

Hoffmann, Lars; Azizi Ghanbari, Shahram; Schott, Franz

**Untersuchung zur Entwicklung von Bildungsstandards. Rekonstruktion eines vorläufigen Kompetenzmodells und der Kompetenzdiagnostik am Beispiel einer Unterrichtseinheit für das Schulfach Ethik**

*Lehrerbildung auf dem Prüfstand 2 (2009) 2, S. 225-243*



Empfohlene Zitierung/ Suggested Citation:

Hoffmann, Lars; Azizi Ghanbari, Shahram; Schott, Franz: Untersuchung zur Entwicklung von Bildungsstandards. Rekonstruktion eines vorläufigen Kompetenzmodells und der Kompetenzdiagnostik am Beispiel einer Unterrichtseinheit für das Schulfach Ethik - In: Lehrerbildung auf dem Prüfstand 2 (2009) 2, S. 225-243 - URN: urn:nbn:de:0111-pedocs-147029

<http://nbn-resolving.de/urn:nbn:de:0111-pedocs-147029>

in Kooperation mit / in cooperation with:

**VEP**

[www.vep-landau.de](http://www.vep-landau.de)

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

**Kontakt / Contact:**

peDOCS  
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft



**Herausgeber**

Rainer Bodensohn, Reinhold S. Jäger und Andreas Frey  
Bürgerstraße 23, 76829 Landau/Pfalz  
Telefon: +49 6341 906 165, Telefax: +49 6341 906 166

**Verlag**

Empirische Pädagogik e. V.  
Bürgerstraße 23, 76829 Landau/Pfalz  
Telefon: +49 6341 906 180, Telefax: +49 6341 906 166  
E-Mail: [info@vep-landau.de](mailto:info@vep-landau.de)  
Homepage: <http://www.vep-landau.de>

**Umschlaggestaltung**

© Harald Baron

**Druck**

DIFO Bamberg

*Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung, werden vorbehalten. Kein Teil des Werks darf in irgendeiner Form (durch Fotografie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung des Verlags reproduziert oder unter Verwendung elektronischer Systeme verbreitet werden.*

ISSN 1867-2779

© Verlag Empirische Pädagogik, Landau 2009

---

---

## Inhalt

---

---

|   |     |
|---|-----|
| <b>Editorial</b> .....  | 201 |
| <b>Originalarbeiten</b>   |     |
| Beck, K.:<br>Implementation von Anreizstrukturen zur Qualitätsverbesserung<br>in der Lehrerbildung .....  | 206 |
| Hoffmann, L., Azizi Ghanbari, S. & Schott, F.:<br>Untersuchung zur Entwicklung von Bildungsstandards: Rekonstruk-<br>tion eines vorläufigen Kompetenzmodells und der Kompetenzdiagnos-<br>tik am Beispiel einer Unterrichtseinheit für das Schulfach Ethik .....  | 225 |
| König, J.:<br>Zur Bildung von Kompetenzniveaus im Pädagogischen Wissen von<br>Lehramtsstudierenden: Terminologie und Komplexität kognitiver<br>Bearbeitungsprozesse als Anforderungsmerkmale von Testaufgaben? .....  | 244 |
| Košinár, J.:<br>Das Konzept „Ganzheitliche Stressprävention“ – Überprüfung der<br>Effektivität eines Trainingsprogramms in der Lehrerausbildung .....   | 263 |
| Rothland, M.:<br>Kooperation und kollegiale Unterstützung im Lehrerberuf.<br>Empirische Befunde zum Berufsbild und zu berufsrelevanten<br>Einstellungen angehender Lehrkräfte .....   | 282 |
| Schubarth, W., Speck, K., Seidel, A. & Wendland, M.:<br>Unterrichtskompetenzen bei Referendaren und Studierenden.<br>Empirische Befunde der Potsdamer Studien zur ersten und zweiten<br>Phase der Lehrerausbildung .....  | 304 |
| Weresch-Deperrois, I., Bodensohn, R. & Jäger, R. S.:<br>Curriculare Standards in der Praxis: Einschätzung ihres Stellenwerts,<br>ihrer Anwendungshäufigkeit, Schwierigkeit und Bedeutung in der<br>Lehrerausbildung und universitären Vorbereitung im Bachelor-Studium<br>der Lehrerbildung – eine Erkundungsstudie ..... | 324 |
| Frey, A. & Jäger, R. S.:<br>Der entwicklungspädagogische Ansatz in der Lehrerbildung .....  | 346 |
| <b>Richtlinien zur Manuskriptgestaltung</b> .....   | 361 |
| <b>Impressum</b> .....  | 368 |

---

---

## Contents

---

---

|  |     |
|--|-----|
| <b>Editorial</b> .....   | 201 |
| <b>Articles</b>  |     |
| Beck, K.:<br>Implementation of a structure of incentives to improve the quality<br>of teacher education .....  | 206 |
| Hoffmann, L., Azizi Ghanbari, S. & Schott, F.:<br>An empirical study towards the development of educational standards:<br>Reconstruction of a preliminary competency model and of the compe-<br>tence diagnostics for a teaching unit in the school subject ethics .....                                       | 225 |
| König, J.:<br>On the modelling of proficiency levels of future teachers' pedagogical<br>knowledge: Can test item terminology and complexity of cognitive pro-<br>cesses be used to describe item difficulty? .....   | 244 |
| Košinár, J.:<br>The concept of "Holistic Stress Prevention" – surveying the effective-<br>ness of a teacher training program .....   | 263 |
| Rothland, M.:<br>Cooperation and Support Among Teachers.<br>Empirical findings about the occupational image and about attitudes of<br>prospective teachers concerning their job .....  | 282 |
| Schubarth, W., Speck, K., Seidel, A. & Wendland, M.:<br>Teaching competencies for trainees and students. Empirical findings of<br>the Potsdam studies on first and second phase of teacher education .....   | 304 |
| Weresch-Deperrois, I., Bodensohn, R. & Jäger, R. S.:<br>Curricular standards in practice: Appraisal of their importance, fre-<br>quency of application, difficulty, relevance in teacher education and pre-<br>paration by the university in the field of BA-teacher education –<br>an explorative study ..... | 324 |
| Frey, A. & Jäger, R. S.:<br>The developmental pedagogical approach in teacher education .....  | 346 |

---



---

## Originalbeitrag

---



---

**Lars Hoffmann, Shahram Azizi Ghanbari und Franz Schott**

### Untersuchung zur Entwicklung von Bildungsstandards: Rekonstruktion eines vorläufigen Kompetenzmodells und der Kompetenzdiagnostik am Beispiel einer Unterrichtseinheit für das Schulfach Ethik

**Zusammenfassung:** Der Artikel dokumentiert die erste empirische Erprobung von ComTrans – einem theoriegeleiteten Ansatz zur Kompetenzdiagnostik und zum kompetenzorientierten Unterricht. Hierfür wurde ein vorläufiges Kompetenzmodell zur „Kompetenz, Handlung ethisch bewerten zu können“ rekonstruiert und in vier Schulklassen mit zwei Lehrerinnen im Ethikunterricht der Oberstufe eines Gymnasiums erprobt. Wir gingen dabei vom aktuellen sächsischen Lehrplan für den Ethikunterricht in der Oberstufe des Gymnasiums aus. Aus untersuchungspraktischen Gründen konnten wir uns nur auf eine Unterrichtseinheit mit drei Doppelstunden beschränken. Inhalt der betreffenden Unterrichtseinheit war, dass die Schülerinnen und Schüler lernen Handlungsentscheidungen zu beurteilen, die Begründbarkeit dieser Urteile zu prüfen und diese Überprüfung auf die ethischen Grundpositionen des kategorischen Imperativs und des größtmöglichen Glücks für die größtmögliche Zahl beziehen zu können. Ein für die Studie entwickelter Nachtest wurde als Klassenarbeit von den Schülern bearbeitet. Zusätzlich wurde dieser Test von uns sowohl kriteriumsorientiert ausgewertet, als auch einer Itemanalyse nach der Klassischen Testtheorie und nach dem Rasch-Modell unterzogen. Die Resultate dieser unterschiedlichen Auswertungsmethoden werden gegenübergestellt und diskutiert.

