

Messunsicherheiten als Ausgangspunkt der Förderung im Umgang mit Daten

Forschungsstand

Die Bewertung der Qualität von Daten und ihre Interpretation sind Kompetenzen, die in schulischen und gesellschaftlichen Kontexten zunehmend an Bedeutung gewinnen (Holmes et al., 2015; Sharma, 2006; Chinn & Malhorta, 2002). Dazu zählen auch die Fähigkeiten Messunsicherheiten einzuschätzen und Messdaten miteinander vergleichen zu können. Genau hier haben Schülerinnen und Schüler häufig große Schwierigkeiten (Kok et al., 2019; Ludwig, 2017; Masnick & Morris, 2008). Gleichzeitig wird in der Schule kaum Wissen über Messunsicherheiten vermittelt (Mörke, 2020, in diesem Band; Priemer & Hellwig, 2018; Kirch, 2010), obwohl ohne die Berücksichtigung dieser keine sicheren Aussagen über Messergebnisse möglich sind (Heinicke, 2012).

Forschungsfrage

Um die Fähigkeiten von Schüler*innen im Umgang mit und bei der Auswertung von Daten zu fördern, wird daher hier zunächst die Fragestellung untersucht, welcher Zusammenhang zwischen dem konzeptuellen Wissen über Messunsicherheiten und der Fähigkeit, Daten fachlich korrekt vergleichen zu können, besteht?

Studiendesign

Zu Förderung des konzeptuellen Wissens über Messunsicherheiten wurde eine Intervention entwickelt. Mit einem Pre-Post Studiendesign wurde die Effektivität dieser Intervention überprüft.

Während der Pre- und Post-Phase wurden den Schüler*innen zwei Datenreihen vorgelegt, in welchen Daten zur Geschwindigkeit zweier Skateboards dargestellt sind, siehe Tab. 1a. Die Schüler*innen wurden dann mit folgendem Item zu einer multiple choice Vermutung mit Begründung aufgefordert: „Welches Skateboard ist auf Basis dieser Daten schneller? Skateboard A, Skateboard B, beide Skateboards sind gleich schnell. Begründe deine Entscheidung ausführlich“.

Nach dieser Datenvergleichsaufgabe wird ein Kompetenztest (nach Schulz et al., 2018) von zehn Items zur Messung der Kompetenz im Umgang mit Messunsicherheiten durchgeführt. Das gleiche Verfahren wird im Post-Test wiederholt, nur gibt es hier neue Daten für die Skateboards, siehe Tab. 1b.

Skateboard A (m/s)	Skateboard B (m/s)
1,502	1,507
1,547	1,558
1,498	1,499
1,519	1,525
1,522	1,552
1,499	1,512

Tab. 1a: Die Daten des Pre-Test,
Mittelwerte: $\bar{v}_A = 1,5145 \text{ m/s}$,
 $\bar{v}_B = 1,5255 \text{ m/s}$

Skateboard A (m/s)	Skateboard C (m/s)
1,502	1,488
1,547	1,539
1,498	1,480
1,519	1,506
1,522	1,533
1,499	1,493

Tab. 1b: Die Daten des Post-Test,
Mittelwerte: $\bar{v}_A = 1,5145 \text{ m/s}$,
 $\bar{v}_C = 1,5065 \text{ m/s}$

Nach dem Pre-Test findet die Intervention auf einer digitalen Plattform statt. Hier werden die Inhalte anhand von sechs Videos, die zwischen ein und drei Minuten dauern, vermittelt. Die Inhalte sind zugespielt auf die Teilkonzepte: *Existenz von Messunsicherheiten*, *Direktes Messen* und *Vergleich von Messwerten* (nach Priemer & Hellwig 2018). Das erste Video zeigt die verschiedenen Ursachen von Messunsicherheiten; das zweite erklärt den Prozess wiederholten Messens; das dritte zeigt die Berechnung eines Mittelwerts; das vierte zeigt die Berechnung einer quantitativen Unsicherheit; das fünfte diskutiert die Bedeutung eines Unsicherheitsintervalls und das sechste erklärt, wie man zwei Messergebnisse miteinander vergleicht.

