Stefan UFER, LMU München

Wer kann es? Interindividuelle Unterschiede beim mathematischen Beweisen – zwischen Annahmen und Evidenz

1 Mathematisches Argumentieren und Beweisen

Dass Argumentieren einen Kern mathematischen Arbeitens in allen Altersgruppen ausmacht ist wohl unbestritten. Konsens scheint zu sein, dass sich mathematisches Arbeiten eben nicht nur auf die Konstruktion formaler, rein deduktiver und Axiomen beruhender Schlussketten beschränkt. Im Gegenteil rückt in der *Philosophy of Mathematical Practice* vermehrt in den Blick, wie eine mathematische Theorie als Gefüge von Axiomen, Definitionen, Sätzen und Beweisen entsteht. Exemplarisch dafür steht die Frage, inwiefern informelles Arbeiten wie das Ableiten von Vermutungen als Teil wissenschaftlich-mathematischer Arbeit gesehen werden kann und soll (z.B. Atiyah et al., 1994). Während einige Autorinnen und Autoren hier explizit zwischen dem informellen Argumentieren und dem „eigentlichen mathematischen Arbeiten" (z.B. im Sinne des Beweisens) trennen (z.B. Balacheff, 1999), sehen andere Potential in einer Verknüpfung explorativer und systematisierender Prozesse im Sinne einer Cognitive Unity (Pedemonte, 2007). Schwarz et al. (2010) unterscheiden diese explorativen Anteile, die beispielsweise das Generieren von Hypothesen umfassen, und systematisierende Anteile, wie das Generieren von logischen Schlussketten innerhalb einer Rahmentheorie, als zwei von drei zentralen Teilprozessen mathematischen Argumentierens. Parallel widmen sich auch empirische Arbeiten beispielsweise der Nutzung von Beispielen durch Mathematikerinnen und Mathematiker bzw. Lernende (Lockwood et al., 2012).

Vor diesem Hintergrund stellt sich die Frage, was eigentlich einen „validen Beweis" als solchen kennzeichnet. Reale Beweise sind ja keine streng formal notierten logischen Schlussketten (sogenannte „Derivations"), oder wie Mac Lane (1981) schreibt: „Most verbal or written mathematical proofs are simply sketches which give enough detail to indicate how a full rigorous proof might be constructed." Wie ausführlich diese "sketches" sein müssen kann zwischen Teilfeldern der Mathematik, aber auch je nach Situation (z.B. Forschung oder Lehre; Sommerhoff & Ufer, 2019) variieren und ist als soziomathematische Norm im jeweiligen Kontext auszuhandeln. Entsprechend kann auch nicht erwartet werden, dass sich „völlig logisch" oder „selbsterklärend" erschließt, was als valider mathematischer Beweis zählt und was nicht, sondern dies ist für Lernende zunächst häufig ein Prozess der Enkulturation. Individuelles Wissen über Normen und Kriterien für akzeptable Beweise sowie über die Funktionen von Beweisen scheint damit eine erste zentrale Voraussetzung für einen erfolgreichen Umgang mit Beweisen (de Villiers, 1991). Schwarz et al. (2010) heben in diesem Sinne das „Aufschreiben" (Inscribing) von Beweisen und Argumentationen als wesentlichen dritten Teilprozess beim mathematischen Argumentieren heraus.

In: Kerstin Hein, Cathleen Heil, Silke Ruwisch & Susanne Prediger (Hrsg.). Beiträge zum Mathematikunterricht 2021. Münster: WTM Verlag. https://doi.org/10.37626/GA9783959871846.0
Online unter https://eldorado.tu-dortmund.de/handle/2003/30630

