

Stephan Pfeiler  
Burkhard Priemer  
Annette Upmeier zu Belzen

Humboldt-Universität zu Berlin

### **Der kritische Umgang mit Primär- und Sekundärdaten im naturwissenschaftlichen Unterricht**

#### **Einleitung**

Der Umgang mit und die Evaluation von Daten ist elementarer Bestandteil von wissenschaftlichem Denken (Klahr, 2002). Um in der Schule die Verarbeitung und Bewertung von Datensätzen zu erlernen, werden in der Regel Daten aus Experimenten verwendet. Diese können von Schülerinnen und Schülern selbst (Primärdaten) oder von der Lehrkraft durchgeführt werden. Die Daten können auch aus der Arbeit anderer, externer Personen stammen, welche nicht mit der Schule assoziiert sind (Sekundärdaten). Es wird vermutet, dass die Behandlung von eigenhändig gewonnenen Daten und den Daten dritter Personen unterschiedliche Wirkungen auf das Lernen von Schülerinnen und Schülern haben (Kanari & Millar, 2004; Osborne, Collins, Ratcliffe, Millar & Duschl, 2003). Die empirische Forschung kann dazu aber bisher keine klaren Ergebnisse vorweisen.

#### **Stand der Forschung**

Die Arbeiten von Hug und McNeill (2008) und Magnusson, Palincsar, Hapgood und Lomangino (2004) bilden den momentanen Forschungsstand zum Vergleich der beiden Datentypen ab.

Laut Hug und McNeill (2008) sind Primärdaten solche, die während einer eigenen Untersuchung gewonnen werden. Damit ist also die Beteiligung der Schülerinnen und Schüler an der Erhebung der Daten der Indikator für die Unterscheidung von Primär- und Sekundärdaten. Sekundärdaten sind Daten, die von anderen Personen gesammelt und zur Verfügung gestellt werden. Dies unterscheidet sich von der Definition von Magnusson et al. (2004), welche Sekundärdaten als Bericht über die Untersuchung einer dritten Person verstehen. Es handelt sich also um Daten, die in einen Text zu einer Untersuchung eingebettet sind. In diesem Text wird ein Phänomen untersucht, welches die Schülerinnen und Schüler ebenfalls behandelt haben.

Hug und McNeill (2008) weisen darauf hin, dass ihre Definition zu kurz gefasst ist. Sie impliziert, dass es eine scharfe Grenze gibt, welche beide Datentypen voneinander abgrenzt. Jedoch verstehen sie die Ausprägung des Datentyps eher als kontinuierlich. Demnach gibt es Faktoren, die diese Ausprägung beeinflussen, in der Literatur jedoch nicht explizit genannt oder spezifiziert werden. Hug und McNeill (2008, S. 1728) nennen z. B. „personal relevance“ als wahrgenommene Eigenschaft von Daten, welche diese haben, wenn sie primär sind. Was damit genau gemeint ist, bleibt jedoch unklar. Wichtigster Faktor ist u. E. die Beteiligung an der Erhebung der Daten. So müssen Schülerinnen und Schüler immer an der Erhebung der Daten durch eigenes Handeln beteiligt sein, damit von Primärdaten die Rede sein darf. Es muss jedoch beachtet werden, dass sich das Ausmaß dieser Beteiligung variieren lässt und so unterschiedlich wirken könnte.

Für das vorliegende Forschungsprojekt sind grundsätzlich vier potentielle Faktoren relevant. 1. das Ausmaß der Beteiligung an der Erhebung der Daten, 2. die Datenquelle bzw. der Autor, 3. der Grad der Aufbereitung der Daten und 4. das Wissen über das Messverfahren aus dem der Datensatz stammt.

In der bisherigen Forschung zu Primär- und Sekundärdaten wurde nicht untersucht, inwiefern sich die Ausprägungen dieser Faktoren auf das untersuchte Konstrukt, z. B. den Lernerfolg, auswirken. Es muss davon ausgegangen werden, dass die bisherigen Ergebnisse

auf das gleichzeitige Zusammenspiel vieler Einzelfaktoren zurückzuführen ist. Hug und McNeill (2008, S. 1747) geben beispielsweise an, dass Sie nicht darauf geachtet haben, dass die verwendeten Datensätze den gleichen Grad an Komplexität aufweisen. Auch wurden Datensätze zu verschiedenen Phänomenen aus verschiedenen Fächern benutzt.

