

Pitt Hild
 Josiane Tardent
 Christoph Gut
 Susanne Metzger

Pädagogische Hochschule Zürich

Projekt ExKoNawi: Typenspezifische Kompetenzprogressionen bei hands-on Testaufgaben

Das im Rahmen des Projektes ExKoNawi (Experimentelle Kompetenzen in den Naturwissenschaften) der Pädagogischen Hochschule Zürich entstandene Kompetenzentwicklungsmodell versucht, Kompetenzprogressionen von 12- bis 15-jährigen Jugendlichen beim Lösen von experimentellen, fächerübergreifenden *Problemtypen* zu beschreiben. Kompetenz wird hier als die auf transferfähigem Wissen basierte Problemlösefähigkeit angesehen (Gott & Duggan, 2002). Im Folgenden werden die Kodierung der mitentwickelten *hands-on* Testaufgaben sowie die Ergebnisse aus der zweiten Pilotierung vorgestellt. Zu jedem Problemtyp wurden jeweils drei Testaufgaben entwickelt und untersucht. Die Vorvalidierung belegt, dass auf der Basis des entwickelten Strukturmodells und der typenspezifischen Progressionsmodellierung Tests mit guten Fitwerten entwickelt werden können, die signifikante Leistungsdifferenzen messen.

Motivation

Das Projekt ExKoNawi ist ein Folgeprojekt des interdisziplinären large-scale Experimentiertests HarmoS und versucht die Validitätsprobleme (lokale Itemabhängigkeiten, kompetenzirrelevante Anforderungen, Generalisierbarkeit und Interpretierbarkeit der Experimentieraufgaben) des Letzteren (Gut, 2012) in zufriedenstellender Weise zu lösen. Längerfristig soll das neu entwickelte Kompetenzentwicklungsmodell auch als Richtmodell für die Diagnose und individuelle Förderung experimenteller Kompetenzen im schweizer Naturwissenschaftsunterricht der Sekundarstufe I dienen (Hild et al., 2014).

Problemlöseansatz

Aufgrund des integrierten naturwissenschaftlichen Unterrichts in der Schweizer Volksschule wird für die Messung experimenteller Kompetenzen ein Modell benötigt, das biologische, chemische und physikalische Aspekte des Experimentierens berücksichtigt. Im Zusammenhang mit Experimentiertests werden zurzeit drei unterschiedliche Ansätze diskutiert: 1) Im *Teilprozessansatz* wird das Experimentieren als Zusammenspiel unterscheidbarer, idealisierter Teilprozesse wie Fragestellung, Hypothesenformulierung, Planung und Durchführung des Experiments, Auswertung der Daten und Reflexion verstanden (vgl. Emden, 2012; Nawrath et al. 2011). 2) Der *Arbeitsweisenansatz* der Biologiedidaktik beruht auf der hypothetisch-deduktiven Erkenntnislogik unterschiedlicher Untersuchungsmethoden wie z. B. Beobachten, Vergleichen, Ordnen oder Experimentieren (vgl. Mayer, 2007; Wellnitz, 2012). Auch hier lassen sich, über alle Untersuchungsmethoden hinweg, gemeinsame Teilprozesse erfassen. 3) Im Gegensatz zu den ersten beiden wird beim *Problemlöseansatz* (vgl. Shavelson & Ruiz-Primo, 1998) das Experimentieren je nach Problemtyp unterschiedlich modelliert. Da die Leistungen der Schülerinnen und Schüler bei *hands-on* assessments stark von der experimentellen Aufgabenstellung abhängen (Shavelson et al., 1993), erhofft man sich von diesem Ansatz, das Problem der lokalen Itemabhängigkeiten und der fehlenden Generalisierbarkeit besser in den Griff zu bekommen.

