B. Havighurst, 1953). Die Zeit des Übergangs kann dabei mit zusätzlichem Leistungsdruck einhergehen, der das Wohlbefinden negativ beeinflussen kann. Unklar ist, ob sich adoleszente Mädchen und Jungen in einer leistungsstarken Umgebung im letzten High School Jahr, vor der Transition in die tertiäre Bildung, in zentralen Wohlbefindensaspekten sowie in der Wahrnehmung des Leistungsdrucks unterscheiden und ob für diese spezifische Gruppe

Unterschiede in den Relationen zwischen Leistungsdruck und Wohlbefinden bei Mädchen und Jungen bestehen.

## 2. Theorie und Forschungsbefunde

### 2.1 Wohlbefinden von Lernenden

Das subjektive Wohlbefinden ist ein multidimensionales Konstrukt (Seligman, 2011), das abdeckt, wie Personen in Bezug auf das Leben allgemein oder einen bestimmten Bereich wie etwa Schule denken und fühlen (z. B. Diener et al., 1999). Es existieren verschiedene theoretische Ansätze zur Beschreibung des Wohlbefindens. Ein weit verbreiteter ist der hedonistische (z. B. Eid & Larsen, 2008; Ryan & Deci, 2001), der zwischen einer kognitiven und affektiven Komponente differenziert (Diener et al., 1999). Die kognitive Komponente umfasst Evaluationen, die affektive kann unterschieden werden in einen positiven und negativen Affekt (Bradburn, 1969; Diener et al., 2013). Beide Komponenten können global oder bereichsspezifisch ausgerichtet sein. Ein globales Beispiel der kognitiven Komponente ist die *Lebenszufriedenheit*, die abbildet, wie zufriedenen Personen mit den Lebensbedingungen beziehungsweise ihrem Leben allgemein sind (z. B. Diener et al., 2013). Im Mittel sind Lernende mit ihrem Leben zufrieden (z. B. OECD, 2017), was auch auf begabte Jugendliche zutrifft (z. B. Bergold et al., 2015). In Bezug auf den spezifischen Bereich Schule stellt die *Schulzufriedenheit*, verstanden als positive schulbezogene Einstellungen, ein zentrales Beispiel für die kognitive Komponente dar (z. B. Morinaj & Hascher, 2019). Studien zeigten, dass (begabte) Lernende im Mittel der Schule gegenüber positiv eingestellt sind (z. B. Morinaj & Hascher, 2019; Zeidner & Schleyer, 1999). *Schulbezogene Erschöpfung* als bereichsspezifischer negativer Affekt gibt Auskunft darüber, wie lustlos und emotional ausgelaugt sich Lernende bedingt durch schulische Anforderungen fühlen (z. B. Salmela-Aro et al., 2009a). Erschöpfung ist eines von insgesamt drei Symptomen des psychologischen Syndroms Burnout und resultiert aus der Wahrnehmung langanhaltenden Stresses (z. B. Maslach et al., 2001). Studien ergaben, dass jugendliche Lernende im Durchschnitt über eine mittelhoch ausgeprägte Erschöpfung beziehungsweise Burnout berichten (z. B. Lee et al., 2021; Salmela-Aro et al., 2009b). Salmela-Aro und Kollegen (2008) konnten darüber hinaus zeigen, dass jugendliche Lernende an Schulen mit im Mittel höherem Leistungsniveau (*academic track*) durchschnittlich erschöpfter waren als Lernende an Schulen mit im Mittel geringerem Leistungsniveau (*vocational track*). Neben der psychischen Facette konstituiert sich das multidimensionale Konstrukt Wohlbefinden zudem aus dem physischen und sozialen Wohlbefinden (WHO, 2014). Ein relevanter Indikator des sozialen Wohlbefindens im schulischen Kontext ist die *Peerzufriedenheit*, die widerspiegelt wie qualitativ die Beziehungen mit Peers sind und inwiefern Lernende soziale Probleme wie zum

Beispiel Ausgrenzung erfahren (z. B. Morinaj & Hascher, 2019). Die Peerzufriedenheit kann somit als Evaluation der Peerbeziehungsqualität verstanden werden. Im Mittel sind jugendliche Lernende mit ihren Peers zufrieden beziehungsweise berichten von Zugehörigkeitsgefühl, einem wichtigen Indikator für Peerzufriedenheit, oder der Abwesenheit sozialer Probleme in der Schule (z. B. Hascher & Hagenauer, 2020; Flick et al., 2022). Für begabte Jugendliche beziehungsweise Jugendliche in leistungsstarken Umgebungen zeigten sich ähnliche Befunde (z. B. Godor & Szymanski, 2017; Stang-Rabrig et al., 2023a). Studien ergaben, dass das Wohlbefinden mit verschiedenen Leistungsindikatoren, dem Lernverhalten sowie mit lernrelevanten Lernendenmerkmalen wie der Motivation assoziiert ist (z. B. Bücken et al., 2018; Howard et al., 2021; Kleinkorres et al., 2020; Stang-Rabrig et al., 2023a; Stang-Rabrig et al., 2023b). Aufgrund der Bedeutsamkeit des Wohlbefindens ist es unerlässlich, Faktoren zu identifizieren, die mit dem Wohlbefinden positiv oder negativ assoziiert sind.

## 2.2 Geschlecht, Leistungsdruck und Wohlbefinden

Verschiedene Faktoren, verortet auf unterschiedlichen Ebenen, nehmen Einfluss auf Individuen, deren Merkmale und Entwicklung (z. B. ökosystemischer Ansatz: Bronfenbrenner & Morris, 2006). Neben zentralen Individualfaktoren wie dem Geschlecht, sind ebenfalls Umgebungsvariablen relevant. Im Mikrosystem, das das Individuum direkt umgibt, sind außer Peers beispielsweise auch die zwei weiteren, wichtigen Sozialisationsinstanzen Familie und Schule (z. B. Berns, 2015) lokalisiert. Nicht nur der schulische Kontext (z. B. Differenzielle Lernmilieus: Baumert et al., 2009), der beispielsweise leistungsstark sein kann, sondern auch der wahrgenommene Leistungsdruck, der durch Eltern und Schule appliziert werden kann, ist von Relevanz. Die Zusammenhänge können dabei mittels der Selbstbestimmungstheorie erklärt werden (z. B. Ryan & Deci, 2020). Hinsichtlich des Geschlechts wird angenommen, dass Mädchen beziehungsweise Frauen im Vergleich zu Jungen beziehungsweise Männern weniger Möglichkeiten haben, ihre psychologischen Grundbedürfnisse nach Autonomie, Kompetenz und Eingebundenheit zu befriedigen, wodurch strukturelle Disparitäten im Wohlbefinden zwischen Mädchen und Jungen beziehungsweise Frauen und Männern resultieren können (z. B. Batz & Tay, 2018). Ausgeübter Leistungsdruck, der Leistungserwartungen sowie -anforderungen widerspiegelt (z. B. Quach et al., 2015), lässt ebenfalls die psychologischen Grundbedürfnisse der Lernenden unberücksichtigt. Druck, der durch Eltern oder die Schule ausgeübt wird, kann vor allem das Bedürfnis nach Autonomie beeinträchtigen, was dann wiederum das Wohlbefinden unterminieren kann (z. B. Ryan & Deci, 2017). Erlebter Leistungsdruck kann zudem eine Fehlpassung zwischen Lernendenbedürfnissen in bestimmten zentralen Entwicklungsphasen und Anforderungen der Umgebung reflektieren (z. B. Stage-environment-Fit-Theory: Eccles et al., 1993).

