

Wollersheim, Heinz-Werner; März, Maren; Schminder, Jan
**Digitale Prüfungsformate. Zum Wandel von Prüfungskultur und Prüfungspraxis
in modularisierten Studiengängen**

Zeitschrift für Pädagogik 57 (2011) 3, S. 363-374

urn:nbn:de:0111-opus-87315



in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen / conditions of use

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.

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.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.
By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Mitglied der Leibniz-Gemeinschaft
Informationszentrum (IZ) Bildung
Schloßstr. 29, D-60486 Frankfurt am Main
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

ZEITSCHRIFT FÜR PÄDAGOGIK

Heft 3

Mai/Juni 2011

■ *Thementeil*

Digitale Medien in der Hochschullehre

■ *Allgemeiner Teil*

Beeinflussen individuelle Werte Motivation und Lernerfolg bei Schule-Freizeit-Konflikten?
Eine experimentelle Untersuchung

Methodische Schwächen bei dem Versuch
Dagmar Hänsels, die Rezeptionsgeschichte
eines „Schulbuchs“ nachzuzeichnen

Bildung als öffentliches Gut und das Problem
der Gerechtigkeit

Inhaltsverzeichnis

Thementeil: Digitale Medien in der Hochschullehre

Manuela Pietraß

Digitale Medien in der Hochschullehre – Einführung in den thematischen Schwerpunkt 307

Claudia de Witt

Kommunikation in Online-Lerngemeinschaften: Digitale Hochschullehre im Spiegel des Pragmatismus 312

Frank Fischer/Karsten Stegmann/Christof Wecker/Ingo Kollar

Online-Diskussionen in der Hochschullehre: Kooperationskripts können das fachliche Argumentieren verbessern 326

Manuel Pietraß

Digitale Präsenz – der didaktische Mehrwert der Mediengestaltung 338

Heidi Schelhowe

Interaktionsdesign: Wie werden Digitale Medien zu Bildungsmedien? 350

Heinz-Werner Wollersheim/Maren März/Jan Schminder

Digitale Prüfungsformate. Zum Wandel von Prüfungskultur und Prüfungspraxis in modularisierten Studiengängen 363

Deutscher Bildungsserver

Linktipps zum Thema „Digitale Medien in der Hochschullehre“ 375

Allgemeiner Teil

Manfred Hofer/Britta Kilian/Claudia Kuhnle/Sebastian Schmid

Beeinflussen individuelle Werte Motivation und Lernerfolg bei Schule-Freizeit-Konflikten? Eine experimentelle Untersuchung 385

Gerhard Eberle

Methodische Schwächen bei dem Versuch Dagmar Hänsels, die Rezeptionsgeschichte eines „Schulbuchs“ nachzuzeichnen 404

Johannes Giesinger

Bildung als öffentliches Gut und das Problem der Gerechtigkeit 421

Besprechungen

Walburga Hoff

Anne Schlüter (Hrsg.): Erziehungswissenschaftlerinnen in der Frauen- und Geschlechterforschung 438

Anna Lenski

Christiane Spiel/Barbara Schober/Petra Wagner/Ralph Reimann (Hrsg.): Bildungspsychologie 441

Heinz-Elmar Tenorth

Ulrich Binder: Das Subjekt der Pädagogik – Die Pädagogik des Subjekts. Das Subjektdenken der theoretischen und der praktischen Pädagogik im Spiegel ihrer Zeitschriften 443

Klaus Zierer

Martin Lehner: Allgemeine Didaktik 445

Dokumentation

Pädagogische Neuerscheinungen 448

Impressum U3

Table of Contents

Topic: Digital Media in Academic Teaching

Manuela Pietraß

Digital Media in Academic Teaching – An introduction 307

Claudia de Witt

Communication in Online Learning Communities: Digital teaching in higher education as reflected by pragmatism 312

Frank Fischer/Karsten Stegmann/Christof Wecker/Ingo Kollar

Online Discussions in Academic Teaching: Cooperation scripts as a means of improving specialist debates 326

Manuela Pietraß

Digital Presence – the didactic surplus value of media design 338

Heidi Schelhowe

Interaction Design: How can digital media be turned into educational media? 350

Heinz-Werner Wollersheim/Maren März/Jan Schminder

Digital Examination Formats. On the changes in the examination culture and examination practice in modular courses of studies 363

Deutscher Bildungsserver

Tips on links relating to the topic of „Digital Media in Academic Teaching“ 375

Beilagenhinweis: Dieser Ausgabe der Z.f.f.Päd. liegt ein Prospekt des Hogrefe Verlags, Göttingen, bei.

Contributions

Manfred Hofer/Britta Kilian/Claudia Kuhnle/Sebastian Schmid

Do individual Values Influence Motivation and Learning Success in case of
Conflicts between School and Leisure Time? An experimental investigation 385

Gerhard Eberle

Methodological Weaknesses in Dagmar Haensel's Attempt to Sketch the History
of the Reception of a „School Book“ 404

Johannes Giesinger

Education as a Public Good and the Problem of Justice 421

Book Reviews 438

New Books 448

Impressum U3

Mitteilung der Redaktion

Nach jahrzehntelanger Mitarbeit für die Zeitschrift für Pädagogik haben Herr Prof. Dr. Andreas Krapp und Herr Prof. Dr. Diether Hopf Ihren Abschied aus dem Herausgeberkreis bekanntgegeben. Die Redaktion bedankt sich im Namen aller Herausgeber und Herausgeberinnen herzlich bei Herrn Krapp und Herrn Hopf für das jahrelange Engagement, die stete Zuverlässigkeit und den besonderen Einsatz für die Zeitschrift für Pädagogik.

