

Standardisierte Prüfungsverfahren in der Erziehungswissenschaft: Kontext, Formen, Konsequenzen¹

1. Einleitung und Problembeschreibung

Durch die im Zuge des Bologna-Prozesses erfolgende gleichzeitige Einführung von Bachelor-Master-Struktur einerseits und der durchgehenden Modularisierung dieser neuen Studiengänge andererseits hat sich die Situation in der akademischen Lehre an den deutschen Hochschulen grundlegend verändert. Die neuen Studiengänge sind nicht mehr nach zu bearbeitenden Themenbereichen gegliedert, sondern sollen an vorab klar auszuweisenden, zu vermittelnden und zu prüfenden Kompetenzen ausgerichtet sein. Durch die Sequenzierung von Modulen und durch die Kumulation der jeweils erworbenen Leistungspunkte sollen am Ende bei den Absolventen die angestrebten komplexen Fähigkeiten angebahnt und möglichst weit entwickelt werden. Diese neue Form des Studierens wird kontinuierlich von Prüfungen begleitet; Studium und Prüfung sind in zeitlicher und sachlicher Hinsicht von Beginn an sehr eng miteinander verkoppelt. Im Alltag von Lehrenden und Lernenden nimmt kontinuierliches Prüfen und Geprüftwerden einen quantitativ und qualitativ sehr viel höheren Stellenwert ein, als dies in der traditionellen Studienorganisation der Fall gewesen ist. Umgekehrt nimmt die inhaltliche und psychologische Bedeutung der großen Abschlussprüfung als dramatischem Endpunkt eines langen, kaum formell vorstrukturierten Studienverlaufs sehr stark ab. Denn heute wird für jeden Absolventen am Ende eines kontinuierlichen Prozesses der Kumulation von Einzelleistungen eine differenzierte Übersicht über seine gesammelten Studienelemente (inklusive der Abschlussarbeit) erstellt, alle Bewertungen werden zur Gesamtnote zusammengeführt etc. – das Studium klingt gewissermaßen aus.

Insbesondere für die geistes- und sozialwissenschaftlichen Fächer bedeutet dies eine vergleichsweise radikale Abkehr von bisherigen akademischen Lehr- und Lerngewohnheiten. Für Lehrende wie Lernende entsteht ein sehr viel höheres Maß an Verbindlichkeit: Die Lehrangebote müssen in ausreichender Zahl sowie in einem genau auf die Studiengänge

¹ Auf Initiative von Ewald Terhart hat der Vorstand der DGfE im Herbst 2007 eine ad hoc-Arbeitsgruppe eingerichtet, die sich mit Fragen der Prüfungspraxis in modularisierten Studiengängen befassen sollte. Als Mitglied des (damaligen) Vorstandes hat Norbert Meder mitgearbeitet. Die Arbeitsgruppe mit Franzjörg Baumgart und Gaja von Sychowski als weiteren Mitgliedern legt hiermit ihr Ergebnis vor; thematisch hat sich die Arbeitsgruppe primär mit standardisierten Prüfungsverfahren im Kontext großer Lehrveranstaltungen (Vorlesungen) befasst. Sie bedankt sich bei Klaus Breuer, Heiner Drerup, Klaus-Peter Horn, Rudolf Tippelt, Klaus-Jürgen Tillmann und Peter Vogel, die Vorfassungen des vorliegenden Papiers kritisch und konstruktiv kommentiert haben. – Die Arbeitsgruppe ist sich der Vorläufigkeit, der thematischen Begrenztheit und des Versuchscharakters ihrer Überlegungen und Argumente bewusst.

Eine kürzere Fassung dieses Textes erscheint in ERZIEHUNGSWISSENSCHAFT 2009.

bzw. die studentische Nachfrage abgestimmten inhaltlichen und zeitlichen Muster angeboten werden, Studienleistungen sind sehr schnell zu bewerten, damit die Möglichkeit der Wiederholung besteht und früh genug klar wird, ob und wie ein Studierender sein Studium im nächsten Semester weiterführen kann, Lehrende wie Lernende müssen sich in den Anforderungen der neuen Studienstruktur zurecht finden etc. Vor allem aber führt dieses neue Regime zu einer Explosion der Lehr- und Prüfungsnachfrage in mehrfacher Hinsicht:

- Traditionell wurde in den geistes- und sozialwissenschaftlichen Studiengängen der Anteil der notwendigen Lehrveranstaltungen durch die Angabe von Semesterwochenstunden (SWS) definiert. Ein beachtlicher Teil dieser SWS wurde jedoch durch bloßes Belegen abgedeckt: Studierende trugen diese Veranstaltungen in die Belegbögen ein. Eine Prüfung der tatsächlichen Teilnahme erfolgte nicht. Obwohl sehr viel mehr belegt als tatsächlich besucht wurde, kam es in manchen thematischen Bereichen zu überfüllten Veranstaltungen. Wenn im neuen Studienregime nun aber der Erwerb von Leistungspunkten auf breiter Front an die Präsenz in Seminaren gebunden wird, sind nun plötzlich alle Studierenden „auch wirklich und immer da“! Ein System, das bereits unter den alten Bedingungen an vielen Stellen überfüllt war, muss unter den neuen Bedingungen in allergrößte, wenn nicht unüberwindbare Schwierigkeiten kommen.
- Die Zahl der zu erwerbenden Leistungsnachweise – seien sie benotet oder unbenotet – lag in den herkömmlichen Studiengängen sehr deutlich unterhalb der formalen Gesamtzahl der zu besuchenden Veranstaltungen. Heute dagegen absolvieren alle Studierenden tatsächlich alle vorgeschriebenen Veranstaltungen, werden dort geprüft und erzeugen damit einen sprunghaften Anstieg der Zahl der Prüfungssituationen, die zudem alle präzise verwaltet und dokumentiert werden müssen. Allerdings: Nur scheinbar erzeugt Modularisierung diese Probleme; faktisch macht sie bestimmte Defizite der traditionellen akademischen Studienorganisation allererst sichtbar und geht sie zugleich in einer konstruktiven, an der Idee der Erzeugbarkeit von Fähigkeiten orientierten Form an. Damit werden eingeschliffene, nicht zuletzt für Lehrende wie Lernende sicherlich hier und da auch bequeme Routinen aufgebrochen. Aber wie dem auch sei: Durch Modularisierung ändert sich das tradierte Selbstverständnis akademischer Lehre jedenfalls in den Geistes- und Sozialwissenschaften grundsätzlich.
- Die Abbruchquote in vielen sozial- und geisteswissenschaftlichen Studiengängen (insbesondere in solchen, die zum Magister führten) war traditionell sehr hoch. Dies führte dazu, dass zwar sehr viele Studierende eingeschrieben waren, der weit überwiegende Teil der Anfänger das Studium jedoch aus unterschiedlichsten Gründen abbrach und insofern auch nie zum Prüfungsfall bei der großen Abschlussprüfung wurde. Die Zahl der potentiellen Kandidaten für die Abschlussprüfung sank im Studienverlauf kontinuierlich. Die neuen Studienstrukturen scheinen jedoch die Abbrecherquote zu senken; auch aus diesem Grunde steigt die Zahl der Prüfungsfälle. Allerdings liegen hierzu unterschiedliche Erfahrungen an den verschiedenen Universitäten vor – eindeutig ist die Tendenz eines Rückgangs der Abbrecherquote nicht. Ebenso muss auf einen Faktor hingewiesen werden, der unabhängig vom Bologna-Prozess gegenwärtig die Prüfungsdichte erhöht: Die breite Einführung von Studiengebühren motiviert die gegenwärtige Studierendengeneration ganz generell zu einem rascheren Ansteuern des Abschlusses.

