

Entwicklung digitaler Simulationsspiele mit integrierten Entscheidungsbäumen zur Förderung der diagnostischen Entscheidungskompetenzen in der sonderpädagogischen Lehrkräfteausbildung

Judith Zellner, Nikola Ebenbeck & Markus Gebhardt

Zusammenfassung

Die Anwendung von pädagogischer Diagnostik benötigt sowohl konzeptuelles als auch strategisches Wissen über diagnostische Entscheidungsprozesse. Um diese Entscheidungsprozesse für die Ausbildung im Lehramt Sonderpädagogik der Hochschule möglichst anwendungsbezogen nachzustellen, wurde ein digitales fallbasiertes Simulationsspiel entwickelt. Ziel ist es, angehende Sonderpädagog:innen im Studium zu ökonomischen diagnostischen Entscheidungen im Rahmen pädagogischer Diagnostik zu befähigen und den Aufbau von Handlungsschablonen zu fördern. Diese Studie untersucht mittels Learning Analytics, inwieweit die Simulationsspiele zur Weiterentwicklung der diagnostischen Entscheidungskompetenz von Studierenden ($N = 30$) beitragen. Untersucht wird, wie treffsicher und effizient Studierende aus fünf Testverfahren den jeweils passenden Test auswählen. Der wiederholte Einsatz der Simulationsspiel führt zu signifikanten Veränderungen in der Entscheidungsheuristik. Studierende treffen ihre Entscheidungen nach der Übung mit durchschnittlich 2,1 Tests und einer Treffsicherheit von 41,39%. Die Einsatzmöglichkeiten und der Mehrwert von digitalen Simulationsspielen wird für die Lehrer:innenausbildung diskutiert.

Schlagworte

Diagnostische Kompetenzen, Lehrer:innenbildung, Simulationen, interaktive Simulationsspiele, Sonderpädagogik, Professionalisierung

Title

Development of digital case-based simulation games with integrated decision trees to improve diagnostic decision-making skills in special education teacher training programs

Abstract

It is essential that teachers possess a comprehensive understanding of the conceptual and strategic aspects of the diagnostic decision-making process. A digital case-based simulation game was developed with the objective of enabling prospective special education students to make economic diagnostic decisions in the context of educational diagnostics and to promote the development of action templates. This study employs learning analytics to ascertain the extent to which the simulation games contribute to the further development of the diagnostic decision-making skills of students ($N = 30$). The study examines the accuracy and efficiency with which students select the appropriate test from five test procedures. Repeated use of the simulation game leads to significant changes in decision-making heuristics. Students make their decisions after the intervention with an average of 2.1 tests and an accuracy of 41.39%. The added value of digital simulation games in teacher training is discussed.

Keywords

Diagnostic skills, teacher training, simulations, interactive learning games, special needs education, professionalization

Inhaltsverzeichnis

1. Pädagogische Diagnostik in Sonderpädagogik und inklusiver Pädagogik
 2. Ansprüche an pädagogisches Handeln im Rahmen pädagogischer Diagnostik
 3. Förderung diagnostischer Entscheidungskompetenz in der Hochschullehre
 4. Fragestellungen
 5. Methode
 - 5.1. Instrument & Durchführung
 - 5.2. Stichprobe
 - 5.3. Auswertung
 6. Ergebnisse
 - 6.1. Selbsteinschätzung der diagnostischen Kompetenzen (DaKI)
 - 6.2. Learning Analytics
 7. Diskussion
 - 7.1. Diskrepanz zwischen Selbsteinschätzung und tatsächlichem Übungsverhalten
 - 7.2. Veränderung der Handlungsschablonen durch Simulationsspiele
 - 7.3. Weiterführende Forschung
 8. Limitierung
- Literatur
Kontakt
Zitation

1. Pädagogische Diagnostik in Sonderpädagogik und inklusiver Pädagogik

Pädagogische Diagnostik ist ein Erkenntnisprozess, um pädagogische Entscheidungen zu treffen und zu begründen (Klauer, 1978). In einer inklusiven Schule dient sie nicht primär der Selektion, sondern stellt die Teilhabe und optimale Entwicklung aller Schüler sicher und verbessert didaktische Entscheidungen (Hähn, Häsel-Weide & Scherer, 2021). Pädagogische Diagnostik als Grundlage für individuelle Förderung wird im deutschen Schulsystem aktuell weniger angewendet, da der Diagnostik weniger rechtliche Bedeutung zugesprochen wird, als beispielsweise in den USA (Prince, Yell & Katsiyannis, 2018). Internationale Studien belegen die Wirksamkeit standardisierter, datengestützter Entscheidungen im inklusiven Schulsystem (Lai, McNaughton, Amituanai-Tolosa, Turner & Hsiao, 2009; Schildkamp & Ehren, 2013; Staman, Timmermans & Visscher, 2017), auch bekannt als "databased decision making" (Wayman, Jimerson & Cho, 2012). Pädagogische Diagnostik in inklusiven Schulen zielt darauf ab, datengestützt individuelle Bildungsangebote und pädagogische Unterstützungsmaßnahmen zur Verfügung zu stellen (Heimlich, 2012). Durch eine enge Verknüpfung zwischen Förderung und Diagnostik werden Ressourcen geschont und zugleich eine optimale Passung zwischen den Lernvoraussetzungen der Schüler:innen und den gewählten Interventionen und Unterstützungsmaßnahmen der Lehrkraft gewährleistet. Die Passung von Unterricht und Bedarf der Schüler:innen durch datengestütztes Monitoring erhöht die Wahrscheinlichkeit von Lernfortschritten (Inspectie van het

Onderwijs, 2010). Führt pädagogische Diagnostik hingegen nicht zu konkreten pädagogischen Maßnahmen, verfügen Lehrkräfte letztlich über geringere Ressourcen für die Förderung und Unterstützung ihrer Schüler:innen (Beckstein & Sroka, 2022). [1]

2. Ansprüche an pädagogisches Handeln im Rahmen pädagogischer Diagnostik

Alle Lehrämter benötigen insbesondere im Bereich der pädagogischen Diagnostik sonderpädagogische Basiskompetenzen (Kultusministerkonferenz & Hochschulrektorenkonferenz, 2015). So zeigen Urhahne und Wijnia (2021) in ihrem Review, dass die Urteilsgenauigkeit von Lehrkräften bei der Beurteilung der Kompetenzen der Schüler:innen variiert. Die 16 untersuchten Studien ergaben, dass akademischen Leistungen akkurater eingeschätzt werden als die Lernmotivation der Schüler:innen. Bei standardisierten Tests tendieren Lehrkräfte dazu, die Leistungen zu überschätzen und zugleich die Schwierigkeit der Testaufgabe zu unterschätzen. Die Passung von Lernaufgabe und Lernstand der Schüler ist jedoch maßgeblich für den Lernerfolg aller Schüler:innen. Daher ist die Schulung der diagnostischen Kompetenz für alle Lehrkräfte von hoher Relevanz und Bestandteil der Lehramtsausbildung sein. Im Studium für allgemeine Schulen wird die pädagogische Diagnostik in wenigen Sitzungen in Seminaren der Erziehungswissenschaft oder der Fachdidaktik behandelt, während sie im Studium für Sonderpädagogik meist in mehreren Veranstaltungen auf den jeweiligen Förderschwerpunkt bezogen fokussiert wird (Gebhardt, Ebenbeck, Jungjohann, Koutsianikouli & Lutz, 2023). Schröter, Schulze und Kuhl (2023) stellten sich die Frage, welche Bildungsinhalte den Kern der Sonderpädagogik als lehrkräftebildendes Fach der Hochschullehre darstellen. In den Modulplänen diverser deutscher Lehramtsstudiengänge bilden Diagnostik und Unterricht dabei den Schwerpunkt, wobei Unterrichten synonym zu Förderung verstanden wird (Schröter et al., 2023). Pädagogische Diagnostik und passgenaue Förderung gilt es auch im Lehramt für allgemeinbildende Schulen zu verankern, um die Lernprozesse aller Schüler:innen effektiv zu unterstützen. [2]

Professioneller Umgang mit erhobenen diagnostischen Daten und den daraus abzuleitenden Informationen erfordert von Lehrkräften neben inhaltlichem Wissen auch praktische Übung, um standardisierte Tests durchzuführen, die gewonnenen Daten auszuwerten, zu evaluieren und aus den Ergebnissen datenbasierte Fördermaßnahmen abzuleiten (Gebhardt, 2023). Die Herausforderungen von Durchführung und Auswertung liegen zumeist in der effizienten und ressourcenschonenden Umsetzung im Schulalltag (Kingston & Nash, 2011; Moser Opitz, Pool Maag & Labhart, 2019). Schwierigkeiten treten beispielsweise bei der Auswahl geeigneter Tests, der Ableitung von Informationen aus den Ergebnissen und der Auswahl passender Fördermaßnahmen auf (Espin, Saab, Pat-El, Boender & van der Veen, 2018). Um Lehrkräfte bei der Durchführung und Auswertung zu unterstützen, existieren bereits viele praxisorientierte Lösungsansätze. Digitale Testinstrumente helfen bei der Durchführung, Auswertung und Dokumentation von Tests (Gebhardt, Diehl & Mühling, 2016). Automatische Testdurchführungen und -auswertungen führen zu einer höheren psychometrischen Qualität, da so eine bessere Objektivität, Ökonomie und Benutzerfreundlichkeit für Lehrkräfte sowie Testfairness für Schüler:innen erreicht werden kann. Zusätzlich werden Übertragungs-, Rechen- oder Auswertungsfehler reduziert und die Auswertung lässt sich im schulischen Alltag deutlich beschleunigen (Liebers, Kanold & Junger, 2019). Digitale Instrumente können Lehrkräfte jedoch bisher nicht dabei unterstützen, notwendige diagnostische Entscheidungen zu treffen. Diese pädagogischen Handlungskompetenzen zu entwickeln und zu schulen ist Zweck des Lehramtsstudiums (Capovilla, Falkenstörfer & Gebhardt, 2023). [3]

