

## **Experimentelle Fähigkeiten unterstützt durch Schüler selbstbeurteilungen diagnostizieren?**

### **Problemstellung**

Physiklehrkräfte sollen die experimentellen Fähigkeiten ihrer Schülerinnen und Schüler fördern. Dazu werden Schülerexperimente eingesetzt. Voraussetzung für eine individuelle Förderung ist eine Diagnose experimenteller Fähigkeiten. Aber wie soll eine Lehrkraft die von den Schülerinnen und Schülern gezeigten experimentellen Fähigkeiten im Unterrichtsalltag zuverlässig diagnostizieren? Die Rahmenbedingungen für eine zuverlässige Diagnostik sind im Unterrichtsalltag ungeeignet. In der Regel experimentieren ca. 30 Schülerinnen und Schüler in Partner- oder Kleingruppenarbeit und es ist nur eine Lehrkraft im Klassenraum. Vorliegende Tests zur Diagnostik experimenteller Fähigkeiten bieten wenig Hilfestellung für eine formative Individualdiagnostik. Experimentaltests mit Realexperimenten oder Simulationen sind in Durchführung und Analyse sehr aufwändig, eine Diagnostik durch schriftliche Tests insbesondere im Bereich der Durchführung von Experimenten wenig valide (u. a. Schreiber, Theyßen & Schecker, 2014a). Im Unterrichtsalltag besteht die Möglichkeit die Versuchsprotokolle als Informationsquelle zur Diagnostik heranzuziehen (s. a. Prozessprotokolle, Emden & Sumfleth, 2012). Klassische Versuchsprotokolle geben gute Informationen über die Planung und die Auswertung eines Versuchs (Schreiber, Theyßen & Schecker, 2014b, S. 200f). Aus den Versuchsprotokollen wird allerdings nicht deutlich, wie die Schülerinnen und Schüler tatsächlich beim Aufbauen und Messen vorgegangen sind. Beobachtungen der experimentellen Vorgehensweisen der Schülerinnen und Schüler durch die Lehrkraft müssen im Unterrichtsalltag in der Regel unsystematisch bleiben.

### **Selbstbeurteilungen beim Experimentieren**

Um zuverlässiger die Fähigkeiten zur Durchführung eines Experiments diagnostizieren zu können, benötigen Lehrkräfte weitere Informationen. Als zusätzliche Informationsquelle bietet sich im Unterrichtsalltag die Schüler selbstbeurteilung an. Diese ist ein wesentlicher Bestandteil selbstregulierten Lernens.

Das auf den schulischen Kontext zugeschnittene Prozessmodell von Zimmerman (2002) gliedert den Prozess des selbstregulierten Lernens in einen Dreischritt. In der Phase der Vorüberlegungen geht es darum, Ziele zu setzen. Daran schließt die Handlungsphase an, in der Lernende sich selbst kontrollieren und beobachten. Durch einen Abgleich der Selbstbeobachtungen mit den eingangs festgelegten Zielen erfolgt dann in der Phase der Reflexion die Selbstbeurteilung. Mit der Selbstbeurteilung sind unterschiedliche Reaktionen verbunden, die dann wieder die Phase der Vorüberlegungen beeinflussen.

Im Kontext des Experimentierens können die Ziele oder Kriterien zur Beurteilung eines Experimentierprozesses durch eine Lehrkraft vorgegeben oder mit den Schülerinnen und Schülern gemeinsam erarbeitet werden. Diese Kriterien werden beispielsweise in einer Checkliste dokumentiert. Während der Handlungsphase wird das Experiment durchgeführt. In der Reflexionsphase reflektieren die Schülerinnen und Schüler auf Basis ihrer Selbstbeobachtungen die Qualität ihrer Performance beim Experimentieren. Dazu beurteilen sie, wie stark die eigene Performance den festgelegten Zielen genügt. Dabei ist die Selbstbeurteilung aufgabenspezifisch und bezieht sich ausschließlich auf die zuletzt selbst beobachtete Performance. Das unterscheidet diesen Typ der Selbstbeurteilung von anderen, wenn beispielsweise Kompetenzausprägungen mit sog. Kompetenzrastern beurteilt werden.

Abbildung 1 veranschaulicht den Prozess des selbstregulierten Lernens bezogen auf das Experimentieren.

