

Tobias Ludwig¹
 Engin Kardaş¹
 Tim Donner²
 Burkhard Priemer³

¹Pädagogische Hochschule Karlsruhe
²Sportschule am Olympiapark, Berlin
³Humboldt-Universität zu Berlin

Messdaten auswerten lernen – adäquater argumentieren? Erste Ergebnisse

Vorarbeiten

Ein Ziel des Physikunterrichts ist es, dass Schüler*innen evidenzbasiert Aussagen über naturwissenschaftliche Inhalte treffen können (Schiepe-Tiska, Schöps, Rönnebeck, Köller & Prenzel, 2012). Eine Möglichkeit die vorangehenden Aspekte zu fördern stellt das Argumentieren dar, dem im naturwissenschaftlichen Unterricht eine große Bedeutung zukommt. Einerseits kann über Argumentationen ein angemessenes Bild z. B. des z. T. auch vorläufigen, diskursiven und unsicheren Charakters naturwissenschaftlicher Erkenntnis vermittelt werden (Driver, Newton, & Osborne, 2000), andererseits kann das Beurteilen und kritische Reflektieren gefördert werden (Osborne & Patterson, 2011). Das naturwissenschaftliche Experiment eignet sich in diesem Fall besonders, da die Schüler*innen Vermutungen, Behauptungen und Erläuterungen mit Beobachtungen sowie Daten in Zusammenhang bringen müssen. Dieser Prozess lässt sich als Argumentation auffassen (Osborne, 2012). Vor diesem Hintergrund wurde in verschiedenen Studien mit insg. 1500 Probanden untersucht, welchen Einfluss personale Faktoren (Fachwissen, situationales Interesse, Kognitionsbedürfnis, Werteinschätzung der Naturwissenschaften) und situationale Faktoren (bspw. reale vs. virtuelle Experimentierumgebung) auf das Argumentieren von Schüler*innen haben. Im Mittelpunkt der Untersuchung stand, wie Lernende auf der Grundlage selbstständig erhobener experimenteller Messdaten argumentieren (Ludwig, 2017; Ludwig, Priemer und Lewalter, 2019). Es wurde erfasst, wie stark die Schüler*innen beim Argumentieren auf „Intuition“, „Expertenwissen“, „Messunsicherheiten“ und „Daten als Evidenz“ zurückgegriffen haben. Lernende mit hoher Zustimmung in der Kategorie „Intuition“ begründen ihre Vermutungen z. B. stark auf Grundlage eines Bauchgefühls, während die Kategorie „Expertenwissen“ auf Expertise oder externe Autoritäten verweist. Die Kategorie „Messunsicherheiten“ beschreibt, wie stark explizit auf die Unsicherheit der erfassten Daten zurückgegriffen wurde. Der Bezug auf Daten bzw. statistische Kenngrößen (wie z. B. Mittelwert) wird durch die Kategorie „Daten als Evidenz“ abgebildet. Es konnte festgestellt werden, dass das Fachwissen der Schüler*innen die Ausprägung der Argumentkategorie „Daten als Evidenz“ begünstigt und die Verwendung von „Intuition“ reduziert. Während sich „Daten als Evidenzen“ positiv auf den Lernerfolg und auf die zeitliche Stabilität und die Richtigkeit einer Hypothese auswirkt, haben sowohl „Intuition“ und „Messunsicherheiten“ negative Auswirkungen auf den Lernerfolg. Erstere Kategorie wirkt sich negativ auf die zeitliche Stabilität des Lernerfolgs aus und letztere negativ auf die Richtigkeit einer Hypothese.