Mit Fokus auf den zentralen Individualfaktor Geschlecht und den Zusammenhang mit Lebenszufriedenheit ist die empirische Befundlage inkonsistent (z. B. Batz & Tay, 2018). Studien ergaben entweder keine Zusammenhänge zwischen Geschlecht und Lebenszufriedenheit (z. B. Chui & Wong, 2016; Terry & Huebner, 1995) oder aber das Mädchen beziehungsweise Frauen im Mittel niedrigere Werte als Jungen beziehungsweise Männer aufwiesen (z. B. Cosma et al., 2023; Moksnes & Espnes, 2013). Für begabte Jugendliche zeigten Bergold und Kollegen (2015), dass Mädchen im Mittel mit ihrem Leben weniger zufrieden waren als Jungen. Hinsichtlich der Schulzufriedenheit beziehungsweise schulbezogenen positiven Einstellungen erbrachten verschiedene Studien, dass diese bei (begabten) Mädchen im Mittel höher ausgeprägt war als bei Jungen (z. B. Hascher et al., 2011; Liu et al., 2016; Löfstedt et al., 2020; Swiatek & Lupkowski-Shoplik, 2000). Auch in Bezug auf Erschöpfung beziehungsweise Burnout bestehen geschlechtsspezifische Unterschiede: Im Mittel berichteten Mädchen von einer höher ausgeprägten Erschöpfung, was auch auf Lernende an Schulen mit im Mittel höherem Leistungsniveau (*academic track*) im Vergleich zu Lernenden an Schulen mit im Mittel geringerem Leistungsniveau (*vocational track*) zutrifft (z. B. Herrmann et al., 2019; Salmela-Aro et al., 2008; Salmela-Aro et al., 2009b). Hinsichtlich der Peer- und Freundschaftszufriedenheit beziehungsweise Abwesenheit sozialer Probleme scheinen keine Unterschiede zwischen (begabten) Mädchen und Jungen zu bestehen (z. B. Aydin et al., 2022; Hascher et al., 2011; für einen Überblick s. Rose & Rudolph, 2006). All diese zentralen Wohlbefindensmerkmale wurden jedoch selten simultan betrachtet.

Mit Blick auf bedeutsame Umgebungsvariablen wie der wahrgenommene, durch Eltern oder Schule applizierte, Leistungsdruck konnten ebenfalls Unterschiede zwischen Mädchen und Jungen gefunden werden. Für elterlichen Leistungsdruck zeigte sich, dass dieser von Mädchen im Mittel höher wahrgenommen wurde als von Jungen (DuongTran et al., 1996). Löfstedt und Kollegen (2020) fanden darüber hinaus, dass geschlechtsbezogene Disparitäten auch in der Wahrnehmung des schulischen Leistungsdrucks bestehen: Mädchen berichteten im Mittel im Vergleich zu Jungen über mehr wahrgenommenem schulischen Leistungsdruck.

Vereinzelte Arbeiten betrachteten zudem Zusammenhänge zwischen dem wahrgenommenen Leistungsdruck und verschiedenen Wohlbefindensaspekten (z. B. Kleinkorres et al., 2023a; Obermeier et al., 2022). Bei einzelner Betrachtung war insbesondere der wahrgenommene elterliche Leistungsdruck negativ assoziiert mit Schulzufriedenheit und sozialer Integration (z. B. Kleinkorres et al., 2023a). Bei gleichzeitiger Betrachtung beider Leistungsdruckarten sowie der Berücksichtigung von Kovariaten wie Leistung, die ebenfalls mit dem Wohlbefinden zusammenhängt, zeigten Studien, dass Zusammenhänge nicht mehr statistisch signifikant waren oder aber kleiner wurden (Kleinkorres et al., 2023a; Stang-Rabrig & Kleinkorres, 2023). Fokussierend auf adolozente leistungsstarke Lernende ergab die

simultane Berücksichtigung des erlebten elterlichen und schulischen Leistungsdrucks, dass der wahrgenommene schulische, nicht aber elterliche, Leistungsdruck statistisch signifikant positiv mit der Entwicklung des erlebten Stresses und der berichteten Prüfungsangst zusammenhing (Stang-Rabrig et al., 2023a). Insgesamt wurden beide Leistungsdruckarten jedoch selten zusammen untersucht (z. B. Kleinkorres et al., 2023a; Stang-Rabrig et al., 2023a).

### 3. Fragestellungen

Das Wohlbefinden von Lernenden ist von hoher Bedeutung sowohl für den schulischen als auch sich anschließenden beruflichen Erfolg. Theoretischen Modellen und empirischen Befunden zufolge bestehen Disparitäten zwischen Mädchen und Jungen in einzelnen Wohlbefindensaspekten sowie in der Wahrnehmung des Leistungsdrucks. Die Befunde beziehen sich jedoch zumeist nicht auf Lernende in einer leistungsstarken Umgebung, die sich zudem in einer bedeutsamen Schulphase, am Ende der High School, beziehen. Unklar ist auch, ob die Zusammenhänge zwischen wahrgenommenem Leistungsdruck und verschiedenen Wohlbefindensaspekten different für Mädchen und Jungen ausfallen. Da in bisherigen Arbeiten oftmals nur ein oder wenige Wohlbefindensaspekte beziehungsweise Leistungsdruckarten betrachtet wurden, wurden in dieser Arbeit zentrale Wohlbefindensaspekte sowie zwei Leistungsdruckarten simultan betrachtet. Vor diesem Hintergrund wurden folgende Forschungsfragen formuliert:

- 1) Wie unterscheiden sich adoleszente Mädchen und Jungen in einer leistungsstarken Umgebung in ihrem Wohlbefinden (Lebenszufriedenheit, Schulzufriedenheit, schulbezogene Erschöpfung, Peerzufriedenheit) sowie in der Wahrnehmung des elterlichen und schulischen Leistungsdrucks?
- 2) Wie hängt der wahrgenommene Leistungsdruck (Eltern, Schule) bei adoleszenten Mädchen und Jungen in einer leistungsstarken Umgebung mit ihrem Wohlbefinden (Lebenszufriedenheit, Schulzufriedenheit, schulbezogene Erschöpfung, Peerzufriedenheit) zusammen?

