

Mehrdeutigkeit in Diagnosesituationen erkennen

Entwicklung, Durchführung und Evaluation eines Fortbildungsmoduls zum Erwerb diagnostischer Kompetenzen

Elif Özel^{1,*}, Timo Leuders^{1,*} & Katharina Loibl^{1,*}

¹ Pädagogische Hochschule Freiburg

* Kontakt: Pädagogische Hochschule Freiburg,
Institut für Mathematische Bildung
Kunzenweg 21, 79117 Freiburg i.Br.
Mail: elif.oezel@stud.ph-freiburg.de,
leuders@ph-freiburg.de,
katharina.loibl@ph-freiburg.de

Zusammenfassung: Dieser Beitrag zeigt auf, wie Lehrkräfte fachdidaktisches Wissen nutzen können, um Fehlkonzepte von Schüler*innen zutreffend zu diagnostizieren – insbesondere in Situationen, in denen ein identifiziertes Fehlermuster in Aufgabenbearbeitungen durch unterschiedliche Fehlkonzepte erklärt werden kann. Im Fokus stehen also mehrdeutige Diagnosesituationen.

Ausgehend von einer idealtypischen Diagnose von Fehlkonzepten in mehrdeutigen Diagnosesituationen bei Aufgabenbearbeitungen zum Größenvergleich von Brüchen und zur Einschätzung von Wahrscheinlichkeiten werden zunächst die Anforderungen an eine Lehrkraft und die zu deren Umsetzung notwendigen diagnostischen Kompetenzen explizit gemacht. Anschließend wird ein entwickeltes Fortbildungsmodul vorgestellt, welches zum Ziel hat, diese Kompetenzen praxisbezogen durch geeignete vignettenbasierte Lernaufgaben systematisch aufzubauen.

Die Wirksamkeit des entwickelten Fortbildungsmoduls wurde in einer experimentellen Studie mit 40 Lehrkräften empirisch überprüft. Die Ergebnisse dieser Studie bestätigen die positive Wirkung des Fortbildungsmoduls (im Vergleich zu einer Kontrollgruppe) auf die diagnostische Kompetenz der Lehrkräfte, Fehlkonzepte in mehrdeutigen Diagnosesituationen zutreffend zu diagnostizieren.

Schlagerwörter: Diagnostik; Kompetenz; Fachdidaktik; Lehrerfortbildung; Mehrdeutigkeit; Sensibilisierung; Fehlkonzept



Dieses Werk ist freigegeben unter der Creative-Commons-Lizenz CC BY-SA 4.0 (Weitergabe unter gleichen Bedingungen). Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Tabellen, Zitate etc.) sind von der CC-Lizenz ausgenommen. Für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen. <https://creativecommons.org/licenses/by-sa/4.0/legalcode.de>

1 Einleitung

Lehrkräfte stehen täglich vor der Herausforderung, die heterogenen Lernvoraussetzungen und Lernprozesse ihrer Schüler*innen zutreffend zu diagnostizieren, um an diese anknüpfend Lehr-Lern-Prozesse zielgerichtet zu initiieren (Helmke, 2010; Schrader, 2013). Ausgangspunkt zutreffender Diagnosen sind fachdidaktische Überlegungen (Loibl et al., 2020; Prediger et al., 2012; van Ophuysen, 2010; Moser-Opitz, 2010): Liegen den Aufgabenbearbeitungen von Schüler*innen fachlich fehlerhafte Denkweisen, sogenannte Fehlkonzepte, zugrunde? Welche Aufgaben sind geeignet, Fehlkonzepte von Schüler*innen aufzudecken (diagnostische Aufgaben, Kron et al., 2021)?

Fehlkonzepte sind der unmittelbaren Wahrnehmung einer Lehrkraft nicht zugänglich und müssen rekonstruktiv, zum Beispiel aus Aufgabenbearbeitungen, erschlossen werden (vgl. Kaufmann & Wessolowski, 2021; Loibl et al., 2020; Prediger & Wittmann, 2009). Besonders herausfordernd für Lehrkräfte sind Aufgabenbearbeitungen, die keine eindeutigen Diagnosen über zugrunde liegende Fehlkonzepte zulassen. Eindeutige Diagnosen sind nicht möglich, wenn ein identifiziertes Fehlermuster durch unterschiedliche Fehlkonzepte erklärt werden kann. In diesem Fall sprechen wir von mehrdeutigen Diagnosesituationen (Rieu et al., 2024b; Häsel-Weide & Prediger, 2017; Tietze, 1988).

Beim Diagnostizieren von Fehlkonzepten ist das fachdidaktische Wissen einer Lehrkraft handlungsleitend. Zentral dabei ist unter anderem das Wissen über Vorgehensweisen und Vorstellungen sowie typische Fehler und ihre Ursachen (Häsel-Weide & Prediger, 2017; Kaufmann & Wessolowski, 2021).

Dieser Beitrag zielt darauf ab,

- 1) ein Fortbildungsmodul zu entwickeln, das die komplexe diagnostische Kompetenz einer Lehrkraft – die Mehrdeutigkeit in Diagnosesituationen zu erkennen – schrittweise aufbaut und
- 2) empirisch zu untersuchen, ob dieses Fortbildungsmodul die diagnostischen Urteile von Lehrkräften verbessert.

Hierzu werden zunächst – ausgehend von einer idealtypischen Diagnose von Fehlkonzepten in mehrdeutigen Diagnosesituationen zum Größenvergleich von Brüchen (Eichelmann et al., 2012) und zur Einschätzung von Wahrscheinlichkeiten (Pratt, 2000) – die Anforderungen an eine Lehrkraft und die zu deren Umsetzung notwendigen diagnostischen Kompetenzen (Kap. 2) sowie mögliche Urteilsfehler (Kap. 3) explizit gemacht. Anschließend wird ein entwickeltes Fortbildungsmodul vorgestellt, welches zum Ziel hat, diese Kompetenzen praxisbezogen durch geeignete, vignettenbasierte Lernaufgaben schrittweise aufzubauen (Kap. 4). Die Wirksamkeit des entwickelten Fortbildungsmoduls wird in einer experimentellen Studie empirisch überprüft (Kap. 5 & 6).

2 Einfluss fachdidaktischen Wissens auf das Erkennen der Mehrdeutigkeit in Diagnosesituationen

Das Ausrichten des Unterrichts an den Lernvoraussetzungen der Schüler*innen setzt zutreffende Diagnosen voraus. Helmke (2022, S. 241) bezeichnet als zentrales Merkmal von Diagnosen die Beurteilung anhand „vorgegebener Kategorien, Begriffe oder Konzepte“. Deren Kenntnis stellt eine wichtige Facette fachdidaktischen Wissens dar und bildet den Referenzrahmen für diagnostische Urteile (z.B. das Wissen über typische Fehler und ihre Ursachen). In diesem Kapitel wird an zwei konkreten Beispielen (Abb. 1) veranschaulicht, wie fachdidaktisches Wissen eine idealtypische Diagnose von Fehlkonzepten in mehrdeutigen Diagnosesituationen unterstützt.

An der idealtypischen Diagnose lassen sich Anforderungen an die Lehrkraft und die zu deren Umsetzung notwendigen diagnostischen Kompetenzen explizit machen.

Versetzen Sie sich in die Situation der abgebildeten Lehrkraft. 															
 BENTE a) $\frac{2}{3} < \frac{3}{5}$ ✓ b) $\frac{3}{4} < \frac{5}{7}$ ✗ c) $\frac{2}{4} < \frac{6}{14}$ ✗ d) $\frac{5}{11} < \frac{10}{25}$ ✗ e) $\frac{3}{13} > \frac{1}{12}$ ✓	 BENTE <table border="1"> <thead> <tr> <th>Augennummer</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> </tr> </thead> <tbody> <tr> <td>Häufigkeit</td> <td></td> <td>1</td> <td>1</td> <td>1</td> <td>1</td> <td>1</td> </tr> </tbody> </table> Was ich so oft: das ist ein postmodern als die anderen Zahlen Versuche deine Entdeckungen zu begründen: Ich denke dass mein Würfel genau ist, weil du i. d. Regel- lich nicht so häufig post- modern	Augennummer	1	2	3	4	5	6	Häufigkeit		1	1	1	1	1
Augennummer	1	2	3	4	5	6									
Häufigkeit		1	1	1	1	1									
Sie haben mit Ihrer Lerngruppe eine Lernstandserhebung durchgeführt mit dem Ziel, das konzeptionelle Verständnis Ihrer Schüler*innen zur Bruchrechnung sichtbar zu machen. Sie stellen fest, dass Bente Aufgaben zum Größenvergleich von Brüchen teilweise falsch löst. Sie wollen herausfinden, ob diesen Lösungen ein Fehlkonzent zugrunde liegt, um beispielsweise passgenaues Fördermaterial bereitzustellen.	Sie haben mit Ihrer Lerngruppe ein Zufallsexperiment durchgeführt mit dem Ziel, das konzeptionelle Verständnis Ihrer Schüler*innen zum statistischen Denken sichtbar zu machen. Sie stellen fest, dass Bentes Entdeckung von der theoretischen Wahrscheinlichkeit der Augenzahlen eines Würfels (Gleichwahrscheinlichkeit) abweicht. Sie wollen herauszufinden, ob ihrer bzw. seiner Entdeckung ein Fehlkonzent zugrunde liegt, um beispielsweise passgenaues Fördermaterial bereitzustellen.														