**Schlagwörter:** Evaluation von Lernveranstaltungen – Kompetenz – Kriteriumsorientierte Tests – Testtheorie

An empirical study towards the development of educational standards: Reconstruction of a preliminary competency model and of the competence diagnostics for a teaching unit in the school subject ethics

**Summary:** The article is documenting the first empirical testing of ComTrans – a theory driven concept for competence diagnostics and competence oriented teaching. Therefore a preliminary model of the ‘competence to evaluate actions ethically’ was reconstructed and tested with the help of two ethics teachers in four classes in a German secondary school. The basis of our competency model was the current curriculum for the school subject ethics in the Free State of Saxonia. Because of practical reasons we were only able to create and test a teaching unit that involved six forty-five minute lessons. In those lessons the pupils learned how to evaluate their action-decisions and how to verify their judges by considering the basic ethical positions of Kant’s categorical imperative and the utilitarian principle the greatest good for the greatest number. At the end of the teaching unit the pupils were given a test, which was developed on the basis of the competency model. Additionally the test was analysed in four ways: criterion referenced, with the classical test theory and with the latent trait models of Rasch and Birnbaum. The results of these analyses are compared and discussed.

**Key words:** competence – course evaluation – criterion referenced tests – test theory

## 1. Einführung in die Problemlage

Die Bildungspolitik der letzten Jahre ist durch eine Veränderung der Steuerungslogik im Bildungswesen gekennzeichnet: Bis vor wenigen Jahren regelten allein (meist inhaltlich überfrachtete) Lehrpläne, was im Unterricht behandelt werden muss; Festlegungen zum Output – also zu dem, was Schüler einer bestimmten Jahrgangsstufe können sollten – wurden im Rahmen dieser Lehrpläne nicht getroffen (Halbheer & Reusser, 2008). Der ‚PISA-Schock‘ brachte schließlich die Erkenntnis mit sich, dass die Steuerung des Bildungswesens stärker als bislang am Output von Bildungsprozessen zu orientieren ist (Klieme et al., 2007).

Während im Zusammenhang mit der veränderten Steuerungslogik im Bildungswesen vor allem die Konsequenzen der Outputorientierung für Bildungsadministration und -forschung herausgestrichen werden – illustrierend seien in diesem Zusammenhang die Stichworte ‚nationale Bildungsstandards‘ und ‚Länder übergreifende Schulvergleichstests‘ genannt – werden die konkreten Folgen dieser Reformbemühungen für die Praxis in Schulen bzw. für den einzelnen Lehrer u. E. bislang verhältnismäßig wenig diskutiert. Dabei ist gerade der einzelne Lehrer von der Einführung von Bildungsstandards stark betroffen; wird er doch verpflichtet, noch umfassender als bisher Rechenschaft über die Qualität seines Unterrichts abzulegen. Reusser und Halbheer (2008) betonen deshalb, dass der einzelne Lehrer dazu befähigt werden muss, einen qualitativ hochwertigen, an Standards orientierten Unterricht zu planen und durchzuführen. Hierfür sollen ihm im Rahmen von Fort- und Weiterbildungsmaßnahmen entsprechende Methoden und Werkzeuge vermittelt werden. Darüber hinaus müssen auch die diagnostischen Kompetenzen des Lehrers weiterentwickelt und verfeinert werden (Reusser, 2006); schließlich muss er in der Lage sein, zuverlässig und valide feststellen zu können, ob seine Schüler die Lernziele seiner Unterrichtseinheiten erreicht haben oder nicht (Oelkers & Reusser, 2008).

Während sich die Bildungsforschung ausgereifter und jahrelang erprobter testtheoretischer Modelle (z. B. Klassische Testtheorie, Rasch-Modell vgl. Hartig & Klieme, 2006) bedienen kann, um Schulleistungstests mit einer hohen psychometrischen Qualität zu entwickeln (Klieme et al., 2007), können Lehrer in der Regel nicht auf solche Werkzeuge zurückgreifen, wenn sie Lernerfolgskontrollen zur Evaluation ihrer Unterrichtseinheiten erstellen wollen, die testpsychologischen Kriterien genügen sollen: Die Gründe hierfür sind zum einen darin zu suchen, dass ihnen im Unterschied zur Bildungsforschung keine ausreichend großen Analytestichproben zur Verfügung stehen, zum anderen ist auch nicht davon auszugehen, dass sie über die hinreichenden teststatistischen Kenntnisse verfügen, um eine Lernerfolgskontrolle auf Grundlage eines testtheoretischen Modells zu erstellen.

Doch wie können Lehrer ohne die Anwendung der Methoden der Testkonstruktion in die Lage versetzt werden, Tests zur Evaluation einzelner Unterrichtseinheiten zu konstruieren, die eine hinreichende Qualität aufweisen? Wir meinen, dass eine systematische Verwendung von Verfahren der Aufgabenanalyse bei der Unterrichtsplanung und bei der Konstruktion von Lernerfolgskontrollen den Lehrern ermöglicht, den Erfolg ihrer Unterrichtseinheiten reliabel und valide zu bilanzieren.

## 2. Kompetenzbegriff, ComTrans und Verfahren der Aufgabenanalyse

Verfahren der Aufgabenanalyse haben in der Psychologie eine lange Tradition und werden auch in anderen Teildisziplinen des Fachs – wie zum Beispiel in der Arbeitspsychologie – zur Entwicklung von Tests verwendet (DuBois, Shalin, Levi & Borman, 1995); in der Pädagogischen Psychologie findet sich die Aufgabenanalyse beispielsweise in Verfahren der Lehrstoffanalyse wieder (Schott, 1975).

Aufbauend auf Vorarbeiten zur Lehrstoffanalyse haben Schott und Ghanbari (2008) im Rahmen von ComTrans (für Competence Transfer) – einem theoriegeleiteten Ansatz zum Kompetenztransfer – ein aus verschiedenen Formen der Aufgabenanalyse bestehendes Vorgehen vorgeschlagen, das es Lehrern ermöglichen soll, auch ohne die Verwendung testtheoretischer Modelle Lernerfolgskontrollen mit einer angemessenen Qualität zu konstruieren.

