

Arbeitskreis: Psychologie und Mathematikdidaktik

Würzburg, 17./18. 10. 2025

Janina Krawitz und Silke Neuhaus-Eckhardt

Die Herbsttagung des Arbeitskreises Psychologie und Mathematikdidaktik fand in diesem Jahr vom 17.–18. Oktober in Würzburg statt und wurde von Silke Neuhaus-Eckhardt organisiert. Insgesamt nahmen 40 Kolleg:innen aus 15 Universitätsstandorten teil. Im Mittelpunkt standen vier empirische Beiträge, die sich mit soziokulturellen Normen, der Bearbeitung offener realitätsbezogener Aufgaben, der Förderung mathematischer Kompetenzen wie dem bayesianischen Denken sowie der KI-gestützten Vorhersage individueller Lernverläufe befassten. Die Tagung knüpfte damit an die Tradition des Arbeitskreises an, empirische und psychologische Fragestellungen in einer intensiven Diskussionsatmosphäre zu beleuchten.

Den Auftakt machte am Freitag Josephine F. Paul (FAU Erlangen-Nürnberg) mit einem Vortrag zur Bedeutung soziokultureller Normen für die Wahrnehmung von Unterricht in Deutschland und Taiwan. Sie zeigte dabei, wie kulturell geprägte Vorstellungen von qualitativ hochwertigem Unterricht systematisch erfasst und mit Methoden wie Epistemischen Netzwerk-Analysen hinsichtlich ihrer Wirkung auf Teacher Noticing untersucht werden können. Im zweiten Beitrag stellten Nils Völlinger und Stanislaw Schukajlow (Universität Münster) ein Forschungsprojekt zu offenen realitätsbezogenen Aufgaben vor. Präsentiert wurden sowohl ein systematisches Literaturreview zu offenen realitätsbezogenen Problemen als auch die Konzeption einer empirischen Studie, die Lösungsprozesse und typische Schwierigkeiten bei Modellierungsaufgaben mit offenem Endzustand untersucht.

Am Samstagmorgen präsentierten Theresa Büchter und Andreas Eichler (Universität Kassel) ein Projekt zur Förderung des bayesianischen Denkens. Im Mittelpunkt stand die Frage, wie Visualisierungen und natürliche Häufigkeiten das bayesianische Denken unterstützen können – ein Thema von hoher gesellschaftlicher Relevanz, etwa im medizinischen oder juristischen Kontext. Die Ergebnisse zeigten deutliche Unterschiede in der Wirksamkeit verschiedener Visualisierungen und Trainings mit absoluten Häufigkeiten und unterstrichen die Bedeutung der kombinierten Nutzung dieser beiden Zugänge. Im letzten Vortrag stellte David Bednorz (IPN Kiel) eine Studie vor, in der mithilfe KI-gestützter Modelle individuelle Lernverläufe von

Schüler:innen in einer digitalen Lernumgebung untersucht wurden. Ziel war es, Lernende mit einem Risiko für pseudostrukturelles Wissen frühzeitig zu identifizieren. Die Studie zeigte eine hohe Sensitivität der Modelle und verdeutlichte, dass frühe Schwierigkeiten im Wissenserwerb prädiktiv für spätere Fehlkonzeptionen sind.

Auch der informelle Austausch kam nicht zu kurz: Während der Kaffeepausen und insbesondere beim gemeinsamen Abendessen am Freitagabend im Wirtshaus am Dom bot sich reichlich Gelegenheit, die Vorträge weiterzudenken und fachliche Gespräche zu vertiefen.

Wir möchten uns bei allen Vortragenden, die sich im Rahmen der Herbsttagung dem intensiven Vortragsformat gestellt haben, herzlich bedanken. Gleichzeitig danken wir allen Teilnehmenden für ihre Fragen, Kommentare und Anregungen. Ein besonderer Dank gilt der Würzburger Arbeitsgruppe für die hervorragende Organisation und die äußerst angenehme Tagungsatmosphäre.

Im Folgenden geben die Vortragenden der Herbsttagung selbst einen kurzen Einblick in ihre Beiträge sowie in die Anregungen, die sie aus den Diskussionen mitnehmen konnten.

