Vanessa Pupkowski[1]
Elke Sumfleth[1]
Maik Walpuski[1]

[1]Universität Duisburg-Essen

Kompetenzmessung in der Chemie und der Einfluss affektiver Faktoren

Theoretischer Hintergrund

Die fachspezifische Erfassung der Kompetenzen von Schülerinnen und Schülern in großen Schulleistungsstudien, wie z. B. bei der Überprüfung der Bildungsstandards, fokussiert in der Regel auf den kognitiven Teil der Weinert'schen Kompetenzdefinition (Klieme & Leutner, 2006). Der affektive Teil der Kompetenzdefinition bleibt weitestgehend unberücksichtigt. Wise und DeMars (2005) konnten jedoch zeigen, dass bei der Messung von Kompetenzen die aktuelle und gemessene Kompetenz nur dann hoch miteinander korrelieren, wenn die Motivation hoch ist. Zudem konnte in weiteren Studien gezeigt werden, dass sich das Interesse positiv auf die Leistung in Testsituationen ausübt (Prenzel, Schütte & Walter, 2007; Liu Bridgemann & Adler, 2012). In dieser Studie wird daher der aufgabenspezifische Zusammenhang zwischen kognitiven und affektiven Faktoren untersucht.

Nach Krapp und Prenzel (1992) wird das Interesse definiert als Personen-Gegenstandbeziehung und lässt sich weiter differenzieren in situationales Interesse, das hervorgerufen wird durch einen spezifischen Gegenstand in einer bestimmten Situation, und individuelles Interesse als stabiles Persönlichkeitsmerkmal. In Testsituationen kann der spezifische Gegenstand eine Aufgabe darstellen. Daher lässt sich das Interesse sowohl auf Personenebene (situationales Interesse), als auch auf der Aufgabenebene (situationale Interessantheit der Aufgabe) darstellen. Für das individuelle Interesse kann nur ein Einfluss des Personenmerkmals auf die Bearbeitung der Aufgaben angenommen werden, während umgekehrt durch die kurzzeitige Auseinandersetzung mit einer Testaufgabe keine Beeinflussung des individuellen Interesses zu erwarten ist.

Nach Krapp (1998) liegt eine Auseinandersetzung mit einem Gegenstand nur dann vor, wenn diesem Gegenstand eine hinreichend bedeutsame Wertschätzung zugeordnet wird. Dies wird im Folgenden als Relevanz operationalisiert. Auch für diesen Faktor lässt sich eine Unterteilung in ein Aufgabenmerkmal (Relevanz der Aufgabe) und ein Personenmerkmal (persönliche Relevanz) operationalisieren.

Ein weiterer affektiver Faktor in diesem Zusammenhang stellt die Leistungsmotivation nach Wigfield & Eccles (2002) dar. Dabei handelt es sich um eine spezifische Form der Motivation, die im Zusammenhang mit der Messung von Kompetenzen in Testaufgaben relevant ist. Das entsprechende Erwartungs-Wert-Modell operationalisiert die Testmotivation und wird beschrieben durch die Aussicht auf Erfolg (Erwartung), die Wichtigkeit und dem zugeschrieben Nutzen (Wert). Auch hier können die Faktoren sowohl auf Aufgabenebene bestimmt werden (Erwartete Lösungswahrscheinlichkeit der Aufgabe und Wert der Aufgabe), sowie auf Personenebene (Erwartung und Wert). Die Erwartungskomponente der Leistungsmotivation ist eng verknüpft mit dem Fähigkeitsselbstkonzept, das wiederum ein Personenmerkmal darstellt.

Die Grundlage für die Messung von Chemiekompetenzen bilden die in den NBS beschriebenen Kompetenzbereiche. Bisherige Untersuchungen konnten zeigen, dass in der Biologie das Lösen von Aufgaben im Kompetenzbereich Bewertung höheres Interesse erzeugt als im Kompetenzbereich Fachwissen (Holstermann & Bögeholz, 2007). An dieser Stelle soll nun untersucht werden, welchen aufgabenspezifischen Einfluss affektive Faktoren auf die Aufgabenschwierigkeit sowie die Personenfähigkeit haben. Zudem wird angenommen, dass dieser Einfluss im Kompetenzbereich Bewertung höher ist als im Kompetenzbereich Fachwissen.

Methode und Design

Zur Untersuchung des Zusammenhangs zwischen kognitiven und affektiven Faktoren wurde ein Testinstrument im Fach Chemie entwickelt, das die Leistung (kognitiv) und den aufgabenbezogenen Einfluss von Motivation und Interesse (affektiv) erhebt.