Wie zu erwarten, hat die Einführung des neuen Studienstrukturen Widerstände erzeugt und mittlerweile zu manchen Enttäuschungen auch bei zunächst Gutwilligen bzw. Interessierten geführt. Die Probleme sind vielfältiger Art: Bereits der Umstellungsprozess selbst verschlingt sehr viele Energien, und im Ergebnis entsteht eine Situation, die zu einer ungewohnten Arbeitsform *und* einer erhöhten Arbeitsbelastung im Bereich von Lehre, Prüfung und Prüfungsverwaltung führt. Zugleich müssen die Studierenden in den auslaufenden (herkömmlichen) Studiengängen zum Abschluss geführt werden. In manchen Bundesländern interferiert der Bologna-Prozess darüber hinaus mit der staatlichen Verordnung neuer Lehramtsprüfungsordnungen, die nicht oder nur teilweise oder in inkohärenter Weise auf Bologna eingestellt sind – ein ständiger Konfliktherd, der zuverlässig zu kontinuierlicher Frustration bei allen Beteiligten führt. Und schließlich: Weil die Umstellung auf konsequente, modularisierte Studienstrukturen forciert betrieben, ja z. T. erzwungen worden ist, zugleich aber eine tatsächliche inhaltliche Abstimmung unterblieb, wurden vielfach ortsspezifische, z. T. sogar innerhalb einer Universität fakultätsspezifische Speziallösungen gefunden, die die Mobilität der Studierenden zwischen den Fächern und zwischen den Hochschulen einschränken.

Diese aus dem Zusammentreffen verschiedener Faktoren resultierende Verdichtung und Verkomplizierung der akademischen Lehre verleitet manche Universitätsangehörige zu einer Verklärung der früheren Zeiten – eine durchaus nachvollziehbare, verständliche Reaktion. Gleichwohl ist dagegen zu halten: Die ‚alte Welt‘ war keineswegs ideal, genau so übrigens, wie es die ‚neue Welt‘ nicht sein wird! Die früheren, sehr offenen Studienstrukturen sind von einem sehr großen Teil der Studierenden nicht bewältigt worden; bestimmte Studierendengruppen wurden dadurch de facto via Selbstselektion gezielt in die Selbsteliminierung geführt. Der Lehrbetrieb war sehr stark an die individuellen Interessen und Motive einzelner Lehrender geknüpft; die Prüfungspraxis war ebenfalls sehr personengebunden, schwer überprüfbar und bewegte sich – nicht zuletzt deshalb – vielfach in der Fiktion von Objektivität und Qualität. Die Studierbarkeit von Studiengängen bestand weithin nur auf dem Papier; die Probleme des praktischen Studienbetriebs wurden auf die Studierenden abgewälzt, durch extrem großzügige Anerkennungs- und Nachbesserungspraxis seitens der Lehrenden verdeckt etc. Allgemeiner betrachtet: Die Kombination von einem an Humboldt orientierten akademischen Selbstverständnis einerseits mit der objektiv gegebenen Situation der chronisch unterfinanzierten Massenuniversität andererseits führte die Lehrenden individuell wie die Universität institutionell in ausweglose Situationen. Diejenigen, die dieses System erfolgreich durchlaufen haben – und die gegenwärtig Lehrenden gehören offensichtlich dazu – können aus ihrer persönlichen Erfahrung heraus nicht ohne Weiteres für das traditionelle System und seine Folgen *insgesamt* sprechen; die in ihm Gescheiterten kommen nicht bzw. melden sich nicht mehr zu Wort. Jedenfalls sollte bei allem Befremden und Ärger über Modularisierung und ihre Folgen der Blick zurück die frühere Situation nicht unkritisch verklären.

Im Folgenden soll kein Beitrag zur grundsätzlichen Debatte über den Sinn und Unsinn der neuen Studienstrukturen geleistet werden; es geht nicht um den akademischen Kulturkampf Humboldt gegen Bologna. Es geht vielmehr darum, angesichts der sowohl in den im

engeren Sinne erziehungswissenschaftlichen Studiengängen als auch in der Lehrerbildung faktisch sehr weit vorangetriebenen Umstellung auf Bachelor, Master und Modularisierung einige Überlegungen zu den Konsequenzen für die Lehre sowie das Prüfen in Erziehungswissenschaft, *insbesondere: für die Durchführung von standardisierten Klausuren in Verbindung mit Vorlesungen* anzustellen. Für die Erziehungswissenschaft wird die Frage der Organisation und Bewältigung der quantitativ gewachsenen und qualitativ gewandelten Prüfungsaufgaben nämlich besonders virulent.² Die Erziehungswissenschaft ist in den letzten Jahrzehnten eine der großen Universitätsdisziplinen geworden und ist es immer noch. Der disziplineigene Diplom-Studiengang hatte sich gut etabliert und ausdifferenziert. Durch ihre Beteiligung an der Lehrerbildung ist sie mit einer extrem hohen Nachfrage nach Lehrveranstaltungen konfrontiert, da alle Lehramtsstudierenden einer Universität Erziehungswissenschaft studieren müssen. Die Erarbeitung und Verabschiedung von Kerncurricula für die erziehungswissenschaftlichen Studiengänge und Studiengangsanteile (seitens der DGfE und vor Ort an zahlreichen Universitäten) soll zu einer klareren, disziplinweit besser abgestimmten Orientierung für die Studiengangsplanung bzw. Angebotsstruktur in der Lehre führen – was wiederum Konsequenzen für die Wissensstruktur der Disziplin selbst hat (s. Abschnitt 2). Die Personalausstattung ist allerdings angespannt bzw. unzureichend, wenn man daran denkt, dass die neuen Studienstrukturen zu einer Explosion der Lehrnachfrage und der Prüfungsfälle führen (s. Abschnitt 6). Modularisierte Studiengänge benötigen aus diesem Grunde eine Neufestsetzung des Curricularnormwertes, auf dessen Basis die Relation zwischen Lehrkapazität und Studierendenzahl festgelegt wird. Selbst dann, wenn angesichts dieser Bedingungen die Personalsituation der Erziehungswissenschaft verbessert wird, umgekehrt die Studierendenzahl abgesenkt wird oder schließlich eine Kombination von Beidem eintritt, sind im Lehrbetrieb einer derart großen Disziplin wie der Erziehungswissenschaft vermehrt *große, mehr oder weniger standardisierte Vorlesungen* notwendig. Dies gilt insbesondere für Einführungs- und Überblicksvorlesungen zu wichtigen Modulbereichen. Solche Vorlesungen können, ja sollten groß sein; dementsprechend kann und muss die Größe der parallel geführten bzw. daran anschließenden spezielleren Seminare klein sein. In solchen großen Vorlesungen wird die Zertifizierung von Leistungen durch Leistungspunkte und Noten in der Regel nur in Gestalt von Klausuren möglich sein.³ Das bedeutet: Die Erziehungswissenschaft muss Formen des Lehrens und Prüfens finden, die für andere große Universitätsdisziplinen seit Jahrzehnten selbstverständlich sind.