Andreas Stylianides (2007) schlägt eine Definition des Begriffs „Beweis“ vor, die diesen sozialen Konsens mit in Betracht zieht. Ein Beweis in seinem Sinne ist eine zusammenhängende Kette von Schlüssen, die (i) auf einem Satz akzeptierter Aussagen (Rahmentheorie) beruht, als valide vereinbarte (z.B. weitgehend deduktive) Schlussformen benutzt und in einer vereinbarten Form (z.B. unter adäquater Nutzung einschlägiger Schreibweisen) kommuniziert werden. Häufig wird noch ein Kriterium der „prinzipiellen Formalisierbarkeit“ mit angelegt, das eine gewissen Anschlussfähigkeit an den Theorieaufbau der wissenschaftlichen Mathematik sicherstellen soll. Diese Definition gibt viel Raum für eine ganze Bandbreite auch „didaktisch orientierter Beweiskonzepte“ (Brunner, 2014; Kempen, 2019). Dieser Beitrag fokussiert auf Beweiskonzepte die in eine relativ gut strukturierte Rahmentheorie eingebettet sind und algebraische Darstellungsformen nutzen um Allgemeingültigkeit zu kommunizieren.

2 Interindividuelle Unterschiede beim Umgang mit Beweisen

2.1 Potential von Studien zu interindividuellen Unterschieden

Anfang der 1980 Jahre legten Sharon Senk und Zalman Usiskin etwa 2700 Lernende der Jahrgangsstufen 7 bis 12 im Cognitive Development and Achievement in Secondary School Projekt Beweisaufgaben zur Geometrie vor (Senk, 1989). Wie auch spätere Untersuchungen (Reiss et al., 2002) zeigen sich insgesamt niedrige Leistungen zum Beweisen, jedoch auch substantielle interindividuelle Unterschiede. Während etwa zwei Drittel der Lernenden durchaus einfache, einschrittige Beweise führen konnten, stellten mehrschrittige Beweise und das Einzeichnen von Hilfslinien eine substantielle Herausforderung dar. Senk (1989) berichtet, dass diese interindividuellen Unterschiede mit der Leistung in einem Test zu den van Hiele Niveaus geometrischen Denkens und einem standardisierten Geometrietest in Verbindung gebracht werden können. Eine klare Trennung der beiden Einflussfaktoren war jedoch nicht möglich.

Die Studie von Senk ist ein erstes Beispiel für die Untersuchung interindividueller Unterschiede zum Beweisen. Auch wenn sie erstmals überhaupt systematische interindividuelle Unterschiede beim Beweisen aufzeigt, muss kritisch gesehen werden, dass die Wirkmechanismen hinter den beiden Erklärungsvariablen nicht genauer beschrieben und alternative Erklärungen für interindividuelle Unterschiede nicht herangezogen werden. Auch konfundierende Variablen, die allen anderen erhobenen Konstrukten zu Grunde liegen könnten, wurden nicht erhoben. Abgesehen von dieser Kritik bieten Untersuchungen zu interindividuellen Unterschieden jedoch die Möglichkeit, gezielt verschiedene Erklärungen für die Ursachen dieser Unterschiede zu kontrastieren. Auch wenn sie nicht wie experimentelle Studien kausale Evidenz liefern können und auch eine Detailanalyse von Wirkmechanismen – wie sie mit qualitativen Fallanalysen möglich wäre – nicht leisten, bieten sie oft doch einen ersten Zugang in Felder, in denen eine große Bandbreite von Erklärungsansätzen diskutiert wird.