## 4. Methode

### 4.1 Stichprobe und Design

Die Analysen basieren auf Angaben von 1,286 Lernenden aus zwei leistungsstarken US-amerikanischen High Schools eines Schulbezirks in Kalifornien (46.7 % weiblich), die an zwei Erhebungszeitpunkten (T1: Klasse 11, T2: Klasse 12) teilnahmen, wobei zur Beantwortung der Forschungsfragen die Daten des zweiten Messzeitpunktes, des letzten High School Jahres, herangezogen wurden. Zu T1 waren die

Lernenden im Mittel 16.40 Jahre alt ( $SD = 0.44$ ). Der sozioökonomische Hintergrund wurde über den International Socio-Economic Index (ISEI; Ganzeboom & Treiman, 2003) operationalisiert. Der höchste ISEI lag bei durchschnittlich  $M = 62.48$  ( $SD = 18.16$ ). Der ethnische Hintergrund der Lernenden verteilte sich prozentual wie folgt: 42.3% *Chinese*, 16.5% *Hispanic*, 10.6% *Korean*, 9.7% *White* and 21.0% *other* (z. B. *Black* or *African American*, *Vietnamese*). Insgesamt wurden 19.0% der Lernenden außerhalb der USA geboren. Der grade point average (GPA; GPA-Range in Kalifornien: 0–4) am Ende der 10. Klasse betrug 3.36 ( $SD = 0.57$ ). Dieser Wert wurde herangezogen, um festzumachen, dass es sich um eine leistungsstarke Umgebung handelte. Der Referenzwert für wohlhabende Schulen lag 2016 bei 3.0, für weniger wohlhabende Schulen bei 2.59 (Gershenson, 2018). Die Daten stammen aus dem Projekt Be Well at School (BELLS), das vom Institut für Schulentwicklungsforschung (IFS) an der TU Dortmund und einem US-kalifornischem Schulbezirk gemeinsam durchgeführt wurde. Im Rahmen der 45-minütigen Online-Erhebung machten die Lernenden Angaben zu den interessierenden Variablen und zur Soziodemografie. Die Teilnahme war freiwillig und setzte das Einverständnis voraus.

## 4.2 Instrumente

### 4.2.1 Wohlbefinden von Lernenden

Verschiedene Aspekte des Wohlbefindens der Lernenden wurden anhand von vier etablierten Skalen erfasst. Die *Lebenszufriedenheit* (Diener et al., 1985) der Lernenden wurde mittels fünf Items gemessen, die global erfassen, wie Lernende ihr Leben einschätzen. Ein Beispielitem lautet: »Ich bin mit meinem Leben zufrieden.«. Die Reliabilität der Skala war gut ( $a_{\text{Gesamt}} = .86$ ,  $a_{\text{Mädchen}} = .88$ ,  $a_{\text{Jungen}} = .84$ ). Angaben erfolgten auf einer siebenstufigen Likert-Skala von 1 = *stimme überhaupt nicht zu* bis 7 = *stimme voll und ganz zu*. *Schulzufriedenheit* (Huebner, 2001) wurde anhand von sieben Items erfragt, die darüber Auskunft geben, wie häufig Lernende gerne in die Schule gehen beziehungsweise gerne dort lernen (z. B. »Ich freue mich darauf, in die Schule zu gehen«). Lernende konnten den Aussagen auf einer vierstufigen Likert-Skala zustimmen von 1 = *niemals* bis 4 = *fast immer*. Die Reliabilität war gut ( $a_{\text{Gesamt}} = .83$ ,  $a_{\text{Mädchen}} = .84$ ,  $a_{\text{Jungen}} = .83$ ). *Schulbezogene Erschöpfung* (Maslach et al., 2016) wurde mittels sechs Items erfasst, die abdecken, ob sich Lernende durch die Schule erschöpft fühlen und der Meinung sind, zu viel für die Schule tun zu müssen. Ein Beispielitem lautet: »Am Ende eines Schultages fühle ich mich erschöpft.«. Die Items konnten entweder *verneint* (0) oder *bejaht* (1) werden. Die Reliabilität der Skala lag in einem zufriedenstellenden Bereich ( $a_{\text{Gesamt}} = .75$ ,  $a_{\text{Mädchen}} = .78$ ,  $a_{\text{Jungen}} = .71$ ). *Peerzufriedenheit* (Huebner, 2001), die darüber Auskunft gibt, wie häufig Lernende mit ihren Peers zufrieden sind, wurde anhand von

neun Items erfasst (z. B. »Meine Mitschüler\*innen sind großartig.«). Den Aussagen konnte auf einer vierstufigen Likertskala von 1 = *niemals* bis 4 = *fast immer* zugestimmt werden. Die Reliabilität war gut ( $a_{\text{Gesamt}} = .85$ ,  $a_{\text{Mädchen}} = .85$ ,  $a_{\text{Jungen}} = .85$ ).

#### 4.2.2 Leistungsdruck

Leistungsdruck wurde aus Sicht der Lernenden in Bezug auf Eltern und Schule gemessen. Zur Erfassung des wahrgenommenen *elterlichen Leistungsdrucks* (Hagenauer, 2011) wurden fünf Items herangezogen (z. B. »Meine Eltern wollen, dass ich intensiv lerne.«), die aufzeigen, wie stark Lernende den von Eltern applizierten Leistungsdruck wahrnehmen. Die Reliabilität der Skala lag in einem guten Bereich ( $a_{\text{Gesamt}} = .86$ ,  $a_{\text{Mädchen}} = .87$ ,  $a_{\text{Jungen}} = .84$ ). Der wahrgenommene *schulische Leistungsdruck* (Daniels, 2008) wurde anhand von vier Items erhoben, die adressieren, wie stark Lernende den von Lehrkräften ausgeübten Leistungsdruck empfinden. Ein Beispielim lautet: »Wenn wir nicht an den Wochenenden lernen, ist es kaum möglich, die Anforderungen zu erfüllen.«. Die Reliabilität war zufriedenstellend ( $a_{\text{Gesamt}} = .74$ ,  $a_{\text{Mädchen}} = .71$ ,  $a_{\text{Jungen}} = .75$ ). Lernende beantworteten die Items zur Erfassung des Leistungsdrucks auf einer vierstufigen Likertskala von 1 = *nicht wahr/stimme überhaupt nicht zu* bis 4 = *wahr/stimme voll und ganz zu*.