² Zur Funktion und Problematik von Hochschulprüfungen, aber auch für praktische Anregungen vgl. Müller, F.H./ Bayer, Chr.: Prüfungen: Vorbereitung – Durchführung – Bewertung. In: Hawelka, B. u.a. (Hrsg.): Förderung von Kompetenzen in der Hochschullehre. Kröning: Asanger 2007, S. 223-237. Zur tiefen- und sozialpsychologischen Betrachtung der Prüfung vgl. Strauss, B.: Die Psychologie des Prüfens und Geprüft-Werdens. In: Kodalle, K.-M. (Hrsg.): Der geprüfte Mensch. (Kritisches Jahrbuch der Philosophie 2005). Würzburg: Königshausen & Neumann 2006, S. 45-55. Zur empirischen Untersuchung der Kommunikation in mündlichen Prüfungen vgl. Meer, D.: Der Prüfer ist nicht der König. Mündliche Abschlussprüfungen in der Hochschule. Tübingen: Niemeyer 1998.

³ Leistungsfeststellungen im Kontext großer Vorlesungen sind im Prinzip auch anders möglich: begleitende Tutorien mit Zusatzaufgaben, Protokolle, Hausarbeiten, kurze mündliche Prüfungen etc. Hiermit wird ebenfalls an verschiedenen Universitäten experimentiert. Allerdings ist der Aufwand sehr viel höher als bei standardisierten Klausuren.

Wir konzentrieren uns im Folgenden im Wesentlichen auf das Problem des Prüfens von erziehungswissenschaftlichem Wissen durch Klausuren. *Damit soll das breite Spektrum anderer Prüfungsformate keineswegs ignoriert oder in seiner Bedeutung geschmälert werden.* Mündliche Prüfungen, Fallanalysen, Praktikumsbericht, Referate, Studienprojekte, Lern- und Forschungswerkstätten, Lerntagebücher, forschendes Lernen etc. – alle diese Elemente gehören zu einem Studium. Aus diesem Grunde legen die neuen Studien- und Prüfungsordnungen in aller Regel auch fest, dass im Verlaufe des gesamten Studiums die Breite der verschiedenen Prüfungsformate von den Studierenden erfahren wird. In aller Regel sind Überblicksvorlesungen am Beginn des Studiums sinnvoll bzw. immer dort, wo bestimmte Grundlagen vermittelt werden müssen. Im Verlauf des Studiums sollten dann zunehmend solche Lehr- und Prüfungsformen eingesetzt werden, die – auf diesem Wissen aufbauend – geeignet sind, weiter bzw. tiefer gehende Fähigkeiten zu vermitteln und zu bewerten. Für die Lehrenden bedeutet dies, diese breite Palette an Prüfungsformen auch in hinreichender Zahl und angemessenem Zeitrhythmus anzubieten.

Die Konzentration auf (standardisierte) Klausuren erfolgt aus zwei Gründen: Erstens ist dieses Prüfungsformat aus verschiedenen Gründen in der Erziehungswissenschaft traditionell bislang eher unüblich, so dass hier eine neue Praxis entstehen muss, über die man sich fachintern verständigen sollte. Zweitens kommt den Einführungs- und Überblicksvorlesungen inklusive der damit verbundenen Prüfungspraxis eine grundlegende Funktion nicht nur innerhalb des Studienverlaufs der Studierenden zu, sondern ebenso auch für die fachintern-systematische Frage nach dem Status des erziehungswissenschaftlichen Wissens. Um diesen auf die Disziplin selbst zurückweisenden Aspekt geht es im folgenden Abschnitt.

2. Status des erziehungswissenschaftlichen Wissens und die Frage seiner Prüfbarkeit

Wenn es im Folgenden um standardisierte, größtenteils aus *Multiple-choice*-Aufgaben zusammengesetzten erziehungswissenschaftliche Klausuren geht, so handelt es sich dabei nicht nur um das technische Problem der raschen Bewältigung wachsender Prüfungszahlen. Es geht um mehr. Es geht letztendlich um den Status, den die Erziehungswissenschaft dem von ihr disziplinar erzeugten und erörterten, entwickelten und tradierten, gelehrt und geprüften Wissen für die eigene Disziplin, aber auch mit Blick auf die Ausbildung für pädagogischen Berufe zuspricht. Und mit dieser Frage tut sich die Erziehungswissenschaft bekanntermaßen sehr schwer, da in den letzten Jahrzehnten unterschiedliche wissenschaftliche Paradigmen oder Denkschulen in dieser Hinsicht sehr unterschiedliche Problembeschreibungen und -lösungen formuliert haben. Immerhin kann man mittlerweile nicht mehr von Alleinvertretungs- und Verdrängungsansprüchen einzelner, gegeneinander kämpfender Paradigmen sprechen. Pluralitätserfahrung, Generationenwechsel und ein die gesamte Disziplin treffender hoher Außendruck haben deutliche Spuren hinterlassen, so dass Einigungen in Richtung auf Kerncurricula und Prüfungsstandards möglich sind.

Die neuen Studienstrukturen und insbesondere deren Konsequenzen für Lehre und Prüfung wirken sich insofern auf inhaltlich-systematische Grundfragen sowie auf das Selbstverständnis der Erziehungswissenschaft aus. Die Disziplin diszipliniert nicht nur die Studie-

renden, sondern auch sich selbst, wenn sie festlegt, was wie mit welchem Standard geprüft wird.

Vorlesungen als legitime Form: Wir hatten bereits darauf hingewiesen, dass in der Erziehungswissenschaft in den 1970er und 1980er Jahren keine breite, ausgearbeitete Kultur der akademischen Vorlesungen existiert hat. Die Abkehr von der Vorlesung hatte unterschiedliche Gründe: Im Kontext der Studentenbewegung erschien sie manchen als eine zu autoritäre, die Studierenden passiv haltende hochschuldidaktische Form. Auch schienen Vorlesungen allzusehr von den Idiosynkrasien und theoretischen Präferenzen der jeweiligen Lehrenden beeinflusst. Mit lernpsychologischen Argumenten wurde die Lerneffektivität von Vorlesungen kritisch beurteilt. Und schließlich: Das erziehungswissenschaftliche Wissen insgesamt wurde als noch zu wenig systematisch gegliedert und abgesichert betrachtet, als dass ein halbwegs fester Wissenscorpus den Grundstock für vertretbare Übersichtsvorlesungen hätte bilden können, die in gesichertes Wissen des Faches einführen – und nicht nur in die Perspektive des jeweils Vortragenden. Viele dieser Vorbehalte sprechen gegen die *schlechte* Vorlesung, nicht aber gegen eine gut vorbereitete und durchgeführte Vorlesung. Als Einführung in die Gesamtstruktur, Methodik, Geschichte und Inhaltlichkeit einer Disziplin, als Übersicht über einen breiteren Gegenstandsbereich ist sie legitim und sinnvoll – auch und gerade dann, wenn (wie etwa bei Lehramtsstudierenden) ein vergleichsweise bescheidenes Zeitbudget für die bildungswissenschaftlichen Studien zur Verfügung steht. Diesen geringen Umfang an Zeit bzw. Leistungspunkten aus diesem Grunde nun gar nicht zu strukturieren bzw. dem Zufall zu überlassen, wäre fahrlässig.