2.2 Erklärungen für interindividuelle Unterschiede beim Beweisen und Argumentieren

Die Ausführungen unter 1. weisen auf eine erste mögliche Ursache für interindividuelle Unterschiede beim Beweisen hin. Relevant erscheint, inwiefern Lernende die (lokalen) Normen und Kriterien für valide Beweise sowie die Funktionen von Beweisen in der Mathematik verstanden haben. Chazan (1993) charakterisiert die Bandbreite entsprechender Vorstellungen von Lernenden anhand einer Interviewstudie. Manche Lernende weisen etwa Einzelbeispiele aus falschen Gründen zurück, z.B. weil das Messen an einer Figur in der Geometrie fehlerbehaftet sei. Andere Lernende schließen die Existenz von Gegenbeispielen auf der Basis eines Beweises nicht aus, weil der Beweis auf Annahmen beruht, deren Status als unklar wahrgenommen wird. In einigen Arbeiten wird dieses komplexe Bild auf eine einzige Fehlvorstellung reduziert, die Stylianides et al. (2017, S. 120) als „key and persistent problem“ identifizieren: „Many students have the misconception that a few confirming examples suffice to prove the truth of a mathematical generalization.” (s.a. EMS, 2011). Andere Arbeiten bezweifeln, dass die beobachteten Probleme wirklich auf eine fundamentale Fehlvorstellung zum Beweiskonzept zurückgeführt werden können. Sie legen nahe, dass es sich tragfähige Beweisversuche handeln könnte, die scheitern „because…the student lacks either the skill or the will“ dies auch umzusetzen (Inglis & Alcock, 2012). Allgemeine Aussagen zu zeigen erfordert eben Wissen über spezifische Prinzipien, wie sie Durand-Guerrier et al. (2011) auch explizit formulieren. Entsprechend ist die Frage, ob nicht Wissen über beweisspezifische Argumentationsstrategien und ihnen unterliegende Prinzipien viel relevanter sein könnte als abstraktes Wissen über die Funktion von Beweisen in der Mathematik.

Im Überblick werden in der Literatur vielfältige weitere Voraussetzungen für ein erfolgreiches Umgehen mit Beweisen diskutiert (s.a. Sommerhoff, 2017). Schoenfeld (1992) nennt beispielsweise vier Bereiche von Voraussetzungen für erfolgreiches mathematisches Problemlösen im Allgemeinen: *Wissen*, das ganz spezifisch mit dem mathematischen Inhaltbereich verknüpft ist; *Heurismen* und *Kontrollmechanismen*, die sehr allgemein über verschiedene Inhaltsbereiche und sogar außerhalb der Mathematik nutzbar sein sollten; und letztlich *Überzeugungen* dazu, was mathematisches Arbeiten ausmacht, wie man an mathematische Aufgaben herangehen sollte. Aus den oben beschriebenen Arbeiten zum *Beweiskonzept* wird weiter klar, dass auch Wissen über dieses ein wesentlicher Einflussfaktor sein könnte. Auch *affektive Merkmale* wie Interesse, Selbstkonzept oder Identität der Lernenden werden wiederholt als relevante Merkmale benannt. Dass Fähigkeiten zum *logischen Schließen* Unterschiede erklären könnten erscheint auf den ersten Blick plausibel und wird z.B. in den genannten Arbeiten von Durand-Guerrier et al. (2011) weiter ausgeführt.

3 Individuelle Ressourcen und interindividuelle Unterschiede

Inwiefern diese sogenannten Ressourcen nun für den Umgang mit Beweisen relevant sind, kann auch davon abhängen, *wie* konkret mit den Beweisen gearbeitet werden muss. Gibt man Lernenden eine (beweisbare) mathematische Aussage vor, mit dem Auftrag sie zu beweisen, so spricht man von *Beweiskonstruktion*. Bewertet werden diese Beweise meist anhand ihrer inhaltlichen Argumentation und weniger anhand ihrer formal korrekten Darstellung. In jedem Fall setzen derartige Studien voraus, dass eine lokale Kultur zu einer definierten Rahmentheorie, akzeptierten Argumenttypen und Normen zum notwendigen Grad formaler Notationen etabliert ist.

