

KARPUSCHEWSKI, Tim & ROTT, Benjamin
Köln

Problemlösen bewerten – Eine Studie zum Einsatz von Beurteilungsrastern zur Bewertung von Problemlöseprodukten

Problemlöseaufgaben für den Mathematikunterricht zeichnen sich u. a. durch die Existenz verschiedener möglicher Lösungsansätze auf unterschiedlichen Leistungsniveaus aus (Holzäpfel et al., 2018, S. 43; Wälti, 2001, S. 10). Die Bewertung von Problemlöseprozessen stellt Mathematiklehrkräfte deshalb vor eine besondere Herausforderung: Wie können die Lernendenprodukte möglichst vergleichbar und objektiv bewertet werden?

Um Lehrkräfte bei der Bewertung von Problemlöseprozessen zu unterstützen, steht diesen bereits eine Auswahl an unterschiedlichen Beurteilungsformen zur Verfügung, die sowohl für summative, als auch für formative Bewertungen einsetzbar sind. Holzäpfel et al. (2018) beschreiben zum Beispiel diverse Beurteilungsformen, wie diagnostische Gespräche, Beobachtungen im Prozess, Problemlösejournale oder auch die traditionelle Klassenarbeit.

Theoretischer Hintergrund und Forschungsstand

Besonders im Hinblick auf die traditionelle Klassenarbeit zur Vergabe einer Endnote stellt sich die Frage, welche Aspekte des Problemlösens in diesen bewertet werden können und wie eine (Punkte-)Bewertung anhand der schriftlichen Ergebnisse erfolgen kann. In der mathematikdidaktischen Forschung werden für die Bewertung schriftlicher Artefakte häufig mehrstufige Kategoriensysteme verwendet, die eine Punktzahl basierend auf dem Grad der Bearbeitung, wie beispielsweise "Keine Bearbeitung erfolgt" oder "Ansatz vorhanden" zuweisen (vgl. Rott, 2013, S. 185, Dreher et al., 2017, S. 241). Diese Bewertungsverfahren eignen sich gut für statistische Auswertungen, jedoch beachten diese oft weniger die curricularen Vorgaben oder beschränken sich auf einzelne Teilkompetenzen des Problemlösens. Für eine mehrdimensionale Betrachtung der Lernendenleistung eignen sich hingegen insbesondere Beurteilungsraster, die auf curricularen Vorgaben oder verschiedenen Teilkompetenzen beruhen (vgl. Winter, 2015, 134-169). Spezifisch für das Problemlösen existieren bereits Beurteilungsraster, die an den vier Einflussfaktoren *Ressources*, *Heuristics*, *Control* und *Belief Systems* nach Schoenfeld (1985) ausgerichtet sind (vgl. Lacher, o. D.; Holzäpfel et al., 2018, S. 215). Alternativ können Beurteilungsraster auch am Prozess des Problemlösens orientiert werden. Beispiele für solche Beurteilungsraster lassen sich für die, mit dem Problemlösen eng verbundene, Kompetenz des Modellierens bei Maaß (2007, 2009) finden. Da die Beforschung von Beurteilungsrastern zum mathematischen Problemlösen noch nicht weit

In: L. Schick, M. Platz & A. Lambert (Hrsg.),
Beiträge zum Mathematikunterricht 2025.

58. Jahrestagung der Gesellschaft für Didaktik der Mathematik. WTM.

<https://doi.org/10.37626/GA9783959873307.0>

fortgeschritten ist, stellen sich im Hinblick auf eine objektive, reliable und valide Bewertung von Problemlöseprodukten die Fragen:

(FF1) Inwieweit beeinflusst die Nutzung eines Beurteilungsrasters die Bewertung schriftlicher Problemlösungen durch Lehrkräfte.

(FF2) Inwiefern unterscheiden sich die Einzelbewertungen hinsichtlich relevanter Aspekte und Phasen im Problemlöseprozess durch verschiedene Lehrkräfte?