Eindeutigkeit des Wissens: Natürlich stellt sich für den Lehrenden die Aufgabe, ein solches breiteres Themenfeld, über das er eine Übersicht zu vermitteln hat, in gut strukturierter, aktuell in der Disziplin prinzipiell verallgemeinerbarer Form darzustellen, verschiedene Positionen fair zum Zuge kommen zu lassen, eigene Bewertungen natürlich nicht auszusparen, aber kenntlich zu machen etc. Die Vermittlung von Übersichtswissen kann und muss *zunächst* eine gewisse Vereindeutigung, Begradigung und Entproblematisierung von Inhalten, Themen und Diskussionsverläufen beinhalten; für Einführungszwecke ist das legitim – im Alltag kann gar nicht anders verfahren werden. Aber auch in Einführungen kann und muss *auf dieser Basis* und dort, wo es angemessen ist, der unsichere, kontrovers diskutierte, unterschiedlich zu bewertende Status des Wissens verdeutlicht werden. Nicht zuletzt gehört es zur Lebendigkeit einer Vorlesung, dass kontroverse Positionen, differierende Sichtweisen, unübliche Denkwege und moralische Dilemmata zur Sprache kommen. Nicht zuletzt dadurch wird der eigenständige Denk- und Beurteilungsprozess der Zuhörer angeregt.

Die Wiederkehr der Vorlesungen, der Einsatz guter Lehrbücher und die Durchführung standardisierter Klausuren sowie vor allem: ein gewisses Maß an disziplininterner Verständigung hierüber kann zur Folge haben, dass sich die traditionell große Vielfalt und Buntheit der Erziehungswissenschaft auf der Ebene ihres Grundlagenwissens reduziert – oder vorsichtiger: dass im Grundlagenbereich der *mainstream* womöglich breiter wird, ohne aber die Existenz von alternativen, innovativen und abweichenden Denkrichtungen

zu negieren oder zu gefährden. Die Verständigung über Grundlagen sowie die darauf aufbauende Ausarbeitung von Prüfungsformen und Bewertungskriterien führt ja keineswegs etwas völlig Neues in den disziplinären Diskurs ein, da schon immer Grundlagen vermittelt und Prüfungsleistungen bewertet worden sind. Der durch die neuen Studienstrukturen ausgelöste Diskussionsprozess führt nun allerdings dazu, *dass solche impliziten, personengebundenen Bewertungsprozeduren expliziert und zumindest im Ansatz (disziplin)öffentlich gemacht werden* – eine doch durchaus positive Entwicklung. Auf der Basis eines stärker vereinheitlichen Grundlagenwissens, das immerhin eine Verlässlichkeit für Anschlüsse bietet, können dann natürlich weitere, ergänzende Vertiefungen, Problematisierungen, Alternativdeutungen etc. entfaltet werden – wie es sich für jede Wissenschaft gehört. Eine solche Entwicklung würde sich in die schon seit längerem diskutierte bzw. geforderte Normalisierung der Erziehungswissenschaft innerhalb des Konzerts der wissenschaftlichen Disziplinen einordnen.

Berufsbezug des Wissens: Selbst wenn der Anteil des konsensfähigen, gesicherten Wissens in der Erziehungswissenschaft höher wäre als er ist, ergibt sich das Problem, welche Wissens Elemente für die Vermittlung einer bestimmten Teilkompetenz ausgewählt werden bzw. welche Wissenskomplexe dazu geeignet sind, eine spätere erfolgreiche Berufsausübung anzubahnen. Nun ist Wissen nur *ein* Basiselement für qualifizierte Berufsausübung, hinzukommen müssen u.a. praktische Erfahrungen und deren Reflexion, eine gewisses habituelles Einschwingen in die Berufskultur etc. Um praktische Erfahrungsbildung und Entwicklung von berufsbezogenen Haltungen kann es in Vorlesungen nur äußerst bedingt und sehr vermittelt gehen. Erneut soll an dieser Stelle auf die hohe Bedeutung anderer Lehr-, Lern- und Prüfungsformen hingewiesen werden; die Differenz der Wissensformen, Erfahrungsbereiche und Kompetenzdimensionen, die bei der Konstruktion von Studiengängen eine Rolle spielen, muss unbedingt Beachtung finden – auch und gerade in Gestalt einer methodisch vielfältigen Lehr- und Prüfungspraxis. Wenn es aber um *Wissen* geht, so stellt sich die Frage, welches Wissen für alle Studierenden als *unabdingbar* oder vorsichtiger als *nützlich* für den Studienprozess, für die Kompetenzanbahnung, für das spätere Handeln und Reflektieren in den verschiedensten pädagogischen Berufsfeldern betrachtet wird, welches Wissen demgegenüber als *wünschenswert* anzusehen ist, ab wann und in welcher Weise *Spezialisierungen* in Richtung auf wählbare Studienprofile zum Zuge kommen etc. In den Sitzungen von Studienkommissionen gehen die Meinungen hierzu meistens deutlich auseinander. Modularisierung zwingt die Erziehungswissenschaft dazu, solche grundsätzlichen Fragen des Verhältnisses von Disziplin und Profession nicht länger als Dauerdiskurs und mit offenem Ende zu erörtern, sondern hierfür konkrete, umsetzbare und überprüfbare Lösungen zu erarbeiten. Auch hier entsteht also eine Rückwirkung des neuen Prüfens auf systematische Fragen der Disziplin.

Anforderungsniveaus: Zum Abschluss von Vorlesungen standardisiert zu prüfen, muss keineswegs eine Verflachung des Anforderungsniveaus oder eine Zurückschneidung des Prüfungsniveaus auf einfaches Reproduzieren von auswendig gelernten Vorlesungsmitschriften, verteilten Skripten oder zur Verfügung gestellten *powerpoint*-Folien bedeuten.

Es muss ein Spektrum von Aufgabenformaten gefunden werden, das in unterschiedlich anspruchsvoller Weise Wissen, Reflexionsfähigkeit, Transferleistungen, Urteilsvermögen etc. abfordert und das je individuell erreichte Niveau auf einer Skala festzuhalten erlaubt – und gleichwohl eine effiziente Form der Durchführung und Auswertung ermöglicht. Die entsprechenden Erfahrungen aus anderen Disziplinen (Medizin, Jura, Ökonomie) müssen genutzt werden, um eine entsprechend anspruchsvolle Mischung verschiedener Aufgabenniveaus und Aufgabenformate zu erreichen. Eine geschickte Konstruktion von Aufgabenformaten ermöglicht es, inhaltliche Ansprüche und effiziente Auswertbarkeit zu erreichen. Der Beweis hierfür kann nur durch entsprechende Aufgabenbeispiele erbracht werden.