Josephine F. Paul, Anika Dreher, Ting-Ying Wang, Feng-Jui Hsieh und Anke Lindmeier (FAU Erlangen-Nürnberg, PH Freiburg, NTNU Taipei): Die Bedeutung soziokultureller Normen, individueller Perspektiven und situationaler Einflüsse für die Wahrnehmung von Unterricht – Ein Vergleich zwischen Deutschland und Taiwan

Lehrkräfte müssen im Unterricht Ereignisse erkennen und deuten, die für eine hohe Unterrichtsqualität relevant sind – beispielsweise, wann explizitere Erklärungen gegeben oder eingefordert werden sollten. Diese Fähigkeit zur professionellen Unterrichtswahrnehmung (Noticing; Sherin et al., 2011) wurde in den vergangenen 20 Jahren insbesondere durch individuelle Dispositionen erklärt (z. B. professionelles Wissen; König et al., 2022). Jedoch vernachlässigt dieser Erklärungsansatz situationale Einflüsse (z. B. Salienz). Auch kann mangelndes Wissen die bei internationalen Vergleichen festgestellten Unterschiede nur unzureichend erklären. Eine treffendere Erklärung scheinen soziokul-



Foto: Silke Neuhäus-Eckhardt

Gruppenfoto der Herbsttagung des AK Psychologie und Mathematikdidaktik, Oktober 2025

turelle Unterschiede zu sein (Goodwin, 1994). Um Noticing also kultursensibel messen zu können, müssen die verschiedenen Vorstellungen zu qualitativ hochwertigem Unterricht expliziert werden. Im TaiGer Noticing Projekt wurden dazu mithilfe von 19 taiwanischen und 17 deutschen Mathematikdidaktikprofessor:innen Unterrichtsqualitätsnormen zum Umgang mit Aufgabenpotenzialen validiert (unter anderem eine deutsche, eine taiwanische und eine geteilte Norm, Paul et al., 2024). Weitere unterrichtliche Aspekte, die den Professor:innen als relevant erschienen, wurden expliziert. Um zu bestimmen, ob die Fokussierung bestimmter Aspekte tatsächlich kulturell statt durch andere Faktoren bedingt war, wurden logistische Regressionen genutzt (Variablen: Herkunftsland, Situation und vier individuelle Charakteristika).

In der zweiten Projektphase wurde das Noticing von 115 taiwanischen und 124 deutschen Sekundarstufenlehrkräften bezüglich des normativen Referenzrahmens untersucht. Weitere Aspekte, die die Lehrkräfte adressierten, wurden induktiv bestimmt. Situationale und kulturelle Unterschiede wurden deskriptiv und mit Epistemischen Netzwerk Analysen (ENA) untersucht. In den Ergebnissen zeichnete sich die Relevanz kulturell-bedingter Normen für Teacher Noticing ab. Beispielsweise nahmen taiwanische Lehrkräfte einen Verstoß gegen die taiwanische Unterrichtsqualitätsnorm deutlich häufiger wahr als deutsche Lehrkräfte. Ähnliches zeigte sich auch für weitere unterrichtliche Aspekte (z. B. häufigere Fokussierung der taiwanischen Teilnehmenden auf Lernendenaktivitäten). Regressionen und ENA machten außerdem situationale Unterschiede sichtbar. Bei den Professor:innen stellte sich zudem eine – zwar weniger starke – Rele-

vanz individueller, erfahrungsbasierter Charakteristika dar.

Kernpunkte der Diskussion und Anregungen

Da im Projekt ein Referenzrahmen mithilfe von Mathematikdidaktikprofessor:innen validiert wurde, wurde darauf verwiesen, dass auch Perspektiven anderer Lehrkräftebildender (z. B. Fachleitungen) einbezogen werden hätten können. Zudem wurde die Frage aufgeworfen, ob sich die Antwortqualität zwischen den beiden Lehrkräftegruppen unterschied. Mit Bezug zur interkulturellen Forschung stellte sich die Frage, welche Normen künftig in Vergleichsstudien angelegt werden sollten – das heißt, ob ein Referenzrahmen aus geteilten Normen oder verschiedene kulturspezifische Referenzrahmen genutzt werden sollten. Daran anknüpfend wurde überlegt, ob es bestimmte Normen gibt, die *besser* als andere sind und ob dies die Entscheidung bedingen sollte, welche Normen in Vergleichsstudien angelegt werden. Abschließend wurde diskutiert, ob und inwiefern *bessere* Normen aus einem Land in einem anderen übernommen werden könnten.