Data Literacy, die Fähigkeiten Daten auf kritische Art und Weise zu sammeln, zu managen und zu bewerten, bildet eine entscheidende Rolle, um Daten zu verstehen, Ergebnisse sachlich zu interpretieren und auf Basis dessen schlussendlich datengestützte Förderentscheidungen zu treffen (Raffaghelli, 2019). Diagnostische Entscheidungen können dann getroffen werden, wenn Data Literacy planvoll mit didaktischem Wissen verknüpft wird, Handlungsabläufe eingeübt werden und Lehrkräfte mit positiver Wirksamkeitserwartung an Entscheidungsfindungsaufgaben herangehen (Goldberg, Schwerter, Seidel, Müller & Stürmer, 2021; Savolainen, Malinen & Schwab, 2022). Diagnostisches Wissen auf inhaltlicher Ebene wird anschließend in Handlungen

übersetzt und in der Praxis angewendet. Die Schulung diagnostischer Entscheidungskompetenz ist dabei häufig an praktische Phasen der Lehramtsausbildung geknüpft (Gebhardt et al., 2023). [4]

3. Förderung diagnostischer Entscheidungskompetenz in der Hochschullehre

In der medizinischen Ausbildung werden computerbasierte Fallsimulationen bereits seit Ende der 1990er Jahre eingesetzt, um praxisnahe Erfahrungen in einem geschützten, deterministischen Rahmen zu vermitteln und diagnostische Kompetenzen im medizinischen Bereich zu schulen (Fischer, 2000). Simulationen ermöglichen sowohl die authentische Darstellung relevanter Probleme aus der Berufspraxis als auch interaktive Bearbeitungsformen praktischer Herausforderungen (Schulz et al., 2017). Dabei können situationsspezifische Faktoren, die sich auf andere Personen außerhalb der lernenden Person beziehen, reguliert werden. Ziel der Bearbeitung von Entscheidungssimulationen liegt in der Anwendung und Vernetzung deklarativen Wissens in komplexen beruflichen Situationen durch die Bearbeitung realitätsnaher Problemstellungen, ohne dass die getroffenen Entscheidungen echte Folgen nach sich ziehen. Die Lernenden beeinflussen in spielbasierten Simulationen die vorliegende Situation, indem sie ihre Handlungen innerhalb des geschützten Rahmens schulen. [5]

Zur Förderung diagnostischer Kompetenzen haben sich Simulationen als wirksame Lernumgebung erwiesen (Chernikova et al., 2020; Heitzmann et al., 2019; Urhahne & Wijnia, 2021). Simulationsbasierte Lernumgebungen können die Arbeit mit Video-Vignetten (Sommerhoff, Codreanu, Nickl, Ufer & Seidel, 2023), face-to-face-Rollenspiele oder digitale Fallbearbeitungen umfassen (Bauer et al., 2022). Heitzmann et al. (2019) entwickelten ein Rahmenmodell zur Förderung diagnostischer Kompetenzen durch Simulationen. Kompetenzen werden dabei als Disposition der diagnostizierenden Personen verstanden. Diese Dispositionen umfassen diagnostisches Wissen, diagnostische Qualität und diagnostische Aktivitäten. Diagnostisches Wissen beinhaltet konzeptionelles und strategisches Wissen, wobei konzeptuelles Wissen Konzepte und deren Beziehungen umfasst und strategisches Wissen mögliche Wege und Heuristiken in der Diagnostik beschreibt. Diagnostische Qualität bezieht sich auf Genauigkeit und Effizienz, die als wichtige Variablen dienen können. Diagnostische Aktivitäten umfassen Handlungen während des diagnostischen Prozesses, wie die Generierung von Beweisen durch Fragen bei der Anamnese. Der Kontext ist ein wichtiger Moderator bei der Beurteilung, weshalb weitere Forschung über diagnostisches Verhalten in spezifischen Kontexten, wie Simulationsumgebungen, gerechtfertigt ist. Eine Meta-Analyse zum simulationsbasierten Lernen komplexer Fertigkeiten hebt die Bedeutung der Authentizität als Moderator hervor (Chernikova et al., 2020). Eine weitere Meta-Analyse zeigt, dass bestimmte Arten der kognitiven Belastung die Leistung beeinträchtigen können (Cook et al., 2012). Blömeke, Gustafsson und Shavelson (2015) nehmen an, dass die Übersetzung einer Kompetenz in eine beobachtbare Handlung sowohl von persönlichen kognitiven und motivationalen Dispositionen als auch von der spezifischen im vorliegenden Fall diagnostischen Situation abhängt. Demzufolge besteht ein Zusammenhang zwischen der Selbsteinschätzung der diagnostischen Kompetenz und dem diagnostischen Verhalten. [6]

Gemäß einer weit verbreiteten Vorstellung werden Entscheidungen umso besser, je mehr Informationen bei der Entscheidungsfindung berücksichtigt werden (Reimer, Hoffrage & Katsikopoulos, 2007). Der Mensch möchte zunächst so viele Informationen wie möglich einholen, bevor eine Entscheidung getroffen wird – auch, wenn das nicht die Treffsicherheit erhöht. Dieses Vorgehen entspricht der menschlichen Intuition, der menschlichen Unsicherheit. Entscheidungen erfolgen dabei entweder ausschließlich intuitiv, nach verinnerlichter Handlungsschablonen oder algorithmisch (Marx & Lange, 2019). Für diagnostische Entscheidungen eignet sich die Vorgehensweise entlang verinnerlichter Handlungsschablonen, wobei man mittels mentaler, auf Erfahrung beruhender Strategien innerhalb kürzester Zeit mit begrenztem Wissen zu tragfähigen Urteilen gelangen kann (Reimer et al., 2007). Untersuchungen von Gigerenzer und Gaissmaier (2011) sowie von Hoffrage und Reimer (2004) zeigten, dass einfache Heuristiken zu sehr guten Entscheidungen führen können. Diese Heuristiken erlauben schnelle Entscheidungen und sind frugal, benötigen also keine umfassende Informationsbasis. Von sicherer Ent-

scheidungsheuristik und einer gefestigten Handlungsschablone kann hinsichtlich der Testwahl im diagnostischen Entscheidungsprozess dann gesprochen werden, wenn zielgerichtet getestet wird, ohne die Entscheidung vorab durch eine umfassende Informationsbasis abzusichern (Reimer et al., 2007). [7]

Bäume als Datenstruktur sind ein fundamentales Verfahren, das effiziente Entscheidungen und Diagnostik ermöglicht und behavioristische Veränderungen im diagnostischen Entscheidungsprozess herbeiführen kann. Die Struktur von Entscheidungsbäumen in Lernumgebungen bildet eine hierarchische Anordnung von Entscheidungsknoten ab, die verschiedene Lernpfade repräsentieren. Jeder Knoten repräsentiert eine Entscheidung oder ein Kriterium, während die Verzweigungen zu neuen Knoten alternative Wege darstellen. Diese Struktur ermöglicht es komplexe Gedankenabläufe und Entscheidungsprozesse zu strukturieren und individuelle Lernpfade sichtbar zu machen. Denkbahnen werden transparent und ermöglichen die Identifikation fehlerhafter Gedankenvorgänge. Digitale Simulationsspiele in Form von Entscheidungsbäumen führen die Lernenden spielbasiert Schritt für Schritt durch den interaktiven Entscheidungsprozess. Aus den Logdaten lassen sich Informationen über das Lern- und Bearbeitungsverhalten der Teilnehmenden generieren. Im sonderpädagogischen Bereich fand das simulationsbasierte Vorgehen auf Basis eines Entscheidungsbaums bisher keinen systematischen Einsatz, obwohl es aufgrund seiner deterministischen und engmaschigen Struktur von Entscheidungsbäumen einen vielversprechenden Ansatz zur Förderung diagnostischer Kompetenzen darstellt. Auch bleibt die Frage offen, wie angehende Lehrkräfte sich in einer solchen simulationsbasierten Lernumgebung bewegen und lernen. [8]