Heinz-Werner Wollersheim/Maren März/Jan Schminder

Digitale Prüfungsformate

Zum Wandel von Prüfungskultur und Prüfungspraxis in modularisierten Studiengängen

Zusammenfassung: Die neue Diskussion um standardisierte IT-gestützte Prüfungen an Hochschulen verweist auf grundsätzliche Fragen des Zusammenhangs von Lehre und Prüfung. In der Logik des Bologna-Prozesses liegt die Orientierung der Prüfung an einem in Kompetenzen beschriebenen Absolventenprofil. Vergleichend wird die Entwicklung im medizinischen Bereich herangezogen, wo die zu prüfenden Wissensbereiche nicht weniger komplex als im erziehungswissenschaftlichen Bereich sind. Der Spezialfall standardisierter Großklausuren wird näher untersucht.

1. Einleitung

Prüfungen, traditionell ein eher vernachlässigtes Thema in der hochschuldidaktischen Landschaft, stehen zurzeit weit oben auf der hochschulpolitischen Agenda. Gründe für diese Entwicklung sind unter anderem der starke quantitative Anstieg der Prüfungsfälle im Zuge der Modularisierung, die Veränderung des Stellenwertes der Prüfung vom einzelnen Leistungsnachweis zur komplexen Modulprüfung, die Veränderung der Bedeutung der Prüfung vom marginalisierbaren singulären Ereignis zum kumulativen Bestandteil des Abschlussexamens und die konsequente Umsetzung des Vieraugenprinzips in Prüfungen. In der Summe dieser Trends ist die Prüfungsbelastung erheblich gestiegen – bei Studierenden, wie auch bei Hochschullehrern. Entlastung sucht man einerseits in einer Verminderung der Prüfungsfälle (KMK, 2009), andererseits in einer Rationalisierung der Prüfungsabläufe durch Standardisierung und Automatisierung. In diesem Zusammenhang setzt man mancherorts große Hoffnung auf die Einführung digitaler Prüfungsformate, die auf den ersten Blick eine Maximierung der beabsichtigten Effizienzsteigerung über standardisierte IT-gestützte Prüfungsformate zu ermöglichen scheinen. Das populistische Zerrbild vom Verfall der hochkulturell und komplex stilisierten universitären Prüfung zur minderwertigen Ankreuzklausur am PC konturiert die Linie der Auseinandersetzung, in der sich Spengler und Taylor einmal mehr gegenüber zu stehen scheinen. Dabei gerät allzu leicht aus dem Blick, dass die Frage der Standardisierung von Prüfungen zunächst nichts mit der Frage des digitalen oder analogen Formats zu tun hat: „Auswahlantwortaufgaben“ – um eine prominente Erscheinungsform standardisierter Aufgaben als Beispiel heranzuziehen – kann man sowohl mit Papier und Bleistift als auch am PC bearbeiten lassen. Darüber hinaus zeigt eine Beschäftigung mit dem Thema „Prüfen“ sehr bald, dass die hochschuldidaktische Deduktionskette vor dem Hintergrund kompetenzorientierter Studienziele anders aufgebaut werden muss als in einem Verständnis von Universität, in dem die (senderseitige) „Vertretung“

eines Faches oder eines komplexen Wissensgebietes „in Forschung und Lehre“ betont wird. Während hier das Hauptgewicht auf der Frage liegt, was die disziplinäre Struktur eines Fachgebietes und seiner Wissensordnung ausmacht, liegt in kompetenzorientierter Perspektive das Hauptgewicht auf der Zielformulierung, was die Absolventen eines Moduls können sollen (Winter, 2009). Modularisierte, auf Kompetenzen abzielende Studiengänge erfordern somit die stärkere Didaktisierung der einzelnen Lehrveranstaltungen, und dies impliziert die Konstruktion von überprüfbaren Lernzielen in Form von Kompetenzbeschreibungen. Dies hat auf den ersten Blick nichts mit der Frage der Standardisierung von Prüfungen zu tun, denn die Kompetenzziele könnten auch individualisierend und formativ geprüft werden. Indirekt ist die Frage der Standardisierung allerdings naheliegend, doch weniger über Effizienzüberlegungen als über die Definition der Lern- und Prüfziele, die für eine Gruppe von Absolventen definiert werden und damit die Standards der Ausbildungsziele definieren. Prüfungen rücken daher vom End- zum Ausgangspunkt hochschuldidaktischer Überlegungen vor (Huber, 2008).

2. Grundsätzliche Fragen

Als eine Ad-hoc-Arbeitsgruppe der DGfE Anfang 2009 einen Bericht (Terhart, Baumgart, Meder & von Sychowski, 2009) zu standardisierten Prüfungsverfahren in der Erziehungswissenschaft vorlegte, ging sie von grundsätzlichen Fragen zum Status des erziehungswissenschaftlichen Wissens und seiner Prüfbarkeit aus, forderte den auch von uns skizzierten inhaltlichen Zusammenhang von Lernzielen einer Lehrveranstaltung und ihren Prüfungszielen ein und setzte sich auf dieser Grundlage mit Aufgabentypen auseinander. Umgehend erhielt sie eine Replik (Stojanov, 2009), deren emotionale Heftigkeit darauf schließen ließ, dass die Autoren der Arbeitsgruppe mit ihren prononcierten Thesen einen empfindlichen Nerv der aktuellen erziehungswissenschaftlichen Diskussion getroffen hatten. Stojanov (2009) bestritt sowohl die Berechtigung des arbeitsökonomischen Arguments als auch die Dignität der Standardisierung, weil er eine Verminderung der von ihm an sich als wertvoll gesetzten disziplinären Pluralität befürchtete.

