

# Testitems zur qualitativen Untersuchung der Ressourcen von Physiklehrkräften beim Bewerten schriftlicher Schülerleistungen in Physik

Markus Sebastian Feser, Dietmar Höttecke, Timo Ehmke  
markus.sebastian.feser@uni-hamburg.de , dietmar.hoettecke@uni-hamburg.de, timo.ehmke@uni.leuphana.de

## DAS PROJEKT

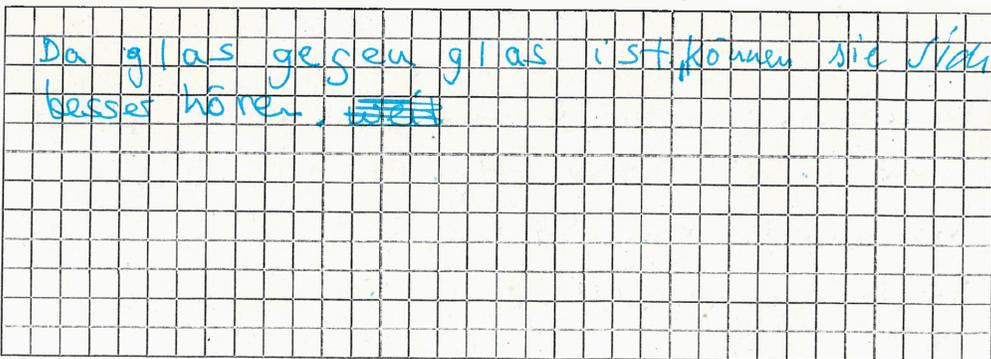
Die Diagnostik schriftlicher Schülerleistungen ist ein wichtiger Teil der täglichen Arbeit von Physiklehrkräften. Wir hegen die begründete Vermutung, dass Physiklehrkräfte dabei ihre Urteile über fachliche und sprachliche Leistungen konfundieren. Wir wissen nicht, ob und ggf. wie sich die Bewertungen über fachliche und sprachliche Schülerleistungen beeinflussen. Zudem liegt keine belastbare empirische Evidenz darüber vor, auf welche Ressourcen Physiklehrkräfte beim Bewerten schriftlicher Schülerleistungen zurückgreifen.



Diese Fragen stellen wir im Projekt "Fachliche und sprachliche Urteilkriterien von Physiklehrkräften". In der Studie wurden zunächst 128 Schüler\_innen Hamburger Gymnasien und Stadtteilschulen gebeten, eine Leistungsaufgabe aus der Akustik zu bearbeiten (siehe unten). Die Schülerlösungen wurden in eine Matrix mit den Modell-Dimensionen fachliche und sprachliche Qualität eingeordnet. Die Einordnung erfolgte mit einem Codiermanual und wurde von insgesamt 8 Rater\_innen abgesichert. Details zu den Entwicklungsschritten der Testitems sind im rechts stehenden Flussdiagramm dargestellt.

## AUFGABE WELTRAUM-SPAZIERGANG

Bei einem Weltraumspaziergang reißt zwischen zwei Astronauten die Funkverbindung ab. Obwohl der eine Astronaut aus Leibeskräften schreit, hört ihn sein Kamerad nicht. Der ältere Astronaut hält seinen in Panik geratenen Kollegen fest und presst seinen Helm an den des Kollegen. Plötzlich kann der jüngere den älteren leise hören. Erkläre beide Phänomene genau!



## LITERATUR

Braaten, M., & Windschitl, M. (2011). Working Toward a Stronger Conceptualization of Scientific Explanation for Science Education. *Science Education*, 95 (4), 639-669.

Biggs, J.B., & Collis, K.F. (1982). *Evaluating the Quality of Learning. The SOLO Taxonomy (Structure of the Observed Learning Outcomes)*. Educational Psychology Series, Academic Press.

Einhaus, E. (2007). *Schülerkompetenzen im Bereich Wärmelehre. Entwicklung eines Testinstruments zur Überprüfung und Weiterentwicklung eines normativen Modells fachbezogener Kompetenzen*. Studien zum Physik- und Chemielemens, Logos Verlag Berlin.