4. Fragestellungen

Digitale Simulationsspiele ermöglichen es, ein tieferes konzeptionelles und strategisches Wissen zu erwerben. Im sonderpädagogischen Lehramtsstudium werden Wissen und Kompetenzen über die Erhebung der Lernausgangslage der Schüler:innen und das Treffen von daraus resultierenden pädagogischen Förderentscheidungen vermittelt. In digitalen Simulationsspielen mit Fallbeispielen und Fragestellungen sowie Entscheidungs- und Antwortmöglichkeiten wird in Anlehnung an das Rahmenmodell von Heitzmann et al. (2019) untersucht, inwiefern die Studierenden gefestigte Handlungsschablonen zur pädagogischen Diagnostik erworben haben. Im Rahmen der Studie wird zunächst die Eigenbewertung der Lernausgangslage der Teilnehmenden anhand eines Fragebogens zur Selbsteinschätzung der diagnostischen Kompetenzen in einem Seminar zur Förderdiagnostik und Förderplanung als Praktikumsbegleitung erfasst. Es wird davon ausgegangen, dass Teilnehmende mit hoher Selbsteinschätzung selbstbewusst an die nachfolgende Übung herangehen und sichere Entscheidungen treffen. Die Teilnehmenden üben im Seminar und wenden datengestützte diagnostische und pädagogische Entscheidungen in Simulationen an, um die von Heitzmann et al. (2019) formulierten Dispositionen diagnostisches Wissen, diagnostische Qualität und diagnostische Aktivität zu vertiefen. [9]

Es werden folgende Forschungsfragen verfolgt: [10]

1. Mit welcher Selbstwahrnehmung ihres diagnostischen Wissens zu Beginn des Seminars nehmen die Studierenden an den Simulationsspielen teil? [11]
2. Welche diagnostische Qualität lässt sich an den Handlungen und Entscheidungen der Studierenden im datengestützten Entscheidungsprozess der digitalen entscheidungsbaumbasierten Simulationsspiele erkennen? [12]
3. Lassen sich durch die wiederholt eingesetzten Simulationsspiele Veränderungen der diagnostischen Aktivität und Handlungsabläufe einer diagnostischen Entscheidungsfindung feststellen? [13]

5. Methode

Die Studie wurde im Rahmen eines praktikumsbegleitenden Seminars durchgeführt, welches die Erstellung und Durchführung eines sonderpädagogischen Förderplans vermittelt. Das verwendete Instrument, Datenset und die Syntax der vorliegenden Studie können unter

https://osf.io/wrgkb/?view_only=c25dcd92391c47739fd05af467a7c592 gefunden werden. Die Datenanalyse fand mit R und der R Studio Umgebung (R Core Team, 2018) statt. [14]

5.1. Instrument & Durchführung

In der Studie werden neu entwickelte fallbasierten Simulationsspiele auf Basis eines Entscheidungsbaums sowie der *Fragebogen zur Selbsteinschätzung der diagnostischen Kompetenz von Lehrkräften in der inklusiven Schule (DaKI, Version 0.2)* eingesetzt. DaKI misst die vier Dimensionen (1) fachdidaktische Diagnostikkompetenzen, (2) Schulleistungs- und Statusdiagnostik, (3) sonderpädagogische Feststell- und Förderdiagnostik und (4) Lernverlaufdiagnostik anhand einer fünfstufigen Likert-Skala (Jungjohann & Gebhardt, 2023). Der Fragebogen dient der Selbsteinschätzung des diagnostischen Wissens. Die individuellen Skalen weisen eine gute interne Kohärenz auf ($\alpha = .92, .82, .92, .93$). Die Werte der Vergleichsstichprobe des Fragebogens aus $N = 252$ Lehrkräften und werdenden Lehrkräften sind alle in der oberen Hälfte der Skala zu verorten, wobei Lehrkräfte sich insgesamt besser einschätzen als Studierende (Jungjohann & Gebhardt, 2023). [15]

Die Studierenden der vorliegenden Studie bearbeiteten den Fragebogen einmalig am Anfang des Semesters als Bestandteil des Seminars individuell durchgeführt und anschließend mit den Studierenden besprochen. Zur Einschätzung der Zuverlässigkeit der Ergebnisse wurde die Trennschärfe der einzelnen Items berechnet. Der Fragebogen dient dazu, vorab einschätzen zu können, wie selbstbewusst die Studierenden an die Simulationsspiele herangehen. Wir gehen davon aus, dass die Studierenden bei hoher Selbsteinschätzung ihrer diagnostischen Kompetenzen, selbstbewusst handeln und sichere Entscheidungen treffen. Der Fragebogen ist kein geeignetes Instrument, um einen Lernzuwachs zu messen, da es sich um eine Selbsteinschätzung der Teilnehmenden handelt. [16]

Die neu entwickelten digitalen Simulationsspiele stellen den diagnostischen Entscheidungsprozess anhand fiktiver Fallbeispiele von Schüler:innen mit Lernschwierigkeiten im Bereich Lesen und Rechtschreiben dar. Ziel ist es, Studierenden eine einfach strukturierte und nutzerfreundliche Lernumgebung bereitzustellen, die sie spielerisch Handlungsabläufe der diagnostischen Entscheidungsfindung üben lässt. Um die Entscheidungsfindung in das Spiel einzubauen, wurde der H5P-Inhaltstyp „Branching Scenario“ (H5P Group, 2019) verwendet, welcher es ermöglicht Entscheidungsbäume in digitale Lernumgebungen zu integrieren und Open Source zur Verfügung zu stellen. Jede Antwortmöglichkeit ist mit einer Zahl hinterlegt. Somit kann das Klick- und Entscheidungsverhalten der Studierenden genau nachvollzogen werden. Die Bearbeitung eines Simulationsspiels nimmt durchschnittlich 5-10 Minuten in Anspruch. Der Aufbau der Fallspiele bleibt über alle Fälle hinweg identisch und wird folgend erklärt. [17]

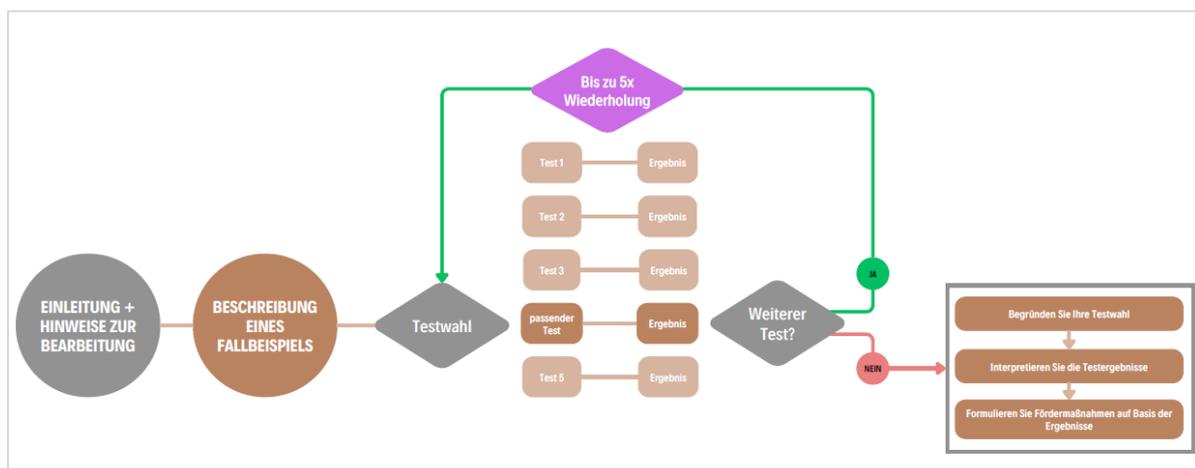


Abbildung 1: Simulationsspiel als Entscheidungsbaum in H5P

Am Anfang des Simulationsspiels bekommen die Studierenden von einer fiktiven Lehrkraft Informationen zu den jeweiligen Schüler:innen des Falls und werden instruiert, eine oder mehrere

diagnostische Verfahren auszuwählen, um möglichst effizient den Lernstand zu erheben und geeignete Förderempfehlungen geben zu können. Es folgt eine detaillierte Fallbeschreibung, bestehend aus Informationen zum allgemeinen Lernverhalten, zur vorliegenden Lernschwierigkeit und zur Motorik oder Wahrnehmung als Distraktoren. Die Studierenden können fünf verschiedene Tests auswählen und erhalten anschließend deren Testergebnisse. Es können maximal fünf Klicks getätigt werden. In jedem Simulationsspiel stehen dieselben fünf Testarten mit jeweils verändertem Wortlaut zur Verfügung: Ein passender standardisierter Test entsprechend der im Fallbeispiel beschriebenen Lese-/Rechtschreibschwierigkeit, standardisierter Test der Vorläuferfähigkeiten, eine informelle Überprüfung der Schwierigkeit (z. B. durch Mithören von Gesprächen), ein standardisierter Test eines Entwicklungsbereichs und standardisierter Intelligenztest. Die Formulierungen der Fallbeschreibungen lassen auf lediglich einen Test schließen, der konkret die dargestellte Schwierigkeit des Kindes überprüft. Alle weiteren Tests sind für die Diagnostik der beschriebenen Schwierigkeiten nicht aussagekräftig und können anhand der Beschreibung ausgeschlossen werden. Drei offene Items schließen jedes Simulationsspiel ab, indem die Studierenden aufgefordert werden, ihre Testwahl zu begründen, die Testergebnisse zu interpretieren und entsprechende Fördermöglichkeiten abzuleiten. Alle Simulationsspiele unterliegen dieser gleichbleibenden Struktur und unterscheiden sich lediglich im verwendeten Wortlaut. [18]

Beispielsätze aus Spiel 1: Die Lehrkraft berichtet, dass Lisa im Unterricht beim Abschreiben von der Tafel viele „Flüchtigkeitsfehler“ unterlaufen. Das Lesen gelingt nur stockend, oft „fantasiert“ sie Wörter in den Text, die da nicht stehen. Sie verwechselt beim Lesen oftmals „ei“ und „ie“. Lisas Mutter übt viel mit ihr, besonders vor Diktaten, aber dazu ist sie immer weniger zu bewegen. [19]