Die Arbeitsgruppe hatte unter anderem standardisierte Prüfungsverfahren hinsichtlich bestehender Vorbehalte, zu stellender Anforderungen und der Analyse gängiger Aufgabentypen untersucht. Anschließend diskutierten die Autoren Möglichkeit und Nutzen einer computergestützten Auswertung, um sich abschließend mit dem Entwurf eines Kompetenzrasters und den in ihm abgebildeten Wissensformen zu beschäftigen. Als einen wesentlichen Vorbehalt thematisierten sie die verbreitete Sorge, dass es bei Einsatz standardisierter Prüfungsverfahren und insbesondere von Multiple-Choice-Aufgaben zu einer unzulässigen Reduktion von Komplexität des Prüfungswissens und seiner Präsentation kommen könnte. So wird verständlich, dass die IT-gestützte MC-Klausur zum Inbegriff des Horrors wird und den Furor der Kontrahenten beflügelt, obwohl beide Kernbestandteile der Auseinandersetzung: Standardisierung und Aufgabenformat, nichts mit der Digitalisierung des Prüfungsformates zu tun haben. IT-gestütztes Prüfen ruft nur in relativ geringem Umfang neue Fragestellungen hervor, sondern ver-

weist unweigerlich auf die Fragen, die allen Prüfungen *unabhängig von ihrem Format* voraus liegen: Was ist das Ziel der Prüfung? Wie zuverlässig ist ein bestimmtes Prüfungsformat zur Ermittlung der Prüfungsziele?

Die Beantwortung dieser Fragen hängt unter anderem davon ab, wie man ein Hochschulstudium und seine Ziele versteht: als akademische Enkulturation in ein universitäres Lehr-Gebiet mit spezifischen Strukturen, Riten und Traditionen oder als eine auf spezifische Kompetenzen zielende und einem Absolventenprofil verpflichtete Hochschulausbildung. In der Logik des Bologna-Prozesses liegt eine Abkehr vom „Enkulturationismus“.

Zurzeit besteht noch keine Einigung über Inhalt und konkrete Gestalt eines Kompetenzprofils für die Absolventen erziehungswissenschaftlicher Studiengänge. Dessen Entwicklung wäre lohnend: Die Lehre kann inhaltlich an Kompetenzziele ausgerichtet werden, und das *gradueller* Erreichen der Kompetenzziele strukturiert korrespondierende Prüfungen. Austausch und Auseinandersetzung über Inhalt und Gestalt des Kompetenzprofils können dazu beitragen, mehr Verbindlichkeit und Klarheit darüber zu schaffen, was man von den Absolventen erwarten darf. Nicht unberechtigt ist die Erwartung, dass dieser Prozess im Innenverhältnis einen wesentlichen Beitrag zur Qualitätsentwicklung und -kontrolle entsprechender Studiengänge und im Außenverhältnis zur Verdeutlichung der Leistungsfähigkeit der Disziplin leisten würde. In Erziehungswissenschaft und korrespondierend im Bereich der Lehrer(aus)bildung ist die Ausgangssituation allerdings besonders schwierig, weil es derzeit keine hinreichend klar strukturierte und vom Alltags- und Bildungswissen ausreichend abgegrenzte Wissensordnung gibt, die zur Formulierung eines Kompetenzprofils herangezogen werden könnte. Das Ziel einer kompetenzbasierten Beschreibung akademischen Lernens ist jedoch mit der Bologna-Reform gesetzt. Und die Folge ist klar: Die kompetenzorientierte Sicht verändert die Perspektive konsequent vom Primat der Lehre auf den Primat des Lernens.

Andere universitäre Fächer haben lange und umfangreiche Erfahrungen mit der Reformulierung dieses Lehr-Lern-Zusammenhangs. Im Bereich Medizin verlagert sich in den letzten zwanzig Jahren der Fokus in der Hochschulausbildung in den englischsprachigen Ländern weg vom „Lehren“ hin zum „Lernen“. Insbesondere im englischsprachigen Raum entwickelten sich lernerzentrierte Studiengänge, deren Module und Kurse sich in Inhalt und Form an definierten Kompetenzen ausrichten sollen (Rust, 2002). Neben der Schaffung einer adäquaten Lernumgebung spielt die Prüfung eine entscheidende Rolle. Sie muss den definierten Lernzielen inhaltlich entsprechen und eine geeignete Form aufweisen (Biggs, 1999).

Prüfungen haben eine sehr stark lernsteuernde Wirkung (Snyder, 1971; Miller & Parlett, 1974). Im medizinischen Bereich gilt der Grundsatz „Assessment drives learning“ (Gesellschaft für Medizinische Ausbildung, 2008). Im Umkehrschluss könnte daraus abgeleitet werden: Was nicht geprüft wird, wird auch nicht gelernt. In gewisser Weise lässt sich der Begriff „verstecktes Curriculum“ auch zur Kennzeichnung dieser Diskrepanz zwischen Gelehrtem und Gelerntem verwenden (Snyder, 1971). Studierende erfassen sehr schnell sowohl die inhaltlichen Anforderungen als auch die durch die Prüfung erfassten kognitiven Bereiche. Auch wenn Lehrende meinen, sie würden

von Ihren Studierenden eigenständiges Denken und Kreativität erwarten, stellen Studierende oft schnell fest, dass eigentlich Auswendiglernen gefordert ist oder zumindest ausreicht (Gibbs, 1992, S. 10). Die Gesellschaft für medizinische Ausbildung stellt daher in ihren Leitlinien für fakultätsinterne Prüfungen fest: „Zugespitzt formuliert sind alle curricularen Reformen und Innovationen wirkungslos, wenn das Curriculum nicht in einer qualitativ hochwertigen Prüfungskonzeption abgebildet wird“ (GMA, 2008, S. 2). Diese Prüfungskonzeption beinhaltet sowohl die summativen Leistungskontrollen als auch die formativen, den kontinuierlichen Lernprozess unterstützenden Prüfungen (Black & Wiliam, 1998; Ecclestone & Pryor, 2003).

Aufschlussreich ist der Blick über die Disziplingrenze vor allem, weil bei aller Unterschiedlichkeit der Wissenskulturen zumindest nicht bestritten werden kann, dass die Wissens- und Prüfungsgebiete mindestens ebenso komplex sind wie im pädagogischen Bereich. Zumindest in Lehramtsstudiengängen, aber auch in anderen stärker anwendungsorientierten Bereichen der Erziehungswissenschaft, kann man sich der berufsfeldspezifischen Definition von Kompetenzen kaum verschließen, so dass Medizin als mögliches Beispiel für die Modellierung solcher Kompetenzen interessant ist.

