

## Welche Schwierigkeiten haben Schülerinnen und Schüler beim Auswerten von Versuchsdaten?

### Theoretischer Hintergrund und Stand der Forschung

Häufig werden naturwissenschaftliche Denk- und Arbeitsweisen als Modi eines naturwissenschaftlichen Problemlösens konzeptualisiert, welches als *Scientific Reasoning* bzw. im deutschen Sprachraum auch als *Wissenschaftliches Denken* begrifflich gefasst wird (Opitz, Heene & Fischer, 2017). An die Arbeiten von Klahr (2000) anknüpfend ist eine Dreiteilung in die drei Komponenten Hypothesen generieren, Untersuchungen planen sowie Daten auswerten üblich. Das kompetente Anwenden dieser Denk- und Arbeitsweisen gilt als Teil des Kompetenzbereichs Erkenntnisgewinnung und dadurch als Lernziel naturwissenschaftlichen Unterrichts (z.B. MSW NRW, 2014).

Das Scientific Reasoning (vgl. Abb. 1) von Schülerinnen und Schülern wird als wissensbasiert konzeptualisiert (z.B. Mayer, 2007). Als relevant werden dabei drei Wissensbereiche angesehen: (1) *Fachwissen* über domänenspezifische Konzepte; (2) *prozedurales Wissen* bzgl. Strategien und Vorgehensweisen spezifischer Denk- und Arbeitsweisen; und (3) *epistemisches Wissen*, welches deklarativ vorliegt und grundlegende Konzepte naturwissenschaftlichen Arbeitens, wie z.B. Gütekriterien, umfasst (Arnold, 2015; Kind, 2013).



Abb. 1 Das naturwissenschaftliche Problemlösen (*Scientific Reasoning*) lässt sich in drei Denk- und Arbeitsweisen einteilen. Für diese wiederum lassen sich einzelne Komponenten ausdifferenzieren (hier für die Auswertung erfolgt). Die Fähigkeit zum adäquaten *Scientific Reasoning* wird dabei vom Wissen in drei Bereichen beeinflusst.

Das Auswerten von Daten ist eine Denk- und Arbeitsweise, die insbesondere im Kontext des Forschenden Lernens hohe Relevanz einnimmt (Hug & McNeill, 2008). Die Arbeitsweise zeichnet sich dadurch aus, dass (z.B. Versuchs-)Daten zunächst aufbereitet, d.h. geeignet ausgewählt und ggf. in eine Repräsentationsform überführt werden, die nähere Analysen erlaubt. Die aufbereiteten Daten werden in der Folge interpretiert, indem sie mit dem fachlichen Vorwissen in Bezug gesetzt und auf Trends und Muster hin untersucht werden. Die interpretierten Daten werden dann genutzt, um Aussagen bzgl. der Fragestellung bzw. der Hypothesen zu machen. Diese Schlussfolgerungen werden verallgemeinert, bzw. es wird eingeschätzt, welche weiteren Informationen oder Untersuchungen es bedarf, um eine solche Generalisierung zu erlauben. Alle vorhergehenden Komponenten werden von der Beurteilung der Gültigkeit der Daten und des experimentellen Designs und Vorgehens begleitet (bspw. Chinn & Malhotra, 2002; Wellnitz & Mayer, 2013).

Schülerinnen und Schüler sind während der Auswertung von Daten mit Schwierigkeiten konfrontiert. Trends aus Daten abzuleiten, fällt vielen Schülerinnen und Schülern schwer (Jeong, Songer & Lee, 2007). Anomale, d.h. dem Vorwissen bzw. den Erwartungen nicht entsprechenden Beobachtungen werden von Schülerinnen und Schülern nicht adäquat behandelt (Toplis, 2007). Auch das Einbeziehen von Messungenauigkeiten und -fehlern in die Auswertung und allgemeiner das Berücksichtigen von Reliabilität und Validität als Gütekriterien stellen eine Schwierigkeit dar (Lubben & Millar, 1996). Dies äußert sich beispielsweise darin, dass vielen Schülerinnen und Schülern die Auswahl geeigneter Versuchsdaten für das Ziehen von Schlussfolgerungen nur eingeschränkt gelingt (McNeill & Krajcik, 2007). Auch sind diese Schlussfolgerungen und Generalisierungen von Schülerinnen und Schülern häufig nicht ausreichend mit Daten belegt (Sandoval & Millwood, 2005).