3. Standardisierte Prüfungsverfahren

3.1 Vorbehalte

Wenn der Eindruck, den man aus Gesprächen mit Kolleginnen und Kollegen gewinnen kann, nicht täuscht, dann setzen sich unter dem Zwang der oben beschriebenen organisatorischen Strukturen der neuen Ausbildung auch im Fach Erziehungswissenschaft standardisierte Prüfungen als übliche Form der Leistungsmessung in großen Vorlesungen, aber auch in Seminaren mit großer Teilnehmerzahl zunehmend durch. Die Expansion der neuen Prüfungsform steht dabei in einem seltsamen Kontrast zu einer verbreiteten Geringschätzung dieses Instrumentariums – selbst bei Lehrenden, die sich unter dem Zwang der Umstände zu seiner Nutzung entschlossen haben.

Im Kern richten sich die Vorbehalte zumeist gegen die Reichweite standardisierter Leistungsmessungen. Die Kritiker stellen in Frage, ob mit solchen Formen, insbesondere mit *Multiple-choice*-Aufgaben, die relevanten Ziele eines Hochschulstudiums bzw. ihrer eigenen Lehrveranstaltungen angemessen erfasst und ein entsprechender Lernerfolg der Studierenden überprüft werden können. Bestenfalls, so die übliche Kritik, könne man mit solchen Verfahren Faktenwissen bzw. Fleiß und Erinnerungsvermögen der Prüfungskandidaten erfassen und bewerten.

Ein Stück weit wird man solchen Einwänden Recht geben müssen: Bestimmte Lernziele und Kompetenzen, die durch ein Hochschulstudium erreicht bzw. gefördert werden sollen, etwa die Steigerung komplexer kognitiver Fähigkeiten oder die Verbesserung wünschenswerter Selbst- und Sozialkompetenzen, lassen sich mit dem Instrumentarium standardisierter Leistungsfeststellungen nicht oder bestenfalls ansatzweise abbilden und erfassen. Ob die Studierenden mit den abgefragten und eventuell nachgewiesenen Wissensbeständen zugleich Arbeits- und Lernstrategien erworben haben oder ob sie erworbene Kenntnisse zur selbstständigen Bearbeitung praktischer Probleme nutzen können – all das lässt sich durch geschlossene Frageformate und vorgegebene Antwortmöglichkeiten nicht hinreichend erfassen.⁴ Und deshalb sollte es unstrittig sein, dass die standardisierten Prüfungs-

⁴ Wir konzentrieren uns im Folgenden auf die Erörterung standardisierter Aufgabenformate, die vornehmlich in Verbindung mit großen erziehungswissenschaftlichen Vorlesungen eingesetzt werden. Wir sind uns dar-

verfahren andere Formen der Lern- und Leistungskontrolle, insbesondere die für geistes- und sozialwissenschaftliche Fächer traditionsreiche, anspruchsvolle Hausarbeit, nicht verdrängen dürfen. Hinsichtlich der traditionellen Prüfungsformen stellt sich allerdings ebenso die Frage, ob und inwieweit sie tatsächlich geeignet sind, komplexe Kompetenzen valide zu erfassen und objektiv zu bewerten.

Ungeachtet der unbestreitbaren Grenzen standardisierter Prüfungsverfahren für die Erfassung komplexer Lernziele bzw. Kompetenzen werden die Möglichkeiten und der Nutzen dieses Instruments in der gegenwärtigen Situation häufig unterschätzt. In der verbreiteten Auffassung, dass in solchen Prüfungen ‚nur‘ Faktenwissen und Begriffe abgefragt werden können, zeigt sich ein doppeltes Missverständnis: Das eine bezieht sich auf den Stellenwert von Faktenwissen für erfolgreiches Lernen und das andere auf die Komplexität des Wissens, das in solchen Tests abgefragt und bewertet werden kann, also auf das Anspruchsniveau der Tests. Gegen eine voreilige Geringschätzung von fachwissenschaftlichen Grundkenntnissen und einer entsprechenden Fachsprache muss daran festgehalten werden, dass sie das Fundament jeder Fachwissenschaft darstellen und die Voraussetzung aller anspruchsvolleren Lernprozesse bei Studierenden des jeweiligen Faches sind. Wenn es gelingt, dieses Fundament erfolgreichen Studiums, das für grundlegend gehaltene Wissen eines Faches bzw. seiner Teildisziplinen, zu definieren, und zugleich der Anspruch erhoben wird, es in entsprechenden Veranstaltungen zu vermitteln, dann gibt es keinen einsichtigen Grund, auf eine Überprüfung der Lernergebnisse der Studierenden in standardisierter Form zu verzichten. Dies kann allerdings auf unterschiedlich anspruchsvolle Weise und auf unterschiedlichem Niveau geschehen. Möglicherweise resultieren viele Vorbehalte gegen standardisierte Prüfungsverfahren, insbesondere in Form von *Multiple-choice*-Tests, aus mangelnden Erfahrungen mit diesem Prüfungsformat oder vielleicht auch aus zufälligen Beispielen, die in Form und Inhalt unterkomplex sind, weil ihre Aufgaben entweder mit dem gesunden Menschenverstand oder durch geschicktes Raten zu lösen sind. Aber diese nicht seltenen Beispiele sind keine Belege für die generelle Untauglichkeit des Instruments, sondern lediglich für eine fehlerhafte Konstruktion im Einzelfall.

3.2 Anforderungen

Die der Arbeitsgruppe vorliegenden – zufälligen und wenigen – Beispiele von standardisierten Abschlussklausuren für Vorlesungen und Seminare vermitteln einen zumindest heterogenen Eindruck, der auf die Schwierigkeiten bei der Entwicklung solcher Prüfungsformate verweist.⁵ Sie sind heterogen hinsichtlich dessen, was sie prüfen und wie sie es

über im Klaren, dass dies innerhalb des breiten Spektrums unterschiedlicher Formen der Leistungserfassung und -beurteilung eine spezifische, insgesamt eher konventionelle, für die neue Lehrsituation in der Erziehungswissenschaft allerdings unausweichliche, aber wenig erprobte Variante darstellt. Wir gehen nicht auf die grundsätzlichen testdiagnostischen Probleme von kompetenzorientierter und zugleich entwicklungsförderlicher Leistungskommentierung und -beurteilung in Schulen und Hochschulen ein (*edumetrics* statt *psychometrics*). Dies würde den Rahmen sprengen.

⁵ Sicherlich wird man in diesem Kontext aus den Erfahrungen anderer Fächer lernen können. Ebenso müsste es hilfreich sein, etwa die Erfahrungen der Fernuniversität Hagen bei der Formulierung und Gestaltung von Frageformaten in Studien- und Lernmaterialien mit erziehungswissenschaftlicher Thematik zu nutzen.

prüfen wollen. Sie unterscheiden sich selbst bei themenähnlichen Bezugsveranstaltungen hinsichtlich der Inhalte, des Anspruchsniveaus der Fragen und der Form der Aufgabenstellung. Wenn man beispielsweise die inhaltliche Auswahl von Prüfungsaufgaben zu allgemeinen Einführungen in das Fach Erziehungswissenschaft miteinander vergleicht, lässt sich eine erhebliche Variationsbreite feststellen. Es gibt thematische Überschneidungen, aber keinen einheitlichen Korpus fachlichen Wissens, das in solchen Veranstaltungen vermittelt wird. Dies liegt zum einen an der je universitätsspezifischen curricularen Struktur des erziehungswissenschaftlichen Lehrangebots, zum anderen an der individuellen Handschrift, den thematischen Präferenzen und entsprechenden Akzentuierungen der jeweiligen Veranstalter.