In einer Anschlussstudie an das BiQua-Projekt haben wir (Ufer et al., 2008) Lernenden der Sekundarstufe I Beweisaufgaben zur Geometrie vorgelegt und unabhängig davon deren konzeptuelles und prozedurales Wissen erhoben sowie Problemlösekompetenzen mit mathematiknahen, aber inhaltsfernen Problemlöseaufgaben. In dieser Studie, wie auch in einer parallelen Studie mit Lernenden in Taiwan, zeigt sich, dass alle drei Ressourcen interindividuelle Unterschiede in der Beweisleistung erklärten, wobei der Zusammenhäng für konzeptuelles Wissen am stärksten war. Sehr ähnliche Ergebnisse finden Chinnappan, Ekanayake & Brown (2012) mit Lernenden der Jgst. 11 in Sri Lanka. In einer Reanalyse von Daten aus dem BiQua-Projekt haben wir (Ufer et al., 2009) erstmals auch das Wissen zum Beweiskonzept berücksichtigt und hierzu Beweisvalidierungsaufgaben herangezogen: Den Lernenden wurden jeweils vier Beweisversuche zu einer geometrischen Aussage vorgelegt, und sie wurden gebeten zu begründen, ob es sich jeweils um einen validen mathematischen Beweis handelt. Über prozedurales Wissen hinaus trug dieses Maß in zwei Teilstudien signifikant zur Erklärung interindividueller Unterschiede beim Beweisen bei. Im Rahmen seiner Promotion hat Sommerhoff (2017) Studierenden der Mathematik Beweisaufgaben zur Analysis vorgelegt, und verschiedene individuelle Ressourcen erhoben. In seinen Studien zeigten sich wiederum signifikante Zusammenhänge zum konzeptuellen und prozeduralen Wissen zum Inhaltsbereich. Außerdem erhob er erstmals sogenanntes mathematisch-strategisches Wissen im Sinne von Weber (2001) und konnte auch hier über die anderen Ressourcen hinaus einen Zusammenhang feststellen. Kein Zusammenhang zeigte sich für Problemlösekompetenzen sowie für andere Maße wie Fähigkeiten zum logischen Schließen mit abstrakten Implikationen. Weiterhin erhob Sommerhoff Wissen zum Beweiskonzept mit Beweisvalidierungsitems, nutzte dabei jedoch im Gegensatz zur vorher genannten Studie Aussagen aus der elementaren Teilbarkeitslehre, um eine Konfundierung von Beweiskonstruktion und -validierung aufgrund des gleichen Inhalts zu vermeiden. In seiner Studie zeigt dieses Maß für das Wissen zum Beweiskonzept über die anderen Ressourcen hinaus keinen Zusammenhang zur Leistung bei der Beweiskonstruktion.

Das *Verstehen von Beweisen* ist eine weitere Anforderung, zu der Ergebnisse zu interindividuellen Unterschieden vorliegen. Hier wird den Lernenden i.d.R. ein (korrekter) Beweis zu einer Aussage vorgegeben und die „Leistung“ wird anhand von Verständnisfragen zum Beweis erhoben. Die Diskussion zum

Beweisverständnis nimmt bisher fast ausschließlich beweisspezifische Lesestrategie, und dabei beispielsweise Selbsterklärungsstrategien als Ursache für interindividuelle Unterschiede in den Blick (z.B. Hodds et al., 2014).

In einer der ersten größeren Studien zu interindividuellen Unterschieden beim Beweisverständnis haben Lin &Yang (2007) taiwanesischen Lernenden der Jgst. 9 und 10 einen Beweis einer geometrischen Aussage vorgelegt, sowie Tests zum konzeptuellen Wissen und zum logischen Schließen. In ihrer Studie klärt vor allem das inhaltliche Wissen, aber auch das logische Schließen interindividuelle Unterschiede im Beweisverständnis auf. Neuhaus und Rach (2019) haben Studierenden der Mathematik einen Beweis des Mittelwertsatzes der Analysis vorgelegt, sowie einen Beweisverständnistest. Weiterhin haben sie Wissen zum Inhaltsbereich mit einem übergreifenden Wissensmaß erhoben sowie allgemeine und beweisbezogene Lesestrategien mit einem Fragebogen. Auch in ihrer Studie erklärt das Wissen zum Inhaltsbereich interindividuelle Unterschiede beim Beweisverständnis, die selbstberichteten Lesestrategien korrelieren zwar teilweise mit dem Beweisverständnis, tragen aber über das inhaltliche Wissen hinaus nicht signifikant zur Erklärung von Unterschieden bei. Alqassab et al. (2018) haben Lehramtsstudierenden der Mathematik einen Beweis einer Aussage zur Elementargeometrie vorgelegt. Sie finden ebenfalls eine signifikante Varianzaufklärung durch Wissen zum Inhaltsbereich. Ein Eye-Tracking-Maß zur Verknüpfung von schriftlichem Beweis und illustrierender Figur erklärt keine Unterschiede im Beweisverständnistest, der in dieser Studie erst *nach* und *getrennt vom* Lesen des Beweises administriert wurde. Für die Korrektheit von Feedback an den fiktiven Autor des Beweises, das *während* des Lesens verfasst wurde, erklärt dieses Maß jedoch durchaus Unterschiede.