Beispielsätze aus Spiel 2: Die Lehrkraft beschreibt, dass Mirco nur stockend liest. Dabei erliest er Wörter durch Zusammenschleifen von Lauten. Er kann die Leserichtung konsequent einhalten. Das Gliedern von Wörtern in Silben sei ihm dabei nicht möglich und häufig „verschluckt“ er wohl die letzten Laute des Wortes. Manchmal gelingt es Mirco, kurze, bekannte Ganzwörter zu erkennen. [20]

Zwischen den beiden Simulationsspielen erhalten die Teilnehmenden mündliches Feedback und diskutieren das Vorgehen im Seminar, um ihr Bearbeitungsverhalten zu optimieren. Die Inhalte, die zwischen dem ersten und zweiten Simulationsspiel behandelt wurden, orientieren sich am Klickverhalten der Teilnehmenden des ersten Simulationsspiels. Im vorliegenden Fall wurde zum einen die Notwendigkeit eines effizienten Entscheidungsverhaltens thematisiert. Zum anderen wurde ergänzend die Relevanz und Gefahren von Intelligenztestungen diskutiert. Zusätzlich zur Reflexion ihres Bearbeitungsverhaltens erhielten die Studierenden die Option, an weiteren vergleichbaren Simulationsspielen ihre diagnostischen Entscheidungskompetenzen zu üben. Diese Übungen waren freiwillig und die Teilnehmerzahl variierte. Aus diesem Grund wurden nur das erste und letzte verpflichtende Simulationsspiel zur Analyse herangezogen. [21]

5.2. Stichprobe

Studierende der Sonderpädagogik ($N = 30$) besuchten einen Kurs zur Praktikumsbegleitung, der im fünften Studiensemester verankert ist (5 männlich, 25 weiblich; $M_{\text{Alter}} = 25.75$, $SD_{\text{Alter}} = 8.99$). Vorab besuchten die Studierenden mehrere Vorlesungen und Seminare zum Thema pädagogische Diagnostik. Im Praktikum erproben die Studierenden ihre diagnostischen Kompetenzen praktisch, indem sie in Zweiergruppen über den Zeitraum eines halben Jahres zwei bis sechs Schüler:innen regelmäßig fördern und diagnostizieren. In der ersten Seminarsitzung beantworteten alle Studierenden den Fragebogen DaKI, gefolgt von sechs Simulationsspielen. Lediglich das erste und letzte Simulationsspiel war von allen Studierenden zu bearbeiten. Die gesamte Übung an Simulationsspielen wurde während der Seminarzeit durchgeführt und erstreckte sich über sechs Wochen von November 2023 bis Januar 2024. [22]

5.3. Auswertung

Learning Analytics (LA) verfolgt das übergeordnete Ziel zu verstehen, wie Lernende lernen (Viberg, Hatakka, Bälter & Mavroudi, 2018). Zwar existiert keine allgemein anerkannte Definition von LA, häufig wird LA jedoch als das Messen, Sammeln und Analysieren von Daten über Lernprozesse, um diese zu verstehen und Lernumgebungen zu optimieren, zusammengefasst (Siemens & Long, 2011). Der Lernprozess und das Lernverhalten der Lernenden steht im Fokus (Ifenthaler & Schumacher, 2016). Diese Methode bietet sich für die Auswertung von digitalen Lernumgebungen an, welche automatisch verschiedene Informationen über die Programmbe- arbeitung der Nutzenden erfassen und so vielfältige Möglichkeiten zur Analyse von Lernprozessen ermöglichen. LA wird insbesondere bei der Analyse von Entscheidungsbäumen empfohlen, da hier die erhobenen Daten ausschlaggebend für das Auswertungsdesign und mitunter individuell und komplex sind (Agarwal, Pandey & Tiwari, 2012; Nisbet, Elder & Miner, 2009). [23]

In dieser Studie analysieren wir das Bearbeitungsverhalten der Lernenden während der neu entwickelten Simulationsspielen zur Förderung der diagnostischen Entscheidungskompetenz. Die Ergebnisse des DaKI vor Beginn des Einsatzes der Simulationsspiele werden auf Skalen- und Itemebene deskriptiv analysiert und als Lernausgangslage der Selbsteinschätzung des diagnostischen (konzeptuellen und strategischen) Wissens in die Analysen miteinbezogen. Die Übung mittels der neu entwickelten digitalen Simulationsspiele werden mit Hilfe von LA ausgewertet. Die generierten Rohdaten liegen als .csv-Datei vor, wobei jeder getätigte Klick als numerischer Wert notiert wird. Die offenen Antworten liegen als Textantworten vor. Zur Beantwortung der Fragestellung zwei und drei werden die diagnostische Aktivität und Qualität anhand der Genauigkeit und Effizienz als wichtige Ergebnisgrößen über die verschiedenen Handlungsschritte hinweg gemessen (Heitzmann et al., 2019). Dabei werden die folgenden Variablen herausgearbeitet und für beide Messzeitpunkte der Simulationsspiele im Vergleich betrachtet: [24]

- Die absolute Anzahl der herangezogenen (= angeklickten) Tests, die jede:r Studierende bei der Bearbeitung des Fallbeispiels für ihren/seinen Entscheidungsprozess genutzt hat. Die Signifikanz der Unterschiede wird mittels t-Test geprüft. [25]
- Der prozentuale Anteil der Studierenden, die den passenden Test durch Anklicken und durch weiteren Einbezug des Testergebnisses in die anschließende Interpretation in Ihren Entscheidungsprozess miteinbezogen haben. [26]
- Die Trefferquote der Studierenden, die sich aus dem prozentualen Anteil des Klicks auf den passenden Test in Bezug auf die absolute Anzahl aller angeklickten Tests pro Simulationsspiel ergibt. [27]
- Der Zeitpunkt, zu dem der passende Test von den Studierenden ausgewählt wird. Es wird angenommen, dass die Studierenden mit dem ersten Klick den Test anklicken, den sie als am treffendsten vermuten. Zudem wird angenommen, dass der Test beendet wird, nachdem die Studierenden ein Ergebnis als ausreichend informativ für die diagnostische Entscheidung wahrnehmen. [28]

6. Ergebnisse

6.1. Selbsteinschätzung der diagnostischen Kompetenzen (DaKI)

Die Studierenden schätzten ihre diagnostischen Kompetenzen leicht negativ bis neutral unter der mittleren Kategorie „stimme teilweise zu“ ein. So waren die Mittelwerte in der fachdidaktische Diagnostikkompetenz $M = 2.91$ ($SD = 0.58$), Schulleistungs- und Statusdiagnostik $M = 2.82$ ($SD = 0.76$), Sonderpädagogische Feststell- und Förderdiagnostik $M = 2.47$ ($SD = 0.64$) und Lernverlaufdiagnostik $M = 2.40$ ($SD 0.83$). Die Ergebnisse der einzelnen Items sind in Tabelle 1 dargestellt. [29]

No	Items	M (SD)	Trennschärfe
	Dimension 1: Fachdidaktische Diagnostikkompetenz	2.91 (0.58)	
1	Ich kann Schüler: innen mit Schwierigkeiten im Lernen identifizieren.	3.80 (0.63)	.35
2	Ich kenne Modelle zur Förderung schriftsprachlicher oder mathematischer Kompetenzen.	3.14 (0.94)	.21
3	Ich kann eine geeignete Förderung für ein Kind auswählen.	2.91 (0.82)	.63
4	Ich kann aus Modellen zum Schriftspracherwerb oder zum Erwerb mathematischer Kompetenzen eine Förderung ableiten.	2.66 (0.75)	.59
5	Ich kann die Passung einer fachspezifischen Fördermethode zu einem Kind einschätzen und begründen.	2.80 (0.93)	.64
6	Ich kann individuelle Materialien für eine gezielte Förderung selbstständig entwickeln.	2.54 (0.92)	.70
7	Ich kann die schriftsprachlichen oder mathematischen Kompetenzen eines Kindes für das Zeugnis einschätzen.	2.74 (1.15)	.73
	Dimension 2: Schulleistungs- und Statusdiagnostik	2.82 (0.76)	
8	Ich kenne mehrere informelle und standardisierte Schulleistungstests.	2.71 (1.13)	.62
9	Ich kenne Normwerte und ihre Bedeutung in der Diagnostik.	3.43 (0.85)	.68
10	Ich wähle ein diagnostisches Verfahren anhand der Gütekriterien aus.	3.23 (1.11)	.70
11	Ich kann informelle und standardisierte Schulleistungstests mit Schüler:innen durchführen und auswerten.	2.89 (1.11)	.67
12	Ich kann Ergebnisse aus standardisierten Leistungstests interpretieren und konkrete Förderziele ableiten.	3.03 (0.95)	.63
13	Ich kann informelle Schulleistungstests (z. B. Test zum Leseverständnis oder zur Erweiterung des Zahlenraums) selbst entwickeln.	2.54 (0.95)	.88
14	Wenn für einen speziellen Fall kein passendes standardisiertes Verfahren verfügbar ist, kann ich ein standardisiertes Verfahren für diese Zweck anpassen und nutzen (z. B. Auswahl von Teilaufgaben).	2.14 (0.85)	.61
	Dimension 3: Sonderpädagogische Feststell- und Förderdiagnostik	2.47 (0.64)	
15	Ich habe viel Erfahrung im Schreiben von sonderpädagogischen Gutachten.	2.46 (1.04)	.87
16	Ich kenne die notwendigen Schritte zur Feststellung des Bedarfs an sonderpädagogischer Unterstützung.	1.26 (0.56)	.33
17	Ich kann Verfahren zur Feststellung eines Bedarfs an sonderpädagogischer Unterstützung durchführen.	2.49 (1.07)	.41
18	Ich kann sonderpädagogische Förderziele formulieren. Ich kenne verschiedene Förderorte (sich inklusive Förderorte) und kann über ihre Vor- und Nachteile beraten.	2.09 (1.01)	.56
19	Ich kann die Erreichung eines sonderpädagogischen Förderziels überprüfen.	2.77 (1.03)	.65
	Dimension 4: Lernverlaufsdiagnostik	2.40 (0.83)	
20	Ich kenne mehrere Verfahren zur Begleitung von Lernprozessen (Formative Diagnostik)	3.63 (0.91)	.80
21	Ich kenne die theoretischen Hintergründe und Konzepte von Lernverlaufsdiagnostik.	2.60 (0.88)	.73
22	Ich kann eine Lernverlaufsdiagnostik durchführen und auswerten.	2.29 (0.93)	.54
23	Ich kann mithilfe von Lernverlaufsgraphen und Ergebnissen der Lernverlaufsdiagnostik den Fortschritt einer Förderung interpretieren.	2.71 (1.10)	.83
24	Ich kann anhand der Daten der Lernverlaufsdiagnostik didaktische Entscheidungen treffen und begründen.	2.40 (1.00)	.83
25	Ich kann eine informelle Lernverlaufsdiagnostik entwickeln.	1.94 (0.97)	.84
26	Ich kann die Effektivität einer Förderung überprüfen.	2.69 (0.93)	.64