Nils Völlinger, Stanislaw Schukajlow, Janina Krawitz, Lieven Verschaffel und Katrin Rakoczy (Universität Münster, Universität Köln, KU Leuven, Universität Gießen): *Offenheit realitätsbezogener Probleme: Ein Literaturreview und eine Studie zu Lösungsprozessen*

Offene Probleme aus der realen Welt spielen eine wichtige Rolle beim Lehren und Lernen von Mathematik. Offene Probleme enthalten oftmals nicht alle zur Lösung der Aufgabe notwendigen Informationen, erläut-

ben multiple Lösungswege sowie mehrere korrekte Lösungen (Schukajlow et al., 2023; Silver, 1995). Im ersten Teil des Vortrags wurde ein systematisches Literaturreview zum aktuellen Stand der Forschung zu offenen realitätsbezogenen Problemen im Mathematikunterricht vorgestellt. Basierend auf einer Volltextanalyse von 67 Zeitschriftenartikeln untersuchte das Review die deskriptiven Merkmale von Studien in diesem Bereich, die Verwendung von Frameworks, die zur Definition der Offenheit realitätsbezogener mathematischer Probleme verwendet werden, die Merkmale und Struktur der verwendeten Aufgaben sowie die in den Studien einbezogenen Konstrukte. Die Ergebnisse zeigten, dass ein Großteil der Studien Aufgaben verwendete, bei denen Informationen zur Lösung der Aufgabe fehlten. Weiterhin adressierten die bisherigen Studien vorwiegend kognitive Konstrukte, während affektive Konstrukte deutlich seltener untersucht wurden. Der zweite Teil des Vortrags widmete sich der Vorstellung eines neuen Forschungsprojektes zu Modellierungsaufgaben mit offenem Endzustand und geschlossenem Anfangszustand. Dabei handelt es sich um Aufgaben, die eine vage formulierte Fragestellung (z. B. Lohnt sich ...?) umfassen, aber alle zur Lösung der Aufgabe benötigten Daten enthalten. Zur Lösung dieses Aufgabentyps müssen verschiedene Faktoren, die den Endzustand der Aufgaben bestimmen, identifiziert und in einem mathematischen Modell in Bezug gesetzt werden. Im Mittelpunkt standen Einblicke in die erste qualitative Studie des Projekts, die sich mit der Analyse von Lösungsprozessen und typischen Schwierigkeiten bei Modellierungsaufgaben mit einem offenen Endzustand befasst.

Kernpunkte der Diskussion und Anregungen

In der anschließenden Diskussion wurde der Punkt aufgegriffen, welche Rolle der Realitätsbezug bei den offenen Aufgaben spielt. Insbesondere wurde angemerkt, dass das kontextbezogene Interesse in zukünftigen Untersuchungen miteinbezogen werden könnte, da dies eine bedeutende Rolle bei der Bearbeitung von realitätsbezogenen Aufgaben spielt. In diesem Zusammenhang wurde auch angemerkt, dass die Operationalisierung des offenen Endzustandes der Modellierungsaufgaben durch vage formulierte Fragestellungen (s. o.) insofern schwierig ist, dass Formulierungen wie „Lohnt sich das?“ kulturell determiniert sein könnten. Auch wurde hier die Frage aufgeworfen, ob hinsichtlich der Identifikation von Faktoren, die den Endzustand der Aufgaben bestimmen, Reihenfolgeeffekte zu erwarten sind. Ein weiterer Diskussionspunkt war, ob es überhaupt möglich ist, Modellierungsaufgaben mit geschlossenem Anfangszustand zu entwickeln.

Es wurde angemerkt, dass dies vermutlich nur für die Vollständigkeit numerischer Daten zu erreichen ist.