Zusammenfassend zeigt sich über die Studien hinweg ein deutlicher Zusammenhang zwischen Leistungen im Umgang mit Beweisen und Wissen zum jeweiligen Inhaltsbereich. Problemlösekompetenzen wurden primär im Schulbereich bei der Beweiskonstruktion untersucht und zeigen dort einen signifikanten Beitrag zur Varianzaufklärung. Bei Studierenden zeigt sich dieser Zusammenhang nicht, was die Frage aufwirft, ob ggf. spezifische Konzeptualisierungen von Problemlösekompetenzen im Kontext des Beweisens im Studium notwendig sind. Als eher inkonsistent zu betrachten sind die Befunde zum Wissen zum Beweiskonzept. Zusammenhänge ließen sich bisher nur in Studien zeigen, die dieses Wissen mit Beweisvalidierungsaufgaben zum gleichen Inhaltsbereich wie die Aufgaben zur Beweiskonstruktion erhoben. Hier wäre mittelfristig eine bessere Konzeptualisierung des Konstrukts und eine alternative Operationalisierung über Beweisvalidierungsaufgaben hinaus wünschenswert. Sehr wenig Evidenz liegt für Fähigkeiten zum logischen Schließen vor, die meist über Aufgaben zum Schließen mit weitgehend dekontextualisierten Implikationen erhoben wurden. Dass sich diese Maße beim Beweisverständnis in der Geometrie, nicht aber bei der Beweiskonstruktion im Studium als prädiktiv erweisen, ist zunächst unerwartet. Auch hier wäre zu prüfen, inwiefern bessere

Konzeptualisierungen von den beim Beweisen notwendigen logischen Fähigkeiten (z.B. ähnlich zu Durand-Guerrier et al., 2011) hier zu einem besseren Verständnis der Zusammenhänge beitragen könnten. Wenig systematisch untersucht wurde weiter der Beitrag affektiv-motivationaler Merkmale, und für die Überzeugungen von Lernenden zur Mathematik liegen beispielsweise aus der Dissertation von Kempen (2019) zwar erste Korrelationsanalysen vor, die jedoch weitere potentiell relevante Ressourcen nicht kontrollieren.

4 Beweisprozesse als Vermittler

In einer Machbarkeitsstudie hat Ottinger (2019) individuelle Ressourcen von Lernenden zum Beweisen in der Zahlentheorie mit einem übergreifenden Test erhoben, bevor die Lernenden eine Aufgabe zum Generieren und Bewiesen einer Vermutung in Partnerarbeit bearbeiteten und anschließend eine individuelle Lösung alleine niederschrieben. Neben der formalen und inhaltlichen Qualität der Vermutung und des Beweises kodierte Ottinger mittels hoch-inferenter Videoratings Merkmale der Arbeitsprozesse, wie sie in der Literatur derzeit stark diskutiert werden (Überblick bei Ottinger, 2019), z.B. inwiefern Beispiele produktiv genutzt werden oder vollständige Argumente im Sinne des Toulmin-Schemas eingebracht werden. Einerseits zeigen die Ergebnisse, dass die Prozesscharakteristika über die individuellen Ressourcen hinaus Unterschiede im abschließenden Leistungsmaß in theoretisch plausibler Weise erklären. Dies spricht dafür, dass Beweisprozesse durchaus einen eigenständigen Beitrag zur Leistung beim Beweisen leisten. Andererseits vermitteln die Prozesscharakteristika in dieser Studie jedoch wenigstens teilweise auch die Einflüsse der Ressourcen auf die Beweisleistung: Nur wer über adäquate Ressourcen verfügt, kann hochwertige Beweisprozesse für die Konstruktion eines validen Beweises nutzen. Dies spricht sowohl dafür bei der Beobachtung von Beweisprozesse, als auch bei deren Unterstützung die individuellen Ressourcen der Lernenden zu berücksichtigen, um diese Zusammenhänge in Beobachtungsstudien trennen zu können.