Tabelle 1: Ergebnisse der Selbsteinschätzung der diagnostischen Kompetenzen

Fünf Items weisen einen $M > 3$ auf und werden somit als überdurchschnittlich betrachtet (Item 1: Ich kann Schüler:innen mit Schwierigkeiten im Lernen identifizieren. Item 2: Ich kenne Modelle zur Förderung schriftsprachlicher oder mathematischer Kompetenzen. Item 9: Ich kenne Normwerte und ihre Bedeutung in der Diagnostik. Item 10: Ich wähle ein diagnostisches Verfahren anhand der Gütekriterien aus. Item 20: Ich kenne mehrere Verfahren zur Begleitung von Lernprozessen (Formative Diagnostik).). Alle weiteren Ergebnisse liegen bei $M < 3$. Zwei Items weisen einen $M < 2$ auf (Item 16: Ich kenne die notwendigen Schritte zur Feststellung des Bedarfs an sonderpädagogischer Unterstützung. Item 25: Ich kann eine informelle Lernverlaufsdagnostik entwickeln.). Diese erfragen Transfer- und Handlungswissen der pädagogischen Diagnostik. [30]

6.2. Learning Analytics

Die Anzahl der Klicks, die Studierende verwendeten, um sich Tests anzusehen, hat sich im Durchschnitt von Simulationsspiel 1 zu Simulationsspiel 2 verringert ($p < .001$, Abbildung 2). In Simulationsspiel 1 werden durchschnittlich 4.33 ($Min = 3$, $Max = 5$, $SD = 0.76$) von 5 möglichen Tests angeklickt. Alle Studierenden wählen dabei mindestens drei Tests aus. In Simulationsspiel 2 ziehen die meisten Teilnehmenden drei oder weniger Tests heran ($M = 2.07$, $Min = 1$, $Max = 5$, $SD = 0.98$). Dennoch gibt es weiterhin Studierende, die alle Tests anklicken. [31]

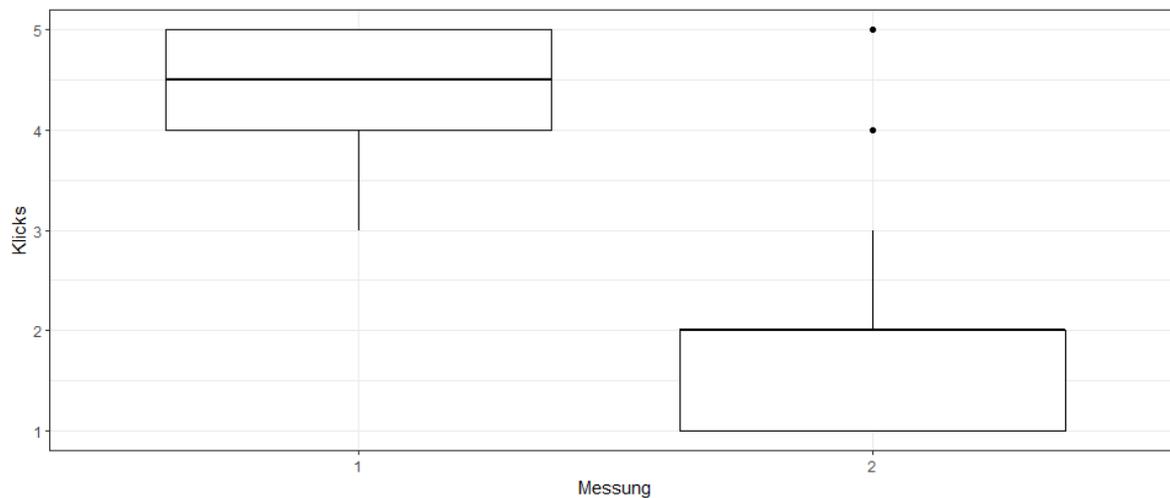


Abbildung 2: Absolute Anzahl der Klicks in Simulationsspiel 1 und Simulationsspiel 2

Die Trefferquote der Studierenden, die sich aus dem prozentualen Anteil des Klicks auf den passenden Test in Bezug auf die absolute Anzahl aller Klicks pro Simulationsspiel ergibt, ist bei Simulationsspiel 2 im Vergleich zu Simulationsspiel 1 deutlich gestiegen (Abbildung 3). Während sich die prozentuale Trefferquote bei Simulationsspiel 1 durchschnittlich bei 17,56% ($Min = 0\%$, $Max = 33,33\%$) bewegt, weist Simulationsspiel 2 eine höhere Trefferquote von durchschnittlich 42,5% ($Min = 0\%$, $Max = 100\%$) auf. 10 Studierende haben eine Trefferquote von 100%, da sie mit nur einem Klick direkt den passenden Test ausgewählt haben. Auch bei Simulationsspiel 2 gab es weiterhin Personen, die nicht den passenden Test ausgewählt haben. Das Simulationsspiel 2 weist dennoch im Vergleich zu Simulationsspiel 1 eine deutlich höhere Trefferquote bei zugleich durchschnittlich deutlich reduzierter Anzahl herangezogener Tests auf. [32]

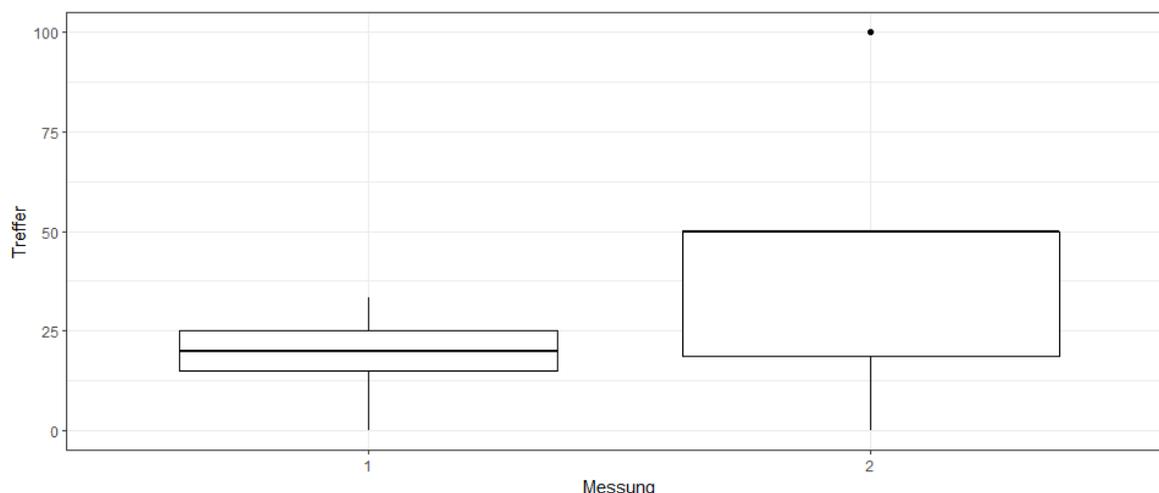


Abbildung 3: Veränderung der prozentualen Trefferquote in Simulationsspiel 1 und Simulationsspiel 2