Ziel dieser Studie

Diese Ergebnisse legen nahe, dass Schüler*innen fachlich adäquater Argumentieren (gemessen an den Kategorien „Daten als Evidenz“ bzw. „Messunsicherheiten“), wenn Kompetenzen im Umgang mit Messdaten und Unsicherheiten vor dem Experiment vorhanden sind. Das Ziel des Projekts ist es daher, Lernumgebungen zu entwickeln, welche diesen thematischen Schwerpunkt aufgreifen. Die Wirksamkeit der Lernumgebungen wird im Hinblick auf a) den Einfluss der Lernumgebungen auf die Ausprägung der Zustimmung zu den Argumentkategorien und b) den Einfluss dieser wiederum auf den Lernerfolg untersucht. Daraus ergeben sich die folgenden Forschungsfragen:

- Führt die Bearbeitung digitaler Lernumgebungen durch die Schüler*innen dazu, dass sie physikalisch richtige Hypothesen aufstellen?
- Lassen sich signifikante Unterschiede in der Ausprägung der Argumentationskategorien der Schüler*innen vor und nach der Intervention nachweisen?
- Welchen Einfluss hat die Ausprägung der Argumentkategorie auf die Richtigkeit der Hypothese?

Methode

Der Entwicklung der Lernumgebungen liegt ein Design-Based-Research-Ansatz zugrunde. Der Prozess ist zirkulär, da die Lernumgebungen in der Praxis eingesetzt, evaluiert und anschließend überarbeitet werden. Im Rahmen dieser Studie wurden zunächst zwei Prototypen entwickelt, welche die „Streuung von Messwerten“ als auch den „Vergleich von Messergebnissen“ thematisieren (Hellwig, 2013). Die Lernumgebungen sind modular einsetzbar und in rund 20-25 Minuten bearbeitbar. Die Umsetzung der Lernumgebung erfolgt digital, um den Zugriff und die Verbreitung zu vereinfachen. Ferner sind interaktive Elemente enthalten, die den Schüler*innen eine individuelle und selbstregulierte Bearbeitung ermöglichen soll. Beispielsweise können die Schüler*innen innerhalb der Lernumgebung selbstständig Messwerte visualisieren und graphische Darstellungen variieren. Weiterhin wird Scaffolding eingesetzt, indem die Schüler*innen eigenständig durch die Programme gehen und je nach Vorgehen ein individuelles Feedback oder eine Hilfestellung erhalten. Bevor die Lernumgebungen hinsichtlich ihrer Wirksamkeit untersucht wurden, wurden mittels Think-Aloud-Studien mögliche Schwierigkeiten in der Bedienung oder Probleme mit dem Layout weitestgehend beseitigt. Als Untersuchungsdesign wurde eine Längsschnittstudie durchgeführt. Der Gesamtumfang beträgt n=75 Schüler*innen aus der achten und neunten Jahrgangsstufe. Die Gesamtdauer der Untersuchung betrug maximal 90 Minuten, wobei die beiden Lernumgebungen in höchstens 45 Minuten bearbeitet wurden. Zu Beginn der Befragung wurde den Schüler*innen ein unbekannter physikalischer Kontext präsentiert (vgl. Ludwig, 2017). Daraufhin wurden die Vermutungen der Lernenden zu diesem Kontext erfasst, wobei anzumerken ist, dass die Schüler*innen aus vorgegebenen Hypothesen auswählen und anschließend ihre Auswahl begründen sollten. Im Anschluss daran wurden Messwerte zu dem physikalischen Kontext präsentiert und nachdem die Schüler*innen ebendiese betrachtet und ausgewertet hatten, wurde erneut die Hypothese inklusive der Begründung abgefragt. Zusätzlich wurden die Zustimmungen zu den Argumentkategorien unter Verwendung eines Fragebogens erfasst. Nun wurden die Lernumgebungen von den Lernenden bearbeitet. Zum Abschluss der Untersuchung wurden den Schüler*innen dieselben Daten wie zuvor präsentiert, erneut die Hypothesen und die dazugehörigen Begründungen aufgenommen sowie die Argumentkategorien erfasst.