Vorliegende Studien haben vor allem Fähigkeiten von jüngeren Lernenden untersucht, wodurch bislang nur eingeschränkte Evidenz vorliegt, die Aussagen über den Förderbedarf von Schülerinnen und Schülern in der Oberstufe erlaubt (Arnold et al., 2014, Vorholzer et al., 2018). In den inhaltlich und prozedural komplexeren experimentellen Untersuchungen der Oberstufe (vgl. MSW NRW, 2014) ist allerdings das Auftreten spezifischer Schwierigkeiten plausibel. Als ein Resultat des in dieser Hinsicht limitierten Forschungsstands liegen bislang wenige Studien vor, die Fördermaßnahmen für Schülerinnen und Schüler in der Sekundarstufe II oder entsprechenden Klassenstufen beim Auswerten experimentell erhobener Daten diskutieren (van Rens, 2014). Zudem ist der Wissensstand zu Schwierigkeiten im Chemieunterricht wenig ausgebaut: Bestehende Befunde sind häufig in benachbarten Domänen (insbesondere der Physik) oder in domänenunspezifischen Settings kontextualisiert.

### **Forschungsfragen**

Aus dem beschriebenen Forschungsstands lässt sich ableiten, dass Bedarf an einer genaueren Betrachtung der Performanz von Chemielernenden der Oberstufe beim Auswerten von Daten besteht. Das hier beschriebene Dissertationsprojekt fokussiert dabei auf die Auswertung von selbst erhobenen Daten (vgl. Hug & McNeill, 2008) mit dem Ziel, Kategorien von Schwierigkeiten zu identifizieren, die die Schülerinnen und Schüler beim kompetenten Durchführung dieses Bereichs des Scientific Reasonings einschränken. Aus dem in Abb. 1 dargestellten Modell des Scientific Reasonings lässt sich darüber hinaus hypothetisch ableiten, dass der Wissensstand der Lernenden bzgl. des prozeduralen, epistemischen und Fachwissens Einfluss auf das Auftreten von bestimmten Schwierigkeiten hat. Entsprechend lassen sich die folgenden zwei Forschungsfragen formulieren.

*FF1:* Welche Schwierigkeiten haben Lernende der Oberstufe bei der Datenauswertung im Fach Chemie?

*FF2:* Welche Zusammenhänge bestehen zwischen dem Auftreten dieser Schwierigkeiten und dem prozeduralen und epistemischen Wissen zum Auswerten von Daten sowie dem Fachwissen der Lernenden?

### **Studiendesign**

Die Forschungsfragen werden beantwortet, indem die Handlungen und die Kommunikation von Lernenden der Oberstufe während der Auswertung selbst erhobener Daten sowie ihre dabei erzielten Ergebnisse analysiert werden. Dazu führen die Lernenden zwei Versuche durch, in welchen sie in Zweiergruppen quantitative Datensätze generieren. Dazu erhalten Sie digitale Sensoren, deren Messwerte mithilfe eines Tabletcomputers dargestellt werden können. Die erhaltenen Datensätze werden von den Schülerinnen und Schülern in der Folge anhand von offenen Aufgabenstellungen, die auf die oben beschriebenen fünf Komponenten

der Datenauswertung abzielen, ausgewertet. Die Lernenden notieren ihre Antworten in Aufgabenheften, die nach Fertigstellung eingesammelt werden.

Ein gegenwärtig in der Pilotierung befindliches Experimentiersetting lässt Schülerinnen und Schüler anhand des Vergleichs der Messwerte von Titrationsen einer Säure-Base-Pufferlösung und einer gleich konzentrierten Kochsalzlösung die chemischen Eigenschaften des betreffenden Puffers untersuchen. Die Versuche sind durch eine vorgegebene Herleitung von Hypothesen und Versuchsplanung, eine angeleitete Versuchsdurchführung und die Verwendung digitaler Sensoren vorstrukturiert, um weitestgehend sicherzustellen, dass es sich bei den zu beobachtenden Schwierigkeiten um genuine Schwierigkeiten bei der Datenauswertung und nicht um Artefakte von fehlerbehaftetem Vorgehen der Lernenden zu früheren Zeitpunkten handelt. Eine weitere, thematisch und strukturell ähnliche Experimentierumgebung wird auf Grundlage der Ergebnisse einer gegenwärtig durchgeführten Pilotstudie ( $N = 50$ ) entwickelt.