Theresa Büchter, Andreas Eichler, Karin Binder, Katharina Böcherer-Linder, Stefan Krauss, Michael Rößner, Nicole Steib und Markus Vogel (Universität Kassel, Universität Paderborn, Universität Freiburg, Universität Regensburg, LMU München, PH Heidelberg): *Wirksame Strategien für das Bayesianische Denken aus einer ganzheitlichen Perspektive*

Bayesianisches Denken ist oft mit Missverständnissen verbunden, die regelmäßig zum Beispiel in der Medizin zu falschen Diagnosen oder zu fehlerhaften Gerichtsentscheidungen in der Rechtsprechung führen (Gigerenzer et al., 2007). Bayesianisches Denken ist in vielen Professionen relevant und repräsentiert einen wichtigen Aspekt der Wahrscheinlichkeitsrechnung. Bei einer typischen Bayesianischen Aufgabe wird ein Risiko (z. B. positiv prädiktiver Wert) auf Basis von drei statistischen Informationen (z. B. Basisrate, Richtig-Positiv-Rate und Falsch-Positiv-Rate) berechnet. Obwohl es viele Forschungsarbeiten zur Performanz bei der Berechnung eines positiv prädiktiven Werts gibt (McDowell & Jacobs, 2017), fehlen noch Forschungsarbeiten, die die optimale Unterstützung des Bayesianischen Denkens durch Visualisierung und Trainings untersuchen (Spiegelhalter et al. 2011; Feufel et al. 2023).

In dem Vortrag haben wir ein Programm zur Förderung des Bayesianischen Denkens vorgestellt. Dazu gehören neben Vorstudien eine längsschnittliche Trainingsstudie, in der wir die beiden bislang erfolgreichsten Strategien – natürliche Häufigkeiten und Visualisierungen – kombiniert und systematisch mit jeweils fünf Experimentalgruppen (vier verschiedene Trainingsgruppen und eine Kontrollgruppe) mit jeweils $n = 260$ Medizinstudierenden und $n = 255$ Jurastudierenden verglichen haben. Die Manipulation zwischen den vier Trainingsgruppen ergab sich durch die Nutzung von natürlichen Häufigkeiten (ja/nein) bzw. einer Visualisierung (Doppelbaum/Einheitsquadrat/Wahrscheinlichkeitsbaum). Weiterhin zählt eine Querschnittsstudie mit $N = 2400$ Studierenden verschiedener Fachrichtungen zu dem Programm. Dort haben wir unterschiedliche Strategien bei der Lösung Bayesianischer Aufgaben untersucht, wobei dabei die Unterstützung durch sieben Visualisierungen und drei Informationsformate systematisch variiert wurde.

Als Ergebnisse präsentierten wir, wie sich die Wirksamkeit der vier Trainings unterscheidet (Steib et al., 2025). Erklärungen für die beobachteten Unterschiede wurden durch qualitative Analysen der Notizen und durch den Verlauf der (fehlerhaften) Strategien ergänzt. Die Analyse von (fehlerhaften) Strategien beim

Lösen Bayesianischer Aufgaben in der Querschnittsstudie ergab einerseits eine Replikation des Format-effekts (natürliche Häufigkeiten vs. Wahrscheinlichkeiten) und andererseits Hinweise darauf, dass korrekte und fehlerhafte Strategien nicht nur vom Informationsformat, sondern auch von der präsentierten Visualisierung abhängen.

Kernpunkte der Diskussion und Anregungen

In der Diskussion wurden aus unterschiedlichen Perspektiven die Implikationen der Ergebnisse für das Lernen über Wahrscheinlichkeiten in der Schule diskutiert. Hierzu wurde diskutiert, inwiefern ein Transfer der Ergebnisse in die Schulpraxis gelingen kann und möglicherweise auch weitere Forschungsarbeiten an der Schule notwendig sind. Dazu wurde auch diskutiert, welcher Grad der Formalisierung bzw. Algorithmisierung in den schulischen Lernprozessen wünschenswert ist und wie man das Wissen über die Nutzung (fehlerhafter) Strategien in den Lernprozess integrieren kann.

