Michael MEYER, Köln

Wann ist ein Argument ein Argument? Erste Ergebnisse eines ExpertInnen-NovizInnen-Vergleichs

Die Bedeutung des Beweisens ist in der mathematischen und mathematikdidaktischen Diskussion unbestritten. Wenn aber die Frage danach auftaucht, was einen Beweis auszeichnet, so gehen die Meinungen auseinander. Der „Idealtyp“ eines mathematischen Beweises mag den Anspruch haben, eine Aussage schrittweise formal deduktiv aus als bekannt vorauszusetzenden Definitionen und Sätzen zu folgern. In der Regel gilt hierbei (stillschweigend) die Annahme, dass dies bis zu den jeweiligen Axiomen hin möglich sei. Eine solche Perspektive auf das Beweisen ist jedoch nicht unproblematisch: Zum einen besitzt allein der Beweis zur Klassifikation der einfachen Gruppen ca. 15000 Seiten (und verwendet immer noch anderswo bewiesene Sätze, s. Dreyfus 2002, S. 16). Zum anderen ist die Bezeichnung „als bekannt vorauszusetzenden“ eine, die von Individuen abhängt. Dass selbst dies in der Mathematik Probleme aufwirft, zeigt die Studie von Inglis et al. (2013), in der 109 MathematikerInnen ein Beweis vorgelegt wird, den 29 als richtig und 80 als falsch (aus unterschiedlichen Gründen) bewerteten (u.a. weil nicht jeder Schritt einzeln gerechtfertigt wurde oder weil Abkürzungen genutzt wurden, die zuvor nicht definiert wurden). Die Studie zeigt also, dass Ansprüche an Beweise unter ExpertInnen sehr unterschiedlich sein können.

Es liegt auf der Hand, dass bei Argumenten im schulischen Mathematikunterricht Vergleichbares gilt. Für Lehrpersonen zeigte dies beispielsweise die Studie von Meyer und Schnell (2016), in welcher verschiedene Bewertungskriterien und Bewertungsprofile aufgezeigt wurden. Neben inhaltlichen Aspekten (z.B. Exaktheit und Vollständigkeit) treten beispielsweise pragmatische Aspekte (z.B. Prägnanz), die sich in den Betrachtungen der Lehrpersonen widersprachen. In der Studie von Healy und Hoyles (2000) wurde deutlich, dass 14- bis 15-jährige SchülerInnen solche Argumente als zu bevorzugen ansahen, die eher formal-symbolisch konstruiert waren.

Es zeigt sich, dass sowohl ExpertInnen als auch NovizInnen unterschiedliche Kriterien für Beweise/Argumente heranziehen. Die hier beschriebene Studie *LerA* (*Le*rnende *r*ezipieren *A*rgumente) fokussiert Bewertungsperspektiven von ExpertInnen und NovizInnen im direkten Vergleich. Es werden Bewertungsperspektiven und -kriterien untersucht, die SchülerInnen (Klassenstufe 8), Studierende, Lehrpersonen (s. Meyer & Schnell 2016) und UniversitätsdozentInnen an Argumente legen. In diesem Beitrag werden erste Befunde des „ExpertInnen (UniversitätsdozentInnen) – NovizInnen (SchülerInnen bzw. Studierende) – Vergleichs“ präsentiert. Den ProbandInnen wurden verschiedene Argumente vorgelegt, die es a) im Schulnotensystem (1-6) zu bewerten und b) diese Bewertung zu begründen galt. Um bereichsspezifische Effekte (u.a. Bauersfeld 1983) erfassen zu können, wurden Argumente zu zwei verschiedenen mathematischen Inhaltsbereichen (Arithmetik und Wahrscheinlichkeitsrechnung) gewählt. Dieser Beitrag dient vorrangig zur Vorstellung der Studie und zur Thematisierung erster Ergebnisse.