3. Digitale Prüfungsformate in der Medizin

Aufgrund der jahrzehntelangen Erfahrung in der Konstruktion, Durchführung und Evaluation standardisierter Prüfungen lässt sich am Beispiel der Medizin zeigen, welchen Qualitätssprung und welche Synergieeffekte die Einführung digital unterstützter Prüfungsformate bieten kann. Dies demonstriert eindrucksvoll das Projekt IMSm (Item Management System Medizin), das seit Juni 2006 von den medizinischen Fakultäten der Charité Berlin, der Universität Heidelberg und der LMU München betrieben wird.

Mit dieser Item-Datenbank kann man Prüfungsinhalte für alle im Staatsexamen Medizin zugelassenen schriftlichen, mündlichen und praktischen Prüfungsformen (Multiple-Choice, Longlist, M2, OSCE, Key-Feature-Fälle)¹ mit einem Web-Interface konzipieren und speichern. Aus der Datenbank kann man anschließend herkömmliche papiergebundene oder auch computergestützte Klausuren mit automatischer Auswertung generieren. Dadurch ist der Aufbau fakultätsinterner und -übergreifender Prüfungsverbünde möglich, die einen Itempool teilen und gleichzeitig ein Reviewnetzwerk für das vorgeschriebene Pre- und Postreview der Klausur darstellen. Die Benutzung der Datenbank bietet damit die Chance, die Prüfungsqualität ganz verschiedener Aufgabenformate bei besserer Nutzung von Ressourcen zu sichern. Das System fand schnell Verbreitung an den medizinischen Fakultäten: Im April 2009 beinhaltete die Datenbank bereits über 20.000 Prüfungsisems aus 12 teilnehmenden Hochschulen (Kompetenzzentrum für Prüfungen in der Medizin, o.J.).

Durch die gemeinsame Nutzung der Klassifizierung von Fragetypen und Prüfungsformen wird ein gemeinsamer Qualitätsmaßstab angelegt, der sich aus der akzeptierten

¹ Zur Erläuterung vgl. Medizinische Fakultät Heidelberg (o.J.).

Prüfungstheorie ergibt. Es werden also nicht Prüfungsinhalte standardisiert im Sinne von „kleinstem gemeinsamem Nenner“, sondern die Prüfungsfragen selbst sind auf die individuelle Lehre an den Hochschulen zugeschnitten und erhalten die Pluralität der Lehre im Fach, weisen aber keine formalen Fehler auf und benachteiligen Studenten nicht so, wie dies ohne einen gemeinsamen Qualitätsmaßstab möglich wäre.

Auch formatives Assessment kann durch Computereinsatz unterstützt werden. Im „Zentrum für virtuelle Patienten“ der Universität Heidelberg werden multimediale Patienten programmiert, mit denen der Medizinstudent wie in einem komplexen digitalen Rollenspiel interagieren muss, um seine Diagnose- und Therapiefähigkeiten zu trainieren, bevor er diese im Krankenhaus erprobt.

Prüfungsformen in den Geisteswissenschaften entbehren oft solcher „hands-on“-Kompetenztests. Hier findet die praktische Anwendung des Methoden- und Fachwissens eher in den Lernritualen eines Seminars statt: Diskussion, Vortrag und die abschließende Hausarbeit sind die tradierten Formen. Selbst das Klausurformat wird (besser: wurde vor Bologna) während des Studiums kaum angewandt, was manche Kandidaten bei der Zwischen- oder Abschlussprüfung vor nicht unerhebliche Schwierigkeiten stellte, weil sie dieses Prüfungsformat wenig trainiert hatten.

Ob nun aber ein Kompetenznachweis aus einer Hausarbeit, einem Vortrag oder auch einer Posterpräsentation besteht; man kann allgemein festhalten: Hier werden von den Lernenden individuelle oder kollektive Artefakte in Text oder Multimediaform geschaffen, die als Dokumentation von Lernprozessen auch Gegenstand von Beurteilungen sein können. Das legt nahe, einen Portfolioansatz bei dem Nachweis von Kompetenzen und Wissen in diesen Fächern einzuschlagen. Um hier Qualität zu gewährleisten, sollten die Werke von mehreren unabhängigen Prüfern und mithilfe eines elaborierten Bewertungsschlüssels oder eines gestuften Erwartungsschemas benotet werden. Auch ist der Nutzen von digitalen Formaten offensichtlich: E-Portfolios sind einfach zu nutzen und sind teilweise schon Bestandteile von gängigen Lern-Management-Systemen. Die Nutzung einer digitalen Lernumgebung unterstützt vor allem die Organisation von Gruppenaufgaben und Teamprojekten und fördert die bewusste Konstruktion komplexer Arbeitsprozesse besser als ihre analoge Entsprechung, vor allem weil sie individuelle Zeit- und Ortseinteilungen bei Lehrenden und Lernenden ermöglicht. Trotz dieser Reduktion von Aufwand durch digitale Hilfsmittel, erfordert jedes formative Assessment mehr Zeit als summative Assessments und die elaborierte Korrektur schriftlicher Hausarbeiten als komplexer Kompetenznachweis einer Modulabschlussprüfung ist bei heutigen Teilnehmerzahlen in Seminaren kaum zu schaffen. Denn ohne ausreichende Konsultation und ausführliches gemeinsames Auswertungsgespräch verfehlt auch die Hausarbeit ihren Zweck, wissenschaftliche Kompetenzen zu üben und zu demonstrieren. So drängt sich das Problem der Durchführbarkeit von Prüfungen ab einer kritischen Größe von Lehrveranstaltungen zunehmend in den Vordergrund und weicht Idealvorstellungen von Prüfungen schnell auf.