Nach Schott und Ghanbari beschreiben Kompetenzen Fähigkeiten, die über einen bestimmten Zeitraum stabil sind, also eine gewisse Nachhaltigkeit besitzen. In ComTrans werden Kompetenzen analog eines Lernziels operationalisiert; sie bestehen erstens aus einer Klasse von Aufgaben, die man ausführen kann, wenn man die betreffende Kompetenz besitzt und zweitens aus dem Kompetenzgrad, der angibt, wie gut man diese Aufgaben ausführen kann. Dabei kann die Angabe des Kompetenzgrads quantitativer Natur (z. B.: „90 % der Mülltrennung muss richtig sein“) oder qualitativer Natur (z. B.: „Batterien und Akkus dürfen in keine Mülltonne geworfen werden“) sein. Eine derartige festsetzende Definition von Kompetenzen ist insofern zweckmäßig, als dass sie in einer Aufgabenmenge resultiert, auf deren Basis sowohl die Vermittlung als auch die Überprüfung von Kompetenzen, welche durch Bildungsprozesse erworben und weiterentwickelt werden sollen, erfolgen kann (Schott & Ghanbari, 2008). Somit wird das Problem einer geeigneten Kompetenzbestimmung oder -operationalisierung (Klieme & Leutner, 2006) auf die Anwendung adäquater Verfahren der Aufgabenanalyse fokussiert (Schott & Ghanbari, 2008).



Innerhalb von ComTrans wird zwischen vier Arten der Aufgabenanalyse unterschieden, die semantische, fachliche<sup>1</sup>, psychologische und empirische Aufgabenanalyse genannt werden (Schott & Ghanbari, 2008). In der semantischen Aufgabenanalyse wird ermittelt, was aus Sicht eines Fachvertreters oder eines Fachexperten mit der betreffenden Kompetenz gemeint ist. In der fachlichen Aufgabenanalyse wird – ebenfalls aus der Sicht eines Experten – festgelegt, aus welchen Aufgaben oder Klassen von Aufgaben die Kompetenz genau besteht und wie diese Aufgaben gelöst werden können. Die psychologische Aufgabenanalyse beschäftigt sich damit, wie bestimmte Aufgaben bzw. eine bestimmte Kompetenz unter der Annahme bestimmter psychologischer Prozesse von einer Person mit bestimmten Eigenschaften gelernt oder bearbeitet werden können. In der empirischen Aufgabenanalyse wird schließlich ermittelt, was eine Person tatsächlich tut, wenn sie die Aufgaben löst, welche die Kompetenz charakterisieren. Der Zusammenhang zwischen diesen vier Arten der Aufgabenanalyse ist in Abbildung 1 dargestellt.

Im Rahmen des ComTrans-Ansatzes wird außerdem empfohlen, zur näheren Bestimmung der einzelnen Aufgaben Aufgabenanalyseverfahren zu verwenden, die bereits häufig empirisch erprobt und evaluiert wurden. Die insbesondere für Lehrer sehr praktikable PLANA-Methode (vgl. Schott, Neeb & Wieberg, 1981) zerlegt Aufgaben in Zustände (Aufgabenstellung und Aufgabenlösung) und Zustandsänderungen (die Operatoren, d. h., wie man von der Aufgabenstellung zur Aufgabenlösung gelangt) und ermöglicht es dadurch, Kompetenzen sowie ihre Strukturen unmissverständlich darzustellen. Ergänzt wird die Beschreibung von Aufgaben als Zustände und Zustandsänderungen durch die Angabe des Kompetenzgrades, der notwendig ist, damit eine Aufgabe als gelöst gilt.

Schlussendlich werden die genannten Arten und Verfahren der Aufgabenanalyse zur Konstruktion einer kriteriumsorientierten Lernerfolgskontrolle genutzt. Diese enthält eine geeignete Auswahl der Aufgabenmenge, welche die im Rahmen einer Unterrichtseinheit vermittelte Kompetenz definiert. Der Test wird ausgewertet, indem geprüft wird, ob ein Schüler die jeweiligen Testaufgaben entsprechend der festgelegten Kompetenzgrade gelöst hat oder nicht.

---

<sup>1</sup> In dem Buch von Schott und Azizi Ghanbari (2008) sprachen wir noch von einer „rationalen“ statt von einer „fachlichen“ Aufgabenanalyse.

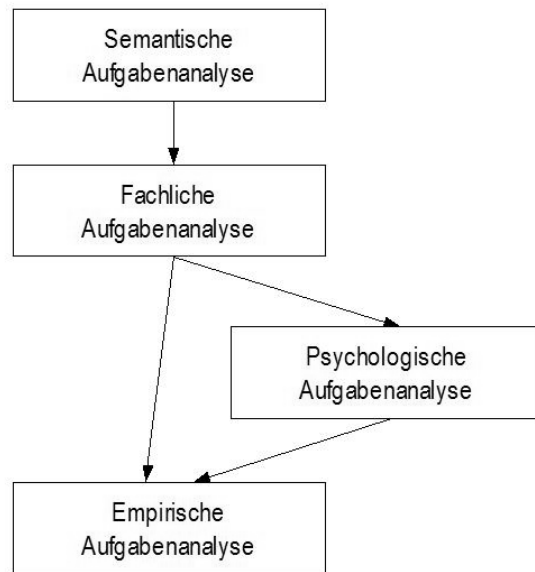


Abbildung 1: Die Abbildung stellt den Zusammenhang zwischen der semantischen, fachlichen, psychologischen und empirischen Aufgabenanalyse nach Schott und Ghanbari (2008) dar.

Eingangs wurde behauptet, dass eine systematische Anwendung der im ComTrans-Konzept beschriebenen Aufgabenanalyseverfahren Lehrer dazu befähigen kann, Lernerfolgskontrollen mit einer angemessenen psychometrischen Qualität zu konstruieren. Um diese Behauptung empirisch zu unterlegen, sollen im Folgenden die Ergebnisse einer Pilotstudie zur Verwendbarkeit von ComTrans vorgestellt werden.

### 3. Was war Gegenstand der Pilotstudie?

In der Pilotstudie wurde unter der Mithilfe zweier Ethiklehrerinnen eines Dresdner Gymnasiums ein vorläufiges Kompetenzmodell zu einer einführenden Unterrichtseinheit für das Schulfach Ethik mit dem Thema „Ethische Bewertung von Handlungen“ rekonstruiert. Dabei wurden die oben skizzierten Verfahren der Aufgabenanalyse verwendet. Um das Vorgehen bei diesen Verfahren besser zu veranschaulichen, sollen die Ergebnisse der Aufgabenanalysen unter 4.1 dargestellt werden.

Darüber hinaus wurde die Unterrichtseinheit auch empirisch erprobt. Eine hierfür – ebenfalls unter der Verwendung von Aufgabenanalyseverfahren – konstru-

ierte Lernerfolgskontrolle wurde sowohl rein kriteriumsorientiert, als auch nach der Klassischen Testtheorie sowie nach dem Rasch- und Birnbaummodell ausgewertet. Im Anschluss wurden die Ergebnisse der verschiedenen Auswertungsmethoden miteinander verglichen (siehe 4.2 und 5.). Eine hohe Übereinstimmung zwischen den Ergebnissen einer rein kriteriumsorientierten Auswertung (mit vorangegangener Aufgabenanalyse) mit den psychometrischen Kennwerten der testtheoretischen Modelle (Klassische Testtheorie, Rasch- und Birnbaummodell) spräche dafür, dass die bloße Anwendung von Verfahren der Aufgabenanalyse bei der Testkonstruktion die Entwicklung einer qualitativ hochwertigen Lernerfolgskontrolle erlaubt, die eine zusätzliche Durchführung von Itemanalysen auf Grundlage testtheoretischer Modelle erübrigt.