Design der Studie

In der zugrundeliegenden Studie wurden potentieller Argumente zu zwei verschiedenen Aufgaben systematisch variiert. Dies geschah auf der theoretischen Grundlage des Argumentschemas von Toulmin (1996). Entsprechend der strukturellen Dimension durch dieses Schema besitzen Argumente verschiedene funktionale Elemente. Die Elemente „Datum", „Konklusion" und „Regel" lassen sich mit denjenigen eines deduktiven Beweises vergleichen. Die Elemente „Ausnahmebedingung" und „Stützung" hingegen verdeutlichen, dass ein (schulmathematisches) Argument mehr Bestandteile beinhalten kann, als ein hochschulmathematischer Beweis (für eine ausführliche Diskussion dieser Aspekte sei auf Meyer 2007 verwiesen). Diese Argumente wurden den ProbandInnen zur Bewertung vorgelegt.

Die funktionalen Elemente potentieller Argumente wurden systematisch variiert. Bei einigen Argumenten fehlte eine Regel, bei anderen wurde eine falsche Regel angewendet. Einige Argumente waren hinsichtlich des Schemas vollständig, andere wiederum enthielten nur eine Konklusion etc. Betrachten wir zwei Beispielargumente zu der Aufgabe aus der Wahrscheinlichkeitsrechnung:

> Zwei Würfel werden geworfen und die Summe wird notiert. Basti gewinnt bei den Augensummen 1, 2, 3, 10, 11 und 12. Derya gewinnt bei 4, 5, 6, 7, 8, 9. Basti beschwert sich: Da habe ich ja gar keine Chance zu gewinnen!
>
> Stimmt das?

Eines von insgesamt 17 Argumenten zu dieser Aufgabe lautete: „Nein, Basti liegt falsch, denn er hat genauso viele Gewinnzahlen wie Derya. Entsprechend haben beide die gleiche Chance zu gewinnen." Die in diesem Argument genutzt Regel besitzt die Ausnahmebedingung, dass die jeweiligen Elementarereignisse gleichberechtigt wären. Dieses wird nicht weitergehend thematisiert, seine Anwendung in der speziellen Situation der Aufgabe ist falsch.

Neben der Angabe bzw. Korrektheit der funktionalen Bestandteile der Argumente wurden in den gegebenen Argumenten zusätzlichen die

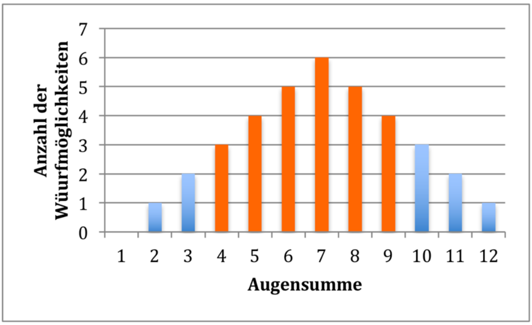
Darstellungsweisen variiert. Beispielsweise wurden die Elementarereignisse bei einigen Argumenten in einer Tabelle oder mittels eines Säulendiagramms dargestellt. Mit dieser Maßnahme sollte beobachtet werden, ob die Darstellungsweise Einfluss auf die Bewertungen hat. Ein Beispiel für ein Argument mit ikonischen Inhalten zeigt die folgende Abbildung:

Das Argument von Murat (Abb. links) zeichnet sich dadurch aus, dass die Anzahl der Zerlegungsmöglichkeiten für die Augensummen (Datum) explizit angegeben werden. Explizite Gewinnwahrscheinlichkeiten für die einzelnen Personen werden nicht angegeben. Die Angabe einer allgemeinen Regel fehlt ebenso.

Murat: Ich habe erst eine Tabelle zu den Wurfmöglichkeiten für jede Augensumme bestimmt:

Augensumme	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Zerlegungs-möglichkeiten	0	1	2	3	4	5	6	5	4	3	2	1

Dann habe ich auf dieser Basis ein Diagramm dazu gezeichnet:



Die orangefarbenen Säulen zeigen die Zerlegungsmöglichkeiten für Derya und die blauen die für Basti. Bei Derya sind alle Säulen höher als die blauen für Basti, deshalb gewinnt Derya häufiger, zumal jedes Einzelergebnis, also jede einzelne Kombination, gleichwahrscheinlich ist. Basti kann zwar sehr viel Glück haben, aber normalerweise gewinnt er nicht. Er beschwert sich also zu Recht.