4. E-Klausuren in Massenveranstaltungen

In universitären Massenveranstaltungen sind standardisierte Klausuren oft der einzige Weg, die mehrdimensionalen Anforderungen (Objektivität, Reliabilität, Validität, Ökonomie) an eine qualitativ hochwertige Prüfung zu erfüllen. Vor dem Hintergrund der Tatsache, dass BA-Abschlüsse (und mithin Modulprüfungen als Teile des sequentiellen Examens) als berufsqualifizierende Abschlüsse gelten, werden an Universitätsprüfungen jetzt auch auf Modulebene die Anforderungen zu stellen sein, vor der Verfassungsnorm des Grundrechts auf Berufswahlfreiheit bestehen zu müssen. Auch deshalb kommt der Konstruktion qualitativ hochwertiger Prüfungen eine gesteigerte Bedeutung zu. Vor dem Hintergrund der bekannten Lern- und Prüfungsziel-Taxonomie: *Kennen, Verstehen, Anwenden und Beurteilen* (Kratwohl, Bloom & Masia, 1976; Guilbert, 1998), lassen sich zumindest auf den beiden Ebenen *Kennen* und *Verstehen* standardisierte Aufgaben konstruieren, die hinsichtlich der Gütekriterien Objektivität, Validität und Reliabilität empirisch hervorragend abgesichert sind. Dabei ist die Frage der Aufgabenkonstruktion zunächst keine Frage des Mediums: Sie betrifft die standardisierte papiergebundene Klausur ebenso wie die digitale Klausur. Auch hier ist ein Blick in die medizinische Prüfungspraxis aufschlussreich. Über einen langen Zeitraum wurden hier Erfahrungen mit standardisierten Aufgaben gesammelt und systematisch im Hinblick auf die genannten Gütekriterien ausgewertet (Möltner, Schellberg & Jünger, 2006). Inzwischen hat sich hier ein Forschungs- und Lehrgebiet etabliert, dessen Ergebnisse über die Grenzen der Disziplin hinaus Beachtung verdienen. Sie sind in Einzelbeiträgen gut dokumentiert und haben in einschlägigen Handbüchern zur Konstruktion und teststatistischen Bewertung von Prüfungen und Prüfungsfragen ihren Niederschlag gefunden (IAWF, 1999; NBME, 2001).

Ein wichtiges Ergebnis, das im Zusammenhang des Streits um die Berechtigung standardisierter Aufgaben zu sehen ist, besteht in der oft unterschätzten Bedeutung des Stimulusformates. Bekanntlich kann man schriftliche Prüfungen sowohl nach ihrem Stimulusformat als auch nach ihrem Antwortformat einteilen. Das Stimulusformat bezieht sich dabei auf die eigentliche Frage, das Antwortformat darauf, wie die Antworten des Prüflings erfasst werden (Cronbach, Gleser, Nanda & Rajaratnam, 1972; Norman, Swanson & Case, 1996). In den Versuchen, die schriftliche Prüfung von Kompetenzen zu verbessern, wurden zahlreiche Fragetypen entwickelt, die jedoch meist nach dem Antwortformat in „offene Fragen“ und „Multiple-Choice-Fragen“ (oder geschlossene Fragen) eingeteilt werden. Bei der Festlegung, welche Kompetenz gemessen wird, spielt das Stimulusformat, also die Art und Weise, wie gefragt wird, jedoch eine essentielle Rolle. So ist die Lösung für eine „Faktenfrage“ in einem Absatz im Lehrbuch oder auf einer Folie in der Vorlesung zu finden. Dieses Denken wird eher als eindimensional eingestuft – unabhängig davon, ob die Antwort selber generiert oder aus vorgegebenen sinnvollen Alternativen ausgewählt werden muss.

Faktenfragen sind, insbesondere in kleinteiligen Klausurformaten, häufig anzutreffen. Ihre Kurzhubigkeit und ihr mangelnder Anwendungsbezug macht sie bei Studierenden eher unbeliebt. Aufgrund ihrer schnellen und einfachen Konstruktion sind sie

bei Prüfern hingegen sehr beliebt. Wir verdeutlichen das Konstruktionsprinzip an einem Beispiel, das bei Lehramtsstudierenden hohes Interesse findet, am Themenkreis Unterrichtsstörungen und entsprechender Präventions- und Interventionsstrategien. In „traditioneller“ Weise lässt sich beispielsweise zur Trainingsraummethode nach Bründel und Simon (2007) folgende Multiple-Choice-Faktenfrage konstruieren:

Frage: Was ist die Grundidee der Trainingsraummethode nach Bründel und Simon (2007)?

Antwortmöglichkeiten (eine aus vier):

1. Die Schülerinnen und Schüler sollen sich ihres Störverhaltens bewusst werden, und dann eine Entscheidung treffen, ob sie im Klassenraum bleiben oder in den Trainingsraum gehen wollen, um mit dem dort anwesenden Lehrer einen Rückkehrplan zu erarbeiten.
2. Die Schülerinnen und Schüler sollen sich ihres Störverhaltens dadurch bewusst werden, dass der Lehrer die Regeln benennt, gegen die sie verstoßen haben und den Schüler zur Strafarbeit in den Trainingsraum schickt, in dem ein Sozialarbeiter die Aufsicht führt.
3. Die Schülerinnen und Schüler entschuldigen sich bei ihren Klassenkameraden; anderenfalls werden sie vom Klassenverband in den Trainingsraum geschickt und dürfen erst zurückkommen, wenn sie zusammen mit dem dort anwesenden Lehrer eine schriftliche Einsichtserklärung und Entschuldigung formuliert haben.
4. Die Schülerinnen und Schüler sollen sich in einem Trainingsraum zusammen mit dem dort anwesenden Lehrer durch körperliche Übungen „Luft verschaffen“, ihre Emotionen abreagieren und dann beruhigt und ausgeglichen in den Klassenverband zurückkehren.

Aufgaben dieses Typs bestehen aus dem Fragestamm und den vier Antwortmöglichkeiten. Mit solchen Aufgaben lassen sich Wissensbestände sicher und valide abprüfen. Eine Schwierigkeit, insbesondere bei dem hier vorliegenden Aufgabentyp „Eins aus vier“, besteht in der Konstruktion der Antwortmöglichkeiten, die eindeutig, plausibel und frei von versteckten Lösungshinweisen sein müssen.