Abbildung 1: Komplexe Diagnosesituation aus dem unterrichtlichen Alltag einer Lehrkraft (eigene Darstellung)

Aus den Lösungen von Bente gilt es nun, Informationen zu identifizieren, die Hinweise auf zugrunde liegende Fehlkonzente liefern (Loibl et al., 2020; Thiede et al, 2015). Das fachdidaktische Wissen zu typischen Fehlern (beim Größenvergleich von Brüchen: Eichelmann et al., 2012; bei der Einschätzung von Wahrscheinlichkeiten: Pratt, 2000) kann diesen Prozess in zweierlei Hinsicht beeinflussen. Zum einen kann das fachdidaktische Wissen die Aufmerksamkeit der Lehrkraft auf die relevanten Informationen in den Lösungen von Bente lenken und somit das Erkennen von Fehlermustern erleichtern (vgl. Wissensgesteuerte Aufmerksamkeitslenkung; van Es & Sherin, 2021):

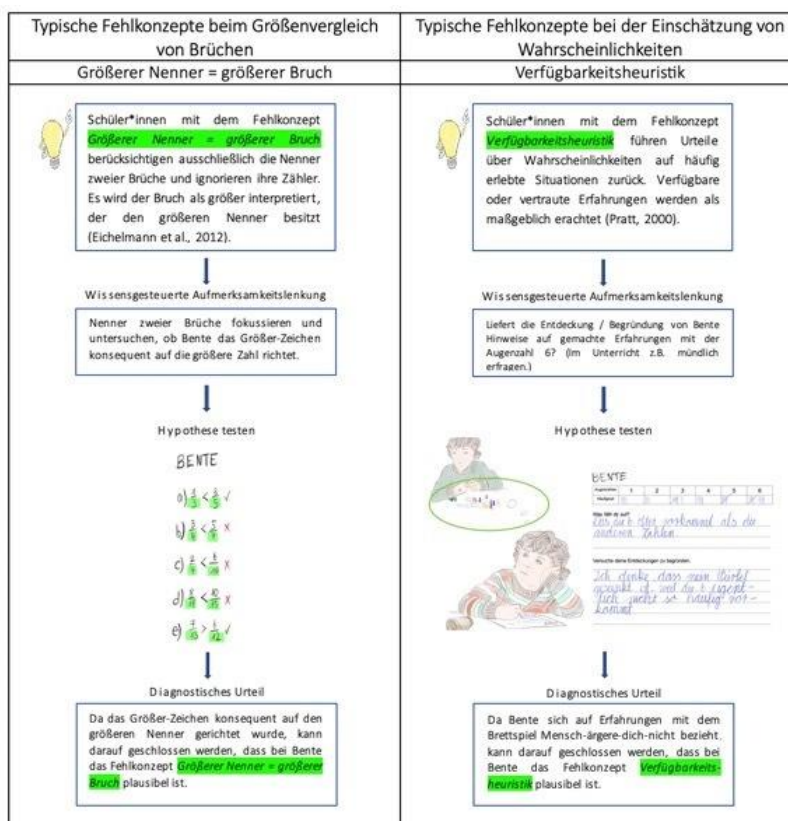


Abbildung 2: Wissensgesteuerte Aufmerksamkeitslenkung auf relevante Informationen der Diagnosesituation (eigene Darstellung)

Zum anderen kann das fachdidaktische Wissen die Bandbreite an möglichen Fehlkonzpten aufzeigen und so den Interpretationsrahmen für die Diagnose festlegen bzw. die Ursachensuche für das identifizierte Fehlermuster organisieren (vgl. Knowledge-based Reasoning; Borko, 2004; van Es & Sherin, 2002). Welche weiteren Fehlkonzpte beim Größenvergleich von Brüchen bzw. bei der Einschätzung von Wahrscheinlichkeiten (neben des Fehlkonzptes in Abb. 2) sind bekannt und welche dieser Fehlkonzpte lassen sich durch die (fehlerhaften) Lösungen von Bente diagnostizieren bzw. anhand dieser Lösungen ausschließen?

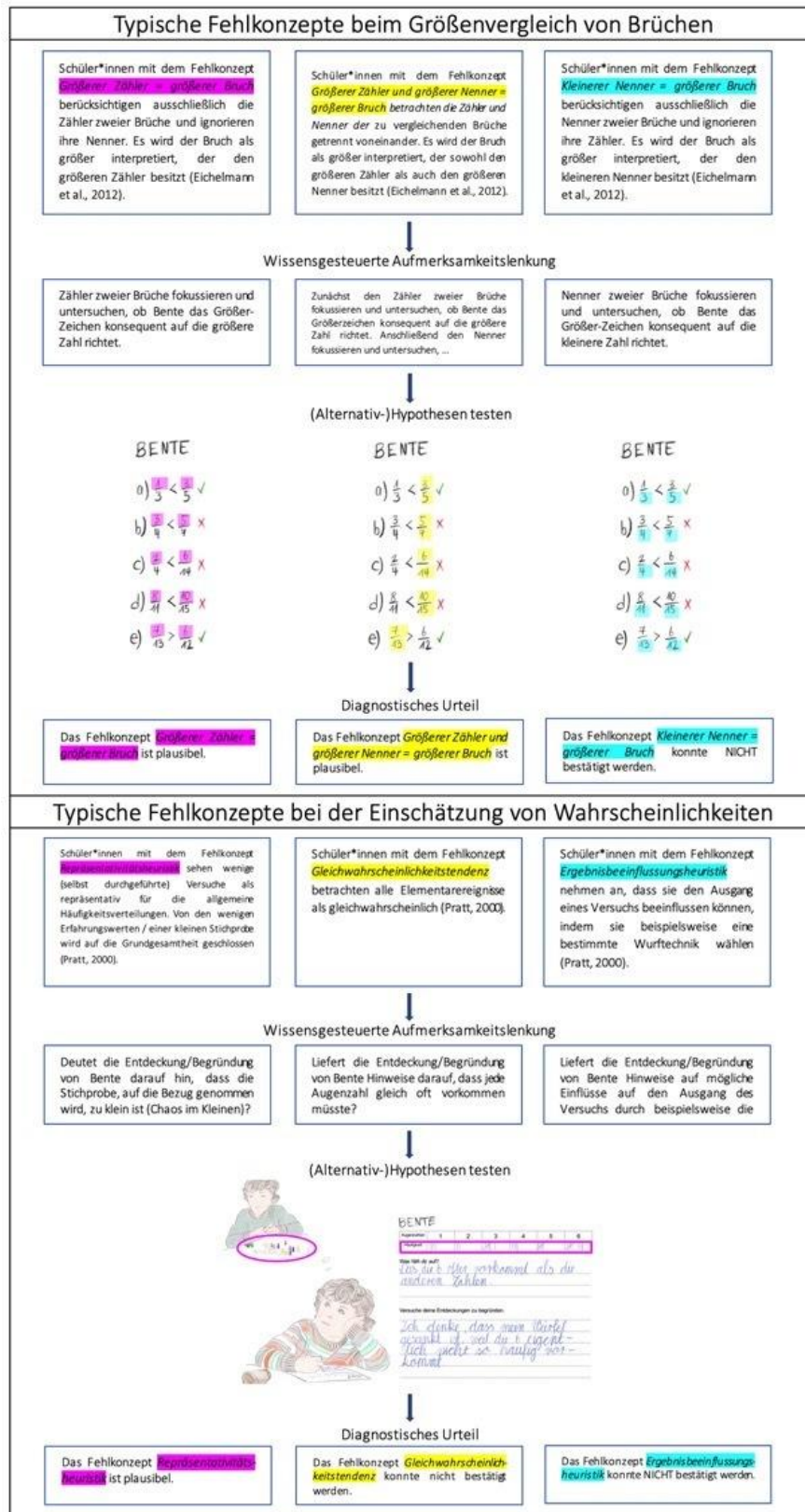


Abbildung 3: Festlegen des Interpretationsrahmens für die Diagnose und systematisches Erproben alternativer Fehlkonzpte (eigene Darstellung)

Aus Abbildung 3 geht hervor, dass Bentes Lösungen keine eindeutige Diagnose über die zugrunde liegenden Fehlkonzepte zulassen. Ein und dasselbe Fehlermuster kann aus unterschiedlichen Fehlkonzepten hervorgehen (Häsel-Weide & Prediger, 2017; Tietze, 1988).

Das Erkennen der Mehrdeutigkeit in Diagnosesituationen stellt hohe Anforderungen an die diagnostischen Kompetenzen einer Lehrkraft.

Lehrkräfte müssen

- Fehlermuster in den Aufgabenbearbeitungen von Schüler*innen identifizieren, indem sie relevante Aspekte und Informationen der Diagnosesituation wahrnehmen (Teilkompetenz 1 = TK 1),
- das Zustandekommen der identifizierten Fehlermuster in den Aufgabenbearbeitungen von Schüler*innen erklären, indem sie zugrunde liegende Denkprozesse auf der Grundlage fachdidaktischen Wissens rekonstruieren (Teilkompetenz 2 = TK 2)
- und die Mehrdeutigkeit der Diagnosesituation erkennen, indem sie trotz Prüfung und Bestätigung eines angenommenen Fehlkonzeptes als Ursache für das identifizierte Fehlermuster, alternative Fehlkonzepte erproben (Teilkompetenz 3 = TK 3).

Werden Lehrkräfte in der unterrichtlichen Praxis mit mehrdeutigen Diagnosesituationen konfrontiert (Abb. 1), so werden in der Regel die oben aufgeführten Teilkompetenzen zugleich eingefordert (Ganzheitlichkeit). Wird beispielsweise die Ursachensuche zu früh abgebrochen, weil alternative Fehlkonzepte nicht erprobt werden (TK 3), bleibt die Mehrdeutigkeit der Diagnosesituation unerkannt. Bleibt die Mehrdeutigkeit unerkannt, kann diese auch nicht (z.B. durch diagnostische Gespräche oder diagnostische Aufgaben; Kron et al., 2021) aufgelöst werden. Es resultieren unzulässige und fehlerhafte diagnostische Urteile.

3 Urteilsfehler beim Erkennen der Mehrdeutigkeit in Diagnosesituationen

Diagnostische Urteile sind mit Unsicherheiten verbunden (z.B. durch die Mehrdeutigkeit von Diagnosesituationen) und daher fehlerbehaftet (Leuders & Loibl, 2021). Rieu et al. (2024a, 2024b) konnten zeigen, dass in mehrdeutigen Diagnosesituationen relevante Informationen, die Hinweise auf zugrunde liegende Fehlkonzepte liefern, nicht ausschöpfend analysiert wurden, obwohl den Lehrkräften das notwendige fachdidaktische Wissen schriftlich zur Verfügung gestellt wurde. Die Informationssuche wurde häufig nach dem ersten identifizierten Fehlkonzept vorzeitig abgebrochen. Da so die Mehrdeutigkeit der Diagnosesituation nicht erkannt wurde, resultierten unzulässige diagnostische Urteile.

3.1 Fachdidaktische Erklärung für Urteilsfehler beim Erkennen der Mehrdeutigkeit in Diagnosesituationen

Aus fachdidaktischer Perspektive könnte eine Begründung für die Vernachlässigung relevanter Informationen beim Diagnostizieren von Fehlkonzepten in mehrdeutigen Diagnosesituationen darin liegen, dass das in der Lehramtsausbildung und im Rahmen beruflicher Fort- und Weiterbildungen für Lehrkräfte erworbene fachdidaktische Wissen (z.B. das Wissen über typische Fehler und ihre Ursachen) nur begrenzt auf die in der unterrichtlichen Praxis zu bewältigenden Herausforderungen (beispielsweise dem Analysieren von fehlerhaften Aufgabenbearbeitungen, die keine eindeutigen Diagnosen über zugrunde liegende Fehlkonzepte zulassen) übertragbar ist.

Das fachdidaktische Wissen zu typischen Fehlern und ihre Ursachen ist für folgende Handlungen relevant (Loibl et al., 2020; Weinert, 2000):

- 1) Es wird angewendet, um Vorhersagen darüber zu treffen, wie Schüler*innen bei bereits bekanntem Fehlkonzepth zukünftige, nicht bearbeitete Aufgaben vermutlich lösen werden und
- 2) um Fehlkonzepthe, die einem identifizierten Fehlermuster in den Aufgabenbearbeitungen von Schüler*innen zugrunde liegen könnten, zu diagnostizieren.

