Anika DREHER, Freiburg & Aiso HEINZE, Kiel

Beweisakzeptanz: Verlassen sich Mathematikerinnen und Mathematiker auf ihre Kolleginnen und Kollegen? Ergebnisse einer internationalen Studie

Beweise spielen eine zentrale Rolle für die Mathematik und sind somit auch für die Mathematikdidaktik von großem Interesse. Dementsprechend ist bereits recht gut erforscht wie Lernende mit mathematischen Beweisen umgehen. Zwar gilt das Hauptaugenmerk dieser Forschung der Konstruktion von Beweisen, doch auch das Validieren von Beweisen wird zunehmend in den Blick genommen (z. B. Sommerhoff & Ufer, akzeptiert; Weber, 2008). Dabei wurde immer wieder betont, dass die Formulierung entsprechender Lernziele auf der Basis eines gründlichen Verständnisses der Praxis von Expertinnen und Experten geschehen sollte (z. B. Inglis & Alcock, 2012). Bezogen auf das Validieren von Beweisen ist als Referenzrahmen also zu untersuchen unter welchen Bedingungen Mathematikerinnen und Mathematiker Beweise akzeptieren. Nutzen sie mathematische Sätze für ihre Arbeit nur dann, wenn sie einen entsprechenden Beweis Schritt für Schritt überprüft haben? Oder verlassen sie sich wie Lernende manchmal auch einfach auf das Urteil anderer? Während diese sogenannte autoritäre Evidenz im schulischen Kontext etwa in Form von Schulbüchern oder Lehrkräften vorkommt, bezieht sie sich im professionellen Bereich eher auf angesehene Fachzeitschriften oder die Annahme, dass bereits genügend Kolleginnen und Kollegen der Community einen Beweis überprüft haben (Weber, Inglis & Mejia-Ramos, 2014). Zwar gibt es Mathematikerinnen und Mathematiker, die für jeden mathematischen Satz, den sie für ihre Arbeit nutzen, einen Beweis im Detail nachvollziehen (Geist, Loewe & van Kerkhove, 2010), aber oft ist dies auch kaum möglich: Manche Beweise sind extrem lang und der hohe Spezialisierungsgrad in der Mathematik führt dazu, dass viele nicht in der Lage sind, Beweise aus einem anderen Spezialgebiet nachzuvollziehen (Auslander, 2008). Möchte man sich in so einem Fall dennoch nicht auf die autoritäre Evidenz verlassen, so könnte eine „partielle“ Überprüfung eines Beweises ein Ausweg sein. Tatsächlich zeigen Ergebnisse erster empirischer Studien, dass Mathematikerinnen und Mathematiker neben autoritärer Evidenz auch weitere Akzeptanzkriterien heranziehen, indem sie die Schlüsselstellen eines Beweises prüfen, von den Kernideen eines Beweises überzeugt sind (Heinze, 2010) oder den Satz für sorgfältig gewählte Beispiele prüfen (Weber et al., 2014). Diese Studien geben jedoch nur wenige Anhaltspunkte zur Beantwortung der Frage, in welchem Ausmaß Mathematikerinnen und Mathematiker sich in ihrer täglichen Arbeit auf diese verschiedenen Akzeptanzkriterien verlassen. Zudem deuten die bisherigen Ergebnisse auf eine beträchtliche

Heterogenität in der mathematischen Community hin (Heinze, 2010; Meija-Ramos & Weber, 2014). Diese Heterogenität könnte ein Ergebnis externer Faktoren (z. B. Statusgruppen, Begutachtungstätigkeit, Teilgebiete der Mathematik) sein oder aber auf verschiedene Typen von Mathematikerinnen und Mathematikern zurückzuführen sein.

Die bisherigen quantitativen Studien zu Akzeptanzkriterien konzentrierten sich hauptsächlich auf Promovierende und Postdocs und haben weniger Professorinnen und Professoren – die eigentliche Expertengruppe – in den Blick genommen. Zudem wurden in der Regel nur nationale Stichproben betrachtet. In Anbetracht der angenommenen Heterogenität erscheint eine breitere Stichprobe jedoch notwendig, um ein hinreichend aussagekräftiges und differenziertes Bild zu erhalten. Außerdem basieren die bisherigen Ergebnisse auf der Analyse von Einzelitems, sodass der Einsatz eines weiter entwickelten Befragungsinstruments wünschenswert ist.

Entsprechend dieses Forschungsbedarfs konzentriert sich unsere Studie auf die folgenden Forschungsfragen:

- Auf welche Kriterien verlassen sich Mathematikerinnen und Mathematiker, um einen Satz mit Beweis in ihrer mathematischen Arbeit als gültig zu akzeptieren?
- Hängen Unterschiede bezüglich Statusgruppe und Begutachtungstätigkeit mit Unterschieden bezüglich Akzeptanzkriterien zusammen?
- Wie lassen sie unterschiedliche Typen von Mathematikerinnen und Mathematiker hinsichtlich ihrer Akzeptanzkriterien charakterisieren?

