

Alexander FRÖHLICH, Sebastian KUNTZE, Anke LINDMEIER, München

Testentwicklung und -evaluation im Bereich von „Statistical Literacy“¹

Parallel zum Konzept der Mathematical Literacy (Deutsches PISA-Konsortium, 2004) fokussiert auch „Statistical Literacy“ auf die Handlungsfähigkeit von Lernenden als partizipierende Bürger(innen):

“‘Statistical Literacy’ is the ability to understand and critically evaluate statistical results that permeate our daily lives—coupled with the ability to appreciate the contribution that statistical thinking can make in public and private, professional and personal decisions.” (Wallman, 1993, S. 1).

Auf der Grundlage von Testentwicklungen charakterisieren Watson and Callingham (2003) Statistical Literacy als komplexes hierarchisch gestuftes Konstrukt. Dabei kann die Grundidee „variation“ gleichsam als eine zentrale verbindende Klammer gesehen werden. So gelang es Watson et al. (2003), auf der Basis des Curriculums von Holmes (1980) einen raschskalierbaren Test für das Verständnis von Variabilität zu entwickeln, das einen breiten Bereich von Statistical Literacy abdeckt. Unter der Perspektive der Bildungsstandards der Kultusministerkonferenz (KMK, 2003) erscheinen solche Befunde von besonderem Interesse. So spielen im Rahmen der Leitidee Daten und Zufall für den Teilbereich der Arbeit mit diagrammartigen Darstellungen und der Reduktion von Information (vgl. Kröpfl, Peschek & Schneider, 2000) insbesondere die Kompetenzen des Verwendens von Darstellungen und des Modellierens eine Rolle für Statistical Literacy. Nach Reading (2002) bildet sich in diesem Teilbereich das Spektrum von Kompetenzunterschieden zwischen Lernenden bei Statistical Literacy vollständig ab. An dieser Stelle setzt eine Teiluntersuchung des Projekts KOMMA (Reiss et al., 2005) an. In KOMMA sollen Kompetenzen von Schülerinnen und Schülern in diesem Teilbereich von Statistical Literacy untersucht und entsprechende rechnergestützte Interventionen evaluiert werden.

Für die Untersuchung von Kompetenzen der Lernenden beim Nutzen von Darstellungen und Modellen in statistischen Kontexten erweist sich gerade das Modell von Curcio (1987) als hilfreich. Dieses Modell, das auch Berührungspunkte zu Untersuchungsansätzen von „reading as a social practice“ (vgl. Luke and Freebody, 1997) aufweist, beschreibt für das Arbeiten mit Daten und Darstellungen die drei Kompetenzstufen 1-„reading the data“, 2-

¹ Dieses Forschungsvorhaben wurde vom Bundesministerium für Bildung und Forschung gefördert (Bew.-Nr. PLI3032).

„reading within the data“ und 3-„reading beyond the data“ (Curcio, 1987). Ein wesentliches Kriterium zur Unterscheidung zwischen diesen Kompetenzstufen ist die Anzahl notwendiger Lösungsschritte beim Nutzen von Darstellungen und Modellen (vgl. Lindmeier, Kuntze & Reiss, eingereicht). Eine Kurzbeschreibung der aus den vorgestellten Ansätzen abgeleiteten Kompetenzstufen für die Kompetenz „Nutzen von Darstellungen und Modellen in statistischen Kontexten“ gibt Tab. 1.