#### 4. Untersuchungsmethoden

##### 4.1 Durchführung der Aufgabenanalysen

###### 4.1.1 Durchführung der Semantischen Aufgabenanalyse

Im Zuge einer semantischen Aufgabenanalyse wurde zunächst bestimmt, was Gegenstand des von uns zu rekonstruierenden Kompetenzmodells zur ‚Kompetenz, Handlungen ethisch bewerten zu können‘ sein sollte: Mit anderen Worten ging es zuallererst darum, genauer zu analysieren, was mit dieser Kompetenz gemeint ist.

Hierfür orientierten wir uns zunächst an den im aktuellen Sächsischen Lehrplan für das Schulfach Ethik in der Sekundarstufe 2 (Sächsisches Staatsministerium für Kultus, 2004) erläuterten ‚allgemeinen fachlichen Zielen, Lernbereichen, -zielen und -inhalten‘. In diesem werden für den Lernbereich ‚Fragen nach dem guten Handeln‘ 48 Unterrichtsstunden veranschlagt. Aus verschiedenen, vor allem praktischen Gründen war es uns allerdings lediglich möglich, eine Unterrichtseinheit zu entwickeln und empirisch zu erproben, welche den Umfang von drei Doppelstunden von jeweils 90 Minuten nicht überstieg. Eine derartig kurze Unterrichtseinheit konnte folglich nur einführenden, Grundlagen schaffenden Charakter haben und somit nicht danach streben, alle im Sächsischen Lehrplan für den Lernbereich genannten Lernziele (Sächsisches Staatsministerium für Kultus, 2004) oder alle in den ‚Einheitlichen Prüfungsanforderungen in der Abiturprüfung Ethik‘ (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2007) erläuterten Anforderungsniveaus abzudecken.

Schließlich entschieden wir, die Inhalte des Kompetenzmodells beziehungsweise die Themen der Unterrichtseinheit an den im Sächsischen Lehrplan unter ‚Kennen ethischer Grundpositionen‘ genannten Lerninhalten (Pflichtethik und utilitaristische Ethik) sowie an den Erläuterungen zum Ziel des ‚Erwerbs morali-

scher Urteilsfähigkeit‘ zu orientierten. Auf dieser Grundlage operationalisierten wir die ‚Kompetenz, Handlungen ethisch bewerten zu können‘ mit dem Lernziel: „Die Schüler können Handlungsentscheidungen beurteilen, die Begründbarkeit dieser Urteile prüfen und diese Überprüfung auf die ethischen Grundpositionen des ‚kategorischen Imperativs‘ und des ‚größtmöglichen Glücks für die größtmögliche Zahl‘ beziehen“.

#### 4.1.2 Fachliche Aufgabenanalyse

Das Ziel der nun folgenden fachlichen Aufgabenanalyse war es festzulegen, welche Mengen von Aufgaben ein Schüler aus fachdidaktischer Sicht lösen können muss, damit davon ausgegangen werden kann, dass sie oder er über die ‚Kompetenz, Handlungen ethisch bewerten zu können‘ verfügt bzw. das Lernziel der Unterrichtseinheit erreicht hat. Diese Aufgaben mussten entsprechend erstellt werden.

Die fachliche Aufgabenanalyse erfolgte auf Grundlage der Analyse von Lehrbüchern für das Schulfach Ethik, für eine noch präzisere Explikation der philosophisch-ethischen Inhalte der Unterrichtseinheit wurde die Enzyklopädie Philosophie und Wissenschaftstheorie (Mittelstraß, 1995) verwendet. Auch die im Projektverlauf zum Projektteam gestoßenen Lehrer halfen bei Fragen zu ethisch-philosophischen Themen. Außerdem wurde der Kontakt zu einem an der Technischen Universität Dresden tätigen Didaktiker für das Schulfach Ethik gesucht.

Schließlich wurde unter Berücksichtigung der Ergebnisse der semantischen Aufgabenanalyse entschieden, Aufgaben zu den Inhaltsaspekten ‚kategorischer Imperativ von Immanuel Kant‘, ‚größtmögliches Glück der größtmöglichen Zahl/Utilitarismus‘ sowie ‚Handlung, Handlungsstruktur und Bewertungsaspekte von Handlungen‘ zu erstellen.

Zu diesen drei Inhaltsbereichen der ‚Kompetenz, Handlungen ethisch bewerten zu können‘ wurden verschiedene Aufgabentypen formuliert: Aufgaben, die lediglich auf die Reproduktion deklarativer Wissensinhalte zielten (Verhaltensaspekt: Wissen); Aufgaben, welche die Anwendung der genannten Inhalte auf konkrete Situationen, die entweder vorgegeben wurden oder selbst erstellt werden mussten oder sogar den Charakter eines moralischen Dilemmas hatten, erforderten (Verhaltensaspekt: Anwenden) sowie Aufgaben, die einen Vergleich der beiden ethischen Grundpositionen des ‚Kategorischen Imperatives‘ und des ‚größtmöglichen Glücks der größtmöglichen Zahl‘ oder sogar eine kritische Auseinandersetzung mit diesen Konzepten beinhalteten (Verhaltensaspekt: Vergleichen).

Um jene Verhaltensaspekte – oder Operatoren – besser bestimmen zu können und um die Voraussetzung für die Erstellung einer kontentvaliden Unterrichtseinheit und Lernerfolgskontrolle zu schaffen, wurden die Aufgaben nach dem PLAN-Verfahren erstellt. Die zu den beiden ethischen Grundpositionen formu-

lierten Wissensaufgaben deckten lediglich die Aspekte ab, die nach Meinung der Lehrer essenziell für ein erstes Grundverständnis dieser moralphilosophischen Ansätze war. Auch die bereits weiter oben genannten Verhaltensaspekte – Wissen, Anwenden und Vergleichen – wurden so gewählt, dass die zu deren Operationalisierung formulierten Aufgaben im Rahmen einer einführenden Unterrichtseinheit mit dem Umfang von drei Doppelstunden zeitlich realisierbar waren.

Schlussendlich mündete der Prozess der Aufgabenkonstruktion in der Fertigstellung einer Aufgabensammlung. Diese umfasste 27 Aufgaben, die wir in die Bereiche bzw. Module ‚Handeln‘, ‚Kategorischer Imperativ‘, ‚Utilitarismus‘ sowie ‚Kritik und Vergleich der ethischen Grundpositionen‘ gliederten. Vier dieser Aufgaben aus dem Modul ‚Handeln‘, welche Aspekte der Bewertung von Handlungen thematisieren, sind in Abbildung 2 dargestellt.