Eine Alternative, die das Prüfungsniveau in Richtung „Verstehen“ verschiebt, besteht in der Konstruktion von Kontextaufgaben. Bei Kontextaufgaben besteht der Fragestamm aus einer längeren Vignette, welche die spezifischen Kontextinformationen enthält, und einer kurzen Frage. Die Antwortmöglichkeiten sollen sehr kurz und prägnant formuliert werden.

Durch die Variation der Vignette verändert sich die richtige Lösung. Im günstigen Fall entfällt die Konstruktion von unterschiedlichen Distraktoren. In unserem Beispiel haben wir als Antwortmöglichkeiten die fünf Prozessschritte vorgesehen, die sich in der Trainingsraummethode ergeben können. In Abhängigkeit von der Vignette ist jeweils eine andere Entscheidung richtig:

Prüfziel: Die Trainingsraummethode nach Bründel und Simon (2007) bei Unterrichtsstörungen anwenden können.

Beispiel Kontextfrage:

Vignette 1): Sie sind Lehrer an einer Gesamtschule in Berlin Mitte. Wegen der häufigen Klagen des Kollegiums über Disziplinprobleme wurde vor zwei Jahren die Arbeit mit der Trainingsraum-Methode nach Bründel und Simon (2007) im Schulkonzept verankert und eingeführt.

Es ist Dienstag, der 3. Mai, 5. Stunde in der 9c. Die Schüler arbeiten seit 15 Minuten in Gruppen an Plakaten zu einer Projektarbeit. Sie gehen herum und beraten die Schüler zur Plakatgestaltung. Die ruhige Arbeitsatmosphäre wird unterbrochen, als Hans beginnt, Zoltan am benachbarten Gruppenarbeitsplatz mit feuchten Papierkugeln zu bewerfen. Alle Schüler unterbrechen ihre Arbeit und schauen zu. Sie sind erstaunt, denn in der Regel kommen die beiden Schüler gut miteinander aus.

Vignette 2): Sie sind Lehrer an einer Gesamtschule in Berlin Mitte. Wegen der häufigen Klagen des Kollegiums über Disziplinprobleme wurde vor zwei Jahren die Arbeit mit der Trainingsraum-Methode nach Bründel und Simon (2007) im Schulkonzept verankert und eingeführt.

Es ist Dienstag, der 3. Mai, 5. Stunde in der 9c. Die Schüler arbeiten seit 25 Minuten in Gruppen an Plakaten zu einer Projektarbeit. Hans streut Zoltan den Inhalt seines Bleistiftspitzers in den Kragen. Zoltan schreit auf, die gesamte Klasse gerät in Unruhe. Sie fragen sich, was mit den beiden los ist, denn in dieser Stunde hat Hans wiederholt Zoltan geärgert. Und obgleich er davor sein Fehlverhalten erkannt und erklärt hat, die Klassenregeln zu beachten, geht es nun schon wieder los.

Vignette 3): Sie sind Lehrer an einer Gesamtschule in Berlin Mitte. Wegen der häufigen Klagen des Kollegiums über Disziplinprobleme wurde vor zwei Jahren die Arbeit mit der Trainingsraum-Methode nach Bründel und Simon (2007) im Schulkonzept verankert und eingeführt.

Heute haben Sie Dienst im Trainingsraum. Zu Ihnen kommt Hans aus der 9c. Er ist ziemlich aufgeregt. Laut dem Laufzettel der Klassenlehrerin hat er wiederholt den Unterricht gestört, weil er einen Mitschüler mit feuchten Papierkugeln beworfen hat. Nachdem er sich beruhigt hat, beginnt Hans Ihnen von dem Vorfall aus seiner Sicht zu erzählen.

Frage: Welcher Prozessschritt der Trainingsraummethode wird in dieser Situation angewendet?

Antworten:

1. Hans fragen, ob er in den Trainingsraum gehen will
2. Hans in den Trainingsraum schicken
3. Hans einen Rückkehrplan erarbeiten lassen
4. Hans nach Hause schicken
5. Hans und seine Eltern zum Gespräch bestellen

Bei Anwendungsfragen wird eine Situation beschrieben. Zur Lösung der Frage müssen die in der Situation gegebenen Fakten innerhalb des gegebenen Kontextes miteinander in Beziehung gesetzt und ausgewertet werden. Die Lösung ist nicht in einem Absatz des Lehrbuches zu finden. Dieses Denken wird als vektoriell eingestuft – und zwar unabhängig vom Antwortformat (Schuwirth & van der Vleuten, 2004; Schuwirth, Verheggen, van der Vleuten, Boshuizen & Dinant, 2001; Norman et al., 1996). Im Umkehrschluss bedeutet das, dass bei einem sorgfältig elaborierten Stimulusformat auch Antwortformate verwendet werden können, die maschinell vollständig auswertbar sind.

Das Vignettenformat hat zudem den Vorteil, dass man durch geringfügige Änderung der Vignette viele Aufgabenvarianten zum gleichen Prüfziel generieren kann und die Frage in seinen Varianten häufiger einsetzen kann, als dies bei Faktenfragen der Fall wäre – auch weil man nicht gezwungen ist, wahr klingende Distraktoren zu erfinden.