In der Lehramtsausbildung und in beruflichen Fort- und Weiterbildungen für Lehrkräfte liegt der Fokus zumeist auf dem Antizipieren von Schülerlösungen (1). Dabei werden zunächst typische Fehlkonzepthe eines Inhaltsbereichs beschrieben und erklärt. Anschließend wird gezeigt, welche Lösungen diese bei der Bearbeitung von Aufgaben hervorbringen. Dass diese Lösungen auch aus anderen Fehlkonzepthen hervorgegangen sein könnten, wird in der Regel nicht thematisiert. So werden etwa bei Rieu et al. (2024b) Fehlkonzepthe beim Größenvergleich von Dezimalzahlen und bei Padberg und Wartha (2017) Fehlkonzepthe beim Größenvergleich von Brüchen nacheinander aufgelistet und jeweils mit einer anderen Aufgabe und deren Bearbeitung illustriert. Obwohl die dort gezeigten fehlerhaften Lösungen durch verschiedene der gelisteten Fehlkonzepthe erklärt werden können, werden diese nur einem einzigen Fehlkonzepth zugeordnet.

(Angehenden) Lehrkräften bleibt dadurch der Erfahrungsraum verschlossen, dass ein und dasselbe Fehlermuster aus verschiedenen Fehlkonzepthen hervorgehen kann. Werden Lehrkräfte in der unterrichtlichen Praxis nun mit mehrdeutigen Diagnosesituationen konfrontiert, muss das erworbene fachdidaktische Wissen eigenständig und häufig unter Zeitdruck in neue Handlungsabläufe (Erkennen der Mehrdeutigkeit; s. Kap. 2) überführt werden (Stark et al., 2001) – eine kognitiv anspruchsvolle Aufgabe, die sich nicht „von selbst“ einstellt und gezielt gefördert werden muss (vgl. Grossman et al., 2009).

Bisher gibt es wenig Forschung dazu, wie der Vernachlässigung relevanter Informationen beim Diagnostizieren von Fehlkonzepthen in mehrdeutigen Diagnosesituationen entgegengewirkt bzw. das Erkennen der Mehrdeutigkeit gezielt trainiert werden kann. Analogien zwischen diagnostischen Urteilen im pädagogischen und medizinischen Bereich – das Bilden diagnostischer Urteile unter Unsicherheit und Urteilsfehler durch eine eingeschränkte Informationsverarbeitung – jedoch lassen vermuten, dass die Forschung zu diagnostischen Kompetenzen einer Lehrkraft von einem Methodentransfer ausgewählter Debiasing-Strategien der Medizin, die zur Reduzierung solcher Urteilsfehler entwickelt wurden, profitieren kann.

Daher werden im nächsten Unterkapitel zunächst die Analogien zwischen diagnostischen Urteilen im pädagogischen und medizinischen Bereich aufgezeigt und anschließend Debiasing-Strategien der Medizin vorgestellt, die sich auf diagnostische Urteile im pädagogischen Bereich übertragen lassen könnten.

3.2 Analogien zwischen diagnostischen Urteilen im pädagogischen und medizinischen Bereich

Auch Ärzt*innen müssen Diagnosen unter Unsicherheit fällen. Ein Symptom kann unterschiedliche Krankheiten zur Ursache haben, ähnlich wie ein Fehlermuster in Aufgabenbearbeitungen von Schüler*innen aus unterschiedlichen Fehlkonzepthen hervorgehen kann. Die Tendenz zur Vernachlässigung relevanter Informationen beim Diagnostizieren mehrdeutiger Symptome ist in der Medizin besser erforscht als in der pädagogischen Diagnostik und kann Anregungen für das Erkennen der Mehrdeutigkeit in Diagnosesituationen geben. Dabei werden verschiedene Prozesse beschrieben, die zu fehlerhaften Urteilen führen können.

Eine vielfach untersuchte und nachgewiesene Ursache für die Vernachlässigung von relevanten Informationen ist der Bestätigungsfehler (Confirmation Bias; Croskerry, 2002). Ärzt*innen suchen ausschließlich nach Informationen, die ihre erste Annahme (Verdachtsdiagnose) als Ursache für den vorliegenden Befund (wahrgenommenes

Symptom) bestätigen und vernachlässigen dabei solche, die diese widerlegen (z.B. Croskerry, 2002; Kovacs & Croskerry, 1999). Infolge der einseitigen und unzureichenden Informationsbasis können unzulässige und fehlerhafte Urteile für Patient*innen resultieren.

Eine weitere Erklärung für die verkürzte Informationssuche ist das Search Satisficing (z.B. Croskerry, 2003). Das Search Satisficing („Erfüllung bei Sucherfolg“; Valentin, 2009) beschreibt das Verhalten von Urteilenden, die die Informationssuche abbrechen, sobald eine identifizierte Ursache die zu beurteilenden Krankheitssymptome zufriedenstellend erklärt (vgl. Hogarth, 1987).

Rieu et al. (2024a, 2024b) konnten zeigen, dass solche Urteilsverzerrungen auch bei Lehrkräften zu einer reduzierten Diagnoseakkuratesse führen. Dabei scheint der Bestätigungsfehler eine größere Rolle als das Search Satisficing zu spielen, da zwar die Art der verarbeiteten Informationen, nicht aber deren reine Menge mit der Diagnoseakkuratesse zusammenhängt.

Da fehlerhafte Urteile in der Medizin schwerwiegende Konsequenzen für Patient*innen nach sich ziehen können, müssen diese kontinuierlich reflektiert und der Diagnoseprozess gegebenenfalls optimiert werden (Faller & Lang, 2019). Dabei führt das bloße Wissen über Urteilsfehler nicht zu deren Reduzierung (Stacy & MacMillan, 1995).

3.3 Debiasing-Strategien – Methoden zur Reduzierung von Urteilsfehlern

Debiasing-Strategien zielen darauf ab, Denkfehler im diagnostischen Prozess zu erkennen, also beispielsweise Vernachlässigungen relevanter Informationen beim diagnostischen Urteilen aufzudecken.

Wirksame Debiasing-Strategien sind zum Beispiel

- a.) die Verifikation der Verdachtshypothese (Ersthypothese) durch die Suche nach Informationen, die die Diagnose bestätigen können (vgl. TK 1: Wahrnehmung relevanter Informationen und TK 2: Rekonstruktion von Ursachen),
- b.) die Falsifikation der Verdachtsdiagnose (Ersthypothese) durch die Suche nach Informationen, die die Diagnose widerlegen können (vgl. TK 1: Wahrnehmung relevanter Informationen und TK 2: Rekonstruktion von Ursachen)
- c.) und der Ausschluss von alternativen Hypothesen, auch zur Abwendung eines gefährlichen Verlaufs (z.B. Gäbler, 2017; Crosskerry et al., 2013; Graber et al., 2005) (vgl. TK 3: Alternative Diagnosen erproben).

Diese Debiasing-Strategien scheinen auf den pädagogischen Kontext übertragbar zu sein. Bei der Entwicklung des Fortbildungsmoduls zum Erkennen der Mehrdeutigkeit in Diagnosesituationen wurde daher auf diese Debiasing-Strategien zurückgegriffen.

Die Entwicklung des Fortbildungsmoduls wird im nächsten Kapitel beschrieben.

4 Entwicklung des Fortbildungsmoduls

Den theoretischen Rahmen zur Entwicklung des Fortbildungsmoduls bildet das *4-Component-Instructional-Design-Modell* (4C/ID) von van Merriënboer (2020). Dabei handelt es sich um ein strukturiertes, präskriptives Modell für die „Entwicklung von Bildungsprogrammen zur Vermittlung komplexer Fähigkeiten bzw. beruflicher Kompetenzen“ (van Merriënboer, 2020, S. 154).

Diagnostische Kompetenzen von Lehrkräften umfassen ein komplexes Bündel an Fähigkeiten: Lehrkräfte sollen Denkprozesse von Schüler*innen rekonstruieren und mit Blick auf empirische Befunde und Theorien einordnen können. Sie sollen unterschiedliche Verfahren zur Diagnose der Denkweisen begründet auswählen und selbst Aufgaben zur Diagnostik entwickeln und deren Güte kritisch hinterfragen können. Sie sollen typische Fehler und Verzerrungen beim Urteilen kennen und ihre eigene Urteilspraxis reflektieren (s. Kap. 2 & 3; Aufschnaiter et al., 2015).

Für den Erwerb solcher komplexen diagnostischen Kompetenzen benötigen Lehrkräfte ausreichende Lerngelegenheiten, in denen sie sowohl fundiertes theoretisches Wissen (fachdidaktisches Wissen) erwerben, als auch Fähigkeiten und Fertigkeiten im Prozess des praktischen Diagnostizierens (prozedurales Können) anbahnen können. Vor diesem Hintergrund erweist sich das Modell von van Merriënboer (2020) als eine geeignete Wahl für dieses Fortbildungsziel.

Zur Gestaltung von Lerngelegenheiten differenziert van Merriënboer (2020) vier Designkomponenten: Lernaufgaben, unterstützende Informationen, prozedurale Informationen und das Üben von Teilaufgaben. Diese werden im Folgenden (Kap. 4.1) näher beschrieben und deren konkrete Umsetzung im Fortbildungsmodul beispielhaft aufgezeigt (Kap. 4.2).

4.1 Designkomponenten des 4C/ID-Modells

Die Designkomponenten des 4C/ID-Modells (van Merriënboer, 2020) sind darauf ausgerichtet, Lerngelegenheiten zu gestalten, die den Erwerb komplexer Fähigkeiten bzw. beruflicher Kompetenzen, etwa den Erwerb diagnostischer Kompetenzen von Lehrkräften, fördern. Zur Strukturierung solcher Lerngelegenheiten werden vier Designkomponenten beschrieben:

- 1) *Lernaufgaben*: Im 4C/ID-Modell erfolgt der Aufbau komplexer Kompetenzen in der Auseinandersetzung mit authentischen (d.h. realitätsnahen) Lernaufgaben. Authentische Lernaufgaben bilden das volle Spektrum an zentralen Anforderungen und Herausforderungen nach, die Lehrkräfte in ihrer beruflichen Tätigkeit (z.B. beim Diagnostizieren von Fehlkzepten) bewältigen müssen. Die zur Bewältigung dieser Anforderungen notwendigen Kompetenzen umfassen (1) fundiertes Wissen zu typischen Fehlern und ihre Ursachen in einem Inhaltsbereich (fachdidaktisches Wissen), (2) die Anwendung von Vorgehensweisen zur Verarbeitung und Umsetzung des fachdidaktischen Wissens im Prozess des praktischen Diagnostizierens (prozedurale Aspekte des Könnens) sowie (3) die Haltung, dass fachdidaktisches Wissen über typische Fehler und ihre Ursachen die Diagnose von Fehlkzepten erleichtern kann (Leuders, 2020; van Merriënboer, 2020). Solche Kompetenzaspekte wirken beim beruflichen Handeln typischerweise zusammen (Ganzheitlichkeit von Lernaufgaben als ein zentrales Prinzip des 4C/ID-Modells; Leuders, 2020). Daher werden diese Aspekte auch in Lernaufgaben miteinander vernetzt vermittelt – fachdidaktisches Wissen wird stets mit seiner praktischen Anwendung dargeboten.