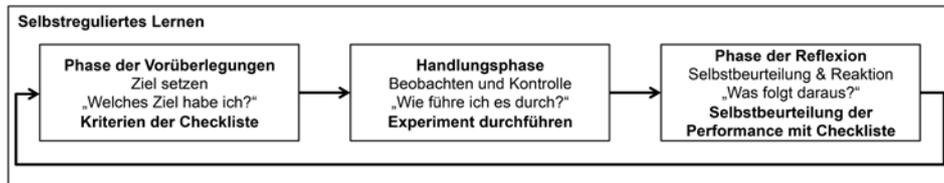


Abb. 1 Selbstreguliertes Lernen nach Zimmerman (2002) und Andrade (2010), hier bezogen auf das Experimentieren lernen (eigene Darstellung)

Außerdem ist davon auszugehen, dass die Fähigkeit zur Selbstbeurteilung von verschiedenen personenbezogenen Variablen beeinflusst wird. Schließlich legen die Komponentenmodelle des selbstregulierten Lernens (Überblick s. Thillmann, 2008, S. 16f) nahe, dass selbstreguliertes Lernen von verschiedenen personenbezogenen Variablen kognitiver, metakognitiver und motivationaler Komponenten beeinflusst wird. Beispielsweise zählen zu der motivationalen Komponente die aktuelle Motivation und die Selbstwirksamkeit, zu der kognitiven Komponente das Fachwissen und zu der metakognitiven Komponente die Erfahrung mit Selbstbeurteilungen.

Um Schüler selbstbeurteilungen zu erheben, bieten sich Checklisten an. Checklisten fordern die Schülerinnen und Schüler dazu auf, ihre eigene Performance beim Experimentieren mit konkreten Bewertungskriterien abzugleichen. Durch die Verwendung von aufgabenspezifischen Checklisten mit Ankerbeispielen wird die Zuverlässigkeit der Beurteilung erhöht (Jonsson & Svingby, 2007).

Zur Genauigkeit der Selbstbeurteilungen im Kontext des Experimentierens liegen bisher nur wenige Befunde vor. Diese deuten auf eine zufriedenstellende Genauigkeit von Selbstbeurteilungen hin (z. B. Stefani, 1994), können jedoch nicht unmittelbar auf die oben dargestellte Problemstellung übertragen werden.

### Ziel und Forschungsfragen

Ziel der hier vorgestellten Studie ist, zu klären, ob Schüler selbstbeurteilungen mit einer Checkliste eine geeignete Informationsquelle zur Diagnostik der Performance beim Experimentieren, insbesondere für das Aufbauen und Messen, darstellen. Dazu ist zu überprüfen, ob die experimentierenden Schülerinnen und Schüler potenziell in der Lage sind, ihre eigene Performance beim Aufbauen und Messen zu beurteilen. Dies führt zur ersten Forschungsfrage:

*F1: Wie hoch ist die Urteilsgenauigkeit der Schülerinnen und Schüler bei der Selbstbeurteilung der eigenen Performance beim Experimentieren?*

Da davon auszugehen ist, dass verschiedene personenbezogenen Variablen die Fähigkeit zur Selbstbeurteilung beeinflussen, sollen in einer zweiten Forschungsfrage Prädiktoren für die Genauigkeit der Schüler selbstbeurteilungen identifiziert werden.

*F2: Welche personenbezogenen Variablen sind Prädiktoren für die Genauigkeit der Schüler selbstbeurteilung?*

### Studie

Die Laborstudie gliedert sich in drei Phasen – Entwicklung, Pilotierung und Hauptstudie. In der Entwicklungsphase sind experimentell lösbare Aufgabenstellungen für Schülerexperimente zu entwickeln, die einerseits typisch für das selbstständige Experimentieren in der Schule sind und andererseits die in den Lehrplänen geforderten experimentellen Fähigkeiten abdecken. Anschließend sind aufgabenspezifische Checklisten