Zweiter Pilottest

Zum Kompetenzentwicklungsmodell von ExKoNawi wurden bis dato folgende vier Problemtypen definiert und analysiert: *kategoriengeleitetes Beobachten*, *skalenbasiertes Messen*, *fragengeleitetes Untersuchen* und *effektbasiertes Vergleichen* (Gut, Hild et al., 2014). Bei jedem Problemtyp wird zwischen mehreren Qualitätsstandards unterschieden die *a priori* aus Theorie und Erfahrung festgelegt (Gut, Metzger et al., 2014) und *a posteriori* mittels Rasch-Skalierung (partial credit model) und einem Vergleich zwischen einer bedingten und unbedingten Kodierung (Metzger et al., 2014) überprüft wurden.

Für die beiden Problemtypen *effektbasiertes Vergleichen* und *fragengeleitetes Untersuchen* (Pilottest 2) konnten jeweils vier Qualitätsstandards (QS) kodiert und bei einer Stichprobe von 330 Schülerinnen und Schüler des Sekundarstufe I ermittelt werden (siehe Tab. 1 und Tab. 2).

Tab.1: Struktur- und Progressionsmodellierung vom Problemtyp *Effektbasiertes Vergleichen*

<i>effektbasiertes Vergleichen</i>	Objekte anhand einer gegebenen Eigenschaft experimentell (ohne direkte Messung) vergleichen	QS 1: Wird ein adäquater qualitativer Vergleich zweier Objekte beschrieben bzw. durchgeführt? (5 Kriterien)
		QS 2: Werden Daten protokolliert und daraus korrekte Schlussfolgerungen (Reihenfolge der drei Objekte) gezogen? (3 Kriterien)
		QS 3: Werden Bedingungen für einen fairen Vergleich genannt? (1 Kriterium)
		QS 4: Wird ein adäquater quantitativer Vergleich durchgeführt? (4 Kriterien)

Tab.2: Struktur- und Progressionsmodellierung vom Problemtyp *Fragengeleitetes Untersuchen*

<i>fragen-geleitetes Untersuchen</i>	korrelative Zusammenhänge zwischen gegebenen Variablen untersuchen	QS 1: Wird ein Zusammenhang (ohne Variablenkontrolle) adäquat untersucht? (1 Kriterium)
		QS 2: Werden Daten ausgewertet und daraus korrekte Schlussfolgerungen gezogen? (2 Kriterien)
		QS 3: Wird der zweite Zusammenhang adäquat untersucht und aus Daten eine korrekte Schlussfolgerung gezogen? (3 Kriterien)
		QS 4: Wird der Variablenkontrollansatz verstanden und angewendet? (2 Kriterien)

Ergebnisse

Qualitätsstandards und typenspezifische Kompetenzprogressionen: Zu *effektbasiertem Vergleichen* sowie zu *fragengeleitetem Untersuchen* konnten bis auf eine Ausnahme (7. Sek A bei *fragengeleitetem Untersuchen*) konsistente Progressionen der Qualitätsstandards erfasst werden (siehe Abb.1 und Abb.2). Die Rater-Übereinstimmung von 63 kodierten Kriterien der Antwortqualitäten ist grösser als 0.78 (bis auf 14 Kriterien ist Cohens Kappa > 0.61).

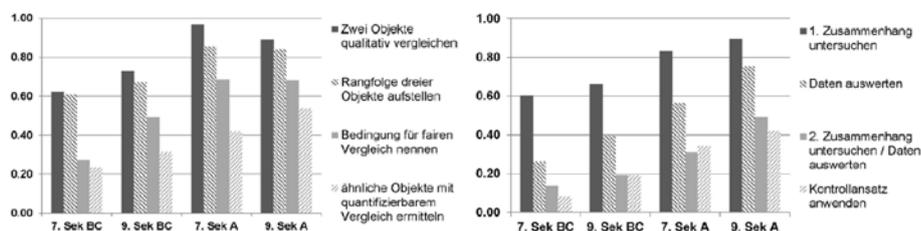


Abb.1: Häufigkeiten der erreichten QS summiert über alle drei Aufgaben zum Problemtyp *effektbasiertes Vergleichen*

Abb.2: Häufigkeiten der erreichten QS summiert über alle drei Aufgaben zum Problemtyp *fragengeleitetes Untersuchen*

Aufgabenvalidität: Die eindimensionale Rasch-Skalierung für die 6 Experimentier-tests (3 Tests zu jedem Problemtyp) ergab sehr gute Itemfit-Werte: Item-Separations-Reliabilität und Trennschärfe zeigen eine gute Zuverlässigkeit der errechneten Itemparameter (siehe Abb. 3).