Die Analyse der Aktivitäten der Lernenden in der Hauptstudie und der dabei auftretenden Schwierigkeiten während der Auswertung dieser zwei Versuche soll durch eine Analyse von Videodaten der Zweiergruppen erfolgen. Um entsprechendes Datenmaterial zu erheben, werden je teilnehmender Klasse bis zu fünf Zweiergruppen ausgewählt, deren Arbeit während der Auswertung des Versuchs videographiert wird (vgl. Arnold et al., 2014). Bei einer Stichprobe von ca.  $N = 100$  (d.h. 50 Zweiergruppen), soll eine Teilstichprobe von  $n = 40$  Schülerinnen und Schülern (d.h. 20 Zweiergruppen) videographiert werden. Darüber hinaus werden von der gesamten Stichprobe die verschriftlichten Antworten auf die Auswertungsaufgaben eingesammelt.

Die Videos sollen in der Folge durch eine weitestgehend induktive qualitative Inhaltsanalyse auf Schwierigkeiten hin gesichtet werden (Mayring, 2015). Ziel ist es dabei, Kategorien von Schwierigkeiten zu identifizieren, die bei den Schülerinnen und Schülern während der Auswertung des Versuchs auftreten. Die Analyse der Videodaten wird dabei turnbasiert durchgeführt. Als Selektionskriterium für das Kodieren eines Turns wird dabei auf die Definition einer „unmittelbar beobachtbaren Schülerschwierigkeit“ von Kechel (2016, S. 57ff., S. 113) zurückgegriffen. Diese nutzt das Verfehlen eines Teilziels der Aufgabe, unerwünschte Handlungen und Ausdrücke von Mühe und/oder negativem Erleben durch die Lernenden zur Identifikation einer Schwierigkeit. Im Zuge der Pilotierung der Experimentiersettings sollen bereits Grundlagen eines Kodiermanuals entwickelt werden.

Um Schlüsse bzgl. der zweiten Forschungsfrage zu ziehen, werden die Probanden in der Hauptstudie auf Grundlage ihrer Leistungen in Tests zu den drei in Abbildung 1 dargestellten Wissensbereichen zu homogenen Zweiergruppen gruppiert. So soll qualitativ geprüft werden, ob die Ausprägung des Vorwissens der Zweiergruppen in bestimmten Wissensbereichen mit dem Auftreten spezifischer Schwierigkeiten in Zusammenhang steht. Darüber hinaus können Analysen bzgl. der Bewertungen der Antworten der Lernenden auf die Auswertungsaufgaben und der Testscores durchgeführt werden. Zur Erhebung des Vorwissens der Lernenden wurden drei Instrumente adaptiert. Das Instrument zum prozeduralen Wissen besteht aus 15 Multiple-Choice-Items, die aus verschiedenen bestehenden Instrumenten (Henke, 2007; Nehring, 2015; Roberts & Gott, 2003; Vorholzer, von Aufschnaiter & Kirschner, 2016) adaptiert wurden. Das Instrument zum epistemischen Wissen wurde aus Arnold (2015) entnommen und um drei Items auf insgesamt 13 Single-Choice-Items gekürzt. Beim Instrument zum Fachwissen handelt es sich um 24 Single-Choice-Items, welche sich mit dem Themenbereich Säuren und Basen befassen und aus einem für den Einsatz in der Oberstufe entwickelten Itempool von Hülsmann (2015) stammen. Parallel zur Pilotierung des Experimentiersettings werden gegenwärtig auch die Testinstrumente erprobt und evaluiert ( $N = 100$ ).

