

Bodo Eickhorst¹
 Martin Dickmann²
 Horst Schecker¹
 Heike Theyssen²
 Knut Neumann³

¹Universität Bremen
²Universität Duisburg-Essen
³IPN Kiel

Messung experimenteller Kompetenz im Large-Scale: Bewertung experimenteller Aufgaben

Einleitung

Im Projekt „Messung experimenteller Kompetenz in Large Scale Assessments“ (MeK-LSA)¹ wurde ein computerbasiertes Testverfahren zur Erfassung experimenteller Kompetenz entwickelt (Theyßen et al., 2013). Die Testentwicklung wurde umfangreich durch empirische Studien zu Fragen der Validität abgesichert (z. B. Dickmann et al., 2013; Dickmann et al. 2014). Im Sommer 2014 wurde der Test im Large-Scale mit 1.200 Schülerinnen und Schülern der 9. Jahrgangsstufe durchgeführt. Mit den Daten soll zum einen die Testreliabilität geprüft und zum anderen die Struktur experimenteller Kompetenz empirisch analysiert werden (Konstruktvalidierung). In diesem Beitrag liegt der Schwerpunkt auf der konstruktbezogenen Auswertung der vorliegenden Testbearbeitungen.

Experimentelle Kompetenz

Als Bezugspunkt für eine Konstruktvalidierung verwenden wir die folgende Arbeitsdefinition experimenteller Kompetenz:

Experimentelle Kompetenz ist eine latente Fähigkeit zur *mindestens intuitiv regelbasierten Planung und Durchführung* von Versuchen, die der Klärung einer physikalischen Fragestellung dienen, sowie zur *methodisch bewussten Auswertung* der damit gewonnenen Daten. In *Performanz* realisiert sich experimentelle Kompetenz, wenn zusätzlich das für den Themenbereich notwendige *Fachwissen* vorhanden ist.

Die Definition orientiert sich an der Aufgliederung des experimentellen Prozesses in drei Bereiche: Planung, Durchführung und Auswertung (vgl. Emden, 2011). Den drei Bereichen werden experimentelle Teilkompetenzen zugeordnet. Dabei liegt der Schwerpunkt auf den handlungsbezogenen Teilkompetenzen, wie dem Aufbauen der Versuchsanordnung und der Durchführung von Messungen, während das Aufstellen von Hypothesen hier nicht zum Kern experimenteller Kompetenz gezählt wird. Die Vorgehensweise bei der Planung, Durchführung und Auswertung sollte zumindest intuitiv regelbasiert sein, da sich aus explorativem Handeln noch keine Kompetenz ableiten lässt (von Aufschnaiter & Rogge, 2010, S. 106). Zusätzlich wird in der Definition die Rolle des Fachwissens berücksichtigt. Dazu zählt neben den fachlichen Begriffen und Zusammenhängen insbesondere die themenbezogene Gerätekunde. Experimentelle Kompetenz soll somit als eigenständiges Konstrukt verstanden werden, welches sich aber nicht unabhängig von Fachwissen realisiert. Die aus dem Large-Scale-Einsatz des Testverfahrens gewonnenen Daten sind die Grundlage von Analysen zur Validierung des so beschriebenen Konstrukts.

Testkonzeption

Der entwickelte Experimentiertest ist in Units gegliedert. Eine Unit besteht aus einer experimentellen Fragestellung und sechs Items, die jeweils Teilkompetenzen abdecken. Zu Beginn eines Items erhalten die Schüler jeweils eine Musterlösung des vorhergehenden Items, z. B.

¹ Finanzierung durch das Bundesministerium für Bildung und Forschung im Programm zur Förderung von Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments (FKZ LSA005).

wird eine Skizze der Versuchsanordnung vorgegeben, wenn ein Versuch funktionsfähig aufgebaut werden soll (Details in Theyßen et al., 2013; Dickmann et al., 2014). Der Test wird vollständig am Computer bearbeitet. Die auf experimentelle Handlungen bezogenen Teilkompetenzen werden dabei mithilfe interaktiver Simulationen von Realexperimenten getestet. Die Bearbeitungen der Probanden können anhand von automatisch erfassten detaillierten Logdaten (Eingaben, Mausaktionen in Simulationen) ausgewertet werden. Aus den Logdaten lassen sich alle Handlungsschritte der Probanden innerhalb der interaktiven Simulationen rekonstruieren und über ein entsprechendes Tool abspielen.