Die Abgrenzung der einzelnen Faktoren ist nicht einfach. Sie beeinflussen sich gegenseitig. Wir vermuten, dass eine Beteiligung an der Erhebung der Daten zu Wissen über das Messverfahren führt, wobei mehr Beteiligung zu mehr Wissen führen sollte (Frischknecht-Tobler & Labudde, 2010). Der Grad der Aufbereitung der Daten wird von der vermeintlichen Expertise der Datenquelle abhängig sein. Je stärker die Daten aufbereitet sind, desto stärker muss darauf geachtet werden, dass der Autor der Daten diese Aufbereitung sachgemäß und richtig durchgeführt hat. Die Schülerinnen und Schüler werden davon ausgehen, dass dies eher der Fall sein wird, wenn sie wissen, dass der Autor die Aufbereitung der Daten beherrscht. Hierbei wird auch deutlich, dass es nicht die objektiven Eigenschaften des Autors sind, die die Ausprägung des Datentyps beeinflussen, sondern die von den Schülerinnen und Schülern wahrgenommenen Eigenschaften dieses Autors.

Laut Sandoval (2004) sollte eine Untersuchung zu epistemologischen Überzeugungen auch beachten, dass es Hinweise auf domänenspezifische Unterschiede zwischen verschiedenen Fächern gibt. Daher wird diese Studie physikalische und biologische Kontexte nutzen.

### **Forschungsfrage**

Es wird vermutet, dass die Autorenschaft der Daten (zur Unterscheidung von Primär- und Sekundärdaten) beeinflusst, wie Schülerinnen und Schüler bezüglich des Beibehaltens oder Verwerfens einer eingangs aufgestellten Hypothese argumentieren. Zur Operationalisierung des Argumentierens wird untersucht, ob Schülerinnen und Schüler Evidenzkriterien für Datensätze anders anwenden, wenn der Typ der Daten variiert wird. Die Forschungsfragen lauten:

- Welchen Einfluss hat die Form der Datenquelle (Primär- oder Sekundärdaten) auf das Verwerfen oder Beibehalten von eingangs aufgestellten Hypothesen und auf die Einschätzung der Evidenz (Qualität) der Daten?
- Gibt es dabei Unterschiede zwischen den Disziplinen Physik und Biologie?

### **Methode**

Als unabhängige Variable wurde einer der Faktoren für die Unterscheidung von Primär- und Sekundärdaten gewählt, die Autorenschaft. Die abhängige Variable ist zum Einen das Beibehalten oder Verwerfen der zu Beginn aufgestellten Hypothese und zum Anderen die zur Begründung der Entscheidung genutzten Evidenzkriterien.

Die Autorenschaft des Datensatzes wurde als Variable gewählt, da die Interaktion zwischen ihr und den anderen Faktoren, z.B. Wissen über das Messverfahren, als eher gering vermutet wird. Die Autorenschaft kann variiert werden. Dabei würde ein Ende des Spektrums einen Autor beschreiben, der sich in sozialer Nähe zu der Kontrollgruppe befindet. Solch ein Datensatz würde zum Beispiel von Mitschülerinnen und -schülern oder der Lehrkraft stammen. Am anderen Ende des Spektrums könnten Daten von Wissenschaftlern liegen. Die Schülerinnen und Schüler sind jedoch der Meinung, dass die Wissenschaftler in den Naturwissenschaften geschult sind. So wird vermutet, dass die Daten von Wissenschaftlern eine gewisse Autorität ausstrahlen (Hug und McNeill, 2008). Methodisch umgesetzt wird dies mit drei Gruppen: 1. mit eigenen Datensätzen, 2. mit Datensätzen von anderen Schülerinnen und Schülern und 3. mit Datensätzen von vermeintlichen Wissenschaftlern.

Zusätzlich soll eine Unabhängigkeit von der Disziplin am Beispiel Physik und Biologie geprüft werden. Insgesamt werden deshalb sechs Gruppen von Schülerinnen und Schülern der Sekundarstufe 1 an der Studie teilnehmen.