Diese inhaltliche Varianz der Prüfungsaufgaben zu ähnlichen oder sogar identischen Themen der Veranstaltungen lässt sich durch innerfachlichen Austausch und Diskurs verringern, aber keinesfalls völlig beseitigen, und das ist auch gut so. Während man über die fachliche Qualität einer Veranstaltung, die Relevanz ihrer thematischen Schwerpunkte und Lernziele vielleicht streiten kann, gilt dies nicht für die Güte einer standardisierten Abschlussklausur. Sie muss die zentrale Anforderung der Validität, der Gültigkeit, erfüllen, also genau das prüfen, was in der jeweiligen Veranstaltung unter bestimmten Zielperspektiven inhaltlich vermittelt werden sollte.

Die Forderung nach Validität oder Gültigkeit der Prüfungen scheint auf den ersten Blick trivial zu sein – für die Prüfungspraxis gilt dies keineswegs. Wenn manche standardisierten Tests den Eindruck vermitteln, dass ihre Aufgaben auch ohne Kenntnis des jeweiligen Veranstaltungsprogramms und seiner spezifischen Lernziele bearbeitet werden können, dann liegt offensichtlich ein Verstoß gegen die grundlegende Forderung nach Validität vor – es sei denn, dass die jeweilige Bezugsveranstaltung keine spezifischen, über das Alltagswissen hinausgehenden Kenntnisse vermittelt hätte.

Soll die Anforderung an die Validität einer Abschlussklausur erfüllt werden, so ist es erforderlich, dass sich die Lehrenden über das, was sie in ihren jeweiligen Veranstaltungen vermitteln wollen, im Klaren sind und ihre generellen Lernziele in Teilziele ausdifferenzieren und operationalisieren können. Dies ist zwar eine Grundanforderung professionellen pädagogischen (lehrenden) Handelns, aber in der Praxis universitärer Lehre eine allzu optimistische Unterstellung. Im Blick auf eine valide Abschlussklausur ist es jedenfalls erforderlich, parallel zur Veranstaltungsplanung und -durchführung die Frageformate für den Abschlusstest zu konzipieren. Dabei ist zu beachten, dass sich die Prüfungsaufgaben auf das gesamte Spektrum der Themen, aber auch auf die unterschiedlichen Ebenen der Lernziele der Veranstaltung beziehen. Insbesondere die zuletzt genannte Forderung impliziert, dass der Schwierigkeitsgrad der Aufgabenstellung variieren, über das bloße Wiedererkennen von Fakten und Informationen hinausführen und anspruchsvollere Lernziele abbilden muss.

Entgegen einem verbreiteten Vorurteil sind auch Auswahlaufgaben (im Sinne von *Multiple-choice*) in der Lage, die angestrebte Unterscheidungsfähigkeit, das Differenzierungsvermögen und die Urteilsfähigkeit, die in der jeweiligen Veranstaltung – themenbezogen – entwickelt werden sollten, zu erfassen. Auch solche geschlossene Aufgabenformate können und müssen sich auf komplexe Sachverhalte, konkurrierende Interpretationen, an-

spruchsvolle theoretische Beschreibungen und ihre jeweiligen praktischen Implikationen beziehen – sofern diese in der vorangegangenen Veranstaltung erarbeitet worden sind. Die Möglichkeiten dieser Prüfungsform enden erst dort, wo die Kompetenz der Studierenden überprüft werden soll, auf der Basis des erworbenen Wissens selbstständig und kreativ Antworten auf neuartige theoretische und praktische Probleme zu entwickeln. Aber das kann auch nicht das Ziel einer einführenden Veranstaltung sein!

Die Differenzierung des Anspruchsniveaus der Aufgabenstellung ist auch im Blick auf eine leistungsgerechte Beurteilung der Studierenden unverzichtbar. Nur so lassen sich gute von schlechten Leistungen, d. h. die jeweils erreichten Kompetenzen der Studierenden voneinander unterscheiden. Ob diese Unterscheidung zuverlässig, ob das Instrument reliabel ist und nicht von Messfehlern und Zufallsantworten verzerrt wird, hängt maßgeblich von der Konstruktion der Frageformate, etwa von der Auswahl gestellter Aufgaben oder der Genauigkeit der Fragestellungen, und der Durchführung und Auswertung ab. Im Idealfall müssten Studierende bei einer reliablen Testkonstruktion und -auswertung ähnliche Prüfungsergebnisse in einem Paralleltest mit gleichem Anspruchsniveau und anderen Aufgaben zur gleichen Veranstaltung erreichen.

3.3 Aufgabentypen

Bei den Aufgabenstellungen in Abschlussklausuren für große Lehrveranstaltungen werden normalerweise geschlossene und halboffene Bearbeitungsformen gewählt. Die nachfolgenden Hinweise orientieren sich an einem hilfreichen Prüfungsleitfaden der Universität St. Gallen von Metzger und Nüesch.⁶ Man kann die bei solchen Klausuren üblichen Prüfungsformate auch in *Auswahlaufgaben* und *Bearbeitungsaufgaben* unterscheiden. Die Auswahlaufgaben lassen sich wiederum in Richtig-/Falsch-Aufgaben, in Mehrfachwahlaufgaben (*Multiple-choice*-Aufgaben) und in Zuordnungsaufgaben differenzieren. Die erste Variante erlaubt nur die Entscheidung der Prüfungskandidaten zwischen „richtig“ und „falsch“ bei einer vorgegebenen Aussage und ist damit besonders anfällig für Zufallsergebnisse. Die zweite verlangt bekanntlich die Identifikation einer – oder mehrerer – richtiger Aussagen in einer Gruppe von Antwortalternativen. Die dritte Variante erfordert die Zuordnung bestimmter, vorgegebener Aussagen zu unterschiedlichen Theorien oder Sachverhalten.

3.3.1 *Multiple-choice*-Aufgaben

In den meisten Abschlussklausuren zu Großveranstaltungen dominiert der zweite Typ von Auswahlaufgaben, also die *Multiple-choice*-Aufgabe. Dieser Aufgabentyp wird zuweilen

⁶ Metzger, Chr./Nüesch, Ch.: Fair prüfen. Ein Qualitätsleitfaden für Prüfende an Hochschulen. (Hochschuldidaktische Schriften, Band 6). Institut für Wirtschaftspädagogik, Universität St.Gallen 2004; Eugster, B./Lutz, L.: Leitfaden für das Planen, Durchführen und Auswerten von Prüfungen an der ETHZ. Verfügbar unter: http://www.diz.ethz.ch/projects/leistungskontrollen/dokumente/leitfaden_PDA_Pruefungen_DiZ-2003.pdf (1.7.2008).

ergänzt durch zwei Varianten von Bearbeitungsaufgaben, nämlich die Vervollständigungs-
aufgaben (Lückentests), bei denen die Studierenden die angemessenen Fachbegriffe oder
prägnanten Fakten in vorgegebene Aussagen einfügen müssen, und Kurzwortaufgaben,
in denen – auf vorgegebenem Raum – Definitionen von Begriffen, Erläuterungen oder be-
gründeten Einschätzungen zu bestimmten theoretischen Positionen oder Sachverhalten
verlangt werden. Die folgenden Hinweise auf eine sachgerechte Konstruktion und Auswer-
tung der bei Abschlussklausuren dominierenden Prüfungsformaten, den Auswahlaufgaben
und Bearbeitungsaufgaben, mögen auf den ersten Blick trivial erscheinen, werden aber in
der Prüfungspraxis keineswegs immer beachtet.