Kang, H., Thompson, J., & Windschitl, M. (2014). Creating Opportunities for Students to Show What They Know: The Role of Scaffolding in Assessment Tasks. *Science Education*, 98 (4), 674-704.

Kendall, M.G. (1948). *Rank Correlation Methods*. Griffin.

Kuckartz, U. (2014). *Mixed Methods. Methodologie, Forschungsdesigns und Analyseverfahren*. Springer VS.

Leisen, J. (2005). Muss ich jetzt auch noch Sprache unterrichten? Sprache im Physikunterricht. *Unterricht Physik*, 3 (87), 4-9.

Lengyel, D., Reich, H.H., & Roth, H.J., Heintze, A., Scheinhardt Stettner, H. (2009). Prozessbegleitende Diagnose zur Schreibentwicklung: Beobachtung schriftlicher Sprachhandlungen in der Sekundarstufe I. In D. Lengyel, H.H. Reich, & H.J. Roth, M. Döll (Eds.), *Von der Sprachdiagnose zur Sprachförderung*. Waxmann.

Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Beltz Verlag.

Tajmel, T. (2011). Wortschatzarbeit im mathematisch-naturwissenschaftlichen Unterricht. *ide*, 35 (1), 83-93.

Thonhauser, J. (2008). Warum (neues) Interesse am Thema „Aufgaben“? In J. Thonhauser (Ed.), *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik* (pp. 13-26). Münster: Waxmann.



## SCHRITTE ZUR ENTWICKLUNG VON TESTITEMS

### Schritt 1:

- Sammeln von Aufgaben aus Klassenarbeiten (Thonhauser 2008)
- Entwicklung und Erprobung von Aufgaben (Braaten & Windschitl 2011, Bigg & Collis 1982, Kang et al. 2014, Leisen 2005, Lengyel et al. 2009, Tajmel 2011)

### Schritt 2:

- Auswahl der Aufgabe „Weltraumspaziergang“ und Erhebung von Schülerlösungen (n=116; 9. Jgst.)
- induktiv/deduktive Kategoriensystementwicklung (Kuckartz 2014, Mayring 2015)  
Ziel: Unterscheidung der fachlichen und sprachlichen Merkmale von Schülerlösungen
- Codierung des Gesamtmaterials
- Systematische Abwandlung einiger Schülerlösungen  
Ziel: Generierung weitere Schülerlösungen, z. B. Lösungen, die auf fachlicher aber nicht auf sprachlicher Ebene gelungen sind

### Schritt 3:

- Prüfung der Intercoderreliabilität mittels Zweitrater\_in (32% des Gesamtmaterials + abgewandelte Schülerlösungen)
- Unterschiedliche Codierungen wurden intensiv diskutiert, um ein gemeinsames Verständnis der Kategorien und deren Anwendung zu erreichen (vorher:  $0.12 \leq \alpha \leq 0.73$ ; nachher:  $0.88 \leq \alpha \leq 0.92$ ).
- Anschließend: Auswahl von 8 kontrastierenden Schülerlösungen als Kandidaten für Testitems

### Schritt 4:

- Ausbildung von 6 Rater\_innen an Codiermanual und Übungsbeispielen
- Anschließend: Gemeinsame Verbesserung des Kategoriensystems mit der Ratergruppe

### Schritt 5:

- Codierung der 8 Schülerlösungen aus Schritt 3 durch alle 6 Rater\_innen mit dem verbesserten Codiermanual
- Anschließend:
  - Bestimmung der Konkordanz der Codierungen der Rater\_innen ( $0.72^{**} \leq W \leq 0.94^{**}$ )
  - Bestimmung der Beurteilungsübereinstimmung bei der Codierung (Einhaus 2007)
  - Schätzung der „wirklichen“ Rangfolge der 8 Schülerlösungen auf fachlicher und sprachlicher Ebene (Kendall 1948)

Ziel: Aus den 8 Schülerlösungen, die „beste“ Komposition von 4 Kontrastfällen für eine Lehrerstudie identifizieren