David Bednorz, Daniel Sommerhoff und Aiso Heinze (IPN Kiel): *Wer lernt wie? Analysen und Vorhersage individueller Lernverläufe in einer digitalen Lernumgebung*

Nationale und internationale Leistungsstudien, wie beispielsweise PISA (OECD, 2023), zeigen, dass der Erwerb mathematischer Begriffe Lernenden häufig unzureichend gelingt. Digitale Lernumgebungen ermöglichen, den Wissenserwerb der Lernenden detailliert nachzuvollziehen und Lernprozesse gezielt zu fördern. Eine zentrale Voraussetzung hierfür ist die Gestaltung von Lernaktivitäten auf Basis expliziter Hypothesen zu Begriffserwerbsprozessen. Theoretische Ansätze wie das Prozess-Objekt-Modell von Sfard (1991) liefern hierfür konzeptuelle Grundlagen, indem sie zwischen zunächst erworbenem operationalem Wissen und darauf aufbauendem strukturalen Wissen unterscheiden. Diese Entwicklung verläuft in drei Phasen: Interiorization, Condensation und Reification. Abweichungen von dieser Struktur werden in der Literatur als pseudostrukturelles Wissen bezeichnet (Litteck et al., 2024; Sfard, 1991).

Im Vortrag wurde eine Machbarkeitsstudie vorgestellt, die darauf abzielte, das Potenzial KI-gestützter Modelle zur Vorhersage von Lernverläufen im regulären Mathematikunterricht zu erkunden. Dafür wurden Daten von $N = 365$ Lernenden aus 15 Klassen über einen Zeitraum von fünf Wochen analysiert. Im Fokus stand die Frage, inwieweit Lernende mit einem Risiko für den Erwerb pseudostrukturellen Wissens auf Basis individueller Lernverläufe erkannt werden konnten

und welche Rolle die Sfard-Phasen dabei spielten. Die Ergebnisse zeigten über alle Phasen hinweg eine sehr hohe Sensitivität ($\geq .90$), es wurden also nahezu alle Lernenden mit Risiko für den Erwerb pseudostrukturellen Wissens korrekt identifiziert. Die Präzision der Modelle, d. h. die Wahrscheinlichkeit, dass Lernende, die laut Modell ein Risiko aufweisen, auch wirklich ein Risiko haben, lag dagegen nur im mittleren Bereich ($\geq .23$). Dies entspricht einer relativ hohen Anzahl falsch positiver Zuordnungen. Diese Asymmetrie zwischen Sensitivität und Präzision ist jedoch für ein diagnostisches Frühwarnsystem prinzipiell funktional: Entscheidend ist, möglichst wenige gefährdete Lernende zu übersehen (hohe Sensitivität), selbst wenn dadurch einige unbedenkliche Fälle zusätzlich markiert werden.

Die Analysen zur Bedeutung der Phasen nach Sfard verdeutlichten zudem, dass insbesondere Merkmale aus der frühen Phase des Differenzenquotienten eine zentrale Rolle für die Risikoerkennung spielen. Besonders deutlich wurde der Einfluss der Condensation-Phase, als in den Modellen getestet wurde, wie sich Veränderungen in der Gewichtung der verschiedenen Phasen auswirken. Wurden Merkmale dieser Phase manipuliert, sank die Sensitivität am deutlichsten. Diese Befunde unterstreichen, dass frühe Schwierigkeiten im Wissenserwerb prädiktiv für den späteren Erwerb pseudostrukturellen Wissens sind und Interventionen bereits in dieser frühen Phase ansetzen sollten.

Kernpunkte der Diskussion und Anregungen

In der anschließenden Diskussion wurden methodische und perspektivische Fragen aufgegriffen, insbesondere zur kontinuierlichen Datenerfassung im schulischen Kontext und zur Abwägung zwischen Machbarkeit, Datenumfang und analytischem Mehrwert. Betont wurde, dass die Studie zwar wenige Lernende, aber sehr viele Datenpunkte pro Person umfasste, was tiefe Analysen individueller Lernverläufe ermöglicht, jedoch die Generalisierbarkeit einschränkt. Zugleich wurde kritisch diskutiert, inwieweit das Training der KI-Modelle unter diesen Bedingungen überhaupt verlässlich gelingen kann. Abschließend wurde betont, dass langfristige Erhebungsformate wertvolle Einblicke in die Identifikation von Lernenden mit Risiko für unzureichenden Wissenserwerb bieten und dass Machbarkeitsstudien eine wichtige Grundlage für die Weiterentwicklung solcher diagnostischen Ansätze darstellen.

Organisatorisches und Ausblick

Silke Neuhaus-Eckhardt (Universität Würzburg) wurde als Sprecherin des Arbeitskreises wiedergewählt.

Die nächste Tagung wird vom 23.–24. Oktober 2026 in München stattfinden.