Ergebnisse

Im Hinblick auf die erste Forschungsfrage wurde analysiert, wie sich die Anzahl der richtigen und falschen Hypothesen im Verlauf der Befragung entwickelt hat. Zum ersten Befragungszeitpunkt wählten 73,3% der Lernenden eine falsche Hypothese. Nachdem die Lernenden die Messwerte betrachtet haben, steigt die Anzahl der falschen Hypothesen auf 78,7%. Die letzte Erfassung der Hypothesen offenbart, dass der Anteil der falschen Hypothesen auf 50,7% zurückgegangen und der Anteil der richtigen Hypothesen auf 49,3% gestiegen ist. Es handelt sich um eine signifikante Änderung zwischen dem zweiten und dritten Erfassungszeitpunkt, was darauf schließen lässt, dass die Intervention den Lernenden dabei hilft, eine physikalisch richtige Hypothese zu formulieren ($\chi^2(1) = 11.6, p < .001$). Im nächsten Schritt wurde untersucht, wie sich die Zustimmung zu den Argumentkategorien vor und nach der Intervention verändert hat. Dafür wurde zunächst ein t-Test für verbundene Stichproben durchgeführt, der einen signifikanten Effekt zwischen den zuvor erwähnten Zeitpunkten in den Kategorien „Intuition“, „Expertenwissen“ und „Messunsicherheiten“ aufzeigt. Bei der

Bestimmung der Effektstärken für die jeweiligen Kategorien wurden eine negative Wirkung auf die Kategorie „Intuition“ ($d=-0.27$, $t(74)=3.65$, $p<.001$) und „Expertenwissen“ ($d=-0.18$, $t(74)=2.15$, $p<.05$) sowie eine positive Wirkung auf den Bereich „Messunsicherheiten“ ($d=0.69$, $t(74)=5.69$, $p<.001$) festgestellt. Zusammenfassend lässt sich sagen, dass die Intervention eine stärkere Zustimmung zur Kategorie „Messunsicherheiten“ fördert und dazu führt, dass die eher nicht-rationalen Argumentkategorien „Expertenwissen“ und „Intuition“ weniger ausgeprägt sind. Abschließend wurde untersucht, welchen Einfluss die Argumentkategorien auf den Lernerfolg haben. Der Lernerfolg stellt eine dichotome Variable dar (richtige oder falsche Hypothese), daher wurde eine multifaktorielle logistische Regression durchgeführt. Darauf aufbauend wurde für jede Kategorie vor und nach der Intervention die Odds Ratio berechnet. Im Fall vor der Intervention lässt sich allenfalls von Tendenzen sprechen, da es in diesem Fall keine signifikanten Ergebnisse gibt. Die Kategorien „Intuition“ ($OR=0.74$, $p=0.30$) und „Daten als Evidenz“ ($OR=0.57$, $p=0.08$) weisen negative Tendenzen bezüglich des Lernerfolgs auf und „Messunsicherheiten“ ($OR=1.72$, $p=0.21$) legt eine positive Tendenz nahe. Die Ausprägung der Argumentkategorie „Expertenwissen“ zeigt keinen Einfluss auf den Lernerfolg. Nach der Bearbeitung der Lernumgebungen verändert sich der Einfluss der Argumentkategorien auf den Lernerfolg. Während bei den Kategorien „Intuition“ ($OR=0.78$, $p=0.29$), „Expertenwissen“ ($OR=0.78$, $p=0.27$) und „Daten als Evidenz“ ($OR=1.03$, $p=0.91$) weiterhin nur von Tendenzen gesprochen werden kann, gibt es einen kleinen signifikanten Effekt in der Kategorie „Messunsicherheiten“ ($OR=1.72$, $p=0.03$). In den restlichen Kategorien liegen erneut negative Tendenzen vor oder sie haben auf den Lernerfolg keinen Einfluss. Die Ausprägung der Argumentkategorie „Messunsicherheiten“ wirkt sich somit positiv auf den Lernerfolg aus.