Leistungsbewertung

Für die Bewertung der Güte der Item-Bearbeitungen wird ein Bewertungsraster benötigt, das über die unterschiedlichen experimentellen Aufgabenstellungen und Themenbereiche (Optik, Mechanik und Elektrizitätslehre) strukturgleich anwendbar ist. Für die unmittelbar handlungsbezogenen Teilkompetenzen „Vorgehensweise planen“, „Versuch funktionstüchtig aufbauen“ und „Messung durchführen und dokumentieren“ wird ein fünfstufiges Raster verwendet. Im Folgenden werden die Stufen am Beispiel „Vorgehensweise planen“ zunächst allgemein erläutert und dann anhand einer konkreten Aufgabenstellung veranschaulicht.

Die unteren beiden Stufen (0 und 1) unterscheiden sich darin, ob in der Planung eine Basisanordnung für die Durchführung des Experiments erkennbar ist oder nicht. Eine Basisanordnung ist dabei eine experimentelle Anordnung, mit der zwar das Phänomen realisierbar ist (z. B. Reihenschaltung von Glühlampen), jedoch keine Messungen möglich sind. In den nächsten Stufen werden Messgeräte für das Experiment eingeplant. Auf der höchsten Stufe werden die Messgeräte für eine genaue Messung aller relevanten Größen angeordnet:

- Stufe 0: eine *Basisanordnung* ist nicht hergestellt
 - Stufe 1: eine *Basisanordnung* ist hergestellt
 - Stufe 2: *messbezogene Ansätze* sind zu erkennen
 - Stufe 3: eine *Messung* im Sinne der Aufgabenstellung ist *prinzipiell* möglich
 - Stufe 4: eine der Aufgabenstellung *adäquate Messung* ist *unmittelbar* möglich
- Dabei sind alle Stufen konsekutiv zu verstehen.

In einer Aufgabe zur Elektrizitätslehre sollen die Schülerinnen und Schüler beispielsweise anhand einer Reihenschaltung mit zwei Glühlampen das 2. Kirchhoffsche Gesetz (Maschenregel) bestätigen. Dazu müssen sie entsprechend der zugrunde liegenden Modellierung des Planens von Experimenten a) geeignete Geräte auswählen, b) eine Skizze anfertigen und c) das Vorgehen beschreiben. An diesem Beispiel lässt sich die Stufenzuordnung für das Item Vorgehensweise planen wie folgt illustrieren:

- Für die Bestätigung des 2. Kirchhoffschen Gesetzes muss in der Planung zunächst ein Stromkreis mit Spannungsquelle und zwei Glühlampen in Reihe erkennbar sein. Dies entspricht der Basisanordnung (Stufe 1).
- Um auch Messungen durchführen zu können müssen Messgeräte integriert werden. Ist erkennbar, dass in der Lösung zumindest *ein* Messgerät zur Messung einer elektrischen Größe verwendet wird, so ist ein messbezogener Ansatz erkennbar (Stufe 2).
- Durch ein richtig eingebautes Spannungsmessgerät ist bereits prinzipiell eine Messung möglich (Stufe 3). Die Aufgabe ließe sich dabei durch das Umstecken der Multimeter vollständig bearbeiten.
- Um die Aufgabe richtig bearbeiten zu können wäre es allerdings sinnvoller die Spannung an den Glühlampen und der Spannungsquelle einzeln abzugreifen. Erst durch diese Anpassung wird von einer *unmittelbar möglichen* Messung gesprochen (Stufe 4).