#### 4.3 Statistische Analysen

Die Analysen wurden in SPSS 29 und *Mplus* 8.1 (Muthén & Muthén, 1998–2017) durchgeführt. Als Voranalyse wurden Messinvarianzanalysen durchgeführt (s. Tabelle 1). Zur Beantwortung der ersten Forschungsfrage wurde eine multivariate Kovarianzanalyse und zur Beantwortung der zweiten Forschungsfrage ein Mehrgruppenvergleich gerechnet. Gruppierungsvariable für den Mehrgruppenvergleich war das Geschlecht der Lernenden. Die Konstrukte wurden latent modelliert. Die Modellgüte wurde anhand des *comparative fit index* (CFI), des *root mean square error of approximation* (RMSEA) und des  $\chi^2$  bewertet (Hu & Bentler, 1999). Residualkovarianzen zwischen Itempaaren wurden zugelassen, um die Modellpassung zu verbessern. Für fehlende Werte (s. Tabelle 2) wurde in SPSS listenweiser Fallabschluss und in *Mplus* die FIML-Option (Graham & Coffman, 2012) genutzt. Der GPA wurde in allen Analysen als Kontrollvariable aufgenommen.

Tabelle 1.: Ergebnisse Messinvarianzanalyse

Modell	$\chi^2$	df	p	CFI/TLI	RMSEA/ SRMR	Vergleich	$\chi^2$	df	p( $\chi^2$ )
<i>Lebenszufriedenheit</i>									
1. Konöğural	10.90	4	.028	.995/.977	.076/.013				
2. Metrisch	16.48	8	.036	.994/.986	.059/.038	2 vs.1	5.58	4	.233
3. Skalar	40.45	12	.000	.981/.968	.089/.065	3 vs.2	23.97	4	.000
<i>Schulzufriedenheit</i>									
1. Konöğural	48.68	18	.000	.990/.977	.058/.026				
2. Metrisch	70.07	24	.000	.985/.974	.061/.054	2 vs.1	21.38	6	.002
3. Skalar	74.97	30	.000	.986/.980	.054/.058	3 vs.2	4.90	6	.557
<i>Erschöpfung</i>									
1. Konöğural	38.87	14	.001	.984/.966	.057/.025				
2. Metrisch	46.94	19	.000	.981/.970	.054/.040	2 vs.1	10.08	5	.073
3. Skalar	84.19	24	.000	.958/.948	.070/.048	3 vs.2	37.25	5	.000
<i>Peerzufriedenheit</i>									
1. Konöğural	622.75	44	.000	.869/.786	.162/.092				
2. Metrisch	625.52	52	.000	.871/.821	.148/.092	2 vs.1	2.77	8	.948
3. Skalar	692.42	60	.000	.857/.829	.145/.098	3 vs.2	66.90	8	.000
<i>Elterlicher Leistungsdruck</i>									
1. Konöğural	11.18	2	.003	.996/.962	.095/.008				
2. Metrisch	14.20	6	.028	.997/.989	.052/.022	2 vs.1	3.03	4	.553
3. Skalar	25.52	10	.004	.994/.987	.055/.028	3 vs.2	11.32	4	.023
<i>Schulischer Leistungsdruck</i>									
1. Konöğural	32.85	4	.000	.968/.903	.119/.031				
2. Metrisch	35.84	7	.000	.968/.945	.090/.038	2 vs.1	2.98	3	.395
3. Skalar	42.60	10	.000	.963/.956	.080/.034	3 vs.2	6.76	3	.080

Tabelle 2: Mittelwerte, Standardabweichungen und Interkorrelationen für die Gesamtschichtprobe sowie getrennt für Mädchen und Jungen

	Gesamt <i>M (SD)</i>	Mäd- chen <i>M (SD)</i>	Jungen <i>M (SD)</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Lebenszufriedenheit <sup>1</sup>	5.09 (1.40)	4.96 (1.43)	5.24 (1.37)	-	<b>.39</b>	<b>-.33</b>	<b>.35</b>	<b>-.15</b>	<b>-.10</b>	<b>.05</b>	<b>.10</b>
(2) Schulzufriedenheit <sup>2</sup>	2.37 (0.59)	2.34 (0.58)	2.40 (0.61)	<b>.38/.39</b>	-	<b>-.49</b>	<b>.37</b>	<b>-.05</b>	<b>-.16</b>	<b>.10</b>	<b>.05</b>
(3) Erschöpfung <sup>3</sup>	0.72 (0.30)	0.78 (0.29)	0.68 (0.29)	<b>-.34/-.29</b>	<b>-.52/-.46</b>	-	<b>-.28</b>	<b>.13</b>	<b>.32</b>	<b>.02</b>	<b>-.15</b>
(4) Peerszufriedenheit <sup>2</sup>	3.01 (0.54)	3.02 (0.54)	3.02 (0.54)	<b>.41/.31</b>	<b>.43/.32</b>	<b>-.28/-.28</b>	-	<b>-.13</b>	<b>-.16</b>	<b>.15</b>	<b>-.01</b>
(5) Elterlicher Leistungsdruck <sup>2</sup>	2.40 (0.83)	2.30 (0.85)	2.50 (0.82)	<b>-.19/-.13</b>	<b>-.08/-.03</b>	<b>.16/.14</b>	<b>-.13/-.13</b>	-	<b>.29</b>	<b>-.07</b>	<b>.13</b>
(6) Schulischer Leistungsdruck <sup>2</sup>	2.74 (0.61)	2.79 (0.59)	2.70 (0.64)	<b>-.13/-.06</b>	<b>-.18/-.15</b>	<b>.30/.32</b>	<b>-.15/-.16</b>	<b>.29/.31</b>	-	<b>.00</b>	<b>-.07</b>
(7) GPA	3.36 (0.57)	3.47 (0.52)	3.28 (0.58)	<b>.10/.05</b>	<b>.19/.06</b>	<b>-.06/.04</b>	<b>.18/.14</b>	<b>-.06/-.04</b>	<b>-.03/.00</b>	-	<b>-.16</b>
(8) Geschlecht <sup>4</sup>	-	-	-	-	-	-	-	-	-	-	-
% fehlende Werte				20.8	20.5	20.8	22.0	20.8	20.6	4.5	1.3

*Anmerkungen.* Über der Diagonalen werden bivariate Korrelationen für die Gesamtgruppe, unter der Diagonalen werden vor dem Schrägstrich die Werte für Mädchen, hinter dem Schrägstrich die der Jungen berichtet. Fett gedruckte Werte verdeutlichen Signifikanz ( $p < .05$ ). <sup>1</sup> = *stimme überhaupt nicht* zu bis *7 = stimme voll und ganz zu*; <sup>2</sup> *1 = niemals/nicht wahr/stimme überhaupt nicht* zu bis *4 = fast immer/wahr/stimme voll und ganz zu*; <sup>3</sup> *0 = nein*, *1 = ja*; <sup>4</sup> *1 = Mädchen*, *2 = Junge*. Angaben zu fehlenden Werten beziehen sich auf die Gesamtgruppe.