Haben Sie Lust bekommen, an unserer Tagung teilzunehmen und mitzudiskutieren? Auf der Homepage des AK können Sie sich in den E-Mail-Verteiler eintragen. Dort finden Sie zudem aktuelle Informationen zu vergangenen und kommenden Tagungen: akpsy.didaktik-der-mathematik.de

Wenn Sie vortragen möchten, melden Sie sich bitte per E-Mail an eine der Sprecherinnen Silke Neuhaus-Eckhardt (neuhaus@dmu.wu.de) oder Janina Krawitz (janina.krawitz@uni-koeln.de). Die Teilnehmenden unserer Herbsttagung interessieren sich vornehmlich für Studien, bei denen die Bezugsdisziplin Psychologie eine Rolle spielt. Bis zu vier Arbeiten, die eher fortgeschritten oder kurz vor dem Abschluss sind, können vorgestellt werden, egal ob es ein Promotionsprojekt, Ausschnitt aus einer laufenden Studie oder eine Arbeit im Publikationsprozess ist. Sie sollten dazu bereit sein, die Arbeiten im Sinne eines ausführlichen Werkstattberichts zur Diskussion zu stellen. Unterjährig wird der AK Psychologie und Mathematikdidaktik voraussichtlich keine weitere planmäßige Aktivität anbieten.

Gemeinsames Literaturverzeichnis

- Feufel, M. A., Keller, N., Kendel, F., & Spies, C. D. (2023). Boosting for insight and/or boosting for agency? How to maximize accurate test interpretation with natural frequencies. *BMC Medical Education*, 23(1), 75.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological science in the public interest*, 8(2), 53–96.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606–633. DOI: [10.1525/aa.1994.96.3.02a00100](https://doi.org/10.1525/aa.1994.96.3.02a00100)
- König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., & Kaiser, G. (2022). Teacher noticing: A systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educational Research Review*, 36, 100453. DOI: [10.1016/j.edurev.2022.100453](https://doi.org/10.1016/j.edurev.2022.100453)
- Litteck, K., Rolfes, T., & Heinze, A. (2025). The structure of knowledge about the concept of derivative – a study investigating a process-object framework. *International Journal of Mathematical Education in Science and Technology*, 56(11), 2164–2184. DOI: [10.1080/0020739X.2024.2397990](https://doi.org/10.1080/0020739X.2024.2397990)
- McDowell, M., & Jacobs, P. (2017). Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychological bulletin*, 143(12), 1273.
- OECD (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD. [10.1787/53f23881-en](https://doi.org/10.1787/53f23881-en)
- Paul, J. F., Dreher, A., Wang, T.-Y., Hsieh, F.-J., & Lindmeier, A. (2024). Culture-specific norms regarding high-quality use of task potential for mathematical learning – Contrasting researchers’ perspectives from Germany and Taiwan. *Journal für Mathematik-Didaktik*, 45(2), 14. DOI: [10.1007/s13138-024-00237-5](https://doi.org/10.1007/s13138-024-00237-5)
- Schukajlow, S., Krawitz, J., Kanefke, J., Blum, W., & Rakoczy, K. (2023). Open modelling problems: Cognitive barriers and instructional prompts. *Educational Studies in Mathematics*, 114(3), 417–438. DOI: [10.1007/s10649-023-10265-6](https://doi.org/10.1007/s10649-023-10265-6)
- Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22(1), 1–36. DOI: [10.1007/BF00302715](https://doi.org/10.1007/BF00302715)
- Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011). Situating the study of teacher noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp, *Mathematics Teacher Noticing*. Routledge.
- Silver, E. A. (1995). The nature and use of open problems in mathematics education: Mathematical and pedagogical perspectives. *ZDM – Mathematics Education*, 95(2), 67–72.
- Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, 333(6048), 1393–1400.
- Steib, N., Büchter, T., Eichler, A., Binder, K., Krauss, S., Böcherer-Linder, K., Vogel, M. & Hilbert, S. (2025). How to teach Bayesian reasoning: An empirical study comparing four different probability training courses. *Learning and Instruction*, 95, 102032.

Janina Krawitz, Universität zu Köln
janina.krawitz@uni-koeln.de

Silke Neuhaus-Eckhardt,
Julius-Maximilians-Universität Würzburg
silke.neuhaus-eckhardt@uni-wuerzburg.de