Lernaufgaben werden in Aufgabenklassen (AK) wachsender Komplexität eingeteilt. Jede Aufgabenklasse besteht aus mehreren Lernaufgaben auf gleichem Komplexitätsniveau.

Lernaufgaben variieren über alle Dimensionen, in denen sich auch diagnostische Aufgaben im realen Unterrichtskontext unterscheiden können. So umfassen beispielsweise die zu diagnostizierenden Fehlkzepten einer Aufgabenklasse die gesamte Bandbreite an möglichen Fehlkzepten (in jeweils unterschiedlichen Lernaufgaben), die in der fachdidaktischen Forschung als typisch identifiziert wurden und daher auch im unterrichtlichen Kontext zu erwarten sind.

Zudem umfassen Lernaufgaben sowohl Routinefähigkeiten (z.B. das Testen von Hypothesen durch systematisches Abgleichen) als auch nicht-routinisierbare Fähigkeiten (z.B. Problemlösen, Denken, Entscheidungen treffen; van Merriënboer, 2020). Solche Fähigkeiten können durch die Bereitstellung von Informationen unterstützt werden, wobei zwischen zwei Arten unterschieden wird: unterstützende Informationen und prozedurale Informationen.

Die Unterstützung der Lehrkräfte wird bei Lernaufgaben einer Aufgabenklasse mit wachsender Expertise sukzessive zurückgenommen, bis sie fähig

sind, die Aufgaben selbstständig, ohne jede Unterstützung oder Anleitung, auszuführen (Scaffolding; van Merriënboer, 2020).

- 2) *Unterstützende Informationen:* Unterstützende Informationen helfen Lehrkräften, die nicht-routinisierten Fähigkeiten einer komplexen Lernaufgabe zu bewältigen, indem sie die notwendige „Theorie“ liefern (z.B. fachdidaktisches Wissen zu typischen Fehlern und ihre Ursachen zu einem Inhaltsbereich). Sie werden entweder vor der Bearbeitung der Lernaufgaben dargeboten („Erst die Theorie, dann die Praxis“) oder sind bei Bedarf während der Bearbeitung abrufbar („Frag nach der Theorie, nur wenn du sie brauchst“; van Merriënboer, 2019).

Da alle Lernaufgaben einer Aufgabenklasse demselben Komplexitätsniveau zugeordnet sind, basieren sie auf denselben Wissensgrundlagen. Entsprechend sind die unterstützenden Informationen für alle Lernaufgaben einer Aufgabenklasse identisch (van Merriënboer, 2019).

- 3) *Prozedurale Informationen:* Prozedurale Informationen unterstützen Lehrkräfte dabei, die Routineaspekte einer komplexen Lernaufgabe effizient zu bewältigen (van Merriënboer, 2020). Routineaspekte werden stets auf die gleiche Weise ausgeführt (z.B. das Testen von Hypothesen durch systematisches Abgleichen). Mit zunehmender Übung werden diese Routinen automatisiert, sodass die prozeduralen Informationen schrittweise reduziert und schließlich entfernt werden können.

Prozedurale Informationen werden während der Bearbeitung von Lernaufgaben bereitgestellt, etwa in Form von „How-to“ oder „Schritt-für-Schritt-Instruktionen“ (van Merriënboer, 2020).

Fehler bei der Ausführung von Routinen können durch korrekatives Feedback ausgeglichen werden. Feedback unterstützt Lehrkräfte darin, mögliche Lücken zwischen ihrem aktuellen sowie dem angestrebten Lern- und Leistungsstand zu identifizieren und zu schließen (Hattie & Timperley, 2007).

- 4) *Üben von Teilaufgaben:* Die Routineaspekte einer Lernaufgabe (z.B. das Testen von Hypothesen durch systematisches Abgleichen) werden in der Regel bereits durch die Bearbeitung der Lernaufgabe selbst (z.B. Diagnostizieren von Fehlkonzepthen) ausreichend eingeübt. Sollte die Lernaufgabe jedoch nicht das erforderliche Maß an Übung bieten oder ein besonders hohes Niveau an Automatisierung der Routineaspekte gefordert sein, können Teilaufgaben dieser Routineaspekte durch gezielte Übungseinheiten separat trainiert werden. Da die im Fortbildungsmodul eingesetzten Lernaufgaben jedoch ein ausreichendes Maß an Übung beinhalten, wird diese Designkomponente des 4C/ID-Modells nicht weiter berücksichtigt.

Im folgenden Unterkapitel (Kap. 4.2) wird aufgezeigt, wie ausgewählte Designkomponenten des 4C/ID-Modells (Lernaufgaben, unterstützende Informationen, prozedurale Informationen) in der Entwicklung des Fortbildungsmoduls „Mehrdeutigkeit in Diagnosesituationen erkennen“ exemplarisch auf den Inhaltsbereich „Einschätzen von Wahrscheinlichkeiten“ angewendet werden.

4.2 Entwicklung des Fortbildungsmoduls

Bei der Entwicklung des Fortbildungsmoduls stand am Anfang die Dekomposition der zu vermittelnden Kompetenz (Zielkompetenz) in ihre konstitutiven Teilkompetenzen als Basis der Erstellung der Lernaufgaben. Diese erfolgte bereits ausführlich in Kapitel 2 und wird an dieser Stelle übernommen. Die identifizierten Teilkompetenzen werden zusätzlich mit den in Kapitel 3.3 vorgestellten Debiasing-Strategien kombiniert. Diese Strategien liefern die Methoden zur Anbahnung der jeweiligen Teilkompetenzen (beispielsweise sollen Lehrkräfte befähigt werden, Fehlermuster in Aufgabenbearbeitungen zu identifizieren (TK 1), indem Sie durch Instruktion (Debiasing-Strategien) geschult

werden, ihre Aufmerksamkeit auf die relevanten Informationen der Diagnosesituation zu lenken).

Schritt 1: Dekomposition der Zielkompetenz in ihre konstitutiven Teilkompetenzen

Die Zielkompetenz

Lehrkräfte können die Mehrdeutigkeit in Diagnosesituationen erkennen.

wird in ihre konstitutiven Teilkompetenzen (TK) zerlegt:

Lehrkräfte können

- TK 1: Fehlermuster in den Aufgabebearbeitungen von Schüler*innen identifizieren, indem sie relevante Aspekte und Informationen der Diagnosesituation wahrnehmen,
- Debiasing-Strategien a und b: Verifikation und Falsifikation von Hypothesen*
- TK 2: das Zustandekommen der identifizierten Fehlermuster in den Aufgabebearbeitungen von Schüler*innen erklären, indem sie zugrunde liegende Fehlkonzpte auf der Grundlage fachdidaktischen Wissens rekonstruieren (Welcher Denkprozess liegt dem identifizierten Fehlermuster zugrunde?),
- Debiasing-Strategien a und b: Verifikation und Falsifikation von Hypothesen*
- TK 3: die Mehrdeutigkeit in Fehlermustern erkennen, indem sie trotz Prüfung und Bestätigung eines angenommenen Fehlkonzptes als Ursache, alternative Fehlkonzpte erproben.
- Debiasing-Strategie c: Ausschluss alternativer Hypothesen*

Die Teilkompetenzen sind hierarchisch angeordnet (gradueller Anstieg der Komplexität zwischen den Teilkompetenzen). Die Zielkompetenz wird mit dem Erwerb der nötigen Teilkompetenzen kumulativ aufgebaut. Kumulatives Lernen kann dann ermöglicht werden, wenn das Kriterium der Adaptivität erfüllt ist und die Fortbildung an Kenntnissen der Lehrkräfte ansetzt (Häsel-Weide & Prediger, 2017). Für den Erwerb der Teilkompetenzen bedarf es entsprechender Lerngelegenheiten. Die Gestaltung der Lerngelegenheiten orientiert sich an den Designkomponenten des 4C/ID-Modells (s. Kap. 4.1).

Schritt 2: Allgemeine Überlegungen zur Umsetzung der Designkomponenten des 4C/ID-Modells im Fortbildungsmodul

Lernaufgaben: Aus den Teilkompetenzen lassen sich für ihren Erwerb Lernaufgaben ableiten. Jede Lernaufgabe in diesem Fortbildungsmodul beinhaltet Elemente einer realen Diagnosesituation aus dem unterrichtlichen Alltag einer Lehrkraft. Diagnostiziert werden Fehlkonzpte von Schüler*innen zur Einschätzung von Wahrscheinlichkeiten, die sich als Fehlermuster in ihren Aufgabebearbeitungen äußern. Dabei wurden alle Fehlkonzpte zur Einschätzung von Wahrscheinlichkeiten hinzugezogen, die in der fachdidaktischen Literatur als typisch für diesen Inhaltsbereich identifiziert wurden (Anhang A im Online-Supplement).

Lernaufgaben sind in Aufgabenklassen (AK) organisiert. Die Komplexität der Lernaufgabe bestimmt, welcher Aufgabenklasse sie zugeordnet wird. Das Fortbildungsmodul gliedert sich in drei Aufgabenklassen wachsender Komplexität.

- AK 1: Testen vorgegebener Hypothesen in eindeutigen Diagnosesituationen (Debiasing-Strategien a und b); Wahrnehmen relevanter Informationen (TK 1),
- AK 2: Aufstellen und Testen von Hypothesen in eindeutigen Diagnosesituationen (Debiasing-Strategien a und b); Rekonstruktion von Ursachen (TK 2),
- AK 3: Aufstellen alternativer Hypothesen in mehrdeutigen Diagnosesituationen (Debiasing-Strategie c); Erproben alternativer Diagnosen (TK 3).

Unterstützende Informationen: Zur Bewältigung der Diagnosesituationen bedarf es an spezifischem fachdidaktischem Wissen über typische Fehler und ihre Ursachen. Dieses Wissen wird den Lehrkräften über die gesamte Fortbildungszeit unterstützend in Form einer Tabelle zur Verfügung gestellt (s. Anhang A im Online-Supplement):

Prozedurale Informationen: Ziel des Fortbildungsmoduls ist die Ausbildung von Routinen, die Lehrkräfte dazu befähigen, die Mehrdeutigkeit in Diagnosesituationen zu erkennen. Die Ausbildung dieser Routinen erfolgt im Fortbildungsmodul mittels Schritt-für-Schritt-Instruktionen. Diese Instruktionen sind integraler Bestandteil der Lernaufgabe und stellen die prozeduralen Informationen dar.