EAP/PV-Reliabilität	0.645
Varianz	0.464
Item-Sep-Reliabilität	0.983
Infit	$0.93 - 1.09, T \leq 0.9$
Outfit	$0.88 - 1.14, T \leq 1.1$
Trennschärfe	0.69 - 0.80

Abb.3: 1-dimensionale Rasch-Skalierung des Pilottests 2

Leistungen der Schülerinnen und Schüler: Werden die Ergebnisse des ExKoNawi-Tests in PISA-Metrics (siehe Abb. 4) umgerechnet, so ergeben sich die in der Abbildung dargestellten Mittelwerte: der Test misst signifikante Unterschiede zwischen den Schulniveaus (A und B/C) und Jahrgangsstufen, vor allem im tiefen Leistungsbereich.

	Sek B/C	Δ	Sek A	alle
Jahrgangsstufe 7	395 (79)	117 *** †	512 (90)	
Δ	39 * †		25 *	
Jahrgangsstufe 9	434 (101)	103 *** ($d = 1.0$)	537 (78)	500 (100) [†]

Abb.4: Leistungen der SuS in PISA-Metrics (Pilottest 2)

Ausblick

Um die Generalisierbarkeit der Performanz über alle Problemtypen hinweg zu steigern, soll der Test auf 6 Items pro Problemtyp verdoppelt werden. Geplant ist eine Hauptstudie mit 900 12- bis 15-jährigen Jugendlichen.

Literatur

- Emden, M. (2012). Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I. Berlin: Logos Verlag.
- Gott, R., & Duggan, S. (2002). Problems with the assessment of performance in practical science: which way now. *Cambridge Journal of Education*, 32 (2), 183-201.
- Gut, C., Metzger, S., Hild, P., & Tardent, J. (2014). Problemtypenbasierte Modellierung und Messung experimenteller Kompetenzen von 12- bis 15-jährigen Jugendlichen. *PhyDid, Beiträge zur DPG Frühjahrstagung*.
- Gut, C., Hild, P., Metzger, S., & Tardent, J. (2014). Projekt ExKoNawi: Modell für hands-on Assessments experimenteller Kompetenzen. In S. Bernholt (Ed.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht*. GDCP, Jahrestagung in München 2013 (S.171-173). Kiel: IPN.
- Gut, C. (2012). Modellierung und Messung experimenteller Kompetenz. Analyse eines large-scale Experimentier-tests. Berlin: Logos Verlag.
- Hild, P., Metzger, S., & Parchmann, I. (2014). Individuelle Förderung experimenteller Kompetenzen mit Lernaufgaben. In S. Bernholt (Ed.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht*. GDCP, Jahrestagung in München 2013. Kiel: IPN.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen, In D. Krüger & H. Vogt (Eds.), *Theorien in der biogiedidaktischen Forschung*, S. 177-186. Berlin: Springer
- Metzger, S., Hild, P., Gut, C., & Tardent, J. (2014). Projekt ExKoNawi: Aufgaben und erste Ergebnisse der hands-on Assessments. In S. Bernholt (Ed.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht*. GDCP, Jahrestagung in München 2013 (S.174-176). Kiel: IPN.
- Nawrath, D., Maiseyenko, V., & Schecker, H. (2011). Experimentelle Kompetenz. Ein Modell für die Unterrichtspraxis. *Physik in der Schule*, 60 (6), 42-48
- Shavelson, R. J., & Ruiz-Primo, M. A. (1998). On the assessment of science achievement CSE Technical Report (Vol. 491). Los Angeles: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing.
- Shavelson, R. J., Baxter, G. P., & Gao, X. (1993). Sampling variability of performance assessments, *Journal of Educational Measurement*, 30, 3, 215-232.
- Wellnitz, N. (2012). Kompetenzstruktur und -niveaus von Methoden naturwissenschaftlicher Erkenntnisgewinnung. Berlin: Logos Verlag.