## Literatur

- Arnold, J., Kremer, K. & Mayer, J. (2014). Understanding Students' Experiments – What kind of support do they need in inquiry tasks? *International Journal of Science Education*, 36 (16), 2719–2749. <https://doi.org/10.1080/09500693.2014.930209>
- Arnold, J. (2015). *Die Wirksamkeit von Lernunterstützungen beim Forschenden Lernen*. Berlin: Logos.
- 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>
- Henke, C. (2007). *Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe. Untersuchung am Beispiel des HIGHSEA-Projekts in Bremerhaven*. Berlin: Logos.
- Hug, B. & McNeill, K. L. (2008). Use of First-hand and Second-hand Data in Science: Does data type influence classroom conversations? *International Journal of Science Education*, 30 (13), 1725–1751. <https://doi.org/10.1080/09500690701506945>
- Hülsmann, C. (2015). *Kurswahlmotive im Fach Chemie. Eine Studie zum Wahlverhalten und Erfolg von Schülerinnen und Schülern in der gymnasialen Oberstufe*. Berlin: Logos.
- Jeong, Heisawn; Songer, Nancy B.; Lee, Soo-Young (2007): Evidentiary Competence. Sixth Graders' Understanding for Gathering and Interpreting Evidence in Scientific Investigations. In: *Research in Science Education* 37 (1), S. 75–97. [10.1007/s11165-006-9014-9](https://doi.org/10.1007/s11165-006-9014-9).
- Kechel, J.-H. (2016). *Schülerschwierigkeiten beim eigenständigen Experimentieren. Eine qualitative Studie am Beispiel einer Experimentieraufgabe zum Hooke'schen Gesetz*. Berlin: Logos.
- Kind, P. M. (2013). Establishing Assessment Scales Using a Novel Disciplinary Rationale for Scientific Reasoning. *Journal of Research in Science Teaching*, 50 (5), 530–560. <https://doi.org/10.1002/tea.21086>
- Klahr, D. (2000). *Exploring Science*. Cambridge/London: MIT Press.
- Lubben, F. & Millar, R. (1996). Children's ideas about the reliability of experimental data. *International Journal of Science Education*, 18 (8), 955–968. <https://doi.org/10.1080/0950069960180807>
- Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (12., überarb. Aufl.). Weinheim: Beltz.
- McNeill, K. L. & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In: Lovett, M. & Shah, P. (Hrsg.), *Thinking with data*. (S. 233–265). New York: Taylor & Francis.
- MSW NRW. (2014). *Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen: Chemie*. Verfügbar unter [https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan/151/KLP\\_GOSt\\_Chemie.pdf](https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan/151/KLP_GOSt_Chemie.pdf)
- Nehring, A. (2015): *Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung*, Berlin: Logos.
- Opitz, A., Heene, M. & Fischer, F. (2017). Measuring scientific reasoning – a review of test instruments. *Educational Research and Evaluation*, 23 (3-4), 78–101. <https://doi.org/10.1080/13803611.2017.1338586>
- Roberts, R. & Gott, R. (2004). A written test for procedural understanding. A way forward for assessment in the UK science curriculum? *Research in Science & Technological Education*, 22 (1), 5–21. <https://doi.org/10.1080/0263514042000187511>
- Sandoval, W. A. & Millwood, K. A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and Instruction*, 23 (1), 23–55. [https://doi.org/10.1207/s1532690xci2301\\_2](https://doi.org/10.1207/s1532690xci2301_2)
- Toplis, R. (2007). Evaluating Science Investigations at Ages 14–16. Dealing with anomalous results. *International Journal of Science Education*, 29 (2), 127–150. <https://doi.org/10.1080/09500690500498278>
- Van Rens, L., Hermarij, P., Pilot, A., Beishuizen, J., Hofman, H. & Wal, M. (2014). Pre-university Chemistry Students in a Mimicked Scholarly Peer Review. *International Journal of Science Education*, 36 (15), 2514–2533. <https://doi.org/10.1080/09500693.2014.895447>
- Vorholzer, A.; Aufschnaiter, C. von; Kirschner, S. (2016): Entwicklung und Erprobung eines Tests zur Erfassung des Verständnisses experimenteller Denk- und Arbeitsweisen. In: *Zeitschrift für Didaktik der Naturwissenschaften* 22 (1), S. 25–41. <https://doi.org/10.1007/s40573-015-0039-3>
- Vorholzer, A., Aufschnaiter, C. von & Boone, W. J. (2018). Fostering Upper Secondary Students' Ability to Engage in Practices of Scientific Investigation. A Comparative Analysis of an Explicit and an Implicit Instructional Approach. *Research in Science Education*, 103 (1). <https://doi.org/10.1007/s11165-018-9691-1>
- Wellnitz, N. & Mayer, J. (2013). Erkenntnismethoden in der Biologie – Entwicklung und Evaluation eines Kompetenzmodells. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 315–345.