Die erste Grundregel gilt für alle Prüfungsaufgaben, ist aber für die Konstruktion von
Mehrfachwahlaufgaben (*Multiple-choice*-Aufgaben) besonders wichtig: Das Ausgangs-
problem bzw. die Frage muss unmissverständlich, sprachlich klar und inhaltlich genau
formuliert werden. Gleiches gilt für die nachfolgenden Antwortangebote, die Items zu der
gestellten Frage. Fragen, die zu allgemein gestellt sind, beispielsweise nach der Differenz
zwischen Erziehung und Sozialisation oder der Definition von Bildung, können streng ge-
nommen nicht eindeutig beantwortet werden ohne Angabe der Theoriereferenzen, die in
der jeweiligen Veranstaltung expliziert wurden. Man kann zwar davon ausgehen, dass dem
Konstrukteur der jeweiligen Aufgaben der Sinn der Aufgabenstellung in der Regel klar
sein dürfte. Bei den Adressaten der Aufgaben ist dies aber keineswegs immer der Fall.
Rückfragen nach Austeilung der Klausurbögen sind deutliche Hinweise auf mangelnde
Klarheit der Aufgabenstellung. Insofern ist es sinnvoll, Kolleginnen und Kollegen⁷ und
möglichst auch Studierenden, Fragen und Auswahlantworten zur Überprüfung der sprach-
lichen Klarheit und inhaltlichen Eindeutigkeit vorzulegen. Im Prinzip müsste sogar ein
entsprechender *pretest* organisiert und ausgewertet werden.

Die zweite Grundregel betrifft Zahl und Art der Prüfungsaufgaben und ihre jeweiligen al-
ternativen Antwortmöglichkeiten. Grundsätzlich gilt, dass die Zufallswahrscheinlichkeit
für das Bestehen einer standardisierten Klausur mit der Zunahme der Prüfungsaufgaben
sinkt. Noch wichtiger für die Verminderung zufällig richtiger Antworten ist allerdings die
Zahl und Formulierung der alternativen Items zu einer bestimmten Frage. Die richtigen
Antworten sollten zu den falschen mindestens im Verhältnis von 1 : 3 stehen, um ein er-
folgreiches Raten zu erschweren, vor allem aber müssen die falschen Antworten bestimm-
ten Kriterien genügen. Wenn sich beispielsweise eine Frage auf die bildungshistorische
Bedeutung Humboldts bezöge, wäre es nicht zweckmäßig, eine Antwort anzubieten, nach
der es sich bei diesem Bildungstheoretiker um den Vordenker nationalsozialistischer Bil-
dungspolitik handelte. Auch die falschen Antwortangebote müssen nämlich eine gewisse
Plausibilität besitzen und in einem möglichst engen Zusammenhang mit der gestellten Fra-
ge stehen.

⁷ Das in manchen Prüfungsordnungen verlangte Vier-Augen-Prinzip bei Prüfungsbewertungen (Prü-
fer/Zweitprüfer) kann bei standardisierten Prüfungsverfahren u.U. in Richtung auf die Abstimmung der Fra-
gen mit einem thematisch einschlägigen, prüfungsberechtigten Lehrenden bewältigt werden.

Der dritte Hinweis betrifft die Auswertung der Antworten zu den einzelnen Fragen und die Bewertungen der Prüfungsleistungen insgesamt. Hinsichtlich der Auswertung der einzelnen Prüfungsaufgaben konkurrieren unterschiedliche Varianten, die sich aus der Frage ergeben, was als richtige Antworten in die Berechnung der Note eingeht. Zum einen kann die Gesamtzahl der zu Recht angekreuzten Antwortalternativen die Grundlage der Bewertung sein. Falsche Antworten werden hingegen nicht berücksichtigt. In diesem Fall ist es erforderlich anzugeben, wie viele richtige Items die jeweilige Aufgabe enthält und eine entsprechende Anzahl von Markierungen vorzuschreiben. Zum anderen kann aber auch jedes einzelne Item zu einer Frage als Ja/Nein-Option interpretiert und eine Entscheidung gegen eine falsche Antwortalternative ebenfalls als richtige Antwort bewertet werden. Dieses Bewertungsverfahren hat den doppelten Effekt, dass sich dadurch die Gesamtzahl möglicher Punkte, aber zugleich die Zufallswahrscheinlichkeit für richtige Antworten erhöht – was bei der Bewertung der Leistung berücksichtigt werden muss.

Hinsichtlich der Bewertung der Prüfungsleistung insgesamt kann man davon ausgehen, dass es sich bei solchen Klausuren in der Regel um eine summative Prüfung handelt, deren Bestehen häufig die Voraussetzung für die Teilnahme an nachfolgenden Modulveranstaltungen darstellt. Dies legt zumindest nahe, dass die Bezugsnorm für die vergebene Note kriteriumsorientiert und nicht gruppenbezogen ist. Damit wird zugleich impliziert, dass eine Mindestzahl der erreichbaren Punkte für das Bestehen der Prüfung die Voraussetzung ist. Wo die Grenze für eine ausreichende Leistung gezogen wird, hängt wiederum von der oben skizzierten alternativen Interpretation richtiger Antworten ab. Werden lediglich die richtig angekreuzten Items gezählt, bilden 50% der richtigen Antworten normalerweise die Untergrenze für mindestens ausreichende Leistungen. Werden hingegen auch die zu Recht unterlassenen Markierungen als richtige Antworten gezählt, sollte diese Grenze auf etwa $\frac{2}{3}$ der möglichen Punkte erhöht werden. – In jedem Fall sollten die Bewertungsmodalitäten für die gesamte Prüfung wie ihre Einzelteile den Studierenden im Voraus bekannt sein.

Der letzte Hinweis bezieht sich auf eine Anforderung, die in der Praxis vermutlich eher selten berücksichtigt wird. Das Instrument der Auswahlaufgaben, das die Leistungen der Studierenden überprüfen soll, muss seinerseits auf seine Leistungsfähigkeit überprüft werden. Wie bereits angedeutet, müsste vor dem erstmaligen Einsatz der Test einem Pretest unterworfen werden. Beabsichtigt man, das so entwickelte Instrument kontinuierlich in Nachfolgeveranstaltungen einzusetzen, dann wird es erforderlich sein, die Ergebnisse des ersten Durchlaufs, d.h. das Antwortverhalten der Studierenden in den einzelnen Aufgabebereichen, hinsichtlich der Gütekriterien Validität und Reliabilität auszuwerten, um zu einer Verbesserung dieses Prüfungsinstruments zu kommen. Varianten sind auch deshalb nötig, da ansonsten die Aufgaben nachfolgenden Kohorten bekannt sind.