5 Zusammenfassung und Ausblick

Im Feld wird eine große Bandbreite von Erklärungsansätzen für interindividuelle Unterschiede beim Beweisen postuliert. Auch wenn die Befundlage noch lückenhaft ist, zeigt sich dich der Einfluss inhaltlichen Wissens stabil über verschiedene Kontexte hinweg. Dies mag nicht sonderlich überraschend erscheinen, interessant ist jedoch der Kontrast zu den anderen Erklärungsansätzen, für die wir bisher keine Evidenz, oder nur schwache bzw. unsystematisch auftretende Zusammenhänge sehen. Diese Ergebnisse deuten darauf hin, dass wir die relevanten Aspekte und vermittelnden Mechanismen für einzelne Ressourcen vielleicht noch nicht ausreichend verstanden haben.

Weiter steht einer großen Bandbreite von potentiell relevanten Prozessmerkmalen in der internationalen Diskussion eine spärliche Befundlage zur Bedeutung dieser

Prozessmerkmale gegenüber. Die bestehenden Befunde weisen einerseits darauf hin, dass hier potentielle Ansatzpunkte für wirksame Interventionen liegen, andererseits aber auch darauf, dass entsprechende Studien ohne Berücksichtigung der individuellen Ressourcen von sehr eingeschränkter Aussagekraft sind.

Für beide Fragekomplexe – die Rolle von individuellen Ressourcen und von Prozessmerkmalen – stellt sich die Frage, inwiefern Zusammenhänge in gleicher oder ähnlicher Stärke in unterschiedlichen Kontexten (Stand im Kompetenzerwerb, Inhaltsbereich, All- oder Existenzaussagen, verschiedene Beweistechniken, verschiedene Arten des Umgangs mit Beweisen) erwartet und nachgewiesen werden können. Sicher ist es hier nicht möglich alle Ressourcen und Prozessmerkmale in allen denkbaren Merkmalskombinationen der Kontexte zu untersuchen. Dennoch wäre es wünschenswert zunehmend über vergleichbare, aber dennoch systematisch variierte Studien zu verschiedenen Kontexten zu verfügen, die aufgrund plausibler theoretischer Modelle unterschiedliche bzw. ähnliche Ergebnismuster erwarten lassen. Dies wäre ein zentraler Schritt, um die beeindruckende Bandbreite von Erklärungsmodellen, die die Mathematikdidaktik in den letzten Jahren entwickelt hat, wissenschaftlich zu systematisieren, aber auch in einen für die Praxis nutzbaren gemeinsamen Rahmen zu setzen.

Literatur

Alqassab, M., Strijbos, J.-W. & Ufer, S. (2018). The impact of peer solution quality on peer-feedback provision on geometry proofs: Evidence from eye-movement analysis. *Learning and Instruction 58*, 182-192.

Atiyah, M. et. al. (1994). Responses to “Theoretical Mathematics”, *Bulletin of the American Mathematical Society 30*(2), 178-207.

Balacheff, N. (1999). Is argumentation an obstacle? *International Newsletter on the Teaching and Learning of Mathematical Proof 5/6*.

Brunner, E. (2014). *Mathematisches Argumentieren, Begründen und Beweisen*. Heidelberg: Springer.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics 24*(4), 359-387.

Chinnappan, M., Ekanayake, M. B. & Brown, C. (2012). Knowledge use in the construction of geometry proof by Sri Lankan students. *International Journal of Science and Mathematics Education*, *10*(4), 865-887.

de Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras 24*, 17-24.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S. & Tanguay, D. (2011). Examining the role of logic in teaching proof. In G. Hanna & M. de Villiers (Hrsg.): *Proof and Proving in Mathematics Education* (S. 369-389). Dordrecht: Springer.