Für die Leistungsmessung wurden Items mit identischen kognitiven Anforderungen (nach Walpuski et al., 2010) zu den Kompetenzbereichen Fachwissen und Bewertung entwickelt. Dieses Testinstrument wurde in einer Pilotstudie eingesetzt und ausgewertet. 128 Items dieses Testinstruments (64 Items zum Fachwissen und 64 Items zur Bewertungskompetenz im offenen und Multiple-Choice-Single-Select Format) wurden anschließend in einer Hauptstudie zusammen mit einem Fragebogen eingesetzt. Die Items des Fragebogens wurden in einem Embedded-Design im Anschluss an jede Aufgabe vorgelegt. Dabei wurden Items zum situationalen Interesse (3 Items) und der Relevanz (3 Items) verwendet, um den aufgabenspezifischen Einfluss des Interesses zu erheben. Zusätzlich wurden Items eingesetzt, die die aufgabenspezifische Motivation erheben (4 Items Wert, 3 Items Erwartung; Boekarts, 2002; Sundre, 2007). Auch die Anstrengungsbereitschaft (5 Items, Sundre, 2007) wurde aufgabenbezogen mit erhoben. Aufgabenunabhängig wurde das individuelle Fachinteresse und das Fähigkeitsselbstkonzept mit je 5 Items erhoben (Kölbach, 2011).

Die Testhefte wurden im Multi-Matrix-Design eingesetzt und über Brücken-Items verlinkt. Jedes Testheft enthielt 2 Aufgaben, eine Aufgabe zum Kompetenzbereich Fachwissen und eine Aufgabe zum Kompetenzbereich Bewertung (je 4 Items) und den Fragebogen im embedded Design. Die Datenerhebung wurde an Schulen in NRW zum Ende des mittleren Schulabschlusses durchgeführt (N = 1899, 51.7 % ♂, M(Alter) = 15.18, SD = 0.87).

Die Auswertung der Daten erfolgte mit Hilfe von IRT Analysen, sowie Korrelationsanalysen und Regressionsmodellen.

Ergebnisse

Für die Analyse wurden nur Datensätze verwendet, bei denen Aufgaben zu beiden Kompetenzbereichen und die dazugehörigen Fragebögen bearbeitet wurden. Daher verringert sich die Stichprobengröße auf N = 1543. Für Aussagen auf Aufgabenebene wurden personenzentrierte Rasch-Analysen durchgeführt, für Aussagen auf Personenebene itemzentrierte Analysen.

Analyse des Kompetenztests

Zunächst wurde eine Dimensionsanalyse durchgeführt. Dabei stellte sich heraus, dass es sich bei dem eingesetzten Kompetenztest um 2 Skalen handelt, die die beiden Kompetenzbereiche Fachwissen und Bewertung abbilden. Eine Überprüfung mit Hilfe des Chi-Quadrat-Tests belegt dies (Δ Deviance = 5635.584, $p \leq .001$). Anschließend wurden die Items der Skalen hinsichtlich ihrer Kennwerte untersucht und es ließ sich zeigen, dass 126 Items gute Item-fit Werte aufzeigen (Fachwissen: N = 63, $0.88 \leq wMNSQ \leq 1.11$, $.21 \leq r \leq .80$; Bewertung: N = 63, $0.85 \leq wMNSQ \leq 1.17$, $.44 \leq r \leq .80$). Die Itemreliabilität liegt in einem sehr guten Bereich (.958), so dass die Itemschwierigkeiten sehr gut geschätzt werden können. Die Personenreliabilitäten ($EAP/PV_{Fachwissen}$ .538; $EAP/PV_{Bewertung}$ .505) liegen in einem eher geringen Bereich. Dies ist darauf zurück zu führen, dass die Schülerinnen und Schüler pro Skala nur 4 Items beantwortet haben.

Ein Vergleich der Itemschwierigkeiten zwischen den Kompetenzbereichen zeigt, dass die Items zum Kompetenzbereich Bewertung signifikant leichter sind als zum Kompetenzbereich Fachwissen ($t(124) = -7.426$, $p \leq .001$, $d = 1.323$).

Analyse des Fragebogens

In einem ersten Schritt wurde eine Dimensionsanalyse durchgeführt, um zu überprüfen, ob die theoretisch angenommenen Skalen Interesse und Motivation empirisch voneinander trennbar sind. Es konnte mit Hilfe eines Chi-Quadrat-Tests gezeigt werden, dass die Daten

ein 4-dimensionales Modell abbilden: Situationales Interesse und Relevanz als Faktoren zum Interessenskonstrukt und Erwartung und Wert als Faktoren des Motivationskonstrukts ($p \leq .001$). Die Itemreliabilität (.965) und auch die Personenreliabilitäten ($.810 \leq EAP/PV \leq .909$) liegen in einem sehr guten Bereich, so dass Itemschwierigkeiten und Personenfähigkeiten sehr gut geschätzt werden können. Die Skala zum Faktor Anstrengungsbereitschaft wurde in einem Eindimensionalen Modell berechnet und zeigt auch zufriedenstellende Kennwerte auf (Itemreliabilität .969, EAP/PV .760). Anschließend wurden die Korrelationen der einzelnen Skalen untereinander bestimmt (siehe Tabelle 1).