Ferner wurden 12 Beispielsituationen konstruiert, in denen sich Personen in einem moralischen Dilemma befanden. Diese Situationen wurden so gewählt, dass die in ihnen beschriebenen Konflikte auch Gegenstand der Erlebniswelt von Schülern der Sekundarstufe 2 hätten sein können.

| Nr. | Aufgabenstellung  | Operator  | Aufgabenlösung  | Kompetenzgrad  | Bemerkgn |
|-----|---|-----------|---|--|----------|
| 6   | Gib ein Beispiel für die Bewertung eines Handlungsmotivs!   | Anwenden  | [das Beispiel ist richtig bzw. angemessen]                | Konkretes Beispiel muss genannt sein.  |          |
| 7   | Gib ein Beispiel für die Bewertung einer Handlung   | Anwenden  | [das Beispiel ist richtig bzw. angemessen]                | Konkretes Beispiel muss genannt sein.  |          |
| 8   | Gib ein Beispiel für die Bewertung eines Handlungsziels   | Anwenden  | [das Beispiel ist richtig bzw. angemessen]                | Konkretes Beispiel muss genannt sein.  |          |
| 9   | Erläutere an einem eigenen Beispiel einen Handlungsverlauf: <ul style="list-style-type: none"> <li>• [für eine Person] oder</li> <li>• [für viele Personen] oder</li> <li>• [für alle Menschen]</li> </ul> bei dem: <ul style="list-style-type: none"> <li>• [das Motiv] und/oder</li> <li>• [die Handlung] und/oder</li> <li>• [das Ziel]</li> </ul> jeweils: <ul style="list-style-type: none"> <li>• [gut] oder</li> <li>• [schlecht]</li> </ul> Ist/sind! | Verstehen | [das Beispiel ist richtig und die Erläuterung angemessen] | Mindestens gelöst werden sollen Aufgaben, bei denen Motiv, Handlung und Ziel jeweils alle gut oder alle schlecht sind und eine Aufgabe, bei der das Motiv gut, aber Handlung und Ziel schlecht sind sowie eine Aufgabe, bei der das Motiv und das Ziel gut sind, aber die Handlung schlecht. |          |

Abbildung 2: Vier Aufgaben aus dem Modul ‚Handeln‘

### 4.1.3 Psychologische Aufgabenanalyse

Im Rahmen der Aufgabenerstellung fiel auf, dass Themen bzw. Ansätze aus dem Gegenstandsbereich der wissenschaftlichen Psychologie sowohl in den Lehrbüchern als auch im Lehrplan überhaupt nicht zu finden sind. Dies verwunderte deswegen etwas, da unseres Erachtens insbesondere die Psychologie wichtige Beiträge zur Erschließung des Lernbereichs ‚Fragen nach dem guten Handeln‘ liefern kann.

Im Ergebnis dieser Überlegungen wurde entschieden, Aufgaben zu einem weiteren (bereits unter 4.1.2 genannten) Themenbereich (‚Handlung, Handlungsstruktur und Bewertungsaspekte von Handlungen‘) zu formulieren. Im Zuge dessen wurde für die Schüler mit der ‚Struktur von Handlungsverläufen‘ (ein Handlungsverlauf besteht aus Motiv, Handlung und dem zugehörigen Handlungsergebnis) ein Wissensschema konzipiert, welches ihnen ermöglichen sollte, die ethischen Grundpositionen des ‚Kategorischen Imperativs‘ sowie des ‚größtmöglichen Glücks der größtmöglichen Zahl‘ und deren Bezug zur zu vermittelnden Kompetenz besser zu verstehen und gegenüberstellen zu können (der kategorische Imperativ bewertet das Handeln von Menschen vor allem hinsichtlich des Motivs, während die utilitaristische Position des größtmöglichen Glücks der größtmöglichen Zahl den moralischen Gehalt einer Handlung an deren Folgen festmacht). Ergänzend wurden weitere Aufgaben zu handlungsrelevanten Themen formuliert (z. B. Wie kann man den Begriff Handlung definieren? Was ist ein Handlungsmotiv? Unter welchen Bedingungen ist man für sein Handeln moralisch verantwortlich?).

### 4.1.4 Empirische Aufgabenanalyse

Im Rahmen der empirischen Aufgabenanalyse wurde von den beiden Ethiklehrerinnen im März 2008 eine einführende, drei Doppelstunden umfassende Unterrichtseinheit zur ‚Kompetenz, Handlungen ethisch bewerten zu können‘ in vier Ethikkursen eines Dresdner Gymnasiums unterrichtet. Insgesamt nahmen 100 Schüler der elften Klasse an diesem Unterricht teil. Grundlage der Unterrichtseinheit waren die 27 Aufgaben und 12 Dilemmata umfassende Aufgabensammlung sowie ergänzende Erläuterungen – beispielsweise hinsichtlich der Projektziele, der Inhalte, die durch die Aufgaben des Aufgabenpools abgedeckt wurden sowie bezüglich weiterer, im Rahmen von ComTrans wichtiger Aspekte. Aus diesen Erläuterungen und aus der Aufgabensammlung wurde ein Skript erstellt, welches den beiden Lehrerinnen als eine Art Lehrerkompetenzmodell zur Vermittlung des im Zuge der semantischen Aufgabenanalyse formulierten Lehrziels dienen konnte. Weitere Festlegungen – beispielsweise zur Art der Aufbereitung des Lernstoffs für den Unterricht – wurden nicht getroffen.

Am Ende der Unterrichtseinheit wurde mittels eines circa 25 Minuten dauernden und 6 Aufgaben bzw. 11 Teilaufgaben umfassenden schriftlichen Nachtests geprüft, ob die Schüler das Ziel der Unterrichtseinheit erreichen konnten. Die Aufgaben dieser kriteriumsorientierten Lernerfolgskontrolle stammten aus der im Zuge der vorangegangenen Aufgabenanalysen erstellten Aufgabensammlung. Die Auswertung der Nachtests erfolgte zunächst entsprechend der im Zuge der Anwendung des PLANA-Verfahrens getroffenen Festlegungen: Für jede Teilaufgabe wurde geprüft, ob die im Kompetenzgrad angegebenen Inhalte genannt wurden. War dies der Fall, so wurde die Teilaufgabe als gelöst gewertet. Somit wurde für jeden Schüler bestimmt, wie viele Teilaufgaben des Nachtests er richtig bearbeiten konnte (= kriteriumsorientierte Auswertung mit vorangegangener Aufgabenanalyse).

## 4.2 Empirische Itemanalyse

Die so ermittelten Antwortmuster waren außerdem Grundlage für die zusätzlich durchgeführten empirischen Itemanalysen.

### 4.2.1 Itemanalyse nach der Klassischen Testtheorie

Zunächst wurden die Aufgaben einer empirischen Itemanalyse nach der Klassischen Testtheorie (KTT) unterzogen. Im Zuge dessen wurden die Schwierigkeitsindizes und Trennschärfen der Items bestimmt. Vier Items wurden von mehr als 80 % der Schülerinnen und Schüler richtig gelöst. Entsprechend der Konventionen von Lienert und Raatz (1994), denen zufolge nach Items außerhalb des Schwierigkeitskorridors zwischen 0.2 und 0.8 nicht Bestandteile einer Testform sein sollten, wurden diese Items aus der KTT-Skala entfernt. Als Maß der internen Konsistenz der resultierenden Skala wurde der Alpha-Koeffizient nach Cronbach berechnet (Lienert & Raatz, 1994; Rost, 1996). Der Skalenwert für jeden Schüler wurde durch Aufsummieren der jeweils richtig gelösten Items gebildet.

### 4.2.2 Dimensionalitätsprüfungen

Im Zuge der Itemanalysen nach dem Rasch- und Birnbaummodell wurde die Dimensionalität der resultierenden Testskalen unter der Verwendung der statistischen Verfahren der Faktorenanalyse sowie der "modified parallel analysis" untersucht. Die Ergebnisse dieser Berechnungen legten nahe, dass im Rahmen der vorliegenden Untersuchung die für die Verwendung der probabilistischen Testmodelle wichtige Voraussetzung der Skaleneindimensionalität erfüllt war.