EMS (European Mathematical Society, 2011). Do theorems admit exceptions? Solid findings in mathematics education on empirical proof schemes. *EMS Newsletter 82*, 50–53.

Hodds, M., Alcock, L. & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education 45*(1), 62-101.

Inglis, M. & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education 43*(4), 358-390.

Kempen, L. (2019). *Begründen und Beweisen im Übergang von der Schule zur Hochschule*. Wiesbaden: Springer.

Lockwood, E., Ellis, A. B., Dogan, M. F., Williams, C. & Knuth, E. (2012). A Framework for Mathematicians' Example-Related Activity When Exploring and Proving Mathematical Conjectures. In L. van Zoest & J. Kratky (Hrsg.): *Proceedings of the 34th Annual Meeting of the North American Chapter of the the International Group for the Psychology of Mathematics Education* (S. 151-158). Klamazoo: Western Michigan University.

Lin, F.-L. & Yang, K.-L. (2007). The reading comprehension of geometric proofs: The contribution of knowledge and reasoning. *International Journal of Science and Mathematics Education*, *5*(4), 729-754.

Mac Lane, S. (1981). Mathematical models: A sketch for the philosophy of mathematics. *The American Mathematical Monthly*, *88*(7), 462-472.

Neuhaus, S. & Rach, S. (2019). Proof comprehension of undergraduate students and the relation to individual characteristics. *Eleventh Congress of the European Society for Research in Mathematics Education*. Utrecht University, Feb 2019, Utrecht, Netherlands.

Ottinger, S. (2019). *Mathematical conjecturing and proving*. Dissertation LMU München.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics 66*(1), 23-41.

Reiss, K., Hellmich, F. & Reiss, M. (2002). Reasoning and proof in geometry: prerequisites of knowledge acquisition in secondary school students. In A.D. Cockburn & E. Nardi (Hrsg.): *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, S. 113–120). Norwich: University of East Anglia.

Schoenfeld, A. (1992). Learning to Think Mathematically: Problem Solving, Metacognition and Sense-making in Mathematics Teaching and Learning. In D. A. Grouws (Hrsg.): *Handbook of Research on Mathematics* (S. 334-370). New York, NY: Simon & Schuster.

Schwarz, B. B., Hershkowitz, R. & Prusak, N. (2010). Argumentation and mathematics. In K. Littleton & C. Howe (Hrsg.): In K. Littleton & C. Howe (Hrsg.). *Educational Dialogues: Understanding and Promoting Productive Interaction* (S. 103–127). London: Routledge.

Senk, S. (1989). Van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education 20*(3), 309-321.

Sommerhoff, D. (2017). *The individual cognitive resources underlying students' mathematical argumentation and proof skills*. Dissertation LMU München.

Sommerhoff, D. & Ufer, S. (2019). Acceptance criteria for validating mathematical proofs used by school students, university students, and mathematicians in the context of teaching. *ZDM 51*(5), 717-730.

Stylianides, A. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education 38*(3), 289-321.

Stylianides, G. & Stylianides, A. (2017). Research-based interventions in the area of proof: the past, the present, and the future. *Educational Studies in Mathematics 96*(2), 119-127.

Ufer, S., Heinze, A. & Reiss, K. (2008). Individual predictors of geometrical proof competence. In Figueras, O., Cortina, J.L., Alatorre, S., Rojano, T., Sepulveda, A. (Hrsg.): *Proceedings of the Joint Meeting of PME 32 and PME-NA XXX* (Vol. 4, S. 361-368). Morelia: PME.

Ufer, S., Heinze, A., Kuntze, S. & Rudolph-Albert, F. (2009). Beweisen und Begründen im Mathematikunterricht. Die Rolle von Methodenwissen für das Beweisen in der Geometrie. *Journal für Mathematikdidaktik 30*(1), 30-54.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics 48*(1), 101-119.