für die Selbstbeurteilung der Performance beim Experimentieren kriteriengeleitet zu entwickeln. Zusätzlich wird eine Begleiterhebung zusammengestellt, die aus etablierten und adaptierten Skalen zu möglichen Prädiktoren der Urteilsgenauigkeit besteht. Schließlich ist als Beurteilungsmaßstab ein Videomaß zu adaptieren, mit dem eine möglichst objektive Beurteilung der Schülerperformance möglich ist. Die Beurteilungskriterien des Videomaßes sind mit den Beurteilungskriterien der Checkliste identisch. An einer Stichprobe von 20 Schülerinnen und Schülern sollen die Denkprozesse bei der Selbstbeurteilung analysiert werden. Dann werden an einer weiteren Stichprobe von 40 Schülerinnen und Schülern Begleiterhebung, Aufgabenstellungen der Schülerexperimente, Checklisten und Videomaß pilotiert. Die Evaluation der Begleiterhebung betrifft die typischen Testgütekriterien. Für die Aufgabenstellungen und aufgabenspezifischen Checklisten ist zu prüfen, ob sie ausreichend Varianz erzeugen. Die beiden Aufgabenstellungen mit der größten Varianz werden für die Hauptstudie ausgewählt. Schließlich ist für das Videomaß die Objektivität der Beurteilungen durch Doppelkodierungen sicherzustellen. Das Ziel der Hauptstudie ist schließlich, die beiden Forschungsfragen zu beantworten. Dazu bearbeiten 120 Schülerinnen und –schüler die Vorerhebung und zwei Schülerexperimente. Die Schülerinnen und Schüler werden bei der Bearbeitung der Schülerexperimente gefilmt und anschließend mithilfe des Videomaßes durch geschulte Rater beurteilt. Nach jedem Schülerexperiment beurteilen die Schülerinnen und Schüler ihre eigene Performance selbst mit einer Checkliste. Die Selbstbeurteilung wird mit der Beurteilung auf Basis des Videomaßes verglichen und damit die Urteilsgenauigkeit für jeden einzelnen Probanden bestimmt (F1). Zur Beantwortung der F2 werden multiple lineare Regressionsanalysen durchgeführt.

### **Erwarteter Ertrag**

Die Studie soll neue Erkenntnisse gewinnen zur Genauigkeit von Schüler selbstbeurteilungen beim selbstständigen Experimentieren. Außerdem sollen Prädiktoren für die Urteilsgenauigkeit der Schülerinnen und Schüler identifiziert werden. Wenn die Ausprägungen der Prädiktoren bei den einzelnen Schülerinnen und Schülern bekannt sind, könnten Lehrkräfte im Unterrichtsalltag Schülerurteile identifizieren, die potenziell genau genug sind, um zur Diagnostik experimenteller Fähigkeiten herangezogen zu werden. Damit würde das Projekt einen Beitrag zur Innovation der Diagnostik im Unterrichtsalltag liefern. Zusätzlich eröffnet es aber auch neue Möglichkeiten zur individuellen Förderung experimenteller Fähigkeiten.

### **Literatur**

- Andrade, H. (2010). Students as the Definitive Source of Formative Assessment: Academic Self-Assessment and the Self-Regulation of Learning. NERA Conference Proceedings 2010, Paper 25.
- Emden, M. & Sumfleth, E. (2012). Prozessorientierte Leistungsbewertung des experimentellen Arbeitens. Zur Eignung einer Protokollmethode zur Bewertung von Experimentierprozessen. *Der mathematische und naturwissenschaftliche Unterricht (MNU)*, 65 (2), 68 – 75.
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review* 2, 130 – 144.
- Schreiber, N., Theyßen, H. & Schecker, H. (2014a). Diagnostik experimenteller Kompetenz: Kann man Realexperimente durch Simulationen ersetzen? *Zeitschrift für Didaktik der Naturwissenschaften* 2014, doi: 10.1007/s40573-014-0017-1
- Schreiber, N., Theyßen, H. & Schecker, H. (2014b). Diagnostik experimenteller Kompetenz: Modell, Testverfahren und Analysemethoden. In A. Fischer, C. Hößle, S. Jahnke-Klein, H. Kiper, M. Komorek, J. Michaelis, V. Niesel & J. Sjuts (Hrsg.): *Diagnostik für lernwirksamen Unterricht*. Hohengehren: Schneider-Verlag, 200 – 214.
- Stefani, L. A. J. (1994). Peer, Self and Tutor Assessment: Relative Reliabilities. *Studies in Higher Education*, 19 (1), 69 – 75.
- Thillmann, H. (2008). *Selbstreguliertes Lernen durch Experimentieren: Von der Erfassung zur Förderung*. Dissertation, Universität Duisburg-Essen.
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 41 (2), 64 – 70.