Diskussion

Es lässt sich zum einen konstatieren, dass die Lernumgebungen dazu geeignet sind, den Umgang mit Daten zu fördern. Denn nach der Intervention erhöht sich der Anteil der fachlich richtigen Hypothese. Zum anderen führen –entsprechend der Implikationen aus den Vorarbeiten– Kompetenzen im Umgang mit Daten und Unsicherheit zu einer höheren Ausprägung lernförderlicher Argumentkategorien: Die Schüler*innen gehen eher dazu über, Messdaten als Evidenz heranzuziehen und auf dieser Grundlage zu argumentieren. Gleichzeitig unterliegt diese Untersuchung gewissen Limitationen. Zum einen handelt es sich bei $n=75$ um eine kleine Stichprobe, so dass hier berichtete nicht-signifikante Tests aufgrund einer geringen Teststärke Effekte ggf. nicht aufdecken konnte (Fehler 2. Art). Aufgrund der Ergebnisse kann zwar davon ausgegangen werden, dass der Umgang mit Daten gefördert wurde, aber eine explizite Kompetenzerfassung in diesem Bereich liegt nicht vor und ist Gegenstand künftiger Studien. Weiterhin wurde zwar darauf geachtet, dass sich der physikalische Kontext vor und innerhalb der Lernumgebung unterscheiden, jedoch bleiben die Untersuchung und damit die Resultate hoch situativ. Dies sollte in zukünftigen Untersuchungen beachtet werden.

Ausblick

Für das weitere Forschungsvorhaben sollen weitere Lernumgebungen zum Umgang mit Messunsicherheiten erstellt werden. Dabei dient ein Sachstrukturmodell (Hellwig, 2013) als Grundlage für die Schwerpunktkompetenzen in den Lernumgebungen. Die Lernumgebungen sollen weiterhin bezüglich ihrer Wirksamkeit untersucht werden, wobei die empirische Evaluation mit einer größeren Stichprobe realisiert werden soll. Neben der reinen Vermittlung der Kompetenzen im Umgang mit Daten und Unsicherheiten sollen zukünftige Untersuchungen und Evaluationen diese Kompetenzen sowie deren Einfluss auf die Argumentation der Lernenden erfassen.

Das Projekt „Förderung der Argumentationsfähigkeit beim Experimentieren im Physikunterricht“ wird gefördert durch die Deutsche Telekom Stiftung.

Literatur

- Driver, R., Newton, P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287-312, [https://doi.org/10.1002/\(SICI\)1098-237X\(200005\)84:3<287::AID-SCE1>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A)
- Hellwig, J. (2013). Messunsicherheiten verstehen: Entwicklung eines normativen Sachstrukturmodells am Beispiel des Unterrichtsfaches Physik (Dissertation). Ruhr-Universität Bochum, Bochum. Retrieved from <http://hss-opus.ub.ruhr-unibochum.de/opus4/frontdoor/index/index/docId/1700>
- Ludwig, T. (2017). Argumentieren beim Experimentieren—Die Bedeutung personaler und situationaler Faktoren (Dissertation). Humboldt-Universität zu Berlin, Berlin. <http://dx.doi.org/10.18452/18408>
- Ludwig, T., Priemer, B., & Lewalter, D. (2019). Assessing Secondary School Students' Justifications for Supporting or Rejecting a Scientific Hypothesis in the Physics Lab, *Research in Science Education*, <https://doi.org/10.1007/s11165-019-09862-4>
- Osborne, J. (2012). The Role of Argument: Learning How to Learn in School Science. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 933–949). https://doi.org/10.1007/978-1-4020-9041-7_62
- Osborne, J., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, 95(4), 627–638. <https://doi.org/10.1002/sce.20438>
- 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>
- Schiepe-Tiska, A., Schöps, K., Rönnebeck, S., Köller, O., & Prenzel, M. (2012). Naturwissenschaftliche Kompetenz in PISA 2012: Ergebnisse und Herausforderungen. In M. Prenzel, E. Sälzer, E. Klieme, & O. Köller (Eds.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (pp. 189–219). Münster: Waxmann.