Methode

Die sekundäre Forschungsfrage bezieht sich insbesondere auf die Konstruktion des Beurteilungsrasters, denn das für diese Studie erstellte Raster orientiert sich an den vier Phasen des Problemlösens nach Pólya (1945): *Verstehen des Problems*, *Ausdenken eines Plans*, *Ausführen des Plans* und *Rückschau halten* in Kombination mit dem darauf aufbauenden Modell von Mason, Burton und Stacey (2010), dass zusätzlich die *Flexibilität* während des Problemlöseprozesses betont und stetige Wechsel zwischen Pólyas vier Phasen erlaubt. Diese fünf Merkmalsklassen wurden mit jeweils zwei bis drei Indikatoren genauer ausformuliert. Die Bewertung der Merkmalsklassen erfolgt auf einer vierstufigen Likert-Skala (++, +, -, --) und den Lehrkräften stehen Kommentarfelder zu jeder Merkmalsklasse zur Verfügung, die zur Bewertung und Rückmeldung an die Lernenden genutzt werden können. Für die Merkmalsklasse *Flexibilität* steht zudem die Option "nicht bewertbar" zur Verfügung, da z. B. bei Lernenden, die instinktiv einen zielführenden Ansatz gewählt haben, keine Anzeichen für Flexibilität im Lösungsprozess erkennbar sein müssen. Am Ende des Bewertungsrasters wird eine Punktzahl für die Aufgabe auf einer Skala von 0 (ungenügend) bis 15 (sehr gut) vergeben. Damit das Beurteilungsraster aufgabenübergreifend eingesetzt werden kann, ist keine Gewichtung der fünf Merkmalsklassen vorgegeben. Diese kann somit individuell an den Lernstand der Schüler*innen sowie den eigenen Unterrichtsinhalt angepasst werden.

In der Erhebung werden die Teilnehmenden zunächst angehalten das Problem, dessen Lösungen später bewertet werden sollten, selbst zu lösen, damit sie einen eigenen Einblick in den Problemraum erhalten, bevor ihnen weitere mögliche Lösungswege vorgestellt werden. Diese Vorbereitung soll den Teilnehmenden das Problem sowie dessen Lösungen insofern näherbringen, wie dies in der Vorbereitung einer Klassenarbeit oder dem Einsatz eines Problems im eigenen Unterricht geschieht. Anschließend werden zwei, randomisiert aus einem Lösungspool ausgewählte, Lernendenlösungen frei bewertet, bevor das Beurteilungsraster vorgestellt und zur anschließenden Bewertung zweier weiterer randomisierter Lösungen durch die Lehrkräfte

eingesetzt wird. Die Erfassung der Daten erfolgt durch das Nutzen des Online-Befragungstools von Unipark. Die Durchführung steht noch aus. Für die Studie soll eine möglichst große Anzahl an Lehrkräften ($N > 100$) der Sekundarstufe I und II aus Deutschland befragt werden. Eine Einschränkung bezüglich der Berufserfahrung, dem Geschlecht oder dem Alter wird nicht vorgenommen.

Pilotierung

In drei Pilotierungsdurchläufen wurde insbesondere die Konzeption der Merkmalsklassen und die Vorbereitung der Teilnehmenden auf die Bewertung der Probleme genauer betrachtet. Aus diesem Grund erfolgte noch kein Vergleich zwischen einer Bewertung mit und ohne Beurteilungsraster. Insbesondere wurde der Frage nachgegangen, ob eine wertungsfreie Besprechung ausgewählter Lernendenlösungen im Vorfeld die Bewertung durch die Teilnehmenden beeinflusst. Für die Bewertung wurden den Teilnehmenden Bearbeitungen von Lernenden einer fünften Klasse zu einem ausgewählten Problem bereitgestellt.

Die zwei ersten Kohorten ($N_1=8$, $N_2=8$) bestanden aus Lehrkräften der Sekundarstufe I und II, die im Rahmen von Fortbildungen zum Problemlösen an der Umfrage teilgenommen haben. Kohorte 3 ($N_3=11$) bestand aus Doktorand*innen des Instituts für Mathematikdidaktik an der Universität zu Köln.