	SI	R	E	W
R	$.976^{***}$			
E	$.661^{***}$	$.629^{***}$		
W	$.847^{***}$	$.827^{***}$	$.773^{***}$	
A	.109	.053	.245	.252

Tab. 1 Korrelationsanalysen der affektiven Faktoren (SI = Situationales Interesse, R = Relevanz, E = Erwartung, W = Wert, *** $p \leq .001$)

Analyse des Einflusses affektiver Faktoren auf die Aufgabenschwierigkeit

Um den Einfluss der affektiven Faktoren auf die Aufgabenschwierigkeiten zu bestimmen, wurden zunächst Korrelationsanalysen zwischen den Aufgabenmerkmalen berechnet. Dabei zeigte sich, dass die Aufgabenschwierigkeiten für den gesamten Kompetenztest mit den Faktoren Situationale Interessantheit ($r = -.411^{*}$), Relevanz der Aufgabe ($r = -.355^{*}$), Erwartete Lösungswahrscheinlichkeit ($r = -.874^{***}$) und Wert der Aufgabe ($r = -.496^{**}$) korrelieren. Die negativen Korrelationen zeigen, dass mit steigender Aufgabenschwierigkeit das Interesse und die Motivation sinken. Betrachtet man die Aufgabenschwierigkeiten getrennt nach Kompetenzbereich, so stellt sich heraus, dass nur noch die erwartete Lösungswahrscheinlichkeit mit der Aufgabenschwierigkeit korreliert (Fachwissen $r = -.571^{*}$; Bewertung $r = -.775^{***}$). Dabei zeigt sich zusätzlich, dass der Zusammenhang zwischen Aufgabenschwierigkeit und erwarteter Lösungswahrscheinlichkeit im Kompetenzbereich Bewertung höher ist als im Kompetenzbereich Fachwissen.

Durch Berechnung von Regressionsmodellen wurde die Varianz der Aufgabenschwierigkeit bestimmt. Dabei zeigt sich, dass sich die Aufgabenschwierigkeit für den gesamten Kompetenztest zu 85 % durch die Faktoren erwartete Lösungswahrscheinlichkeit ($\beta = -1.219^{***}$) und Wert der Aufgabe ($\beta = 0.447^{***}$) aufklären lässt. Getrennt nach Kompetenzbereichen zeigt sich, dass sich die Aufgabenschwierigkeit im Kompetenzbereich Fachwissen zu 33 % durch den Faktor erwartete Lösungswahrscheinlichkeit ($\beta = -0.571^{*}$) aufklären lässt, im Kompetenzbereich Bewertung zu 80 % durch die Faktoren erwartete Lösungswahrscheinlichkeit ($\beta = -1.316^{***}$) und Wert der Aufgabe ($\beta = 0.702^{**}$). Bei Betrachtung der hohen Varianzaufklärung ist zu beachten, dass bei der Konstruktion der Items gleiche kognitive Anforderungen nach dem ESNaS-Modell für alle Items gewählt wurden.

Analyse des Einflusses affektiver Faktoren auf die Personenfähigkeit

Für die Analyse wurde ein Regressionsmodell berechnet, bei dem die Personenmerkmale als Prädiktoren eingesetzt wurden und die Personenfähigkeit als abhängige Variable. Es lässt sich hier zeigen, dass die Faktoren situationales Interesse ($\beta = 0.057^{*}$), Erwartung ($\beta = 0.266^{***}$) und Anstrengungsbereitschaft ($\beta = 0.208^{***}$) die Varianz der Personenfähigkeit zu 20 % aufklären können. Das individuelle Interesse und das Fähigkeitsselbstkonzept zeigen keine weitere Aufklärung der Personenfähigkeit auf.

Literatur

Boekaerts, M. (2002). The On-Line Motivation Questionnaire: A self-report instrument to assess students' context sensitivity. *New Directions in Measures and Methods, 12*, 77-120.

Holstermann, N., & Bögeholz, S. (2007). Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften, 13*, 71-86.

Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik, 52*, 876-903.

Krapp, A. (1998). Entwicklung und Förderung von Interesse im Unterricht. *Psychologie in Erziehung und Unterricht*, 45, 186-203.

Krapp, A., & Prenzel, M. (Hrsg.). (1992). *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung*. Münster: Aschendorff.

Kölbach, E. (2011). *Kontexteinflüsse beim Lernen mit Lösungsbeispielen*. Berlin. Logos.

Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring Learning Outcomes in Higher Education: Motivation Matters. *Educational Researcher, 41*(9), 352-362.

Prenzel, M., Schütte, K., & Walter, O. (2007). Interesse an den Naturwissenschaften. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 107-124). Münster: Waxmann.

Sundre, D. L. (2007). *The Student Opinion Scale (SOS): A measure of examinee motivation. Test Manual*. Harrisonbourg (VA): The Center for Assessment & Research Studies.

Walpuski, M., Kauertz, A., Kampa, N., Fischer, H.E., Mayer, J., Sumfleth, E. & Wellnitz, N. (2010). ESNaS – Evaluation der Standards für die Naturwissenschaften in der Sekundarstufe I. In A. Gehrmann, U. Hericks & M. Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle – Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 171-184). Bad Heilbrunn: Klinkhardt.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In Weinert, F. E. (Hrsg.), *Leistungsmessungen in Schulen* (S. 17-31). Weinheim und Basel.

Wigfield, A. & Eccles, J.S. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, (53), 109-132.

Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment. Problems and potential solutions. *Educational Assessment, 10* (1), 1-17.