## 5. Ergebnisse

### 5.1 Deskriptiva

Die Ergebnisse der deskriptiven Analysen (s. Tabelle 2) verdeutlichen für die Gesamtstichprobe zum Beispiel, dass die Lernenden im Mittel mit ihrem Leben zufrieden waren. Der Mittelwert wich statistisch signifikant vom theoretischen Skalenmittelwert ab,  $t(1017) = 24.84$ ,  $p < .001$ ,  $d = .78$ . Gleiches galt für die Substichprobe der Mädchen,  $t(471) = 14.58$ ,  $p < .001$ ,  $d = .70$ , und Jungen,  $t(1017) = 20.59$ ,  $p < .001$ ,  $d = .88$ . Des Weiteren zeigte sich beispielsweise sowohl in der Gesamt- als auch in den beiden Substichproben, dass die Lebenszufriedenheit statistisch signifikant positiv mit der Schul- und Peerzufriedenheit sowie negativ mit der schulischen Erschöpfung zusammenhing.

### 5.2 Unterschiede im Wohlbefinden und im wahrgenommenen Leistungsdruck

Die multivariate Kovarianzanalyse ergab Unterschiede zwischen Mädchen und Jungen für die kombinierten abhängigen Variablen, Wilks-L = 0.95,  $F(6, 940) = 9.08$ ,  $p < .001$ , partielles  $h^2 = .06$ . Auch die Kovariate GPA wurde signifikant, Wilks-L = 0.97,  $F(6, 940) = 5.71$ ,  $p < .001$ , partielles  $h^2 = .04$ . Zwischen Mädchen und Jungen bestanden konkret im Mittel Unterschiede in der subjektiven Lebenszufriedenheit, der Schulzufriedenheit, der schulbezogenen Erschöpfung sowie dem wahrgenommenen elterlichen und schulischen Leistungsdruck (s. Tabelle 3). Im Vergleich zu Mädchen waren Jungen im Mittel zufriedener mit ihrem Leben und der Schule und nahmen einen höheren durch Eltern applizierten Leistungsdruck wahr. Mädchen hingegen berichteten im Mittel im Vergleich zu Jungen über eine höhere schulische Erschöpfung und einen als höher erlebten schulischen Leistungsdruck. Die Effekte sind insgesamt als klein einzustufen (Cohen, 1988).

Tabelle 3: Unterschiede zwischen Mädchen und Jungen

Variable	Mädchen	Jungen	$F(1, 945)$	$h_p^2$	1-b
Lebenszufriedenheit	4.96 (1.43)	5.24 (1.37)	11.58***	.01	-
Schulzufriedenheit	2.34 (0.58)	2.40 (0.61)	3.84*	.01	-
Erschöpfung	0.78 (0.29)	0.68 (0.29)	23.03***	.02	-
Peerzufriedenheit	3.02 (0.54)	3.02 (0.54)	0.59	-	.12
Elterlicher Leistungsdruck	2.30 (0.85)	2.50 (0.82)	11.55***	.01	-
Schulischer Leistungsdruck	2.79 (0.59)	2.70 (0.64)	5.39*	.01	-

Anmerkung. \*\*\*  $p < .001$ , \*  $p < .05$ .

### 5.3 Zusammenhang Leistungsdruck und Wohlbefinden

Die Zusammenhänge zwischen den unterschiedlichen Leistungsdruckarten und den verschiedenen Wohlbefindensaspekten sind für Mädchen und Jungen in Abbildung 1 dargestellt. Der Modellfit war zufriedenstellend,  $q^2 = 2909.53$ ,  $df = 1266$ ,  $p < .05$ ,  $CFI = .90$ ,  $RMSEA = .05$ . Der Mehrgruppenvergleich ergab sowohl für Mädchen als auch für Jungen, dass der wahrgenommene elterliche Leistungsdruck negativ mit der Lebenszufriedenheit zusammenhing. Der Unterschied in der Zusammenhangsstärke bei Mädchen und Jungen war statistisch nicht signifikant,  $Dq^2 = 2.11$ ,  $Ddf = 1$ ,  $p = .15$ . Darüber hinaus zeigte sich in der Gruppe der Jungen, dass der erlebte elterliche Leistungsdruck positiv mit der Schulfriedenheit assoziiert war. Sowohl bei Mädchen als auch bei Jungen stand der wahrgenommene schulische Leistungsdruck in einem negativen Zusammenhang mit der Schul- und Peerzufriedenheit sowie in einem positiven Zusammenhang mit der berichteten schulbezogenen Erschöpfung. Die Zusammenhangsstärke war in beiden Gruppen ähnlich beziehungsweise gleich für die Assoziationen zwischen schulischem Leistungsdruck und Erschöpfung beziehungsweise Peerzufriedenheit. Für die Stärke des Zusammenhangs zwischen dem wahrgenommenen schulischen Leistungsdruck und der Schulfriedenheit zeigte sich, dass diese nicht statistisch signifikant unterschiedlich für Mädchen und für Jungen war,  $Dq^2 = 0.39$ ,  $Ddf = 1$ ,  $p = .53$ . Die Varianzaufklärung war insgesamt gering bis mittel (Cohen, 1998).

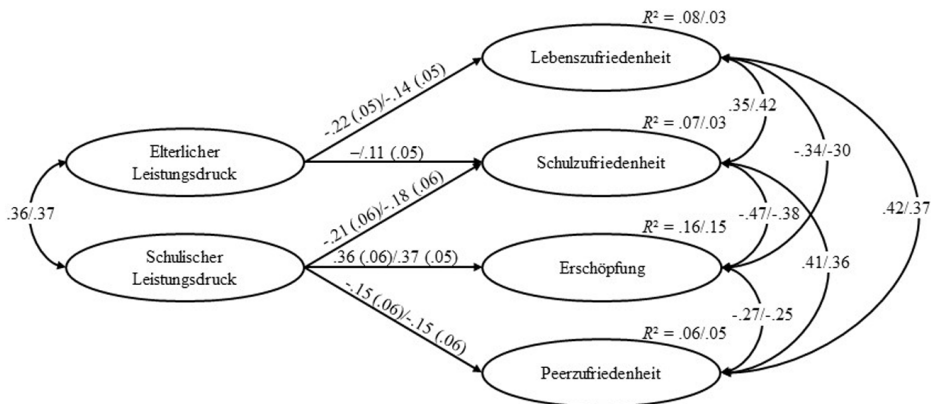


Abbildung 1. Zusammenhänge zwischen den beiden wahrgenommenen Leistungsdruckarten und den verschiedenen Wohlbefindensaspekten getrennt für Mädchen und Jungen.