Die Zusammensetzung aller Tests, die von den einzelnen Studierenden für die Entscheidungsprozesse verwendet wurden, verändert sich deutlich. Bereits in Simulationsspiel 1 wählen 80% der Studierenden den passenden Test aus. 60% aller Studierenden und damit 75% derjenigen, die den passenden Test angeklickt und die Ergebnisse sahen, berücksichtigten diese in der Interpretation. Bei Simulationsspiel 1 wurden mehr Distraktoren angeklickt im Vergleich zum passenden Test. Davon wurde zu 86,67% ein Test zur Überprüfung der Vorläuferfähigkeiten, ebenfalls zu 86,67% eine informelle Überprüfung der vorliegenden Schwierigkeiten, zu 76,67% ein Test zur Überprüfung der Motorik oder Wahrnehmung und zu 20% ein standardisierte Intelligenztest ausgewählt. Zehn Studierende haben alle verfügbaren Tests angeklickt und dabei mit dem ersten angezeigten Test begonnen und mit dem letzten Test geendet. Bei Simulationsspiel 2 wählten 70% aller Studierenden den passenden Test aus. Obwohl Distraktoren weiterhin angeklickt wurden, ging die Häufigkeit dieser Testauswahl zurück. Der Test zur Überprüfung der Vorläuferfähigkeiten wurde zu 50% angeklickt, die informelle Überprüfung der vorliegenden Schwierigkeiten zu 30%, der Test zur Überprüfung der Motorik oder Wahrnehmung zu 40% und der standardisierte Intelligenztest zu 10%. Alle 21 Studierende, die den passenden Test in Simulationsspiel 2 auswählten, berücksichtigten die Testergebnisse in ihrer Interpretation. [33]

Der Zeitpunkt, zu dem der passende Test ausgewählt wurde, unterscheidet sich vom ersten zum letzten Simulationsspiel. Der passende Test wurde in Simulationsspiel 2 von 23,33% der Studierenden mit dem ersten Klick ausgewählt. 56,67% beendeten das Simulationsspiel, nachdem sie die Testergebnisse des passenden Tests gesehen haben. Bei 10 Studierenden war der erste Klick auch der letzte Klick. In Simulationsspiel 1 hingegen klickte sich die Mehrheit der Studierenden der Reihe nach durch alle verfügbaren Tests. Somit variierte der Zeitpunkt, zu dem der passende Test in Simulationsspiel 1 angeklickt wurde. Er wurde jedoch selten mit dem ersten (6,67%) oder mit dem letzten Klick (13,33%) ausgewählt. [34]

7. Diskussion

Digitale Simulationsspiele ermöglichen es, ein tieferes konzeptionelles und strategisches Wissen zu erwerben. Ein interaktives digitales Simulationsspiel zur diagnostischen Entscheidungskompetenz für angehende Lehrkräfte wurde entwickelt und mit $N = 30$ Studierenden der Sonderpädagogik überprüft. [35]

7.1. Diskrepanz zwischen Selbsteinschätzung und tatsächlichem Übungsverhalten

Die Studierenden schätzen ihre diagnostischen Kompetenzen zu Beginn der Intervention leicht negativ bis neutral konstant über die einzelnen Dimensionen ein ($M = 2.91$ ($SD = 0.58$), $M = 2.82$ ($SD = 0.76$), $M = 2.47$ ($SD = 0.64$), $M = 2.40$ ($SD = 0.83$)). Die Gesamtwerte liegen leicht unter denen der Vergleichsgruppe des Fragebogens (Jungjohann & Gebhardt, 2023). Nur einzelne Items wie Item 1 „Schwierigkeiten im Lernen identifizieren“, Item 9 „kenne Normwerte“ und Item 21 „kenne mehrere Verfahren zur Begleitung von Lernprozessen“ stimmten die Stu-

dierenden zu. Dies deutet darauf hin, dass die Studierenden zwar über allgemeines Vorwissen durch universitäre Veranstaltungen und schulische Erfahrung anhand der Praktika verfügten, jedoch kein Kompetenzerleben in Bezug auf die Anwendung von Diagnostik erfahren haben. Item 16 „Ich kenne die notwendigen Schritte zur Feststellung des Bedarfs an sonderpädagogischer Unterstützung“ wurde somit kaum zugestimmt. Das digitale Simulationsspiel wurde entwickelt, um diesen Kompetenzbereich zu verbessern und Selbstwirksamkeit aufzubauen. [36]

Die Studierenden unterschätzen ihre diagnostischen Handlungskompetenzen. Die beiden Items der Selbsteinschätzung mit den niedrigsten Mittelwerten $M < 2$ (Item 16 „kenne die notwendigen Schritte zur Feststellung des Bedarfs an sonderpädagogischer Unterstützung“ und Item 25 „kann eine informelle Lernverlaufsdagnostik entwickeln“) erfragen Transfer- und Handlungswissen der pädagogischen Diagnostik. In der Anwendung im ersten Simulationsspiel zeigten die Studierenden hingegen von Anfang an ein passendes Klickverhalten. Bereits 80% der Studierenden wählten in Simulationsspiel 1 unter anderem den passenden Test aus. Die Selbsteinschätzung hat daher nur beschränkt Aussagekraft über das tatsächliche Können der Studierenden. Eine mögliche Erklärung für diese vorsichtige Einschätzung der eigenen Kompetenz könnte sein, dass Wissen über die konkrete Handlung und positive Erfahrungen in der Anwendung fehlen (Savolainen et al., 2022). [37]

7.2. Veränderung der Handlungsschablonen durch Simulationsspiele

Zwar wählten die Studierenden im Simulationsspiel den passenden Test aus, bezogen aber auch die anderen Informationsquellen (durchschnittlich getätigte Klicks in Spiel 1: $M = 4.33$ von 5) mit ein und hatten dadurch ein zeitintensives und ressourcenstrapazierendes Klickverhalten. Daraus ergibt sich eine mittlere Trefferquote von 17,56% ($Min = 0\%$, $Max = 33,33\%$) bei zugleich vielen Klicks ($M = 4.33$). Dies kann entweder als neugieriges Klickverhalten interpretiert werden oder spricht für eine größere Unsicherheit der Studierenden im diagnostischen Entscheidungsprozess (Espin et al., 2018; Moser Opitz et al., 2019). Im letzten Simulationsspiel zeigten die Studierenden ein selektiveres Entscheidungsverhalten und verinnerlichte Handlungsschablonen (gestiegene Trefferquote auf durchschnittlich 42,5% ($Min = 0\%$, $Max = 100\%$) bei niedriger Anzahl Klicks ($M = 2.07$)). Der Zeitpunkt, zu dem der passende Test von den Studierenden ausgewählt wird, zeigt in Simulationsspiel 2 eine systematische Handlungsschablone, während in Simulationsspiel 1 keine spezifische Tendenz erkennbar ist. In Simulationsspiel 1 klickte sich die Mehrheit der Studierenden durch alle verfügbaren Tests. Somit variierte der Zeitpunkt, zu dem der passende Test in Simulationsspiel 1 angeklickt wurde, er wurde jedoch selten mit dem ersten (6,67%) oder mit dem letzten Klick (13,33%) ausgewählt. In Simulationsspiel 2 hingegen wurde der passende Test von 23,33% der Studierenden bereits mit dem ersten Klick ausgewählt. 56,67% beendeten das Simulationsspiel, nachdem sie die Testergebnisse des passenden Tests gesehen haben. Bei 10 Studierenden war der erste Klick auch der letzte Klick. Das Ergebnis des passenden Tests wurde vermutlich als ausreichend informativ für die diagnostische Entscheidung wahrgenommen und konsistent zur Ableitung von Fördermaßnahmen genutzt. Distraktoren wurden im letzten Simulationsspiel deutlich seltener angeklickt. Die Phase der Interpretation und Ableitung konkreter Fördermaßnahmen wurde somit schneller und mit höherer Sicherheit und Effizienz erreicht. Das zielgerichtete Entscheidungsverhalten weist darauf hin, dass Studierende konzeptuelle Handlungsschablonen und Entscheidungsheuristiken verinnerlicht haben (Gigerenzer & Gaissmaier, 2011). [38]

Die Entscheidungsbaumstruktur der Simulationsspiele erlaubt das Sichtbarmachen von Handlungs- und Gedankenvorgängen einer diagnostischen Entscheidungsfindung. So ist in Simulationsspiel 1 erkennbar, dass alle Studierenden mindestens 3 Tests anklickten ($M = 4.33$ $Min = 3$, $Max = 5$, $SD = 0.76$), was entweder als Unsicherheit und Absicherung oder als Akkuratheit, keine relevante Information zu übersehen verstanden wird. Ein Drittel klickte sich beim ersten Test beginnend durch alle Tests. Dies deutet darauf hin, dass die diagnostische Aktivität der Studierenden zu diesem Zeitpunkt keiner strategischen, planvollen Handlungsschablone unterliegt. Ebenso lässt sich am Klickverhalten feststellen, dass 86,67% der Studierenden zu Beginn nicht-standardisierte Tests auswählten, was darauf hinweisen kann, dass ihnen diese Verfahren

eher vertraut sind als standardisierte Verfahren. Die Betrachtung der Logdaten liefert somit vielfältige Informationen über das Bearbeitungsverhalten der Studierenden (Chernikova et al., 2020). Auf diese Weise können durch die Übung an Simulationsspielen fehlerhafte Entscheidungsprozesse identifiziert und Entscheidungen hinsichtlich der Tragfähigkeit durch gezielte Hinweise optimiert werden. In der vorliegenden Übung erfolgte dies anhand von formativem Feedback zwischen den Messzeitpunkten. Veränderungen im Klickverhalten bei Simulationsspiel 2 lassen sich auf diese Hinweise zurückführen. So ging beispielsweise der prozentuale Anteil angeklickter nicht-standardisierter Tests von 86,67% bei Simulationsspiel 1 zurück auf 30% beim letzten Simulationsspiel. Ungeklärt ist, in welchem Maße diese Anpassung Effekt von sozial erwünschtem Verhalten ist, oder aus einer Einsicht der Studierenden resultiert. [39]