Module umfassen in der Regel mehrere Einzellehrveranstaltungen. Modulprüfungen sind daher grundsätzlich Komplexprüfungen und unterscheiden sich dadurch wesentlich von traditionellen Abschlussprüfungen von Einzellehrveranstaltungen. Werden sie als Komplexprüfungen angelegt, so ist vorher festzulegen, wie der Blueprint der Prüfung aussehen soll: Aus wie vielen und wie gearteten Fragen soll die gesamte Prüfung bestehen und welcher Anteil soll auf welches in der Prüfung vertretene Wissensgebiet entfallen. Die Komplexprüfung impliziert, dass mehrere Aufgabensteller an der Formulierung der Gesamtprüfung mitwirken. Damit wird automatisch eine wesentliche Forderung der Rechtsprechung erfüllt, wonach bei Multiple-Choice-Klausuren die typische Prüfertätigkeit in die Aufgabenstellung vorverlagert ist und die Wahrung des Zweiprüferprinzips deshalb durch kollaborative Aufgabenerstellung sichergestellt werden muss (OVG Münster, 2006). Einen wesentlichen Beitrag zur Qualitätsentwicklung erreicht man, wenn die von den einzelnen Autoren erstellten Aufgaben einem gemeinsamen Review unterzogen werden. Dieses Aufgabenreview schafft Transparenz und sichert die Standards, die im Kollegium für die Formulierung der Prüfungsaufgaben festgelegt werden. Insbesondere dieser Kommunikationsprozess fehlte bislang häufig in traditionellen Hochschulprüfungen, wodurch subjektivistische Prüferstrategien begünstigt wurden. Bezogen auf die Konstruktion von MC-Prüfungsfragen haben sich im Review folgende einfache Regeln bewährt: 1. Kann die Frage auch beantwortet werden, wenn die Antwortalternativen abgedeckt sind? 2. Bildet die Frage ein einziges Lernziel ab? 3. Ist die Frage positiv formuliert? 4. Enthält die Fragestellung verdeckte Lösungshinweise? 5. Ist die Frage der Zielgruppe angemessen? 6. Ist die Frage relevant? Regel eins schließt Frageformate aus, die dem Muster: „Welche der nachfolgenden Aussagen sind

richtig?“ folgen. Zusammen mit der zweiten Regel wird die Entwicklung fokussierter Fragen gestützt und die inhaltliche Beziehung zwischen Lernstoff und Prüfungsfrage gesichert. Regel drei bezieht sich auf die Vermeidung nicht prüfungsrelevanter Schwierigkeiten: Die kognitive Auflösung der Verneinung in einer Frage testet eher sprachlich-kognitive Fähigkeiten als das inhaltliche Lernziel. Die vierte Regel bezieht sich auf die *testwiseness* der Probanden und muss hier nicht weiter erläutert werden. Diese vier Regeln beziehen sich auf formale Kriterien, die letzten beiden vor allem auf die Zuverlässigkeit (Regel 5) und Gültigkeit (Regel 6) der Prüfungsfrage. Transparente Lern- und korrespondierende Prüfziele erleichtern die Beantwortung der letzten beiden Fragen. Dieses Vorgehen der Erstellung eines Blueprints anhand der Lernziele, sorgfältige Konstruktion der Fragen und Qualitätssicherung durch ein formales und inhaltliches Review wird auch als direkte Validierung bezeichnet (Schuwirth et al., 2001).

Die Aufgabenkonstruktion für standardisierte Prüfungen ist aufwendig. Der Aufwand für die Erstellung einer einzelnen Prüfung kann durch die Bildung von Fragenpools und die gemeinsame Nutzung von Prüfungsdatenbanken erheblich reduziert werden; die automatisierte Auswertung wirkt sich ebenfalls günstig auf die Ressourcenbilanz aus.

5. Ausblick

Es gibt eine Fülle von Möglichkeiten, die Fähigkeiten von Studierenden auch digital unterstützt zu bewerten, sowohl summativ als auch formativ. Entscheidend für die Qualität der Prüfung – unabhängig von der Frage digital oder analog – sind Planung, inhaltliche Gestaltung und prozessuale Durchführung der Prüfung.

Wichtig sind bei der Auswahl der Prüfungsformen das Prüfungsziel und die vorhandenen Ressourcen sowie Einstellungen von Prüflingen und Prüfern zur Durchführung der Prüfungen aufeinander abzustimmen. Prüfungen, die zwar gut durchführbar sind, aber keinem konkreten, hochschuldidaktisch begründeten Prüfungs-, Lern- und Ausbildungsziel dienen, sind genauso problematisch wie Prüfungen, die zwar ein elaboriertes und hochschuldidaktisch klares Lern- und Prüfungsziel verfolgen, aber mangels Ressourcen oder fundamentaler kultureller Widerstände im Fach nur unvollkommen oder in Teilen durchgeführt werden können, und damit ihren Zweck verfehlen.

Die höheren Kosten, die mit anspruchsvollen digitalen Prüfungsformaten verbunden sind, lassen sich zum Teil durch geschickte hochschuldidaktische Arrangements verringern und gleichzeitig mit einem „Ausbildungs-Mehrwert“ aufladen: Mehrheitlich geht man in der Diskussion unausgesprochen davon aus, dass die Korrektur von Prüfungsleistungen entweder maschinell oder durch Hochschullehrer geschieht, also ein Teacher-Assessment stattfindet. Die Entwicklung von Peer- und Self-Assessments, sei es zur Information über den eigenen Leistungsstand, zur Abbildung des Lernzuwachses über größere Abschnitte des Studiums oder die produktive Auseinandersetzung mit Beiträgen aus der Peergroup trägt zur Ausbildung höherer Kompetenzniveaus bei. Aus studentischer Initiative heraus wurde beispielsweise in der Berliner Charité der Progress-

Test etabliert – ein Test, der spontan verfügbares Wissen testet. Dieser formative Test findet zu Beginn eines jeden Semesters statt und ist für alle Studierenden aller Semester gleich. Die Ergebnisse geben sowohl Aufschluss über den individuellen longitudinalen Wissenszuwachs über das Studium als auch über die individuelle Leistung im Vergleich zur Kohorte.

Die Entwicklung solcher innovativer, intelligenter und effizienter Prüfungen können durch digitale Prüfungsformate ungleich wirkungsvoller unterstützt werden als dies in papiergebundener Form möglich wäre.