Anmerkungen. Nur signifikante Pfade abgetragen. Werte vor Schrägstrich Mädchen, nach Schrägstrich Jungen. Koeffizienten sind standardisiert, Standardfehler in Klammern berichtet. Der Übersichtlichkeit halber Kontrollvariable nicht eingezeichnet.

## 6. Diskussion

Der Beitrag ging der Frage nach, ob bei adolescenten Mädchen und Jungen in einer leistungsstarken Umgebung am Ende der High School als einem bedeutsamen Life-Event Unterschiede im Wohlbefinden sowie wahrgenommenen Leistungsdruck bestehen. Darüber hinaus wurde untersucht, ob die Zusammenhänge zwischen wahrgenommenem Leistungsdruck und Wohlbefinden für Mädchen und Jungen unterschiedlich ausfallen. Hierzu wurden verschiedene zentrale Aspekte des Wohlbefindens sowie zwei Leistungsdruckarten simultan betrachtet.

In Einklang mit theoretischen Annahmen und bisheriger Forschung zeigten sich auch für adolescenten Lernende in einem leistungsstarken US-amerikanischen Kontext statistisch signifikante Unterschiede zwischen Mädchen und Jungen in der Lebenszufriedenheit, Erschöpfung sowie dem elterlichen und schulischen Leistungsdruck (z. B. Bergold et al., 2015; Hermann et al., 2019; Löfstedt et al., 2020). Erwartungskonträr war, dass Jungen im Mittel über eine höhere Schulzufriedenheit berichteten als Mädchen, was an der Spezifität der Stichprobe liegen könnte. Ein leistungsstarker Kontext kommt einer stärkeren Wettbewerbssituation gleich, wobei Frauen Wettbewerbssituationen stärker meiden (z. B. Niederle & Vesterlund, 2007) und sich in solchen Situationen eventuell weniger wohlfühlen könnten.

Erwartungskonform zeigten sich negative Zusammenhänge zwischen dem wahrgenommenen Leistungsdruck und positiven Wohlbefindensaspekten und positive Assoziationen mit negativen Wohlbefindensaspekten (z. B. Kleinkorres et al., 2023a; Stang-Rabrig et al., 2023a). Die Zusammenhänge waren ähnlich für Mädchen und Jungen, bis auf die positive Assoziation zwischen wahrgenommenem elterlichen Leistungsdruck und Schulzufriedenheit, die nur in der Gruppe der Jungen bestand. Da die Relation bivariat negativ und zudem nicht statistisch signifikant ausfiel, sollte diese jedoch nicht überinterpretiert werden. Da Leistungsdruck ein externer Faktor ist, der das Verhalten der Lernenden kontrollieren und deren Autonomie beschränken kann, sind die Zusammenhänge mittels Selbstbestimmungstheorie erklärbar (z. B. Ryan & Deci, 2020).

### 6.1 Limitationen und Stärken

Bei der Ergebnisinterpretation sind Limitationen zu berücksichtigen. Eine ist die Spezifität der Stichprobe zum Beispiel in Bezug auf den mittleren GPA, der vergleichsweise hoch ausfiel (z. B. NCES, 2009). Eine weitere Einschränkung ist, dass nur zwei Leistungsdruckarten untersucht wurden, nicht jedoch der personale Druck. Zudem wurde die Multidimensionalität des Konstrukts Wohlbefinden nicht vollumfänglich berücksichtigt. Allerdings wurden sowohl kognitive als auch affektive sowie positive und negative Aspekte des Wohlbefindens betrachtet, was eine Stärke der Arbeit ist. Eine weitere Stärke ist die Stichprobengröße. Des Weiteren wurde

auf Lernende in einer bedeutsamen, prägenden Lebensphase, zum Ende der Schulzeit, fokussiert und der Kenntnisstand in Bezug auf Wohlbefinden für eine besondere Gruppe an Lernenden in high-achieving schools, die auch als »at-risk group« bezeichnet werden (z. B. Luthar et al., 2020), substanziell erweitert.

## 6.2 Implikationen für Forschung und Praxis

Zukünftige Forschung sollte das Konstrukt Wohlbefinden in seiner vollen Komplexität erfassen. Des Weiteren wäre es von Interesse, auch den Leistungsdruck zu erfassen, den Lernende sich selbst machen (z. B. DuongTran et al., 1996) und die drei Leistungsdruckarten miteinander zu vergleichen. Auch könnten bei leistungsstarken Lernenden Kompositionseffekte betrachtet werden und wie diese mit dem Wohlbefinden zusammenhängen (z. B. Belfi et al., 2012). Für die Praxis bedeuten die Ergebnisse, dass sowohl Eltern als auch Lehrkräfte über die möglichen negativen Effekte von Leistungsdruck auf Wohlbefinden informiert werden sollten. Aufgrund der Relevanz des Wohlbefindens sollten sowohl im außer- als auch schulischen Kontext Situationen geschaffen werden, die dem Wohlbefinden förderlich und zum Beispiel der Befriedigung der psychologischen Grundbedürfnisse dienlich sind (z. B. Ryan & Deci, 2020). Dazu zählt im schulischen Bereich beispielsweise autonomieunterstützendes Lehrkraftverhalten, das positiv mit unterschiedlichen Wohlbefindensaspekten zusammenhängt (z. B. Diseth & Samdal, 2014; Kleinkorres et al., 2023b).

## Literatur

- Aydın, Y. E., Çok, R., Dengiz, G. & Tunç, E. (2022). Investigation of Peer Relationships of Gifted and Non-Gifted High School Students. *Kastamonu Education Journal*, 30(1), 236–250. <https://doi.org/10.24106/kefdergi.685221>
- Batz, C. & Tay, L. (2018). Gender differences in subjective well-being. In E. Diener, S. Oishi & L. Tay (Hrsg.), *Handbook of Well-Being* (S. 358–372). DEF Publishers.
- Baumert, J., Maaz, K., Stanat, P. & Watermann, R. (2009). Schulkomposition oder Institution – was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die Deutsche Schule*, 101, 33–46.
- Belfi, B., Goos, M., De Fraine, B. & Van Damme, J. (2012). The effect of class composition by gender and ability on secondary school students' school well-being and academic self-concept: A literature review. *Educational Research Review*, 7(1), 62–74. <https://doi.org/10.1016/j.edurev.2011.09.002>
- Bergold, S., Wirthwein, L., Rost, D. H. & Steinmayr, R. (2015). Are gifted adolescents more satisfied with their lives than their non-gifted peers? *Frontiers in Psychology*, 151995.
- Berns, R. M. (2015). *Child, family, school, community: Socialization and support*. Cengage Learning.
- Bradburn, N. M. (1969). *The Structure of Psychological Well-Being*. Aldine.

- Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Hrsg.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (S. 793–828). Oxford University Press.
- Bücker, S., Nuraydin, S., Simonsmeier, B. A., Schneider, M. & Luhmann, M. (2018). Subjective well-being and academic achievement: A meta-analysis. *Journal of Research in Personality, 74*, 83–94. <https://doi.org/10.1016/j.jrp.2018.02.007>
- Chui, W. H. & Wong, M. Y. (2016). Gender differences in happiness and life satisfaction among adolescents in Hong Kong: Relationships and self-concept. *Social Indicators Research, 125*, 1035–1051. <https://doi.org/10.1007/s11205-015-0867-z>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Taylor and Francis.
- Cosma, A., Abdrakhmanova, S., Taut, D., Schrijvers, K., Catunda, C. & Schnohr, C. (2023). *A focus on adolescent mental health and wellbeing in Europe, central Asia and Canada. Health behaviour in school-aged children international report from the 2021/2022 survey. Volume 1*. World Health Organization. <https://iris.who.int/handle/10665/373201>
- Daniels, Z. (2008). *Entwicklung schulischer Interessen im Jugendalter*. Waxmann.
- Diener, E., Emmons, R. A., Larsen, R. J. & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71–75.
- Diener, E., Inglehart, R. & Tay, L. (2013). Theory and validity of life satisfaction scales. *Social Indicators Research, 112*, 497–527. <https://doi.org/10.1007/s11205-012-0076-y>
- Diener, E., Suh, E. M., Lucas, R. R. & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*(2), 276–302. <https://doi.org/10.1037/0033-2909.125.2.276>
- Diseth, Å. & Samdal, O. (2014). Autonomy support and achievement goals as predictors of perceived school performance and life satisfaction in the transition between lower and upper secondary school. *Social Psychology of Education, 17*(2), 269–291. <https://doi.org/10.1007/s11218-013-9244-4>
- DuongTran, Q., Lee, S. & Khoi, S. (1996). Ethnic and gender differences in parental expectations and life stress. *Child and Adolescent Social Work Journal, 13*, 515–526. <https://doi.org/10.1007/BF01874304>
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C. & Mac Iver, D. (1993). Development during adolescence: The impact of stage–environment fit on young adolescents’ experiences in school and in families. *American Psychologist, 48*(2), 90–101. <https://doi.org/10.1037/0003-066X.48.2.90>
- Eid, M. & Larsen, R. J. (2008). *The Science of Subjective Well-Being*. Guilford Press.
- Flick, L., Dawes, M., Brian, A., Monsma, E. & De Meester, A. (2022). Relationships among peer-relatedness, self-confidence, peer victimization, social anxiety and school satisfaction in American high school students. *Physical Education and Sport Pedagogy, 1–15*. <https://doi.org/10.1080/17408989.2022.2135692>
- Ganzeboom, H. B. G. & Treiman, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Hrsg.), *Advances in cross-national comparison: A European working book for demographic and socio-economic variables* (S. 159–193). Kluwer Academic Press.

- Gershenson, S. (2018). *Grade inflation in high schools (2005–2016)*. Thomas B. Fordham Institute.
- Godor, B. P. & Szymanski, A. (2017). Sense of belonging or feeling marginalized? Using PISA 2012 to assess the state of academically gifted students within the EU. *High Ability Studies*, 28(2), 181–197. <https://doi.org/10.1080/13598139.2017.1319343>
- Graham, J. W. & Coffman, D. L. (2012). Structural equation modeling with missing data. In R. H. Hoyle (Hrsg.), *Handbook of structural equation modeling* (S. 277–295). Guilford Press.
- Hagenauer, G. (2011). *Lernfreude in der Schule*. Waxmann.
- Hascher, T. & Hagenauer, G. (2020). Swiss adolescents' well-being in school. *Schweizerische Zeitschrift für Bildungswissenschaften*, 42(2), 367–390. <https://doi.org/10.24452/sjer.42.2.5>
- Hascher, T., Hagenauer, G. & Schaffer, A. (2011). Wohlbefinden in der Grundschule. *Erziehung und Unterricht*, 161 (3/4), 381–392.
- Hascher, T., Morinaj, J. & Waber, J. (2018). Schulisches Wohlbefinden: Eine Einführung in Konzept und Forschungsstand. In K. Rathmann & K. Hurrelmann (Hrsg.), *Leistung und Wohlbefinden in der Schule: Herausforderung Inklusion* (S. 66–82). Beltz Juventa.
- Havighurst, R. J. (1953). *Human development and education*. Longmans & Green.
- Herrmann, J., Koeppen, K. & Kessels, U. (2019). Do girls take school too seriously? Investigating gender differences in school burnout from a self-worth perspective. *Learning and Individual Differences*, 69, 150–161. <https://doi.org/10.1016/j.lindif.2018.11.011>
- Howard, J. L., Bureau, J. S., Guay, F., Chong, J. X. & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science*, 16(6), 1300–1323. <https://doi.org/10.1177/1745691620966789>
- Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Huebner, S. (2001). *Manual for the multidimensional students' life satisfaction scale*. Department of Psychology, University of South Carolina, Columbia.
- Kleinkorres, R., Stang, J. & McElvany, N. (2020). A longitudinal analysis of reciprocal relations between students' well-being and academic achievement. *Journal for Educational Research Online*, 12, 114–165. <https://doi.org/10.25656/01:20975>
- Kleinkorres, R., Stang-Rabrig, J. & McElvany, N. (2023a). Comparing parental and school pressure in terms of their relations with students' well-being. *Learning and Individual Differences*, 104, 102288. <https://doi.org/10.1016/j.lindif.2023.102288>
- Kleinkorres, R., Stang-Rabrig, J. & McElvany, N. (2023b). The longitudinal development of students' well-being in adolescence: The role of perceived teacher autonomy support. *Journal of Research on Adolescence*, 1–18. <https://doi.org/10.1111/jora.12821>
- Lee, M. Y., Cho, S., Huy, V. N. & Lee, S. M. (2021). A multilevel analysis of change in emotional exhaustion during high school: Focusing on the individual and contextual factors. *Current Psychology*, 40, 5648–5657. <https://doi.org/10.1007/s12144-020-00869-z>
- Liu, W., Mei, J., Tian, L. & Huebner, E. S. (2016). Age and gender differences in the relation between school-related social support and subjective well-being in school among stu-