### 4.2.3 Itemanalyse nach dem Rasch- und dem Birnbaummodell

Die Schätzung der Item- und Personenparameter des Rasch- und Birnbaummodells erfolgte unter Verwendung des Programmpakets „ltm“ für die Statistiksoft-

ware R (Rizopulos, 2006). Im Zuge von Item-Fit-Analysen wurde festgestellt, dass die Passung einiger Testitems mit den ermittelten Modellen sehr unbefriedigend war: Aus diesem Grund wurde die aus der Rasch-Analyse entstandene Testskala um drei Items auf acht reduziert, die aus Birnbaum-Analyse resultierende Skala bestand aus sechs Testaufgaben.

#### 4.2.4 Bestimmung der Korrelationen zwischen den Testskalen

Um die Angemessenheit der infolge der Aufgabenanalysen entwickelten kriteriumsorientierten Lernerfolgskontrolle zu untersuchen, wurde die Anzahl der in diesem Test richtig gelösten Punktwerte (= kriteriumsorientierte Auswertung mit vorangegangener Aufgabenanalyse) mit den Skalenwerten der Schüler auf der KTT-Skala und mit den anhand des Rasch- und Birnbaummodells für jeden Schüler errechneten Personenparameter verglichen. Diese Gegenüberstellung der verschiedenen Auswertungsmethoden erfolgte, indem die Höhe der Produkt-Moment-Korrelationen zwischen den im Nachtest erreichten Punktwerten (= kriteriumsorientierte Auswertung mit vorangegangener Aufgabenanalyse) und den drei „Fähigkeitsscores“ auf den im Ergebnis der empirischen Itemanalysen ermittelten Testskalen (KTT, Rasch und Birnbaum) berechnet wurde.

### 5. Ergebnisse

#### 5.1 Ergebnisse der kriteriumsorientierten Auswertung

Die Verteilung der Anzahl der von den Schülern gelösten Testaufgaben ist in Abbildung 3 dargestellt. Im Durchschnitt lösten die Schüler 8.06 Testaufgaben richtig ( $s = 1.94$ ). Lediglich 23 % der Schüler konnten 10 oder 11 Aufgaben richtig bearbeiten.



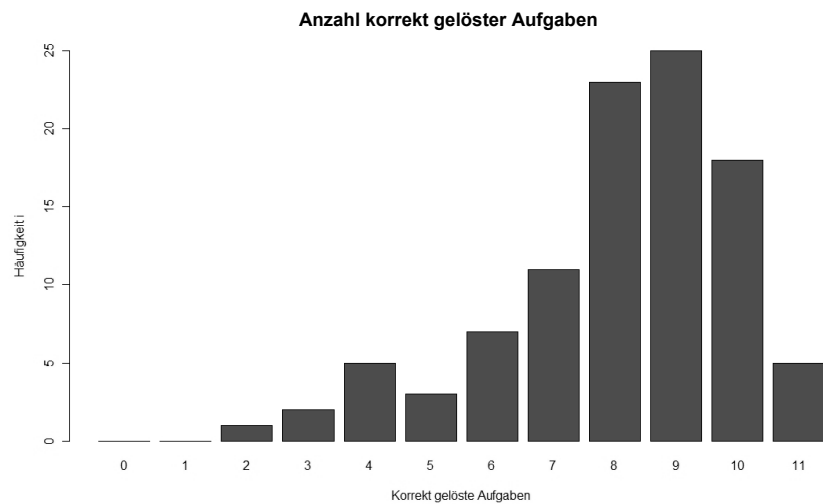


Abbildung 3: Verteilung der Anzahl der von den Schülern korrekt gelösten Aufgaben

## 5.2 Empirische Itemanalyse nach der Klassischen Testtheorie

Die für die Items ermittelten Schwierigkeitsindizes und Trennschärfekoeffizienten sind in Abbildung 4 dargestellt. Für vier Items wurde eine Itemschwierigkeit größer ,80 ermittelt. Diese Items wurden aus der KTT-Skala entfernt. Aufgrund der Entfernung dieser Items bildete die KTT-Skala nicht mehr alle Inhaltsfacetten des Aufgabenpools ab und wies somit eine geringe Kontenvalidität auf: Um auf dieser Skala den maximalen Punktwert zu erreichen, musste ein Schüler beispielsweise nicht erläutern können, unter welchen Bedingungen Menschen für ihr Handeln verantwortlich sind (Item 4).

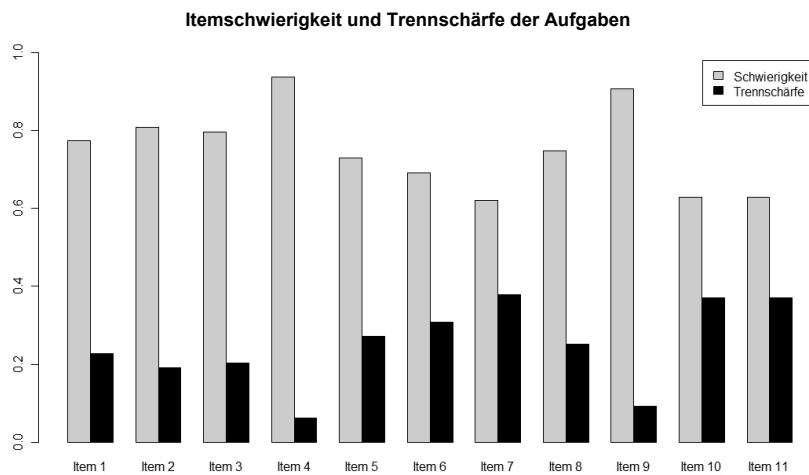


Abbildung 4: Die Schwierigkeitsindizes und Trennschärfekoeffizienten der Test-items

Für die resultierende Skala wurde ein Cronbach-Alpha-Koeffizient von .51 errechnet, der Alphakoeffizient für alle 11 Items des Nachtests betrug .56.

Der Umstand, dass sich die interne Konsistenz der Testskala nach dem Entfernen von vier Items nur geringfügig veränderte, legitimierte also aus psychometrischer Sicht die Reduktion der Skala auf sieben Items. Angesichts der Tatsache, dass mit dem Test unterschiedliche Facetten der Kompetenz, ‚Handlungen ethisch bewerten zu können‘ erfasst wurden, erschien die interne Konsistenz der Testskala durchaus akzeptabel.

### 5.3 Empirische Itemanalyse nach dem Rasch- und Birnbaummodell

Die durchgeführten Item-Fit-Analysen ergaben, dass drei Items nicht den Anforderungen des Rasch- und fünf Items nicht den Anforderungen des Birnbaum-Modells genügten. Diese Items wurden aus den ermittelten Modellen entfernt; folglich war auch die Kontenvalidität der Rasch- und Birnbaum-Skalen reduziert. Auf weitere Ergebnisse der durchgeführten Rasch- und Birnbaum-Analysen kann im Rahmen dieses Artikels leider nicht eingegangen werden. Exemplarisch sind die ermittelten ICC-Graphen in den Abbildungen 5 und 6 dargestellt.