Tab. 1: Nutzen von Darstellungen und Modellen in statistischen Kontexten - Kompetenzstufen

Kompetenzstufe I	Einschrittiges Nutzen einer Darstellung oder Arbeiten innerhalb eines gegebenen Modells (z.B. Ablesen eines gegebenen Werts aus einem Diagramm, Vervollständigen eines gegebenen Diagramms bei gegebener Datentabelle)
Kompetenzstufe II	Zwei- oder mehrschrittiges Nutzen von Darstellungen oder Wechsel zwischen zwei gegebenen Modellen (z.B. Vergleichen von Daten unter Einschluss eines transformierenden Schrittes oder von Begriffswissen)
Kompetenzstufe III	Mehrschrittiges Nutzen von Darstellungen einschließlich der Nutzung eines nicht gegebenen Modells (z.B. eigene Modellierungsaktivitäten zur Unterstützung einer kumulativen Interpretation von in Diagrammen gegebenen Daten)

Zur empirischen Evaluation des entwickelten Tests und zur Untersuchung von Kompetenzen von Schülerinnen und Schülern interessieren also die folgenden Forschungsfragen: *Ist es möglich, den Test mit dem eindimensionalen Rasch-Modell zu beschreiben und werden die theoretisch abgeleiteten Kompetenzstufen empirisch bestätigt? Welche Kompetenzen können bei den untersuchten Schülerinnen und Schülern beobachtet werden?*

Untersuchungsdesign

Zum oben beschriebenen Kompetenzstufenmodell wurde ein Test entwickelt, dessen 10 Items wie in Tab. 2 dargestellt den drei Kompetenzstufen zugeordnet werden konnten. Drei dieser Items stammten aus TIMSS und dienten als Anker-Items für die jeweiligen Kompetenzstufen. Ein Beispiel-Item ist in Lindmeier, Kuntze und Reiss (eingereicht) abgedruckt.

Tab. 2: Überblick über Items und Kompetenzstufen

Kompetenzstufe	I	II	III
Items	A1, A2	A3, A4, A5	A6,A7
Anker-Items aus TIMSS	T1	T2	T3

Der Test wurde in zwei Gymnasien eingesetzt. Insgesamt nahmen N=187 Lernende (davon 99 Schülerinnen) teil, wovon 112 aus der 5. Jahrgangsstufe (davon 52 Schülerinnen) und 75 aus der 8. Jahrgangsstufe stammten (davon 47 Schülerinnen). Die Antworten wurden dichotom codiert. Daraus ergibt sich ein Gesamtscore von maximal 10 Punkten.

Ergebnisse

Die Modellannahme der Rasch-Skalierbarkeit konnte mithilfe eines parametrischen Bootstraps mit 300 Iterationen und einer Genauigkeit von

0,00075 belegt werden. Die Item- und Personenparameter auf der Logit-Skala sind in Abb. 1 dargestellt. Die angenommenen Kompetenzstufen konnten so weitgehend bestätigt werden. Einzig das TIMSS-Item T3 erwies sich als leichter als die übrigen Items auf Kompetenzstufe III und lag mit einem Parameter von 0,99057 nahe bei dem schwierigsten Item aus Kompetenzstufe II.

Insgesamt schnitten die Lernenden sehr gut ab. Die Verteilung des Gesamtscores zeigt eine deutliche Verschiebung nach rechts (Abb. 2). Mit einer Kategorisierung der Schülerleistungen analog zu den Kompetenzstufen erreichen 72,7 Prozent einen Score von 8 Punkten oder mehr, was Kompetenzstufe III entspricht. Dabei zeigen sich Unterschiede in Bezug auf die untersuchten Klassenstufen (Lernende auf Kompetenzstufe III in Klasse 5: 58,9%, in Klasse 8: 93,3%).