- dents. *Social Indicators Research*, 125, 1065–1083. <https://doi.org/10.1007/s11205-015-0873-1>
- Löfstedt, P., García-Moya, I., Corell, M., Paniagua, C., Samdal, O., Välimaa, R., Lyyra, N., Currie, D. & Rasmussen, M. (2020). School satisfaction and school pressure in the WHO European region and North America: An analysis of time trends (2002–2018) and patterns of co-occurrence in 32 countries. *Journal of Adolescent Health*, 66(6), 59–69. <https://doi.org/10.1016/j.jadohealth.2020.03.007>
- Luthar, S. S., Suh, B. C., Ebbert, A. M. & Kumar, N. L. (2020). Students in high-achieving schools: Perils or pressures to be »standouts.« *Adversity and Resilience Science*, 1, 135–147. <https://doi.org/10.1007/s42844-020-00009-3>
- Maslach, C., Jackson, S. E. & Leiter, M. P. (2016). *Maslach Burnout Inventory Manual*. Mind Garden Inc.
- Maslach, C., Shaufeli, W. B. & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, 52, 397–422. <https://doi.org/10.1146/annurev.psych.52.1.397>
- Moksnes, U. K. & Espnes, G. A. (2013). Self-esteem and life satisfaction in adolescents – gender and age as potential moderators. *Quality of Life Research*, 22, 2921–2928. <https://doi.org/10.1007/s11136-013-0427-4>
- Morinaj, J. & Hascher, T. (2019). School alienation and student well-being: A cross-lagged longitudinal analyses. *European Journal of Psychology of Education*, 34, 273–294. <https://doi.org/10.1007/s10212-018-0381-1>
- Muthén, L. K. & Muthén, B. O. (1998–2017). *Mplus – statistical analysis with latent variables. User's guide* (Version 8.). Muthén & Muthén.
- NCES – National Center for Education Statistics. (2009). *America's high school graduates. Results of the 2009 NAEP high school transcript study*. U.S. Department of Education. [https://nces.ed.gov/nationsreportcard/pdf/studies/2011462.pdf?\\_ga=1.239818315.931572072.1438358434](https://nces.ed.gov/nationsreportcard/pdf/studies/2011462.pdf?_ga=1.239818315.931572072.1438358434)
- Niederle, M. & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3), 1067–1101. <https://doi.org/10.1162/qjec.122.3.1067>
- Obermeier, R., Schlesier, J. & Gläser-Zikuda, M. (2022). Differences in students' scholastic well-being induced by familial and scholastic context. *British Journal of Educational Psychology*, 92(3), 994–1010. <https://doi.org/10.1111/bjep.12484>
- OECD – Organisation for Economic Cooperation Development. (2017). *PISA 2015. Results (Volume III): Students' Well-Being*. OECD Publishing.
- Quach, A. S., Epstein, N. B., Riley, P. J., Falconier, M. K. & Fang, X. (2015). Effects of parental warmth and academic pressure on anxiety and depression symptoms in Chinese adolescents. *Journal of Child and Family Studies*, 24, 106–116. <https://doi.org/10.1007/s10826-013-9818-y>
- Rose, A. J. & Rudolph, K. D. (2006). A review of sex differences in peer relationships processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin*, 132(1), 98–131. <https://doi.org/10.1037/0033-2909.132.1.98>

- Ryan, R. M. & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52, 141–166. <https://doi.org/10.1146/annurev.psych.52.1.141>
- Ryan, R. M. & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publishing.
- Ryan, R. M. & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Salmela-Aro, K., Kiuru, N., Leskinen, E. & Nurmi, J. (2009a). School Burnout Inventory (SBI) reliability and validity. *European Journal of Psychological Assessment*, 25, 48–57. <https://doi.org/10.1027/1015-5759.25.1.48>
- Salmela-Aro, K., Kiuru, N. & Nurmi, J.-E. (2008). The role of educational track in adolescents' school burnout: A longitudinal study. *British Journal of Educational Psychology*, 78, 663–689. <https://doi.org/10.1348/000709908X281628>
- Salmela-Aro, K., Savolainen, H. & Holopainen, L. (2009b). Depressive symptoms and school burnout during adolescence: Evidence from two cross-lagged longitudinal studies. *Journal of Youth and Adolescence*, 38, 1316–1327. <https://doi.org/10.1007/s10964-008-9334-3>
- Seligman, M. E. (2011). *Flourish: A Visionary New Understanding of Happiness and Well-Being*. Free Press.
- Stang-Rabrig, J. & Kleinkorres, R. (2023). Multikriteriale Bildungszielerreichung am Ende der vierten Jahrgangsstufe. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 215–229). Waxmann.
- Stang-Rabrig, J., Schwerter, J., Witmer, M. L. & McElvany, N. (2023a). Beneficial and negative factors of the development of students' well-being in educational context. *Current Psychology*, 42, 31294–31308. <https://doi.org/10.1007/s12144-022-04067-x>
- Stang-Rabrig, J., Vogel, S. N. T., Forciniti, M. & McElvany, N. (2023b). Students' school success in challenging times: importance of central personal and social resources during the COVID-19 pandemic. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-023-00739-9>
- Swiatek, M. A. & Lupkowski-Shoplik, A. (2000). Gender differences in academic attitudes among gifted elementary school students. *Journal for the Education of the Gifted*, 23(4), 360–377. <https://doi.org/10.1177/0162353200023004>
- Terry, T. & Huebner, E. S. (1995). The relationship between self-concept and life satisfaction in children. *Social Indicators Research*, 35, 29–52.
- Van Petegem, K., Aelterman, A., Rosseel, Y. & Creemers, B. (2007). Student perception as moderator for student wellbeing. *Social Indicators Research*, 83(3), 447–463. <https://doi.org/10.1007/s11205-006-9055-5>
- WHO – World Health Organization. (2014). *Basic Documents: Forty-eighth Edition*. World Health Organization.
- Yang, Q., Tian, L., Huebner, E. S. & Zhu, X. (2019). Relations among academic achievement, self-esteem, and subjective well-being in school among elementary school students: A

longitudinal mediation model. *School Psychology*, 34(3), 328–340. <https://doi.org/10.1037/spq00.00292>

Zeidner, M. & Schleyer, E. J. (1999). The effects of educational context on individual difference variables, self-perceptions of giftedness, and school attitudes in gifted adolescents. *Journal of Youth and Adolescence*, 28, 687–703. <https://doi.org/10.1023/A:1021687500828>