Zur quantitativen Auswertung der Daten wurden die Notendifferenzen zwischen den NovizInnen und den ExpertInnen betrachtet und hinsichtlich verschiedener Aspekte ausgewertet. Im Folgenden werden erste einzelne Ergebnisse zur Bewertung der Argumente aus dem Bereich der Wahrscheinlichkeitsrechnung betrachtet.

Ausgewählte erste Ergebnisse

In der Studie zeigte sich, dass SchülerInnen der Klassenstufe 8 (n=37, Gymnasialklassen) Argumente, die ausschließlich aus korrekten Konklusionen bestanden, deutlich besser bewerteten als die ExpertInnen. Die mittlere Abweichung von der „Normnote“ (die gemittelte Note der ExpertInnen) war für diese Argumente signifikant höher als bei anderen Argumenten. Die mündlichen Begründungen der SchülerInnen bestätigten diesen Eindruck, insofern mehrfach eine Übereinstimmung zur eigenen Meinung als positiv für die eigene Benotung des Arguments angesehen wurde, dies auch in solchen

Momenten in denen den SchülerInnen bewusst war, dass das Argument eigentlich kein Argument sein konnte. Generell zeigten die SchülerInnen, dass es ihnen leichter fiel, einen korrekten Beweis als korrekt zu klassifizieren, als einen falschen oder unvollständigen Beweis als falsch (s. auch Reiss et al. 2002).

Hinsichtlich der Darstellungsweise (eher symbolisch oder eher ikonisch) zeigte sich in der vorliegenden Studie verschiedene kleinere Effekte. Das Vorhandensein einer Tabelle bei einem Argument äußerte sich darin, dass den SchülerInnen das Bewerten des Arguments leichter zu fallen schien, insofern hier die Standardabweichung der Bewertungen besonders niedrig war. Bei nicht-validen Argumenten die formal-symbolische Darstellungen enthielten, war zudem die Standardabweichung innerhalb der Bewertungen der männlichen Schüler signifikant höher als für nicht-valide Argumente, die eher ikonisch dargestellt waren. Der beobachtete Effekt der Überbewertung symbolischer Notationen (s. oben) zeigt sich hier also in Ansätzen erneut.

Literatur

Bauersfeld, H. (1983). Subjektive Erfahrungsbereiche als Grundlage einer Interaktionstheorie des Mathematiklernens und -lehrens. In H. Bauersfeld et al. (Hrsg.), *Lernen und Lehren von Mathematik* (S. 1-56). Köln: Aulis.

Dreyfus, T. (2002). Was gilt im Mathematikunterricht als Beweis? In W. Peschek (Hrsg.), *Beiträge zum Mathematikunterricht* (S. 15-22). Hildesheim: Franzbecker.

Healy, L. & Hoyles, C. (2000). A study of Proof Conceptions in Algebra. *Journal for Research in Mathematics Education*, 31/4, 296-428.

Inglis, M., Mejia-Ramos, J.P., Weber, K. & Alcock, L. (2013). On Mathematicians' different Standards when evaluating elementary proofs. *Topics in Cognitive Science*, 5, 270-282.

Meyer, M. (2007). *Entdecken und Begründen im Mathematikunterricht. Von der Abduktion zum Argument*. Hildesheim: Franzbecker.

Meyer, M. & Schnell, S. (2016). Was ist ein „gutes" Argument? Bewertung von Schülerargumenten durch Lehrkräfte. *Beiträge zum Mathematikunterricht*. https://eldorado.tu-dortmund.de/handle/2003/35512 (27.10.2018).

Reiss, K., Hellmich, F. & Thomas, J. (2002). Individuelle und schulische Bedingungsfaktoren für Argumentationen und Beweise im Mathematikunterricht. In M. Prenzel & J. Doll (Hrsg.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (S. 51-64). Weinheim: Beltz.

Toulmin, S. (1996). *Der Gebrauch von Argumenten*. Weinheim: Beltz.