Schritt 3: Konkrete Umsetzung der Designkomponenten des 4C/ID-Modells im Fortbildungsmodul

Im Folgenden wird jeweils eine Lernaufgabe pro Aufgabenklasse exemplarisch dargestellt.

Lernaufgabe 1 der Aufgabenklasse 1 (Abb. 4) zielt darauf ab, Lehrkräfte zu befähigen, Fehlermuster in Aufgabenbearbeitungen von Schüler*innen zu identifizieren, indem sie durch Instruktion geschult werden, ihre Aufmerksamkeit auf die relevanten Informationen der Diagnosesituation zu lenken (TK 1).

Die Komplexität der Lernaufgabe ist niedrig, da die zu analysierenden Aufgabenbearbeitungen eindeutig einem Fehlkonzzept zugeordnet werden können. Das Fehlkonzzept ist bekannt und muss ausschließlich überprüft (belegt oder widerlegt) werden.

Das Prüfen des Fehlkonzzeptes wird Schritt für Schritt angeleitet und orientiert sich an der in Kapitel 3.3 vorgestellten Debiasing-Strategien (a), Hypothesen zu verifizieren, indem nach diesen bestätigenden Informationen gesucht wird. Lernaufgabe 1 hat damit das Höchstmaß an instruktionaler Unterstützung.

Aufgabenklasse 1, Lernaufgabe 1: Testen vorgegebener Hypothesen, Wahrnehmen relevanter Informationen

Versetzen Sie sich in die Situation der Lehrkraft. Sie haben mit Ihrer Lerngruppe ein Zufallsexperiment durchgeführt mit dem Ziel, das konzeptionelle Verständnis Ihrer Schüler*innen zum statistischen Denken sichtbar zu machen.

Um herauszufinden, ob Alex bereits tragfähige Konzepte zum statistischen Denken aufgebaut hat, analysieren Sie ihre bzw. seine Aufschriebe zum Würfelexperiment, die im Unterricht entstanden sind.



NOAH

Ausgang	1	2	3	4	5	6
Häufigkeit	12	11	10	11	12	14

Was fällt dir auf?

*das was oft raus kommt und unterschätzt
die Ergebnisse sind*

Versuche deine Entdeckungen zu begründen:

*weil ich immer mehr gewürfelt habe
manchmal schnell und manchmal
langsam und deswegen kommen
manchmal verschiedene Ergebnisse*

Nach Pratt, 2000

Sie stellen fest, dass Noahs Entdeckung (Was fällt dir auf?) von der theoretischen Wahrscheinlichkeit der Augenzahlen eines Würfels (Gleichwahrscheinlichkeit) abweicht.

Hypothese: Sie haben die Begründung zu Noahs Entdeckung gelesen und vermuten, dass dieser das Fehlkonzent **Ergebnisbeeinflussungsheuristik** zugrunde liegt.

Prüfen Sie diese Hypothese. Gehen Sie dabei wie folgt vor:

- Gehen Sie zunächst davon aus, dass die Hypothese wahr ist. Welche Begründung würden Sie von Noah zu ihrer bzw. seiner Entdeckung erwarten?
Notieren Sie Ihrer Erwartung entsprechend eine möglichen Begründung:
- Vergleichen Sie die erwartete Begründung mit der tatsächlichen Begründung von Noah. Stimmen diese (inhaltlich) überein?

Kreuzen Sie an:

- ☐ Die erwartete Begründung deckt sich (inhaltlich) mit der tatsächlichen Begründung von Noah. Die Hypothese, dass Noah dem Fehlkonzent **Ergebnisbeeinflussungsheuristik** unterliegt, bleibt weiterhin plausibel.
- ☐ Die erwartete Begründung stimmt mit der tatsächlichen Begründung von Noah nicht (konsequent) überein. Die Hypothese, dass Noah dem Fehlkonzent **Ergebnisbeeinflussungsheuristik** unterliegt, wird somit widerlegt.

Abbildung 4: Testen vorgegebener Hypothesen, Wahrnehmen relevanter Informationen (eigene Darstellung)

Lernaufgabe 1 der Aufgabenklasse 2 (Abb. 5) zielt darauf ab, Lehrkräfte zu befähigen, das Zustandekommen von identifizierten Fehlermustern in Aufgabenbearbeitungen von Schüler*innen zu erklären, indem zugrunde liegende Fehlkonzent auf der Grundlage fachdidaktischen Wissens rekonstruiert werden (TK 2).

Die Komplexität der Lernaufgaben nimmt zu, da die zu analysierenden Aufgabenbearbeitungen zwar eindeutig einem Fehlkonzent zugeordnet werden können, das zugrunde liegende Fehlkonzent jedoch nicht mehr bekannt ist und ergründet werden muss.

Hierzu müssen Hypothesen über mögliche Fehlkonzent aufgestellt und getestet werden. Durch das Aufstellen mehrerer Hypothesen wird neben der Verifikation von Hypothesen (Lernaufgabe der Aufgabenklasse 1) auch deren Falsifikation trainiert. Die Falsifikation von Hypothesen, also das Suchen nach einer Hypothese widerlegenden Informationen, orientiert sich an der in Kapitel 3.3 vorgestellten Debiasing-Strategie (b).

Das Prüfen angenommener Fehlkonzent wird in dieser Aufgabenklasse vorausgesetzt und nicht mehr instruiert.

Aufgabenklasse 2, Lernaufgabe 1: Hypothesen aufstellen und testen, Ursachen rekonstruieren

Auch Alex hat 30-mal gewürfelt. Sie erinnern sich an ein Unterrichtsgespräch mit Alex. Alex' Erklärungen zum Würfelexperiment gaben Hinweise darauf, dass sie bzw. er noch keine fachlich tragfähigen Konzepte zum statistischen Denken aufgebaut haben könnte.

Um auszuschließen, dass Alex einem Fehlkonzent unterliegt, untersuchen Sie ihre bzw. seine Aufschriebe zum Würfelexperiment, die im Unterricht entstanden sind.



ALEX

Augenzahlen	1	2	3	4	5	6
Häufigkeit						

Was fällt dir auf?
Die 3 wurde bei mir am meisten gewürfelt.

Versuche deine Entdeckungen zu begründen.
Selbst wenn ich noch 100 Mal würfeln dürfte, glaube ich, dass die 3 immer noch vorne wäre.

Nach Pratt, 2000

1. Untersuchen Sie, ob der Begründung von Alex ein Fehlkonzent zugrunde liegen könnte. Gehen Sie dabei wie folgt vor:
 - a. Welche Begründung ist zu erwarten, wenn das Fehlkonzent *Verfügbarkeitsheuristik* vorliegt?
 - b. Welche Begründung ist zu erwarten, wenn das Fehlkonzent *Repräsentativitätsheuristik* vorliegt?
 - c. ...
2. Welche dieser Fehlkonzente deckt sich (inhaltlich) mit der tatsächlichen Begründung von Alex?

Aus meinen Analysen hat sich ergeben, dass Alex ...

Abbildung 5: Hypothesen aufstellen und testen, Ursachen rekonstruieren (eigene Darstellung)

Lernaufgabe 1 der Aufgabenklasse 3 (Abb. 6) zielt darauf ab, eine Sensibilität für die Mehrdeutigkeit von Diagnosesituationen zu schaffen, indem sie Lehrkräfte auffordert, neben dem angenommenen Fehlkonzent (Verifikation und Falsifikation von Hypothesen) auch immer die gesamte Bandbreite alternativer Fehlkonzente zu erproben.

Die Komplexität der Lernaufgabe ist hoch, da das zu treffende diagnostische Urteil mit Unsicherheiten behaftet ist. Anders als in den Lernaufgaben der Aufgabenklassen 1 und 2 sind die Informationen, die zum Testen von Hypothesen genutzt werden, nicht mehr eindeutig. Eindeutige Informationen bestätigen oder widerlegen Hypothesen. In den Lernaufgaben der Aufgabenklasse 3 sind alle Informationen, die in der Diagnosesituation zu finden sind, mehrdeutig und bestätigen unterschiedliche Hypothesen.

Um kein Fehlkonzent zu übersehen, um also Alternativhypothesen auszuschließen, muss die gesamte Bandbreite möglicher Fehlkonzente getestet werden. Der Ausschluss alternativer Hypothesen orientiert sich an der in Kapitel 3.3 vorgestellten Debiasing-Strategie (c).

Das Rekonstruieren von Fehlkonzenten und deren Prüfung wird vorausgesetzt (Teilkompetenzen 1 und 2). Das Maß an Unterstützung und Anleitung wird vollständig zurückgenommen.

Aufgabenklasse 3, Lernaufgabe 1: Alternativhypothesen erproben, Mehrdeutigkeit von Fehlermustern erkennen

Fehlermuster in Lösungen von Schüler*innen lassen nicht immer eindeutige Diagnosen über zugrunde liegende Fehlkonzepte zu. So kann ein und dasselbe Fehlermuster aus unterschiedlichen Fehlkonzepten hervorgehen. Wird die Ursachensuche zu früh abgebrochen und werden alternative Fehlkonzepte nicht überprüft, können fehlerhafte diagnostische Urteile resultieren.

Auch Bente hat 30-mal mit einem Würfel gewürfelt und seine Entdeckung aufgeschrieben und begründet.



BENTE		1	2	3	4	5	6
Augenrollen							
Ergebnis							

Was fällt dir auf?
das, da 6 öfter vorkommt als die anderen Zahlen

Versuche deine Entdeckungen zu begründen.
Ich denke, dass mein Würfel gerollt ist, weil die 6 eigentlich nicht so häufig vorkommt

Nach Pratt, 2000

Untersuchen Sie, welche Fehlkonzepte den Entdeckungen von Bente zugrunde liegen können.

Aus meinen Analysen hat sich ergeben, dass Bente ...

Abbildung 6: Alternativhypothesen erproben, Mehrdeutigkeit von Fehlermustern erkennen (eigene Darstellung)

5 Evaluation des Fortbildungsmoduls

Das Fortbildungsmodul zur Förderung der diagnostischen Kompetenz, die Mehrdeutigkeit in Diagnosesituationen zu erkennen, wurde mit 40 Lehrkräften erprobt. Die Lehrkräfte unterrichten Mathematik an Gemeinschaftsschulen und verfügen über eine durchschnittliche Berufserfahrung von 8,9 Jahren ($SD = 3,09$).