Untersuchungsdesign und Stichprobe

Um diese Forschungsfragen zu beantworten, wurde eine Onlinebefragung mit Hilfe von „Unipark“ erstellt. Der Fragebogen wurde von $N = 229$ Mathematikerinnen und Mathematikern beantwortet, die Teilnehmende von Workshops des namhaften mathematischen Forschungsinstituts Oberwolfach der letzten Jahre waren und damit angesehene Mitglieder der internationalen Forschungsgemeinschaft sind. Die meisten Teilnehmenden haben einen Lehrstuhl inne und die große Mehrheit hat bereits mehrfach für Fachzeitschriften begutachtet. Entsprechend der Forschungsfragen wurden die Teilnehmenden nach ihrer Auffassung gefragt, unter welchen Bedingungen sie einen mathematischen Satz in ihrer mathematischen Arbeit als gültig akzeptieren. Die Datenerhebung erfolgte durch Bewerten von 20 Aussagen auf einer sechsstufigen Likert-Skala mit den Endpunkten „völlige Ablehnung“ und „völlige Zustimmung“. Dabei wurde die Aussage „In meiner mathematischen Arbeit nehme ich an, dass ein mathematischer Satz gültig ist, wenn…“ jeweils mit den oben genannten Arten von Kriterien vervollständigt:

- eigenes Prüfen „Schritt für Schritt“ (z. B. „… ich einen gegebenen Beweis Schritt für Schritt geprüft habe.“),
- eigenes Prüfen „partiell“ (z. B. „… ich überzeugt bin, dass die Kernideen eines gegebenen Beweises korrekt sind.“ oder „… der Satz für alle Beispiele, die ich kenne, gültig ist.“),
- autoritäre Evidenz „Fachzeitschriften“ (z. B. „… der Satz mit Beweis in einer Fachzeitschrift mit Begutachtung veröffentlicht wurde.“),
- autoritäre Evidenz „genügend Mathematikerinnen und Mathematiker“ (z. B. „… ich weiß, dass schon lange ein Beweis verfügbar ist und von vielen Mathematikerinnen und Mathematikern geprüft wurde.“).

Ausgewählte Ergebnisse

Da in diesem Rahmen nur eine knappe Zusammenfassung von ausgewählten Ergebnissen möglich ist, verweisen wir für eine ausführlichere Darstellung auf Dreher und Heinze (2018). Auf Basis einer konfirmatorischen Faktorenanalyse konnte das theoretisch angenommene vierdimensionale Modell entsprechend der vier oben genannten Arten von Akzeptanzkriterien bestätigt werden. Es konnten damit vier Skalen mit akzeptabler bis guter Reliabilität gebildet werden. Die Mittelwerte dieser Skalen zeigen, dass die Mathematikerinnen und Mathematiker im Mittel eine hohe Zustimmung zu dem Akzeptanzkriterium eigenes Prüfen „Schritt für Schritt“ und eine mittlere Zustimmung zu den anderen drei Arten von Akzeptanzkriterien angaben. Für die zweite Forschungsfrage wurde der Einfluss externer Faktoren auf die Akzeptanzkriterien untersucht. Beim Vergleich von Postdocs mit Inhaberinnen bzw. Inhabern von Lehrstühlen zeigten sich keine signifikanten Unterschiede hinsichtlich der beiden Kriterien zum eigenen Prüfen, jedoch zeigten Postdocs eine signifikant höhere Zustimmung zu den autoritären Evidenzen „Fachzeitschriften“ und „genügend Mathematikerinnen und Mathematikern“ (jeweils mittlere Effektstärken). Zudem stimmten Mathematikerinnen und Mathematiker ohne Reviewerfahrung signifikant und substanziell stärker autoritärer Evidenz „Fachzeitschriften“ zu als solche, die bereits mindestens dreimal für Fachzeitschriften begutachtet haben. Um die dritte Forschungsfrage beantworten zu können, wurde eine hierarchische Clusteranalyse auf Basis der vier Skalen durchgeführt. Diese ergab drei Cluster mit unterschiedlichen Antwortmustern: Ein Cluster besteht aus einer Minderheit ($n = 20$), die als ihr einziges Akzeptanzkriterium die eigene schrittweise Prüfung angab. Die anderen beiden Cluster sind fast gleich groß und zeigten die gleiche mittlere Zustimmung zum Kriterium eigenes Prüfen „partiell“. Sie unterscheiden sich jedoch deutlich hinsichtlich ihrer Zustimmung zu autoritärer Evidenz.

Insgesamt zeigt sich also, dass die Mehrheit der Mathematikerinnen und Mathematiker sich nicht nur auf eigene schrittweise geprüfte Beweise verlässt, sondern durchaus auch andere Akzeptanzkriterien heranzieht.

Literatur

Auslander, J. (2008). On the roles of proof in mathematics. In B. Gold & R. A. Simons (Hrsg.), *Proofs and other dilemmas: Mathematics and philosophy* (S. 61–77). Washington, DC: Mathematical Association of America.

Dreher, A. & Heinze, A. (2018). Mathematicians‘ criteria for accepting theorems and proofs – An international study. In E. Bergqvist, M. Österholm, C. Granberg & L. Sumpter (Hrsg.), *Proceedings. of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, S. 363–370). Umea (Sweden): PME.

Geist, C. Löwe, B., & Van Kerkhove, B. (2010). Peer review and testimony in mathematics. In B. Löwe & T. Müller (Hrsg.), *Philosophy of Mathematics: Sociological Aspects and Mathematical Practice* (S. 155-178). London: College Publications.

Heinze, A. (2010). Mathematicians’ individual criteria for accepting theorems and proofs: An empirical approach. In G. Hanna, H. N. Jahnke, & H. Pulte (Hrsg.), *Explanation and proof in mathematics: Philosophical and educational perspectives* (S. 101–111). New York: Springer.

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

Mejía-Ramos, J. P. & Weber, K. (2014). How and why mathematicians read proofs: further evidence from a survey study. *Educational Studies in Mathematics, 85*, 161–173.

Sommerhoff, D. & Ufer, S. (akzeptiert). Acceptance Criteria for Validating Mathematical Proofs Used by Pupils, University Students, and Mathematicians in the Context of Teaching. Erscheint in *ZDM – Mathematics Education*.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*, 431-459.

Weber, K., Inglis, M., & Mejia-Ramos, J. P. (2014). How mathematicians obtain conviction: implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist, 49*(1), 36-58