Nach jedem Video gibt es zwei bis vier Übungsaufgaben, wobei die Schüler*innen Feedback sowohl bei falschen als auch bei richtigen Antworten auf ihre Antworten bekommen.

Ergebnisse der Pilotierung

Bei der Pilotierung dieser Intervention haben $n = 61$ Schüler*innen aus zwei zehnten und zwei elften Klassen an drei Gymnasien in Berlin teilgenommen.

Die Verteilungen der Vermutungen in der Datenvergleichsaufgabe der Pre- und Post-Tests der Schüler*innen sind in Tab. 2 dargestellt. Diejenigen Schüler*innen, die in ihrer Vermutung das Skateboard mit dem kleinsten Mittelwert der Geschwindigkeit auswählten (Skateboard A im Pre-Test und Skateboard C im Post-Test), haben die Daten falsch interpretiert. Dies zeigte sich in ihren Begründungen: „Das Skateboard C ist schneller, da das arithmetische Mittel kleiner ist als das vom Skateboard A.“. Diese Schüler*innen haben entweder nicht verstanden, dass eine kleinere Geschwindigkeit ein langsamerer Skateboard bedeutet, oder sie haben die Daten vielleicht als z.B. Zeitangaben interpretiert.

Vermutung:	größter Mittelwert	kleinster Mittelwert	gleichschnell
Pre-Test	72 % (Skateboard B)	20 % (Skateboard A)	8 %
Post-Test	67 % (Skateboard A)	23 % (Skateboard C)	10 %

Tab. 2: Die Vermutungsverteilung der Datenvergleichsaufgabe

Die Begründungen sind doppelt kodiert mit einer guten bis ausgezeichneten Interoderreliabilität, $\kappa = 0,82$. Hier stellt sich heraus, dass die Schüler*innen sich im Post-Test sehr stark (70 %) auf einen Mittelwertvergleich beziehen, sogar noch stärker als im Pre-Test (58 %). Der Paarvergleich von Einzelwerten reduziert sich im Post-Test (10 %) im Vergleich zum Pre-Test (23 %). Der Bezug auf ein quantitatives Maß von Messunsicherheiten ändert sich nicht signifikant vom Pre-Test (5 %) zum Post-Test (8 %).

Der Score-Unterschied des Kompetenztests zwischen Pre- und Post-Erhebung ist signifikant mit einem mittelstarken Effekt ($p < 0,001$, $r = 0,37$). Im Durchschnitt werden im Post-Test (im Vergleich zum Pre-Test) 1,7 Aufgaben mehr korrekt gelöst, siehe Abb. 1.

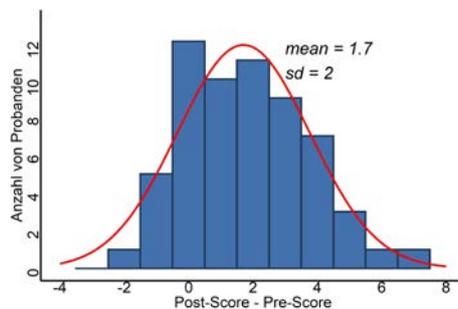


Abb. 1: Die Änderung der Kompetenztestscores zwischen Pre- und Post-Test

Diskussion und Ausblick

Die Ergebnisse zeigen, dass der Transfer von konzeptuellem Wissen, so wie in den Kompetenztests erfasst, auf die Fähigkeit Daten fachlich korrekt zu vergleichen, keinen positiven oder sichtbaren Einfluss hat.

Andererseits zeigen ähnliche Aufgaben im Kompetenztest wie z.B. eine Aufgabe, in der die Schüler*innen ein Messergebnis von $g = (9,3 \pm 0,3) \text{ m/s}^2$ mit ein Referenzwert von $g_{\text{ref}} = 9,806 \text{ m/s}^2$ vergleichen sollen, dass diese durch 52 % der Schüler*innen richtig beantwortet werden. Eine Aufgabe, wo zwei grafisch dargestellte Messergebnisse miteinander verglichen werden, wird sogar durch 75 % der Schüler*innen richtig beantwortet.