Die drei Durchführungen unterschieden sich in der Anzahl der zu bewertenden Lösungen, sowie der erfolgten Vorbesprechung. Zunächst wurde den Teilnehmenden die Problemstellung ausgehändigt, die diese im Anschluss eigenständig bearbeiteten. Anschließend wurden den Teilnehmenden unterschiedliche Lösungen des Problems als Bewertungsgrundlage präsentiert. Nach der Besprechung zweier zu bewertender Lösungen erfolgt abschließend das Bewerten von insgesamt vier bis sieben Lernendenlösungen mit Nutzung des Beurteilungsrasters.

Ergebnisse und Konsequenzen für das Studiendesign

Im Zuge der Pilotierung des Bewertungsrasters erfolgte das Prüfen und Überarbeiten dessen anhand der von Winter (2015, S. 164-165) formulierten Anforderungen an gute Beurteilungsraster:

- Die Merkmalsklasse der Flexibilität wurde in der Pilotierung häufig als nicht bewertbar angesehen. Es wurde eine entsprechende Auswahloption hinzugefügt.

- Die zuvor verwendete dreistufige Likert-Skala (+, o, -) für die Niveaustufen wurde durch die vierstufige Variante ersetzt. Mehr als vier Abstufungen lassen sich in der Praxis oft nicht unterscheiden (Winter, 2015, S. 164), weshalb auf das Hinzufügen einer fünften Option verzichtet wurden.

Zudem erfolgten weitere Änderungen am Beurteilungsraster, die aus den Rückmeldungen der Teilnehmenden und den Ergebnissen der Pilotierungsrunden resultieren:

- Die Kommentarfelder zu den einzelnen Merkmalsklassen wurden hinzugefügt und dabei ein abschließendes Kommentarfeld entfernt.
- Das Vergeben einer Endnote wurde durch das Vergeben einer Punktzahl ausgetauscht, da bei einem Einsatz eines Problems in einer Klassenarbeit in der Regel keine Einzelnoten für Teilaufgaben vergeben werden.
- In der Vorbesprechung werden keine zu bewertenden Lernendenlösungen eingesetzt, da dies zu einer Verzerrung der erhaltenen Daten führen kann. Die Vorbesprechung wurde weiter expliziert, da einige Teilnehmende Rückmeldungen zum Nichtverständnis einzelner Lösungen gaben.

Literatur

- Dreher, U., Holzäpfel, L., Leuders, T. & Stahnke, R. (2017). Problemlösen lehren lernen – Effekte einer Lehrerfortbildung auf die prozessbezogenen mathematischen Kompetenzen von Schülerinnen und Schülern. *Journal für Mathematik-Didaktik*, 39, 227-256
- Holzäpfel, L., Lacher, M., Leuders, T. & Rott, B. (2018). *Problemlösen lehren lernen: Wege zum mathematischen Denken*. Klett Kallmeyer.
- Lacher, M. (o. D.). *Noten fürs Problemlösen?!: Problemlösekompetenzen diagnostizieren, rückmelden und bewerten*. <https://www.friedrich-verlag.de/friedrich-plus/sekundarstufe/mathematik/modellieren-problemloesen/noten-fuers-problemloesen/>
- Maaß, K. (2007). *Mathematisches Modellieren: Aufgaben für die Sekundarstufe I*. Cornelsen.
- Maaß, K. (2009). *Mathematikunterricht weiterentwickeln*. Cornelson Scriptor.
- Mason, J., Burton, L. & Stacey, K. (2010). *Thinking Mathematically*. Pearson Education Limited.
- Pólya, G. (1945). *How to Solve It: a new aspect of mathematical method*. Princeton University Press.
- Rott, B. (2013). *Mathematisches Problemlösen: Ergebnisse einer empirischen Studie*. WTM-Verlag.
- Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando: Academic Press.
- Wälti, B. (2001) *Problemlösen mach Schule: Anregungen zum Mathematikunterricht auf der Sekundarstufe I*. Klett & Balmer.
- Winter, F. (2015). *Lerndialog statt Noten: Neue Formen der Leistungsbeurteilung*. Beltz.