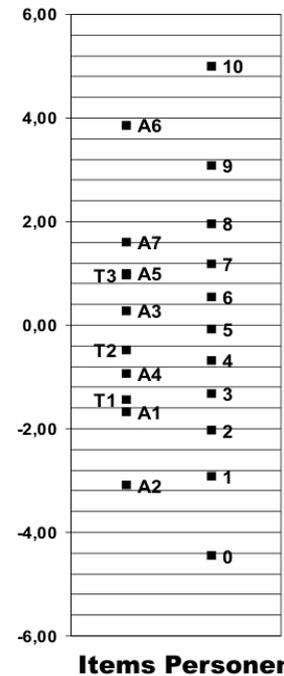


Abb. 1: Item- und Personenparameter

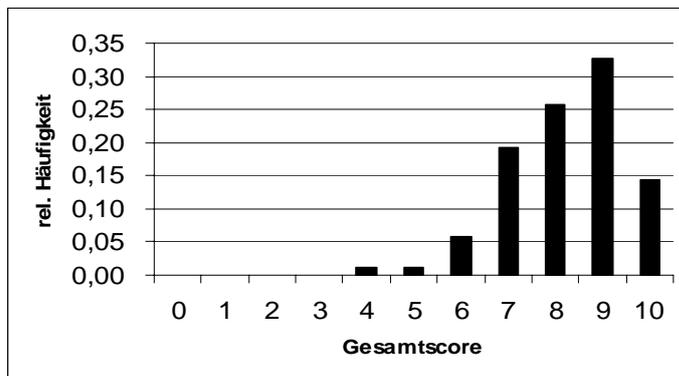


Abb. 2: Verteilung des Gesamtscore

Gesamt-score	5. Klasse	8. Klasse
4	2	0
5	2	0
6	10	1
7	32	4
8	31	17
9	27	34
10	8	19

Diskussion

Die Untersuchung bestätigt das angenommene Kompetenzmodell zum Nutzen von Darstellungen und Modellen in statistischen Kontexten. Sowohl die Eindimensionalität als auch das Kompetenzstufenmodell konnten somit empirisch bestätigt werden. Allerdings sollte bei einer Weiterentwicklung des Tests Deckeneffekten insbesondere für die gymnasiale 8.

Jahrgangsstufe vorgebeugt werden. Dabei könnte der Aspekt der statistischen Variabilität (vgl. Watson et al., 2003) stärker betont werden.

Literatur

- Curcio, F.R. (1987). Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education*, 18(5), 382-393.
- Deutsches PISA-Konsortium (Hrsg.). (2004). PISA 2003. Münster: Waxmann.
- Holmes, P. (1980). Teaching statistics 11-16. Berkshire: Schools Council and Foulsham Educ.
- KMK (Kultusministerkonferenz). (2004). Bildungsstandards im Fach Mathematik für den mittleren Schulabschluss, München: Wolters Kluwer.
- Kröpfl, B., Peschek, W., & Schneider, E. (2000). Stochastik in der Schule: Globale Ideen, lokale Bedeutungen, zentrale Tätigkeiten. *mathematica didactica*, 23(2), 25-57.
- Lindmeier, A., Kuntze, S. & Reiss, K. (eingereicht). Representations of data and manipulations through reduction – competencies of German secondary students. IASE.
- Luke, A., & Freebody, P. (1997). Shaping the social practices of reading. In S. Muspratt, A. Luke & P. Freebody (Hrsg.), *Constructing Critical Literacies* (S. 185-225). Cresskill, NJ: Hampton.
- Reading, C. (2002), Profile for statistical understanding, In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Retrieved January 20, 2007, from http://www.stat.auckland.ac.nz/~iase/publications/1/1a4_read.pdf
- Reiss, K., Dorfmeister, J., Pekrun, R., Kuntze, S., Lindmeier, A., Appel, T., Rupp, F. & Schalk, M. (2005). Kompendium Mathematik (KOMMA) - Implementation und Evaluation eines computergestützten Instruments zum Mathematikunterricht in den Klassen 8 und 9. [Posterpräsentation im Rahmen des Forschungssymposium des Lehrerbildungszentrums der LMU München am 19.11.2005].
- Wallman, K. (1993). Enhancing Statistical Literacy: Enriching our Society, *Journal of the American Statistical Association*, 88(421), 1-8.
- Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2(2), 3-46.
- Watson, J.M., Kelly, B.A., Callingham, R.A., & Shaughnessy, J.M. (2003). The Measurement of School Students' Understanding of Statistical Variation. *International Journal of Mathematics Education in Science and Technology*, 34(1), 1-29.