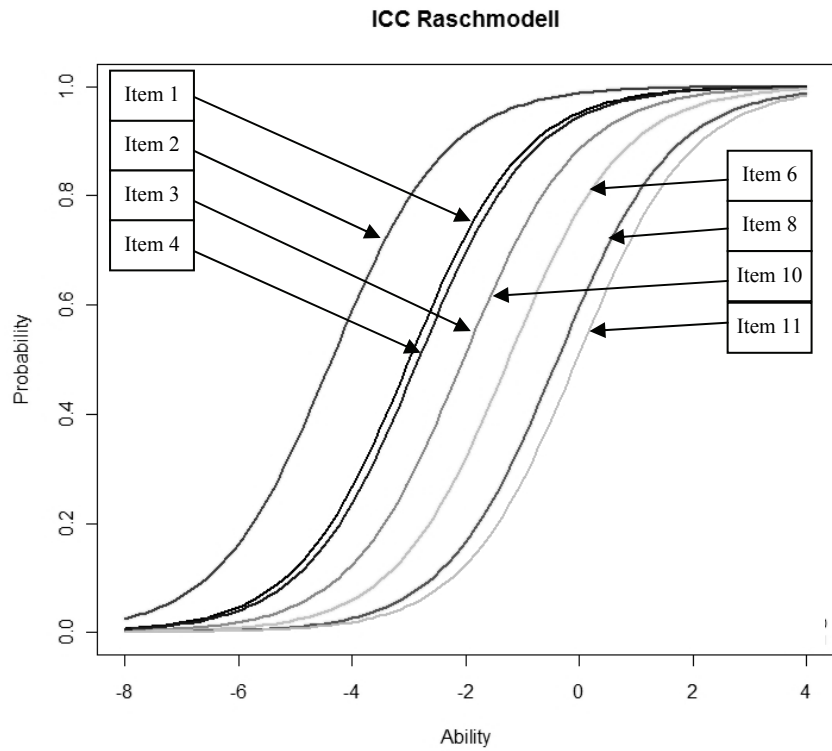


Abbildung 5: Die mittels des Rasch-Modells ermittelten ICC-Graphen der Items

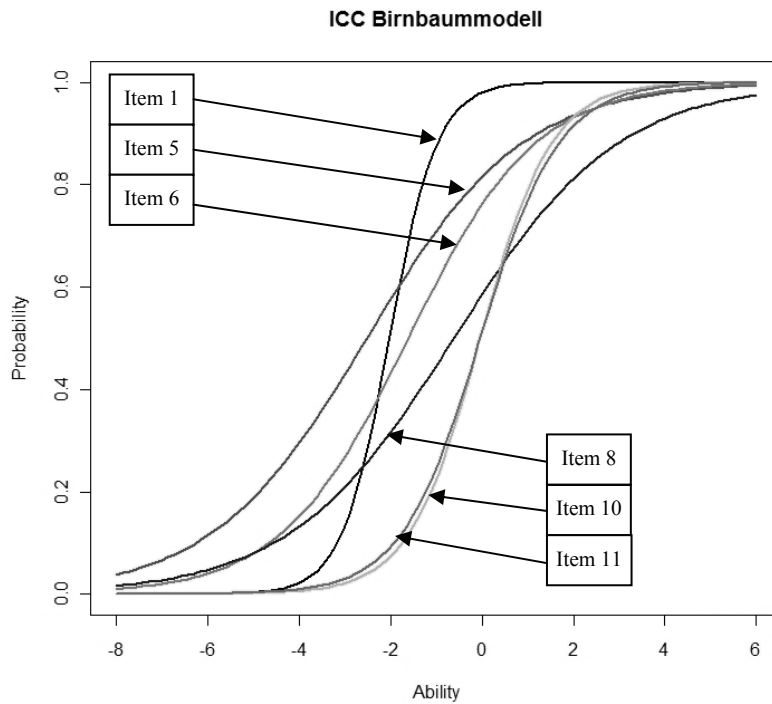


Abbildung 6: Die ICC-Graphen der Items, die mittels des Birnbaum-Modells berechnet wurden

#### 5.4 Korrelationen zwischen den Testskalen

Schlussendlich wurden die von den Schülern im kriteriumsorientierten Nachtest (mit vorangegangener Aufgabenanalyse) erreichten Summenwerte mit den Skalenwerten der KTT-Skala und den mittels des Rasch- und Birnbaum-Modells ermittelten Personenparametern verglichen. Wie anhand von Tabelle 1 ersichtlich ist, fielen die hierfür berechneten Korrelationskoeffizienten allesamt sehr hoch und statistisch signifikant aus.

Tabelle 1: Die berechneten Korrelationskoeffizienten zwischen den von den Schülern auf der kriteriumsorientierten Skala erreichten Summenwerten mit den Skalenwerten der KTT-Skala und den Personenparametern des Rasch- und Birnbaum-Modells

| Korrelation r mit den Skalenwerten der KTT-Skala | Korrelation r mit den Personenparametern des Rasch-Modells | Korrelation r mit den Personenparametern des Birnbaum-Modells |
|--|--|---|
| .96 (p < .001)                                   | .90 (p < .001)   | .88 (p < .001)  |

## 6. Fazit

Wie bereits eingangs erläutert, erfordert die Hinwendung zu einer mehr am Output orientierten Steuerung des Bildungssystems, dass auch der einzelne Lehrer in der Lage sein muss, den Erfolg seiner Unterrichtseinheiten zuverlässig und valide zu bilanzieren. Während Schulleistungstests in der Bildungsforschung auf der Basis testtheoretischer Modelle entwickelt werden, haben Lehrer aber in der Regel weder die fachlichen Voraussetzungen noch die notwendigen Ressourcen, um Leistungstests zur Evaluation ihrer Unterrichtseinheiten mittels dieser Modelle zu entwickeln. Einen möglichen Ausweg könnte ein im Rahmen von ComTrans beschriebenes Vorgehen bieten, mit Hilfe dessen Lernerfolgskontrollen anhand von Aufgabenanalysen und ohne die Anwendung komplizierter statistischer Verfahren erstellt werden können.

Dieses Vorgehen wurde in der hier skizzierten Pilotstudie empirisch erprobt. Dabei wurde die Kompetenz, ‚Handlungen ethisch bewerten zu können‘ in vier Ethikkursen eines Dresdner Gymnasiums im Rahmen einer einführenden Unterrichtseinheit vermittelt. Der Vergleich der von den Schülern in der kriteriumsorientierten Lernerfolgskontrolle (mit Aufgabenanalyse) erreichten Summenwerte mit den Fähigkeitswerten auf den Testskalen, die aus zusätzlich durchgeführten empirischen Itemanalysen resultierten, zeigte sehr hohe korrelative Übereinstimmungen. Dies kann als Beleg für die Angemessenheit des in ComTrans beschriebenen Vorgehens zur Konstruktion von Lernerfolgskontrollen zur Bilanzierung von Unterrichtseinheiten gewertet werden. Einschränkend sei allerdings erwähnt, dass für eine abschließende Beurteilung und Weiterentwicklung dieses Ansatzes noch weitere Untersuchungen (z. B. mit inhaltlich umfangreicheren Unterrichtseinheiten und größeren Schülerstichproben) notwendig sind. In diesen Untersuchungen sollten auch die von den unterrichteten Lehrern vorgenommenen Benotungen Berücksichtigung finden und mit den Ergebnissen unserer Testauswertungen verglichen werden. Zur weiteren empirischen Fundierung des ComTrans-

Konzepts sollte außerdem erfasst werden, inwiefern und in welchem Ausmaß die Lehrer bei der Vermittlung der Unterrichtseinheit auch tatsächlich die Aufgaben des Aufgabenpools nutzten bzw. wie kompetenzorientiert der Unterricht der Lehrer war. Eine hohe positive Korrelation zwischen den Testergebnissen der Schüler und der Kompetenzorientierung des Unterrichts würde für die Angemessenheit und Praktikabilität vom ComTrans sprechen.