Diese Diskrepanz ist erstaunlich.

Eine Erklärung dazu ist, dass der Transfer von konzeptuellem Wissen auf die Fähigkeit Daten korrekt miteinander zu vergleichen, eine zu große Anforderung darstellt. Hierfür können sprachliche oder inhaltliche Ursachen verantwortlich sind. Sprachlich wird in der Intervention und im Kompetenztest z.B. häufig über die „Verträglichkeit“ von Messergebnissen gesprochen, wo hingegen es sich in der Datenvergleichsaufgabe um „schneller“ und „gleich schnell“ handelt. Inhaltlich kann es sein, dass die Schüler*innen zu wenig geübt haben und auf ihren Automatismus Mittelwerte miteinander zu vergleichen, zurückgreifen.

Um die Diskrepanz weiter zu untersuchen, wird eine Studie mit Interviews geplant. Hier sollen Schüler*innen befragt werden, was sie während der Intervention gelernt haben, was sie für den Vergleich von Messreihen als relevant erachten und wie das ihre Vermutung beeinflusst. Auf Grundlage dieser Ergebnisse wird die Intervention weiterentwickelt.

Fazit dieser Pilotstudie ist, dass für die Gestaltung einer Lehrumgebung zum Umgang mit Messunsicherheiten die Konzepte und Aufgaben sehr genau ausgewählt und eindeutig formuliert werden müssen. Somit entstünde eine gute Passung zwischen Konzepten und Datenvergleichsaufgaben, was zu einem erfolgreichen Transfer zwischen beiden führen kann.

Literatur

- Chinn, C. A. & Malhotra, B. A. (2002). Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks. *Science Education*, 86(2), 175–218. <https://doi.org/10.1002/sce.10001>
- Heinicke, S. (2012). *Aus Fehlern Wird Man Klug: Eine Genetisch-Didaktische Rekonstruktion des Messfehlers*. Logos Verlag Berlin GmbH.
- Holmes, N. G., Wieman, C. E. & Bonn, D. A. (2015). Teaching critical thinking. *Proceedings of the National Academy of Sciences*, 112(36), 11199–11204. <https://doi.org/10.1073/pnas.1505329112>
- Kirch, S. A. (2010). Identifying and resolving uncertainty as a mediated action in science: A comparative analysis of the cultural tools used by scientists and elementary science students at work. *Science Education*, 94(2), 308–335. <https://doi.org/10.1002/sce.20362>
- Kok, K., Priemer, B., Musold, W. & Masnick, A. (2019). Students' conclusions from measurement data: The more decimal places, the better? *Physical Review Physics Education Research*, 15(1), 010103. <https://doi.org/10.1103/PhysRevPhysEducRes.15.010103>
- Ludwig, T. (2017). *Argumentieren beim Experimentieren in der Physik—Die Bedeutung personaler und situationaler Faktoren* (Humboldt-Universität).
- Masnick, A. M. & Morris, B. J. (2008). Investigating the Development of Data Evaluation: The Role of Data Characteristics. *Child Development*, 79(4), 1032–1048. <https://doi.org/10.1111/j.1467-8624.2008.01174.x>
- Moerke, P. (2020). Messunsicherheiten im Physikunterricht—Befragung von Lehrkräften in Baden-Württemberg. In: S. Habig (Hrsg.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Wien 2019. Universität Duisburg-Essen.
- Priemer, B. & Hellwig, J. (2018). Learning About Measurement Uncertainties in Secondary Education: A Model of the Subject Matter. *International Journal of Science and Mathematics Education*, 16(1), 45–68. <https://doi.org/10.1007/s10763-016-9768-0>
- Schulz, J., Priemer, B. & Masnick, A. (2018). Development of an Assessment Tool to Probe Student's Understanding of Measurement Uncertainties. *Electronic Proceedings of the ESERA 2017 Conference*, 1458–1467. Dublin, Ireland: Dublin City University.
- Sharma, S. V. (2006). High School Students Interpreting Tables and Graphs: Implications for Research. *International Journal of Science and Mathematics Education*, 4(2), 241–268. <https://doi.org/10.1007/s10763-005-9005-8>