Einige Studierende zeigten bis zuletzt keine Veränderungen in ihrem diagnostischen Entscheidungsverhalten und klickten sich konsequent durch alle Tests. Dieses Verhalten lässt sich mit der menschlichen Intuition zur Absicherung erklären (Reimer et al., 2007). Die Veränderung dieser Intuition erfordert möglicherweise bei diesen Personen mehr Zeit und weitere Übung, welche in der späteren sonderpädagogischen Schulpraxis gegeben ist. Die Umsetzung von Wissen in beobachtbares, angepasstes Entscheidungsverhalten ist von situationsspezifischen Bedingungen sowie weiteren persönlichen Dispositionen abhängig (Blömeke et al., 2015). Die Teilnehmenden übten an einer identischen ausgewählten diagnostischen Situation. Die situationsspezifischen Bedingungen werden somit in ihrer Komplexität reduziert, um diesen Einflussfaktor auf das Übungsverhalten zu reduzieren. Die persönlichen kognitiven und motivationalen Dispositionen bleiben jedoch bestehen und können somit als Erklärung für heterogene Übungserfolge herangezogen werden. Die Studierenden übten zusätzlich zum ersten und letzten Simulationsspiel freiwillig und in stark variierender Teilnehmerzahl an weiteren Simulationsspielen. Die Ergebnisse bestätigen die Annahme, dass pädagogische Diagnostik erst dann wirksam eingesetzt werden kann, wenn Pädagog:innen mit positiven Erwartungen an den Entscheidungsprozess herangehen (Goldberg et al., 2021). [40]

7.3. Weiterführende Forschung

Zukünftig werden die Simulationsspiele authentischer und explorativer weiterentwickelt, um eine höhere Eigenaktivität der Studierenden zu erzielen. In der Praxis erfolgen diagnostische Entscheidungen häufig situativ im Klassenraum, intuitiv und primär durch eigene Aktivität der diagnostizierenden Person (Schildkamp & Kuiper, 2010). Obwohl die Authentizität durch die Struktur von Entscheidungsbäumen gegeben ist, wird sie durch das reine Klickverhalten nicht realistisch abgebildet. Zudem sind die Simulationsspiele primär auf eine ausgewählte Situation, nämlich der Simulation von Statusdiagnostik ausgerichtet, und bilden somit nicht die gesamte Bandbreite der pädagogischen Diagnostik ab. Im nächsten Entwicklungsschritt werden die Simulationsspiele durch Elemente der formativen Diagnostik ergänzt werden, um auch den Prozess und die Evaluation von pädagogischen Entscheidungen nachzubilden. Ein weiterer Forschungsansatz berücksichtigt die Zeit als Einflussgröße beim Treffen von Entscheidungen. Die Übungsumgebung der vorliegenden Studie in H5P erfasst keine Zeitdaten der Teilnehmenden. Für zukünftige Studien ist die Erhebung und Analyse von Zeitdaten als Komponente der Entscheidungsfindung vorgesehen. Schließlich ist es für weiterführende Forschung wichtig, auch die offenen Items der Interpretation und Formulierung konkreter Fördermaßnahmen qualitativ auf inhaltlicher Ebene auszuwerten. Dadurch werden zusätzliche Einblicke in die Art und Weise erhofft, wie Lernende die diagnostischen Aufgaben bearbeiten und welche spezifischen Kompetenzen dabei gefördert werden. [41]

8. Limitierung

Eine Limitierung der vorliegenden Studie besteht darin, dass die Simulationsspiele nur zu zwei Übungszeitpunkten durchgeführt wurden. Um aussagekräftigere Ergebnisse zu erzielen, ist es notwendig, die Simulationsspiele in einer größeren Stichprobe mit häufigeren Übungsmöglichkeiten zu evaluieren und dabei das unterschiedliche Vorwissen der Teilnehmenden zur Diagnostik zu berücksichtigen. Zur Evaluation der Lernwirksamkeit wurde kein spezifisches Instrument eingesetzt. Die Schlüsse basieren auf der Analyse des Übungsverhaltens. Zur Absiche-

rung der Lernwirksamkeit kann in weiterführenden Studien ein Kontrollgruppenexperiment herangezogen werden. [42]

Literatur

- Agarwal, S., Pandey, G. N. & Tiwari, M. D. (2012). Data mining in Education: Data Classification and Decision Tree Approach. *International Journal of e-Education, e-Business, e-Management and e-Learning*, 2(2), 140–144. doi: [10.7763/IJEEEE.2012.V2.97](https://doi.org/10.7763/IJEEEE.2012.V2.97)
- Bauer, E., Sailer, M., Kiesewetter, J., Schulz, C., Gurevych, I., Fischer, M. R. et al. (2022). Learning to Diagnose Students' Behavioral, Developmental, and Learning Disorders in a Simulation-Based Learning Environment for Pre-Service Teachers. In F. Fischer & A. Opitz (Hrsg.), *Learning to Diagnose with Simulations. Examples from Teacher Education and Medical Education* (S. 97–107). Cham: Springer. doi: [10.1007/978-3-030-89147-3_8](https://doi.org/10.1007/978-3-030-89147-3_8)
- Beckstein, G. & Sroka, S. (2022). Ein Erfahrungsbericht aus Bayern – Mobiler Sonderpädagogischer Dienst Sprache, Lernen, emotionale und soziale Entwicklung. In M. Gebhardt, D. Scheer & M. Schurig (Hrsg.), *Handbuch der sonderpädagogischen Diagnostik. Grundlagen und Konzepte der Statusdiagnostik, Prozessdiagnostik und Förderplanung (Version 1.0)* (Regensburger Beiträge zur Inklusions- und Sonderpädagogik, Bd. 4, S. 773–782). Regensburg: Universität Regensburg. doi: [10.5283/epub.53149](https://doi.org/10.5283/epub.53149)
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond dichotomies. Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. doi: [10.1027/2151-2604/a000194](https://doi.org/10.1027/2151-2604/a000194)
- Capovilla, D., Falkenstörfer, S. & Gebhardt, M. (2023). Die Begleitung sonderpädagogischer Praktika aus akademischer Sicht. *Praktika im Lehramt für Sonderpädagogik. Zeitschrift für Heilpädagogik*, 74(10), 446–449.
- 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. doi: [10.3102/0034654320933544](https://doi.org/10.3102/0034654320933544)
- Cook, D. A., Brydges, R., Hamstra, S. J., Zendejas, B., Szostek, J. H., Wang, A. T. et al. (2012). Comparative effectiveness of technology-enhanced simulation versus other instructional methods. a systematic review and meta-analysis. *Simul Healthc.*, 7(5), 308–320. doi: [10.1097/SIH.0b013e3182614f95](https://doi.org/10.1097/SIH.0b013e3182614f95)
- Espin, C. A., Saab, N., Pat-El, R., Boender, P. D. M. & van der Veen, J. (2018). Curriculum-Based Measurement progress data: Effects of graph pattern on ease of interpretation. *Zeitschrift für Erziehungswissenschaft*, 21(4), 767–792. doi: [10.1007/s11618-018-0836-9](https://doi.org/10.1007/s11618-018-0836-9)
- Fischer, M. R. (2000). CASUS - An authoring and learning tool supporting diagnostic reasoning. *Zeitschrift für Hochschuldidaktik*, 1(1), 87–98.
- Gebhardt, M. (2023). *Pädagogische Diagnostik. Leistung, Kompetenz und Entwicklung messen, bewerten und für individuelle Förderung interpretieren (Version 0.2)* (Regensburger Beiträge zur Inklusions- und Sonderpädagogik, Bd. 3). Regensburg: Universität Regensburg. doi: [10.5283/epub.54450](https://doi.org/10.5283/epub.54450)
- Gebhardt, M., Diehl, K. & Mühling, A. (2016). Online-Lernverlaufsmessung für alle SchülerInnen und Schüler in inklusiven Klassen. *Zeitschrift für Heilpädagogik*, 67(10), 444–453.
- Gebhardt, M., Ebenbeck, N., Jungjohann, J., Koutsianikouli, G. & Lutz, S. (2023). *Praktikumskonzept zum studienbegleitenden sonderpädagogischen Praktikum im Schwerpunkt Lernen*, Universität Regensburg. Verfügbar unter: https://www.uni-regensburg.de/assets/humanwissenschaften/lernbehindertenpaedagogik/Praktikumskonzept_FS_Lernen_Regensburg_23.06.23_final.pdf
- Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62, 451–482. doi: [10.1146/annurev-psych-120709-145346](https://doi.org/10.1146/annurev-psych-120709-145346)
- Goldberg, P., Schwerter, J., Seidel, T., Müller, K. & Stürmer, K. (2021). How does learners' behavior attract preservice teachers' attention during teaching? *Teaching and Teacher Education*, 97. doi: [10.1016/j.tate.2020.103213](https://doi.org/10.1016/j.tate.2020.103213)
- H5P Group. (2019). *Branching Scenario with H5P (Interactive Learning Resource)*, H5P. Verfügbar unter: <https://h5p.org/branching-scenario>