Eingangs wurde erläutert, dass die einzelnen Lehrer dazu befähigt werden müssen, reliabel und valide zu diagnostizieren, inwiefern ihre Schüler die Lernziele einer Unterrichtseinheit erreicht haben. Angesichts dieser Forderung erscheint es u. E. nach unbefriedigend, lediglich die Adäquanz des von uns beschriebenen Vorgehens zur Konstruktion von Lernerfolgskontrollen quasi-experimentell zu belegen; vielmehr sollten die von uns gewonnen Erkenntnisse und Erfahrungen bei der Anwendung von Verfahren der Aufgabenanalyse zur Planung, Durchführung und Evaluation von Unterrichtseinheiten auch in die Lehrpraxis an den Schulen einfließen. Aus diesem Grund planen wir, eine Trainingseinheit zu entwickeln, mit der die Diagnosekompetenz von Lehrern und Lehramtsstudenten verbessert werden soll, die aus unserer Sicht bislang in der Aus- und Fortbildung von Lehrern eher vernachlässigt wurde. Im Rahmen dieses Kompetenztrainings werden die Teilnehmer die oben skizzierten Verfahren der Aufgabenanalyse kennen lernen, an vorgegeben Unterrichtsthemen durchführen, üben und lehrpraktisch erproben. Dabei soll insbesondere aufgezeigt werden, welche Vorteile mit einer systematischen Anwendung von Verfahren der Aufgabenanalyse im Hinblick auf eine lehrzielvalide Planung, Durchführung und Evaluation von Unterrichtseinheiten verbunden sind, so dass die Trainingsteilnehmer motiviert werden, diese Verfahren in ihrem Schulalltag einzusetzen. Um das Kompetenztraining auch fest in das System der Aus- und Fortbildung von Lehrern implementieren zu können, ist es natürlich notwendig, dessen Akzeptanz und Wirksamkeit empirisch zu belegen. Aus diesem Grund soll das Kompetenztraining durch ein umfangreiches Evaluationssystem flankiert werden, das dem Nachweis, der Sicherung und der Verbesserung der Qualität der Trainingseinheit dienen wird.

### Literatur

- DuBois, D. A., Shalin, V. L., Levi, K. R. & Borman, W. C. (1995). A cognitively-oriented approach to task analysis and test development. Arlington: Office of Naval Research.
- Halbheer, U. & Reusser, K. (2008). Outputsteuerung, Accountability, Educational Governance – Einführung in Geschichte, Begrifflichkeiten und Funktionen von Bildungsstandards. Beiträge zur Lehrerbildung. Zeitschrift zu Theorie und Praxis der Aus- und Weiterbildung von Lehrerinnen und Lehrern, 3, 253-266.

- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127-143). Berlin: Springer.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Überarbeitete Fassung des Antrags an die DFG auf Einrichtung eines SPP. Verfügbar unter [http://kompetenzmodelle.dipf.de/images/antrag\\_spp\\_kompetenzdiagnostik\\_ueberarbeitet.pdf](http://kompetenzmodelle.dipf.de/images/antrag_spp_kompetenzdiagnostik_ueberarbeitet.pdf) [30.01.2008].
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2007). *Bildungsforschung Band 1. Zur Entwicklung nationaler Bildungsstandards – Expertise*. Bonn: Bundesministerium für Bildung und Forschung (BMBF).
- Lienert, G. A. & Raatz, U. (1994). *Testaufbau und Testanalyse* (5. Auflage). Weinheim: Beltz, PVU.
- Mittelstraß, J. (1995). *Enzyklopädie Philosophie und Wissenschaftstheorie* (In 4 Bänden). Mannheim: Bibliographisches Institut.
- Oelkers, J. & Reusser, K. (unter Mitarbeit von E. Berner, U. Halbheer & S. Stolz) (2008). *Qualität entwickeln – Standards sichern – mit Differenz umgehen*. *Bildungsforschung Band 27*. Bonn: Bundesministerium für Bildung und Forschung. Verfügbar unter: <http://www.bmbf.de/publikationen/2713.php> [13.04.2009].
- Reusser, K. (2006). Konstruktivismus – vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur. In M. Baer, M. Fuchs, P. Füglistner, K. Reusser & H. Wyss (Hrsg.), *Didaktik auf psychologischer Grundlage. Von Hans Aebli kognitionspsychologischer Didaktik zur modernen Lehr- und Lernforschung* (S. 151-168). Bern: h.e.p.-verlag.
- Reusser, K. & Halbheer, U. (2008). Bildungsstandards als Ausgangspunkt für Unterrichtsentwicklung. *Zeitschrift zu Theorie und Praxis der Aus- und Weiterbildung von Lehrerinnen und Lehrern*, 3, 304-317.
- Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analysis. *Journal of Statistical Software*, 17, 1-25.
- Rost, J. (1996). *Testtheorie, Testkonstruktion*. Bern: Huber.
- Sächsisches Staatsministerium für Kultus (2004). *Lehrplan Gymnasium. Ethik*. Verfügbar unter: [http://www.sachsen-macht-schule.de/apps/lehrplandb/downloads/lehrplaene/lp\\_gy\\_ethik.pdf](http://www.sachsen-macht-schule.de/apps/lehrplandb/downloads/lehrplaene/lp_gy_ethik.pdf) [22.11.2007].
- Schott, F. (1975). *Lehrstoffanalyse – Ein Beschreibungssystem zur Analyse von Inhalt und Verhalten bei Lehrzielen*. Düsseldorf: Schwann.
- Schott, F. & Azizi Ghanbari, S. (2008). Kompetenzdiagnostik, Kompetenzmodelle, kompetenzorientierter Unterricht. Zur Theorie und Praxis überprüfbarer Bildungsstandards. *ComTrans – ein theoriegeleiteter Ansatz zum Kompetenztransfer als Diskussionsvorlage*. Münster: Waxmann.

Schott, F., Neeb, K.-E. & Wieberg, H.-J. W. (1981). Lehrstoffanalyse und Unterrichtsplanung – eine praktische Anleitung zur Analyse von Lehrstoffen, Präzisierung von Lehrzielen, Konstruktion von Lehrmaterialien und Überprüfung des Lernerfolges. Braunschweig: Westermann.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2007). Beschlüsse der Kultusministerkonferenz. Einheitliche Prüfungsanforderungen in der Abiturprüfung Ethik. Beschluss vom 1.12.1989 i. d. F. vom 16.11.2006. München: Luchterhand/Wolters Kluwer.

*Anschrift der Autoren:*

Dipl.-Psych. Lars Hoffmann, Dr. Shahram Azizi Ghanbari und Prof. Dr. Franz Schott, TU Dresden, Institut für Pädagogische Psychologie und Entwicklungspsychologie, Mommsenstr. 13, 01062 Dresden, E-Mail: lars.hoffmann@gmx.net; schott@psychologie.tu-dresden.de; shahram.azizighanbari@mailbox.tu-dresden.de