Für die Kodierung der einzelnen Aufgaben werden auf Grundlage der allgemeinen Stufenbeschreibung konkrete Kriterien für die jeweilige Unit formuliert. Die Bewertungskriterien erfordern zum Beispiel eine genaue Überprüfung, ob entsprechende Geräte auch ausgewählt wurden. Die Abbildungen 1 und 2 zeigen Beispiele aus den Daten der Large-Scale Erhebung. Die Stufenzuordnung wird nicht allein anhand der dort dargestellten Skizzen getroffen. In die Bewertung geht ebenso ein, welche Geräte ausgewählt wurden und wie der Proband das geplante Vorgehen beschreibt.

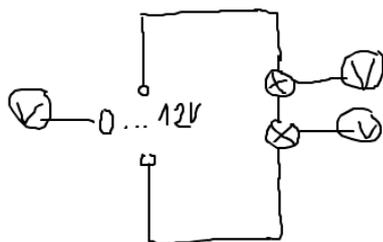


Abb. 1: Beispiel für eine Planung auf Stufe 2. Neben der Basisanordnung (Reihenschaltung) ist durch die Voltmeter ein messbezogener Ansatz erkennbar.

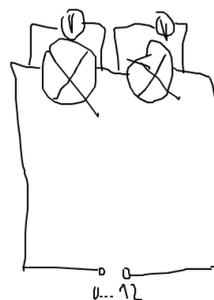


Abb. 2: Beispiel für eine Planung auf Stufe 3. Zwar lässt sich hier die Spannung richtig messen, für eine unmittelbare Messung fehlt allerdings ein Messgerät für die Gesamtspannung.

Anhand der hier beschriebenen grundlegenden Qualitätsstufeneinteilung werden für jede Unit Kriterien erstellt und auf Objektivität überprüft.

Ausblick

Eine zentrale Frage der Konstruktvalidierung im Rahmen des Projekts MeK-LSA ist die Eigenständigkeit des Konstrukts experimentelle Kompetenz. Dazu ist der Zusammenhang zwischen dem Fachwissen und der Leistungsbewertung in den Skalen des Experimentiertests zu untersuchen. Da laut Definition des Konstrukts die Performanz vom Fachwissen abhängt, wäre eine höchstens mittlere Korrelation ein Indikator für experimentelle Kompetenz als eigenständiges Konstrukt. Ein weiterer Indikator wäre, wenn unter Beschränkung der betrachteten Fälle auf Probanden, deren Fachwissen über einem bestimmten Schwellenwert liegt, eine höhere Reliabilität der Skalen für die Experimentierkompetenz erreicht wird.

Literatur

- Dickmann, M., Eickhorst, B., Theyßen, H., Neumann, K., Schecker, H., & Schreiber, N. (2014). Measuring Experimental Skills in Large-Scale Assessments: Developing a Simulation-Based Test Instrument. In C. P. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Hrsg.), E-Book Proceedings of the ESERA 2013 Conference: Science Education Research For Evidence-based Teaching and Coherence in Learning. Part 11 (Robin Millar & Jens Dolin). Nicosia, Cyprus: European Science Education Research Association.
- Dickmann, M., & Theyßen, H. (2013). Curriculare Validität von Units zur Messung experimenteller Kompetenz. In S. Bernholt (Hrsg.), Inquiry-based Learning – Forschendes Lernen. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Hannover 2012 (S. 587–589). Kiel: IPN.
- Emden, M. (2011). Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. Berlin: Logos-Verlag.
- Schreiber, N. (2012). Diagnostik experimenteller Kompetenz. Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells. Berlin: Logos-Verlag.
- Theyßen, H., Schecker, H., Neumann, K., Dickmann, M., & Eickhorst, B. (2013). Messung experimenteller Kompetenz in Large Scale Assessments. In S. Bernholt (Ed.), Inquiry-based Learning – Forschendes Lernen (S. 596–598). Kiel: IPN.
- von Aufschnaiter, C., & Rogge, C. (2010). Wie lassen sich Verläufe der Entwicklung von Kompetenz modellieren? Zeitschrift Für Didaktik Der Naturwissenschaften, 16, 95–114.