Da nicht das Diagnostizieren spezifischer Fehlkonzepte bei einem Lerngegenstand, sondern allgemein das Erkennen der Mehrdeutigkeit in Diagnosesituationen und das Aufstellen und Testen alternativer Hypothesen gefördert werden soll, unterscheiden sich der Prä- und Posttest inhaltlich vom Fortbildungsmodul. Während im Prä- und Posttest aufgefordert wird, Aufgabenbearbeitungen zum Größenvergleich von Brüchen zu analysieren, werden im Fortbildungsmodul Aufgabenbearbeitungen zur Einschätzung von Wahrscheinlichkeiten analysiert. So wird verhindert, dass die Ergebnisse des Fortbildungsmoduls memoriert und im Posttest abgerufen werden können. Die im Bereich der Wahrscheinlichkeiten erworbenen Kompetenzen (Teilkompetenzen 1 bis 3, s. Kap. 4.2) müssen eigenständig auf den Bereich der Brüche übertragen werden. So wird geprüft, ob das Fortbildungsmodul eine situationsübergreifende Kompetenz fördert.

5.1 Durchführung des Fortbildungsmoduls

Die Lehrkräfte wurden mit typischen Fehlern und ihren Ursachen beim Größenvergleich von Brüchen vertraut gemacht (fachdidaktisches Wissen; s. Anhang B im Online-Supplement). In derselben Woche wurden die Lehrkräfte gebeten, das vermittelte fachdidaktische Wissen schriftlich wiederzugeben. Die Aufforderung lautete: „*Sie haben einen kurzen Input zu typischen Fehlern und ihren Ursachen beim Größenvergleich von Brüchen erhalten. An welche können Sie sich erinnern? Benennen und charakterisieren Sie diese.*“ Alle teilnehmenden Lehrkräfte konnten die relevanten Fehlkonzepte korrekt be-

nennen und beschreiben. Hierdurch wurde sichergestellt, dass das erforderliche fachdidaktische Wissen vorhanden und ein fehlendes Erkennen mehrdeutiger Diagnosesituationen im Posttest auf Urteilsverzerrungen zurückzuführen ist.

In der darauffolgenden Woche nahmen die Lehrkräfte an einem Prätest teil, welcher dazu aufforderte, jeweils fünf (fehlerhafte) Lösungen von insgesamt drei Schüler*innen zum Größenvergleich von Brüchen zu analysieren, Fehlermuster zu identifizieren und Diagnosen über zugrunde liegende Fehlkonzepte zu formulieren. Dabei kann jede Lösung aus drei verschiedenen Fehlkonzepten (Größerer Nenner = größerer Bruch, Größerer Zähler = größerer Bruch, Sowohl größerer Zähler als auch größerer Nenner = größerer Bruch) resultieren (mehrdeutige Diagnosesituation). Die folgende Abbildung zeigt beispielhaft das angewendete Format (Testvignette):

Noah		Analysieren Sie die Lösungen von Noah. Was fällt Ihnen auf? Woran liegt das?
a)	$\frac{3}{4} < \frac{5}{7}$	✗
b)	$\frac{1}{3} < \frac{3}{5}$	✓
c)	$\frac{7}{14} > \frac{6}{12}$	✗
d)	$\frac{8}{11} < \frac{10}{15}$	✗
e)	$\frac{2}{4} < \frac{6}{14}$	✗

Wie sicher sind Sie sich mit Ihrer Analyse?
(0 bis 100 %)

Abbildung 7: Im Prä-Posttest angewendetes Format (eigene Darstellung)

Eine Woche nach dem Prätest wurden die Lehrkräfte zufällig in zwei gleich große Gruppen (Kontroll- und Trainingsgruppe) aufgeteilt. Die Kontrollgruppe ($N = 20$) bekam erneut den Auftrag, dieselben Lösungen der Schüler*innen der letzten Woche (Prätest = Posttest) zu analysieren, Fehlermuster zu identifizieren und Diagnosen über zugrunde liegende Fehlkonzepte zu formulieren.

Die Trainingsgruppe ($N = 20$) nahm vor dem erneuten Diagnoseauftrag an dem oben beschriebenen Fortbildungsmodul zur Diagnose von Fehlkonzepten beim Einschätzen von Wahrscheinlichkeiten in mehrdeutigen Diagnosesituationen (s. Kap. 4.2) teil.

5.2 Evaluation der Wirksamkeit des Fortbildungsmoduls

Zur Untersuchung der Wirksamkeit des Fortbildungsmoduls auf das Erkennen der Mehrdeutigkeit in Diagnosesituationen wurden die im Prä- und Posttest gebildeten Urteile der Lehrkräfte analysiert. Dabei wurde die Anzahl der von ihnen als Ursache diagnostizierten Fehlkonzepte erfasst. Die folgenden Beispiele illustrieren die Vorgehensweise.

Die Lehrkraft SB24 hat die Mehrdeutigkeit des identifizierten Fehlermusters in den Lösungen von Alex, Bente und Noah vollständig erkannt (Abb. 8). Für jedes diagnostizierte Fehlkonzept erhält sie einen Punkt. Da die Lösungen der untersuchten Schüler*innen jeweils aus drei unterschiedlichen Fehlkonzepten hervorgehen können, erzielte SB24 insgesamt 9 von 9 möglichen Punkten.

Alex	Bente	Noah
SB24: „Das Größerzeichen hat die Öffnung immer bei den größeren Zahlen. Es ist unklar, ob es am Nenner , Zähler oder beidem liegt. Das Kind hat das Konzept von Brüchen jedoch nicht verstanden.“	SB24: „Bente hat das Konzept von Brüchen wahrscheinlich nicht verstanden. Das Größerzeichen zeigt immer bei den größeren Zahlen den größeren Bruch an. Es ist unklar, ob es auf den Zähler , den Nenner oder beide zurückzuführen ist.“	SB24: „Noah hat das Größerzeichen immer zu der Seite geöffnet, auf der die größeren Zahlen stehen. Es ist unklar, ob es am Zähler , Nenner oder an beidem liegt. Das Konzept der Brüche wurde nicht verstanden.“
Anzahl identifizierter Fehlkonzpte: 3	Anzahl identifizierter Fehlkonzpte: 3	Anzahl identifizierter Fehlkonzpte: 3
Gesamtpunktzahl: 9		

Abbildung 8: Auswertungsmethode der Prä-und Posttestergebnisse_SB24 (eigene Darstellung)

Die Lehrkraft CR25 hat die Mehrdeutigkeit des Fehlermusters in den Lösungen von Alex, Bente und Noah nicht erkannt (Abb. 9). Pro Schüler*in identifizierte sie jeweils nur eines von möglichen drei Fehlkonzpten. Insgesamt erreichte CR25 somit 3 von maximal 9 Punkten.

Alex	Bente	Noah
CR25: „Alex hat die Fehlvorstellung, dass wenn der Nenner größer ist auch der Bruch im Allgemeinen größer sein muss .“	CR25: „Bente denkt ebenso wie Alex, dass wenn der Nenner größer ist der Bruch auch größer sein muss als der andere Bruch mit dem niedrigeren Nenner.“	CR25: „ Nenner größer = Bruch größer Gleich wie bei Alex & Bente → alle haben dieselbe Fehlvorstellung“
Anzahl identifizierter Fehlkonzpte: 1	Anzahl identifizierter Fehlkonzpte: 1	Anzahl identifizierter Fehlkonzpte: 1
Gesamtpunktzahl: 3		

Abbildung 9: Auswertungsmethode der Prä-und Posttestergebnisse_CR25 (eigene Darstellung)

Die erzielten Punkte (Anzahl der diagnostizierten Fehlkonzpte) im Prätest wurden mit den erzielten Punkten im Posttest verglichen. Da eine Normalverteilung der Daten nicht gegeben ist (getestet anhand des Kolmogorov-Smirnov-Tests), wurde der Mann-Whitney U-Test für nicht normalverteilte Daten verwendet.

6 Ergebnisse

Während sich in der Anzahl der diagnostizierten Fehlkonzpte bei den Lehrkräften der Kontrollgruppe im Prä-Posttest-Vergleich keine signifikanten Unterschiede zeigten, diagnostizierten Lehrkräfte der Trainingsgruppe im Posttest nach der Fortbildung signifikant mehr Fehlkonzpte als im Prätest ($U = 349.0$, $Z = 4.343$, $p = 0.001$), die sie als Ursache für das identifizierte Fehlermuster in den Lösungen der untersuchten Schüler*innen benannten (Abb. 10).

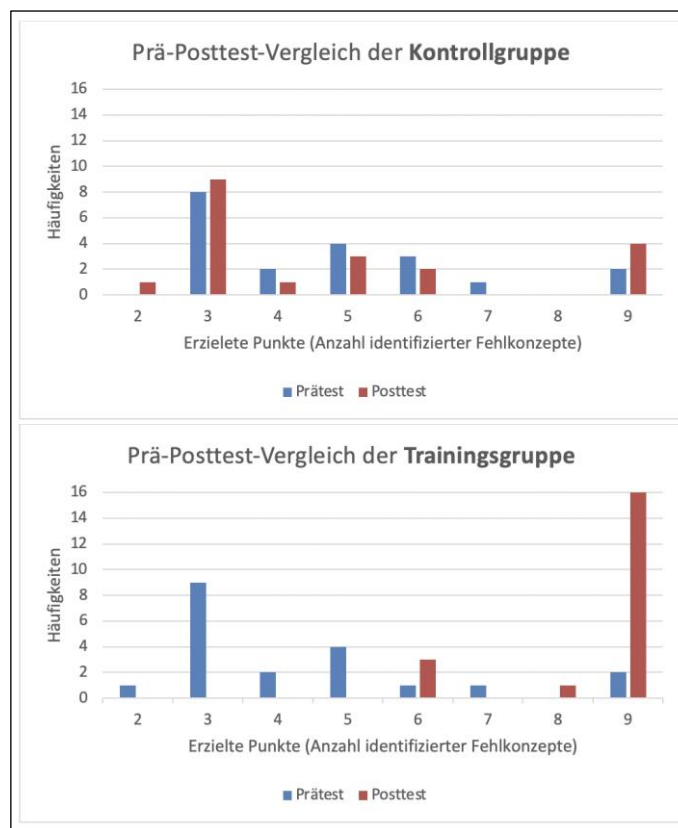


Abbildung 10: Verteilung der Prä-Posttest-Ergebnisse der Kontrollgruppe (oben) und der Trainingsgruppe (unten) (eigene Darstellung)

Die Lehrkräfte der Kontrollgruppe identifizierten im Posttest durchschnittlich 4,8 Fehlkonzepte, wohingegen die Lehrkräfte der Trainingsgruppe durchschnittlich 8,5 Fehlkonzepte (von max. 9) identifizierten (Abb. 11). Das Fortbildungsmodul erzielte eine große Effektstärke ($r = 0,69$) auf die Anzahl der identifizierten Fehlkonzepte.