Literatur

- Biggs, J. (1999). What the Student Does: teaching for enhances learning. *Higher Education Research & Development*, 18(1), 57-75.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Bründel, H., & Simon, E. (2007). *Die Trainingsraummethode* (2. Aufl.). Weinheim/Basel: Beltz Verlag.
- Cronbach, L., Gleser, G., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Generalizability for scores and profiles*. New York: Wiley.
- Ecclestone K., & Pryor, J. (2003). 'Learning Careers' or 'Assessment Careers'? The Impact of Assessment Systems on Learning. *British Educational Research Journal*, 29(4), 473-488.
- Gesellschaft für Medizinische Ausbildung, GMA-Ausschuss Prüfungen & Kompetenzzentrum Prüfungen Baden-Württemberg & Fischer M. (Korrespondenzautor) (2008). Leitlinie für Fakultäts-interne Leistungsnachweise während des Medizinstudiums: Ein Positionspapier des GMA-Ausschusses Prüfungen und des Kompetenzzentrums Prüfungen Baden-Württemberg. *GMS Z Med Ausbild*, 25(1), Doc74.
- Gibbs, G. (1992). *Improving the Quality of Student Learning*. Bristol: TES.
- Guilbert, J. J. (1998). *Educational handbook for health personnel. WHO Offset Publication No. 35* (6th ed. revised and updated). Geneva: World Health Organisation.
- Huber, L. (2008). „Kompetenzen“ prüfen? In S. Dany, B. Szczyrba & J. Wildt, (Hrsg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 12-26). Bielefeld: W. Bertelsmann Verlag.
- IAWF (1999). *Institut für Aus-, Weiter- und Fortbildung, Medizinische Fakultät Universität Bern: Kompetent prüfen. Handbuch zur Planung, Durchführung und Auswertung von Facharztprüfungen*. Bern: o.A.
- KMK (2009). *Die Beschlüsse der KMK zu Prüfungen auf der 328. Plenarsitzung am 10. Dezember 2009*.
<http://www.kmk.org/presse-und-aktuelles/meldung/ergebnisse-der-328-plenarsitzung-der-kultusministerkonferenz-am-10-dezember-2009.html> [23.11.2010].
- Kompetenzzentrum für Prüfungen in der Medizin (o.J.). *Item Management System (IMS)*. https://www.ims.de/joomla/index.php?option=com_content&view=article&id=46%3Amehr-als-20000-items&catid=19%3Anews&Itemid=44&lang=de [23.11.2010].
- Kratwohl, D. R., Bloom, B. S., & Masia, B. B. (1975). *Taxonomie von Lernzielen im affektiven Bereich*. Weinheim/Basel: Beltz Verlag.
- Medizinische Fakultät Heidelberg (o.J.). Kompetenzzentrum für Prüfungen in der Medizin/Baden-Württemberg – Komp Med.
<http://www.medizinische-fakultaet-hd.uni-heidelberg.de/Kompetenzzentrum-fuer-Pruefungen-Komp-Med.3081.0.html> [23.11.2010].

- Miller, C. M. I., & Parlett, M. (1974). *Up to the Mark: a study of the examination game*. Guildford: Society for Research into Higher Education.
- Möltner, A., Schellberg, D., & Jünger, J. (2006). Grundlegende quantitative Analysen medizinischer Prüfungen. *GMS Z Med Ausbild.*, 23(3).
- NBME National Board of Medical Examiners (2001). *Constructing Written Test Questions for the Basic and Clinical Sciences*. Philadelphia: NBME.
- Norman, G. R., Swanson, D. B., & Case, S. M. (1996). Conceptual and methodological issues in studies comparing assessment formats. *Teaching and Learning in Medicine*, 8(4), 208-216.
- OVG Münster (2006). *Beschluss vom 04.10.2006*. 14 B 1035/06.
- Rust, C. (2002). The impact of assessment on student learning – how can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices. *Active Learning in higher education*, 3(2), 145-158.
- Schuwirth, L. W. T., Verheggen, M. M., van der Vleuten, C. P. M., Boshuizen, H. P., & Dinant, G. J. (2001). Do short cases elicit different thinking processes than factual knowledge questions do? *Med Educ.*, 35(4), 348-56.
- Schuwirth, L. W. T., & van der Vleuten, C. P. M. (2004). Different written assessment methods: what can be said about their strengths and weaknesses? *Medical Education*, 38(9), 974-979.
- Snyder, B. R. (1971). *The Hidden Curriculum*. Cambridge, MA: MIT Press.
- Stojanov, K. (2009). Standardisierte Prüfungsverfahren und der Status erziehungswissenschaftlichen Grundlagenwissens. Replik auf Terhart, Baumgart, Meder und von Sychowski. *Erziehungswissenschaft*, 20(39), 9-16.
- Terhart E., Baumgart, F., Meder, N., & von Sychowski, G. (2009). *Standardisierte Prüfungsverfahren in der Erziehungswissenschaft: Kontext, Formen, Konsequenzen. Ergebnis der ad-hoc-Arbeitsgruppe*
http://www.dgfe.de/news/news_item.2009-01-09.6376095549 [23.11.2010].
- Winter, M. (2009). *Das neue Studieren. Chancen Risiken Nebenwirkungen der Studienstrukturreform; Zwischenbilanz zum Bologna-Prozess in Deutschland*. Wittenberg: HoF Verlag.

Abstract: The recent debate on standardized IT-supported examinations at universities refers to fundamental questions concerning the relation between teaching and examining. It is within the logic of the Bologna Process that examinations are oriented by a graduate profile described through competences. The development within the medical field is used as a comparative example, – the areas of knowledge to be examined being hardly any less complex than those of the educational-scientific field. The special case of standardized comprehensive written examinations is investigated more closely.

Anschrift der Autorin/der Autoren

Prof. Dr. Heinz-Werner Wollersheim, Universität Leipzig, Erziehungswissenschaftliche Fakultät, Allgemeine Pädagogik, Jahnallee 59, 04109 Leipzig, Deutschland
 E-Mail: wollersheim@uni-leipzig.de

Dr. Maren März, Dieter Scheffner Fachzentrum für medizinische Hochschullehre und evidenzbasierte Ausbildungsforschung, Assessment-Bereich, Charité Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Deutschland
 E-Mail: maren.maerz@charite.de

Jan Schminder M. A., Universität Leipzig, Erziehungswissenschaftliche Fakultät, Allgemeine Pädagogik, Jahnallee 59, 04109 Leipzig, Deutschland
 E-Mail: jschmind@uni-leipzig.de