- Hähn, K., Häsel-Weide, U. & Scherer, P. (2021). Diagnosegeleitete Förderung im inklusiven Mathematikunterricht der Grundschule - Professionalisierung durch reflektierte Handlungspraxis in der Lehrer*innenbildung. *Qfi - Qualifizierung für Inklusion*, 3(2). doi: [10.21248/qfi.72](https://doi.org/10.21248/qfi.72)
- Heimlich, U. (2012). *Gemeinsam von Anfang an. Inklusion für unsere Kinder mit und ohne Behinderung* (Kinder sind Kinder, Bd. 38). München: Reinhardt.
- Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M. et al. (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. doi: [10.14786/flr.v7i4.384](https://doi.org/10.14786/flr.v7i4.384)
- Hoffrage, U. & Reimer, T. (2004). Models of bounded rationality: The approach of fast and frugal heuristics. *management revue*, 15(4), 437–459. doi: [10.5771/0935-9915-2004-4-437](https://doi.org/10.5771/0935-9915-2004-4-437)
- Ifenthaler, D. & Schumacher, C. (2016). Learning analytics im Hochschulkontext. *WiSt – Wirtschaftswissenschaftliches Studium*, 45(4), 176–181. doi: [10.15358/0340-1650-2016-4-176](https://doi.org/10.15358/0340-1650-2016-4-176)
- Inspectie van het Onderwijs. (2010). *Opbrengstgericht werken in het basisonderwijs. Een onderzoek naar opbrengstgericht werken bij rekenen-wiskunde in het basisonderwijs*, Inspectie van het Onderwijs. Verfügbar unter: <https://www.onderwijsinspectie.nl/documenten/rapporten/2010/11/01/opbrengstgericht-werken-in-het-basisonderwijs>
- Jungjohann, J. & Gebhardt, M. (2023). *Fragebogen zur Erfassung diagnostischer Kompetenz von Lehrkräften in der inklusiven Schule (DaKI, Version 0.2)*. Regensburg: Universität Regensburg.
- Kingston, N. & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37. doi: [10.1111/j.1745-3992.2011.00220.x](https://doi.org/10.1111/j.1745-3992.2011.00220.x)
- Klauer, K. J. (1978). Perspektiven Pädagogischer Diagnostik. In K. J. Klauer (Hrsg.), *Handbuch der pädagogischen Diagnostik* (Bd. 1, S. 3–14). Düsseldorf: Schwann.
- Kultusministerkonferenz & Hochschulrektorenkonferenz. (2015). *Lehrerbildung für eine Schule der Vielfalt. Gemeinsame Empfehlung von Hochschulrektorenkonferenz und Kultusministerkonferenz. Beschluss der Kultusministerkonferenz vom 12.03.2015/Beschluss der Hochschulrektorenkonferenz vom 18.03.2015*. Verfügbar unter: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_03_12-Schule-der-Vielfalt.pdf
- Lai, M. K., McNaughton, S., Amituanai-Tolosa, M., Turner, R. & Hsiao, S. (2009). Sustained acceleration of achievement in reading comprehension: The New Zealand experience. *Reading Research Quarterly*, 44(1), 30–56. doi: [10.1598/RRQ.44.1.2](https://doi.org/10.1598/RRQ.44.1.2)
- Liebers, K., Kanold, E. & Junger, R. (2019). Digitale Lernstandsanalysen in der inklusiven Grundschule? In S. Bartusch, Klektau, C., Simon, T., S. Teumer & A. Weidemann (Hrsg.), *Lernprozesse begleiten. Anforderungen an pädagogische Institutionen und ihre Akteur*innen* (S. 209–221). Wiesbaden: Springer VS. doi: [10.1007/978-3-658-21924-6_16](https://doi.org/10.1007/978-3-658-21924-6_16)
- Marx, D. & Lange, P. (2019). Entscheidungsfindung in der Akut- und Notfallmedizin. *Notfallmedizin up2date*, 14(1), 71–87. doi: [10.1055/a-0757-9089](https://doi.org/10.1055/a-0757-9089)
- Moser Opitz, E., Pool Maag, S. & Labhart, D. (2019). Förderpläne: Instrument zur Förderung oder "bürokratisches Mittel"? Eine empirische Untersuchung zum Einsatz von Förderplänen. *Empirische Sonderpädagogik*, 11(3), 210–224. doi: [10.25656/01:17780](https://doi.org/10.25656/01:17780)
- Nisbet, R., Elder, J. & Miner, G. (2009). *Handbook of statistical analysis and data mining. Applications*. San Diego: Academic press.
- Prince, A. M. T., Yell, M. L. & Katsiyannis, A. (2018). Andrew F. v. Douglas County School District (2017): The U.S. Supreme Court and Special Education. *Intervention in School and Clinic*, 53(5), 321–324. doi: [10.1177/1053451217736867](https://doi.org/10.1177/1053451217736867)
- R Core Team. (2018). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing. Verfügbar unter: <https://www.R-project.org/>
- Raffaghelli, J. E. (2019). Developing a framework for educators' data literacy in the European context: Proposal, implications and debate. *EDULEARN, Proceedings 2019*, 10520–10530. doi: [10.21125/edulearn.2019.2655](https://doi.org/10.21125/edulearn.2019.2655)
- Reimer, T., Hoffrage, U. & Katsikopoulos, K. (2007). Entscheidungsheuristiken in Gruppen. *NeuroPsychoEconomics*, 2(1), 7–29.

- Savolainen, H., Malinen, O.-P. & Schwab, S. (2022). Teacher efficacy predicts teachers' attitudes towards inclusion – a longitudinal cross-lagged analysis. *International Journal of Inclusive Education*, 26(9), 958–972. doi: [10.1080/13603116.2020.1752826](https://doi.org/10.1080/13603116.2020.1752826)
- Schildkamp, K. & Ehren, M. (2013). From “Intuition”-to “Data”-based Decision Making in Dutch Secondary Schools? In K. Schildkamp, M. K. Lai & L. Earl (Hrsg.), *Data-based Decision Making in Education. Challenges and Opportunities* (Studies in Educational Leadership, S. 49–67). Dordrecht: Springer. doi: [10.1007/978-94-007-4816-3_4](https://doi.org/10.1007/978-94-007-4816-3_4)
- Schildkamp, K. & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, 26(3), 482–496. doi: [10.1016/j.tate.2009.06.007](https://doi.org/10.1016/j.tate.2009.06.007)
- Schröter, A., Schulze, S. & Kuhl, J. (2023). Was lernt man da eigentlich? Der Lerngegenstand der Sonderpädagogik als lehrerinnenbildendes Fach. *Zeitschrift für Heilpädagogik*, 74(2), 83–91.
- Schulz, C., Sailer, M., Kiesewetter, J., Meyer, C. M., Gurevych, I., Fischer, F. et al. (2017). *Fallsimulationen und automatisches adaptives Feedback mittels Künstlicher Intelligenz in digitalen Lernumgebungen*, e-teaching.org. Verfügbar unter: https://www.researchgate.net/publication/320616057_Fallsimulationen_und_automatisches_adaptives_Feedback_mittels_Kunstlicher_Intelligenz_in_digitalen_Lernumgebungen
- Siemens, G. & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5), 31–40.
- 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. doi: [10.1016/j.learninstruc.2022.101689](https://doi.org/10.1016/j.learninstruc.2022.101689)
- Staman, L., Timmermans, A. & Visscher, A. J. (2017). Effects of a Data-Based Decision Making Intervention on Student Achievement. *Studies in Educational Evaluation*, 55, 58–67. doi: [10.1016/j.stueduc.2017.07.002](https://doi.org/10.1016/j.stueduc.2017.07.002)
- Urhahne, D. & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review*, 32. doi: [10.1016/j.edurev.2020.100374](https://doi.org/10.1016/j.edurev.2020.100374)
- Viberg, O., Hatakka, M., Bälter, O. & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in human behavior*, 89, 98–110. doi: [10.1016/j.chb.2018.07.027](https://doi.org/10.1016/j.chb.2018.07.027)
- Wayman, J. C., Jimerson, J. B. & Cho, V. (2012). Organizational considerations in establishing the Data-Informed District. *School Effectiveness and School Improvement*, 23(2), 159–178. doi: [10.1080/09243453.2011.652124](https://doi.org/10.1080/09243453.2011.652124)

Kontakt

Judith Zellner, Ludwig-Maximilians-Universität München, Fakultät für Psychologie und Pädagogik, Leopoldstraße 13, 80802 München
E-Mail: judith.zellner@edu.lmu.de

Zitation

Zellner, J., Ebenbeck, N. & Gebhardt, M. (2024). Entwicklung digitaler Simulationsspiele mit integrierten Entscheidungsbäumen zur Förderung der diagnostischen Entscheidungskompetenzen in der sonderpädagogischen Lehrkräfteausbildung. *Qfl - Qualifizierung für Inklusion*, 6(2), doi: [10.21248/Qfl.162](https://doi.org/10.21248/Qfl.162)

Eingereicht: 05. März 2024

Veröffentlicht: 29. November 2024



Dieser Text ist lizenziert unter einer [Creative Commons Namensnennung - Keine Bearbeitungen 4.0 International Lizenz](https://creativecommons.org/licenses/by-nd/4.0/).