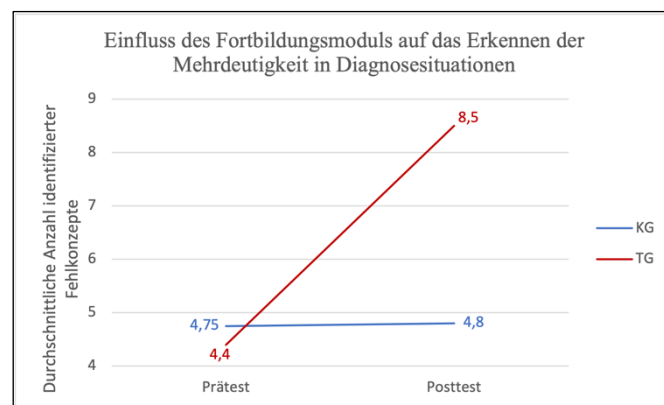


Abbildung 11: Durchschnittliche Prä-Posttest-Ergebnisse der Kontrollgruppe und der Trainingsgruppe im Vergleich (eigene Darstellung)

7 Diskussion

Nur 4 von 40 Lehrkräften erkannten im Prätest die Mehrdeutigkeit der identifizierten Fehlermuster in den Lösungen der untersuchten Schüler*innen zum Größenvergleich von Brüchen vollständig und identifizierten alle drei zugrunde liegenden Fehlkonzepte als mögliche Ursachen. Die Ergebnisse des Prätests zeigen, dass das Diagnostizieren von Fehlkonzepten in mehrdeutigen Diagnosesituationen eine kognitiv anspruchsvolle Aufgabe ist, die gezielt gefördert werden muss (vgl. Grossman et al., 2009; Rieu et al., 2024a, 2024b).

Das vorgestellte Fortbildungsmodul zielt darauf ab, die Kompetenz von Lehrkräften, die Mehrdeutigkeit in Diagnosesituationen zutreffend zu erkennen, zu fördern. Dabei stand insbesondere die Sensibilisierung der Lehrkräfte für die Mehrdeutigkeit von Fehlermustern in Aufgabenbearbeitungen durch Debiasing-Strategien (umgesetzt in den Lernaufgaben) im Vordergrund.

Die Wirksamkeit des Fortbildungsmoduls wurde mit 40 Lehrkräften erprobt. Die Ergebnisse bestätigen die positive Wirkung des Fortbildungsmoduls auf die diagnostische Kompetenz der Lehrkräfte, die Mehrdeutigkeit in Diagnosesituationen zu erkennen (s. Dekomposition der Zielkompetenz in Kap. 4.2). Vor dem Fortbildungsmodul diagnostizierten nur 2 von 20 Lehrkräften der Trainingsgruppe alle möglichen Fehlkonzepte als Ursache für die identifizierten Fehlermuster in den Aufgabenbearbeitungen der drei Schüler*innen zutreffend, während nach der Fortbildung 16 von 20 Lehrkräften alle möglichen Fehlkonzepte zutreffend diagnostizierten. In der Kontrollgruppe zeigte sich keine signifikante Verbesserung in der Anzahl der diagnostizierten Fehlkonzepte im Prä-Posttest-Vergleich (2 zutreffende Diagnosen von insgesamt 20 im Prätest, 4 zutreffende Diagnosen von 20 im Posttest).

In Kapitel 2 wurde an einer idealtypischen Diagnose aufgezeigt, wie fachdidaktisches Wissen – also das Wissen über typische Fehler und ihre Ursachen – das Diagnostizieren von Fehlkonzepten in mehrdeutigen Diagnosesituationen unterstützt: Es lenkt die Aufmerksamkeit auf relevante Informationen der Diagnosesituation und hilft bei der Organisation der Ursachensuche durch Festlegen eines Interpretationsrahmens. Die Ergebnisse des Fortbildungsmoduls sind jedoch nicht allein durch das fachdidaktische Wissen zu erklären. Obwohl allen Lehrkräften (auch den Lehrkräften in der Kontrollgruppe) das notwendige fachdidaktische Wissen vermittelt und alle Lehrkräfte im Test alle vermittelten typischen Fehler und ihre Ursachen zum Größenvergleich von Brüchen wiedergeben konnten, unterschieden sich die Diagnoseleistungen der Trainingsgruppe signifikant von den Leistungen der Kontrollgruppe im Posttest. Zum einen unterstreicht dieses Ergebnis die positive Wirkung des Fortbildungsmoduls auf die diagnostische Kompetenz der Lehrkräfte, zum anderen zeigt es, dass die bloße Vermittlung des fachdidaktischen Wissens nicht ausreicht. Auch die Sensibilisierung für die Mehrdeutigkeit von Diagnosesituationen und insbesondere die Vermittlung von Strategien, diese Mehrdeutigkeit zu erkennen (Teilkompetenzen 1 bis 3), ist notwendig.

Rieu et al. (2024a, 2024b) kommen zu ähnlichen Ergebnissen: Obwohl Lehrkräften beim Analysieren von fehlerhaften Aufgabenbearbeitungen zum Größenvergleich von Dezimalzahlen das notwendige fachdidaktische Wissen schriftlich zur Verfügung stand, brachen viele die Informationssuche häufig nach dem ersten identifizierten Fehlkonzept frühzeitig ab. Da so die Mehrdeutigkeit der Diagnosesituation nicht erkannt wurde, resultierten unzulässige diagnostische Urteile.

Vergleichbare Ergebnisse finden sich in der medizinischen Diagnostik, die Analogien zur pädagogischen Diagnostik aufweist (s. Kap. 3.2). Dort zeigt sich, dass das bloße Wissen über Urteilsfehler nicht zu deren Reduzierung führt (Stacy & MacMillan, 1995). Zur Reduzierung solcher Fehler werden Debiasing-Strategien (korrektive Strategien;

z.B. Heuer, 1999) eingesetzt. Das 4C/ID-Modell bietet hier eine Strukturierung zur Konzeption solcher Fortbildungen (van Merriënboer & Sweller, 2010; van Merriënboer, 2020).

Debiasing-Strategien der Medizin, Hypothesen zu verifizieren, zu falsifizieren und alternative Hypothesen auszuschließen, die insbesondere zur Überwindung des Bestätigungsfehlers und des Search Satisficings und damit der Vernachlässigung relevanter Informationen beim Bilden diagnostischer Urteile entwickelt wurden (z.B. Gäbler, 2017; Croskerry et al., 2013), haben sich in dem vorliegenden Fortbildungsmodul auch in der pädagogischen Diagnostik bewährt.

Hervorzuheben ist zudem die situationsübergreifende Anwendbarkeit der im Fortbildungsmodul vermittelten Kompetenz, die Mehrdeutigkeit in Diagnosesituationen zu erkennen. Das Fortbildungsmodul erfolgte im Inhaltsbereich „Einschätzen von Wahrscheinlichkeiten“, der Prä- und Posttest im Inhaltsbereich „Größenvergleich von Brüchen“. So wird verhindert, dass die Ergebnisse des Fortbildungsmoduls memoriert und im Posttest abgerufen werden können. Die im Bereich der Einschätzung von Wahrscheinlichkeiten erworbenen Kompetenzen (Teilkompetenzen 1 bis 3) mussten eigenständig auf den Bereich der Brüche übertragen werden. Die signifikante Verbesserung der Trainingsgruppe ist daher nicht auf Erinnerungseffekte zurückzuführen, sondern bestätigt die Förderung der situationsübergreifenden Kompetenz, die Mehrdeutigkeit in Diagnosesituationen zu erkennen.

Das Ausrichten unterrichtlicher Aktivitäten an den Lernvoraussetzungen der Schüler*innen erfordert nach dem Erkennen, die Auflösung der Mehrdeutigkeit in Diagnosesituationen. Hierfür kamen im Fortbildungsmodul diagnostische Aufgaben (Kron et al., 2021) zum Einsatz, die eindeutige Rückschlüsse auf zugrunde liegende Fehlkonzepte zulassen. Anders als das Erkennen der Mehrdeutigkeit in Diagnosesituationen ist das Auswählen und Konzipieren solcher diagnostischen Aufgaben inhaltspezifisch. Entsprechend sind keine Transfereffekte zwischen verschiedenen Inhaltsbereichen zu erwarten. Die Kompetenz, die Mehrdeutigkeit in Diagnosesituationen aufzulösen wurde aus diesem Grund nicht evaluiert.

Ein nächster Schritt in der Weiterentwicklung des Fortbildungsmoduls bestand darin, Lehrkräften die Möglichkeit zu geben, mit authentischen Lösungen ihrer eigenen Schüler*innen zu arbeiten. Dieser praxisnahe Zugang könnte die Relevanz und Nachhaltigkeit des Trainings zusätzlich erhöhen, da er die Anwendung der erlernten Debiasing-Strategien im eigenen Unterrichtskontext ermöglicht. Entsprechende Erweiterungen wurden beispielsweise in der Fortbildungsreihe „Daten und Zufall für alle: Passgenau und mit digitalen Werkzeugen unterrichten“ umgesetzt. Das dafür entwickelte Material steht unter <https://hse-heidelberg.de/daten-zufall-fobi> zur Verfügung.

Zusammengefasst bietet das vorgestellte Fortbildungsmodul, das sich an den 4C/ID-Designprinzipien nach van Merriënboer orientiert und Debiasing-Strategien der Medizin zur Überwindung der Vernachlässigung relevanter Informationen integriert, eine fundierte und wirksame Grundlage, Lehrkräfte auf die anspruchsvolle Aufgabe vorzubereiten, die Mehrdeutigkeit in Diagnosesituationen zutreffend zu erkennen.

Literatur und Internetquellen

- Borko, H. (2004). Professional Development and Teacher Learning: Mapping the Terrain. *Educational Researcher*, 33 (8), 3–15. <https://doi.org/10.3102/0013189X033008003>
- Croskerry, P. (2002). Achieving Quality in Clinical Secision Making: Cognitive Strategies and Detection of Bias. *Academic Emergency Medicine*, 9, 1184–1204. <https://doi.org/10.1197/aemj.9.11.1184>
- Croskerry, P. (2003). Cognitive Forcing Strategies in Clinical Decision Making. *Annals of Emergency Medicine*, 41 (1), 110–120. <https://doi.org/10.1067/mem.2003.22>

- Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive Debiasing 2: Impediments to and Strategies for Change. *British Medical Journal Quality and Safety*, 22 (2), 65–72. <https://doi.org/10.1136/bmjqs-2012-001713>
- Eichelmann, A., Narciss, S., Schnaubert, L. & Melis, E. (2012). Typische Fehler bei der Addition und Subtraktion von Brüchen – Ein Review zu empirischen Fehleranalysen. *Journal für Mathematik-Didaktik*, 33 (1), 29–57. <https://doi.org/10.1007/s13138-011-0031-5>
- Faller, H. & Lang, H. (Hrsg.). (2019). *Medizinische Psychologie und Soziologie*. Springer. <https://doi.org/10.1007/978-3-662-57972-5>
- Gäbler, M. (2017). Denkfehler bei diagnostischen Entscheidungen. *Wiener Medizinische Wochenschrift*, 167, 1–10. <https://doi.org/10.1007/s10354-017-0570-6>
- Graber, M.L., Franklin, N. & Gordon, R. (2005). Diagnostic Error in Internal Medicine. *Archives of Internal Medicine*, 165 (13), 1493–1499. <https://doi.org/10.1001/archinte.165.13.1493>
- 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. <https://doi.org/10.1177/016146810911100905>
- Häsel-Weide, U., & Prediger, S. (2017). Förderung und Diagnose im Mathematikunterricht – Begriffe, Planungsfragen und Ansätze. In M. Abshagen, B. Barzel, J. Kramer, T. Riecke-Baulecke, B. Rösken-Winter & C. Selter (Hrsg.), *Basiswissen Lehrerbildung: Mathematik unterrichten mit Beiträgen für den Primar- und Sekundarstufenbereich* (S. 167–181). Kallmeyer.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77 (1), 81–122. <https://doi.org/10.3102/003465430298487>
- Helmke, A. (2010). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Kallmeyer.
- Helmke, A. (2022). *Unterrichtsqualität und Professionalisierung: Diagnostik von Lehr-Lern-Prozessen und evidenzbasierte Unterrichtsentwicklung*. Klett/Kallmeyer.
- Heuer Jr, R.J. (1999). Analysis of Competing Hypotheses. *Psychology of Intelligence Analysis*, 95–110.
- Hogarth, R.M. (1987). *Judgement and Choice: The Psychology of Decision*. John Wiley & Sons.
- Kaufmann, S. & Wessolowski, S. (2021). *Rechenstörungen: Diagnose und Förderbausteine*. Klett/Kallmeyer.
- Kovacs, G. & Croskerry, P. (1999). Clinical Decision Making: An Emergency Medicine Perspective. *Academic Emergency Medicine*, 6 (9), 947–952. <https://doi.org/10.1111/j.1553-2712.1999.tb01246.x>
- Kron, S., Sommerhoff, D. & Achtner, M. (2021). Selecting Mathematical Tasks for Assessing Student's Understanding: Pre-service Teachers' Sensitivity to and Adaptive Use of Diagnostic Task Potential in Simulated Diagnostic One-to-one Interviews. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.604568>
- Leuders, T. (2020). Kohärenz und Professionsorientierung in der universitären Lehrerbildung. Hochschuldidaktische Impulse durch das 4C/ID-Modell. In J. Kreutz, T. Leuders & K. Hellmann (Hrsg.), *Professionsorientierung in der Lehrerbildung. Kompetenzorientiertes Lehren nach dem 4-Component-Instructional-Design-Modell*, (S. 7–24). Springer Fachmedien.
- Leuders, T. & Loibl, K. (2021). Beyond Subject Specificity – Student and Teacher Thinking as Sources of Specificity in Teacher Diagnostic Judgments. *RISTAL – Research in Subject-matter Teaching and Learning*, 4, 60–70.
- Loibl, K., Leuders, T. & Dörfler, T. (2020). A Framework for Explaining Teachers' Diagnostic Judgements by Cognitive Modeling (DiaCoM). *Teaching and Teacher Education*, 9 (3), 1–10. <https://doi.org/10.1016/j.tate.2020.103059>

- Moser Opitz, E. (2010). Diagnose und Förderung: Aufgaben und Herausforderungen für die Mathematikdidaktik und die mathematikdidaktische Forschung. In A.M. Lindmeier & S. Ufer (Hrsg.), *Beiträge zum Mathematikunterricht*, (S. 11–18). WTM.
- Padberg, F. & Wartha, S. (2017). *Didaktik der Bruchrechnung* (5. Aufl.). Springer Spektrum. <https://doi.org/10.1007/978-3-662-52969-0>
- Pratt, D. (2000). Making Sense of the Total of Two Dice. *Journal of Research in Mathematical Education*, 31 (5), 602–625. <https://doi.org/10.2307/749889>
- Prediger, S. & Wittmann, G. (Hrsg.). (2009). Falsch bringt weiter – Aus Fehlern lernen. *Praxis der Mathematik in der Schule*, 51 (27), 1–8.
- Prediger, S., Tschierschky, K., Wessel, L. & Seipp, B. (2012). Professionalisierung für fach- und sprachintegrierte Diagnose und Förderung im Mathematikunterricht: Entwicklung und Erprobung eines Konzepts für die universitäre Fachlehrausbildung. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 17 (1), 40–58.
- Rieu, A., Loibl, K. & Leuders, T. (2024a) How Context Influences the Processing of Relevant Information and Judgment Accuracy – The Role of Information Restriction in Judgment Processes in Diagnosing Misconceptions. *Frontiers in Psychology*, 15, 15:1405756. <https://doi.org/10.3389/fpsyg.2024.1405756>
- Rieu, A., Leuders, T. & Loibl, K. (2024b). Urteilsverzerrungen beim Diagnostizieren von Fehlkonzepten bei Dezimalbrüchen. *Journal für Mathematik-Didaktik*, 45 (7). <https://doi.org/10.1007/s13138-024-00231-x>
- Schrader, F.-W. (2013). Diagnostische Kompetenz von Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 31 (2), 154–165. <https://doi.org/10.36950/bzl.31.2013.9646>
- Stacy, W. & MacMillan, J. (1995). Cognitive Bias in Software Engineering. *Communications of the ACM*, 38 (6), 57–63. <https://doi.org/10.1145/203241.203256>
- Stark, R., Gruber, H., Mandl, H. & Hinkofer, L. (2001). Wege zur Optimierung eines beispielbasierten Instruktionsansatzes: Der Einfluss multipler Perspektiven und instruktionaler Erklärungen auf den Erwerb von Handlungskompetenz. *Unterrichtswissenschaft*, 29 (1), 26–40.
- Thiede, K.W., Brendefur, J.L., Osguthorpe, R.D., Carney, M.B., Bremner, A., Strother, S., Oswalt, S., Snow, J.L., Sutton, J. & Jesse, D. (2015). Can Teachers Accurately Predict Student Performance? *Teaching and Teacher Education*, 49, 36e44. <https://doi.org/10.1016/j.tate.2015.01.012>
- Tietze, U.-P. (1988). Schülerfehler und Lernschwierigkeiten in Algebra und Arithmetik – Theoriebildung und empirische Ergebnisse aus einer Untersuchung. *Journal für Mathematik-Didaktik*, 9, 163–204. <https://doi.org/10.1007/BF03339290>
- Valentin, A. (2009). Entscheidungen in kognitiv komplexen Situationen. *Intensivmedizin*, 46, 308–312. <https://doi.org/10.1007/s00390-009-0048-x>
- Van Es, E.A. & Sherin, M.G. (2002). Learning to Notice: Scaffolding New Teachers' Interpretations of Classroom Interactions. *Journal of Technology and Teacher Education*, 10 (4), 571–596.
- Van Es, E.A. & Sherin, M.G. (2021). Expanding on Prior Conceptualizations of Teacher Noticing. *ZDM – Mathematics Education*, 53, 17–27. <https://doi.org/10.1007/s11858-020-01211-4>
- Van Merriënboer, J.J.G. & Sweller, J. (2010). Cognitive Load Theory in Health Professional Education: Design Principles and Strategies. *Medical Education*, 44 (1), 85–93. <https://doi.org/10.1111/j.1365-2923.2009.03498.x>
- Van Merriënboer, J.J.G. (2019). Das Vier-Komponenten Instructional Design (4C/ID) Modell. In H. Niegemann & A. Weinberger (Hrsg.), *Lernen mit Bildungstechnologien. Konzeption und Einsatz digitaler Lernumgebungen*. Springer. https://doi.org/10.1007/978-3-662-54373-3_8-1

- Van Merriënboer, J.J.G. (2020). Das Vier-Komponenten Instructional Design (4C/ID) Modell. *Handbuch Bildungstechnologie: Konzeption und Einsatz digitaler Lernumgebungen*, (S. 153–170). Springer. https://doi.org/10.1007/978-3-662-54368-9_8
- Van Ophuysen, S. (2010). Professionelle pädagogisch-diagnostische Kompetenz – eine theoretische und empirische Annäherung. In N. Berkemeyer, W. Bos, H.G. Holtappels, N. McElvany & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (Bd. 16, S. 203–234). Juventa.
- von Aufschnaiter, C., Cappell, J., Dübbelde, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., Sträßer, R. & Wolgast, A. (2015). Diagnostische Kompetenz: Theoretische Überlegungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik*, 61 (5), 738–758. <https://doi.org/10.25656/01:12723>
- Weinert, F.E. (2000). Lehren und Lernen für die Zukunft – Ansprüche an das Lernen in der Schule. *Pädagogische Nachrichten Rheinland-Pfalz*, 2, 1–16.

Förderhinweis

Das Fortbildungsmodul entstand im Rahmen des vom Ministerium für Wissenschaft, Forschung und Kunst des Landes Baden-Württemberg geförderten Forschungs- und Nachwuchskollegs „Diagnostische Kompetenzen von Lehrkräften (DiaKom II): Diagnostische Urteilsprozesse als Informationsverarbeitung und die Bedeutung von Personen- und Situationsmerkmalen“.

Beitragsinformationen

Zitationshinweis:

Özel, E., Leuders, T. & Loibl, K. (2025). Mehrdeutigkeit in Diagnosesituationen erkennen. Entwicklung, Durchführung und Evaluation eines Fortbildungsmoduls zum Erwerb diagnostischer Kompetenzen. *HLZ – Herausforderung Lehrer*innenbildung*, 8 (1), 457–480. <https://doi.org/10.11576/hlz-7268>

Eingereicht: 04.06.2024 / Angenommen: 10.07.2025 / Online verfügbar: 07.11.2025

ISSN: 2625–0675



Dieses Werk ist freigegeben unter der Creative-Commons-Lizenz CC BY-SA 4.0 (Weitergabe unter gleichen Bedingungen). Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Tabellen, Zitate etc.) sind von der CC-Lizenz ausgenommen. Für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen. <https://creativecommons.org/licenses/by-sa/4.0/legalcode.de>

English Information

Title: Accurately Diagnosing Misconceptions in Ambiguous Diagnostic Situations: Development, Implementation and Evaluation of a Training Program to Support Diagnostic Competencies

Abstract: This article demonstrates how teachers can use pedagogical content knowledge to accurately identify students' misconceptions, particularly in cases where an identified error pattern in problem-solving can be explained by different underlying misconceptions. Thus, we focus on ambiguous diagnostic situations.

First, an ideal diagnostic judgment process is presented, using pedagogical content knowledge to accurately identify misconceptions in ambiguous diagnostic situations in the context of fraction comparison and probability. Relevant competencies of teachers are derived. To foster these competencies, a training program is developed using vignette-based learning tasks.

The effectiveness of the developed training program was empirically tested in an experimental study with 40 teachers. The results of this study confirm the positive effect of the training program (compared with a control group) on teachers' diagnostic competencies in accurately identifying misconceptions in ambiguous diagnostic situations.

Keywords: diagnostics; competencies; subject didactics, teacher education; ambiguity; awareness; misconception