Die Untersuchung beginnt mit einer Lerneinheit zur Anwendung von Evidenzkriterien auf Daten. Diese werden aus den „Concepts of Evidence“ von Gott, Duggan und Roberts (2014) abgeleitet. Die Lerneinheit stellt sicher, dass alle Teilgruppen die gleichen Fähigkeiten zur Einschätzung der Qualität der Daten haben. Auf diese Lerneinheit folgt ein Test, um die Wirkung der Lerneinheit zu kontrollieren. Darauf folgen experimentelle Tätigkeiten in Physik und Biologie, die empirische Daten erzeugen. Die Eigenschaften der Datensätze aus Physik und Biologie müssen aufeinander abgestimmt sein. Sie müssen sich in Umfang, Komplexität und Grad der Verarbeitung gleichen. Die Experimente werden ferner so gewählt, dass ihre Ergebnisse mit alltagsbasierten Schülervorstellungen konfliktieren. Ein physikalisches Experiment, welches sich in diesem Zusammenhang bewährt hat, ist das Fadenpendel und der Zusammenhang von schwingender Masse und Periodendauer (Ludwig & Priemer, 2013). Einen biologischen Kontext könnte der Wasserhaushalt von Pflanzen darstellen. Die Schülerinnen und Schüler stellen eine Hypothese auf, führen das Experiment durch und nehmen im Anschluss auf Basis der verschiedenen Datensätze Stellung zu ihrer Hypothese. Ein Großteil der Schülerinnen und Schüler wird zu Beginn der Studie eine Hypothese formulieren, die nicht von den Daten gestützt wird. Dadurch wird ein stärkerer Drang zur Beschäftigung mit dem Phänomen erreicht (Kircher, 2010). Der Posttest erfasst den Wechsel bzw. das Beibehalten der Hypothese sowie das Argumentieren der Probanden. Hierzu wird auf Arbeiten von Ludwig und Priemer (2013) zurückgegriffen.

### **Abschluss**

Diese Studie soll helfen, die Unterscheidung zwischen Primär- und Sekundärdaten zu präzisieren. Erst auf Basis dieser Präzisierung kann eine Aussage über die Wirkung dieser verschiedenen Datenquellen auf das Lernen von Schülerinnen und Schülern getroffen werden.

### **Literatur**

- Frischknecht-Tobler, U., & Labudde, P. (2010). Beobachten und Experimentieren. In P. Labudde (Hrsg.), Fachdidaktik Naturwissenschaft 1. - 9. Schuljahr. Haupt Verlag.
- Gott, R., Duggan, S. & Roberts, R. (2014). Concepts of Evidence. Understanding Scientific Evidence. URL: <http://community.dur.ac.uk/rosalyn.roberts/Evidence/cofev.htm> (besucht am 22. 08. 2014).
- Hug, B. & McNeill, K. L. (2008). Use of First-hand and Second-hand Data in Science: Does data type influence classroom conversations? In: International Journal of Science Education 30.13, 1725–1751
- Kanari, Z. & Millar, R (2004). Reasoning from data: How students collect and interpret data in science investigations. In: Journal of Research in Science Teaching 41.7, 748–769
- Kircher, E. (2010). Methoden im Physikunterricht. In E. Kircher, R. Girwidz, & P. Häußler (Eds.), Physikdidaktik. Springer Berlin Heidelberg
- Klahr, D. (2002). Exploring Science – The Cognition and Development of Discovery Processes. The MIT Press
- Ludwig, T. & Priemer, B. (2013). Secondary School Students' Reasoning from Anomalous Data. Vorgetragen auf NARST Annual International Conference, Puerto Rico.
- Magnusson, S. J., Palincsar, A. S., Hapgood, S., & Lomangino, A. (2004). How Should Learning Be Structured in Inquiry-based Science Instruction?: Investigating the Interplay of 1St- and 2Nd-hand Investigations. In Proceedings of the 6th International Conference on Learning Sciences, 318–325. Santa Monica, California: International Society of the Learning Sciences.
- Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What “ideas-about-science” should be taught in school science? A Delphi study of the expert community. In: Journal of Research in Science Teaching 40.7, 692–720
- Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. In Science Education, 89(4), 634–656