3.3.2 Bearbeitungsaufgaben

Im Gegensatz zu den unterschiedlichen Typen von „Auswahlaufgaben“, bei denen sich die Studierenden zwischen vorgegebenen richtigen und falschen Antworten entscheiden müssen, verlangen „Bearbeitungsaufgaben“ die selbstständige Wiedergabe oder – weitgehend – die Anwendung gelernter Informationen auf neue Sachverhalte, auf Fallbeispiele oder theoretische Positionen. Bei solchen Aufgaben geht es nicht – wie im Fall der Auswahl-aufgabe – um die Identifikation richtiger Informationen und deren Unterscheidung von falschen Antwortalternativen, sondern um die Darstellung, Erläuterung und Anwendung fachspezifischer Begriffe theoretischer Aussagen und fachbezogener Informationen. Bei diesen „Bearbeitungsaufgaben“ lassen sich drei Varianten idealtypisch voneinander unterscheiden, die im Anspruchsniveau sehr stark variieren können.

Die erste Variante stellen *Vervollständigungsaufgaben* (Lückentests) dar. Sie verlangen von den Studierenden, in vorgegebenen Aussagen fehlende Begriffe korrekt einzusetzen oder fehlende Sachinformationen zu ergänzen. Mit diesem Aufgabenformat kann also die korrekte Verwendung fachspezifischer Begriffe und die Kenntnis sachbezogener Wissensbestände überprüft werden.

Während die Antwortmöglichkeiten bei solchen Vervollständigungsaufgaben normiert sind und idealiter nur eine Antwort richtig ist, wächst bei der zweiten Variante der Bearbeitungsaufgaben die Varianz möglicher Antworten zumindest hinsichtlich ihrer äußeren Form. Bei diesem Aufgabentyp (*Kurzantworten*) werden kurze, in den Formulierungen „offene“ Antworten der Studierenden zu Sachverhalten und theoretischen Erklärungen verlangt, die in der vorangegangenen Veranstaltung erarbeitet worden sind. Auch bei diesem Aufgabentyp ist es von entscheidender Bedeutung, dass die Fragestellung eindeutig und vor allem eingeschränkt ist. Nur so lassen sich in begrenzter Zeit und auf vorgegebenem Raum angemessene Antworten formulieren und bei der Bewertung die guten von schlechten Antworten unterscheiden. Es wäre deshalb beispielsweise problematisch, bei einer solchen Aufgabe eine Unterscheidung der Begriffe Erziehung, Bildung und Sozialisation zu verlangen. Es wäre zumindest erforderlich, eine in der vorangegangenen Veranstaltung erarbeiteten Referenztheorie einschränkend hinzuzufügen. Genauso müsste eine Frage nach den möglichen Ursachen des schlechten Abschneidens deutscher Schülerinnen und Schüler in den internationalen Schulvergleichsforschungen präzisiert werden. Sie könnte sich – wahlweise – auf außerschulische Faktoren oder auf spezifische Merkmale des deutschen Bildungssystems oder auf unterrichtsbezogene Variablen beziehen, aber keineswegs auf das gesamte Spektrum möglicher Ursachen. Dies würde nämlich die Grenzen einer „Kurzantwortaufgabe“ sprengen und eine objektive Beurteilung der Leistung zumindest erheblich erschweren.

Die dritte Variante der Bearbeitungsaufgaben, die auch im Rahmen größerer Lehrveranstaltungen eingesetzt werden kann, lässt sich als *Anwendungs- oder Transferaufgabe* bezeichnen. Bei dieser Prüfungsvariante können den Studierenden kurze Fallbeispiele oder

pointierte Aussagen zu bestimmten in der Veranstaltung erarbeiteten Sachverhalten vorgelegt werden. Die Studierenden erhalten die Aufgabe, die vorliegenden Fälle unter einer bestimmten theoretischen Perspektive und in der entsprechenden Fachsprache zu interpretieren bzw. zu kommentieren. So könnten beispielsweise alltägliche familiäre Interaktionen mit dem Instrumentarium des operanten Konditionierens interpretiert oder kritische Kommentare zu einseitigen Aussagen etwa zu den PISA-Ergebnissen verfasst werden. In beiden Fällen würde es sich um Aufgaben handeln, die über das Erinnern und die Reproduktion fachlichen Wissens hinausgehen und sich dem Kompetenzbereich der Anwendung theoretischen Wissens auf neue Sachverhalte nähern.

Die Vorzüge von *Bearbeitungsaufgaben* gegenüber *Auswahlaufgaben* liegen auf der Hand. Gleiches gilt aber leider auch für ihre Nachteile, die sowohl die Ökonomie wie die Objektivität der Auswertung solcher Prüfungsaufgaben betreffen. Während die Auswertungsobjektivität bei den Auswahlaufgaben kein ernstes Problem darstellt und die Auswertung auf der Grundlage des Lösungsschlüssels Hilfskräften übertragen oder sogar maschinell erfolgen kann, erfordert die Beurteilung der beiden zuletzt genannten Varianten der Bearbeitungsaufgaben einen erheblich höheren Aufwand – sowohl hinsichtlich der Vorbereitung wie der Durchführung der Bewertung. Für diese Aufgabenformate muss der „Erwartungshorizont“ für die optimalen Antworten auf die jeweiligen Fragen präzise festgelegt werden, müssen die erwarteten Antwortelemente identifiziert und für die zusammenfassende Bewertung gewichtet werden.

Und selbst wenn diese Anforderung optimal erfüllt und entsprechende Musterlösungen für die einzelnen Aufgaben erstellt werden, lässt sich die nachfolgende Korrektur nicht umstandslos an Hilfskräfte delegieren. Um eine zufrieden stellende Auswertungsobjektivität zu gewährleisten, ist es erforderlich, die Anwendung der Musterlösungen an exemplarischen Fällen zu erproben und die daraus resultierenden Bewertungen miteinander zu besprechen und abzustimmen.

Anspruchsvolle standardisierte Prüfungsformen, so lässt sich zusammenfassend sagen, verlangen in ihrer Konstruktion, Durchführung und Auswertung erhebliche Anstrengungen. Aber was wäre angesichts der Anforderungen der neuen Studienstruktur die Alternative?

4. Aufgaben im Kontext technologisch unterstützten Prüfens

Die zentrale Frage dieses Abschnitts ist, inwiefern und inwieweit Prüfungsfragen durch Einsatz neuer Technologien, d.h. durch Einsatz von Computern und Internet, automatisiert ausgewertet werden können.⁸ Es geht hier nicht darum, ob strukturierte Aufgabenformate wie *Multiple-choice*, Lückentext oder Zuordnung besser oder schlechter sind als Freitextaufgaben. Es geht vielmehr um die Frage, ob computergestützte Aufgaben die Arbeit der Klausurbewertung erleichtern können. Insofern steht Arbeitseffizienz im Fokus der folgenden Betrachtungen.

Auch Objektivität, Validität und Reliabilität von Aufgaben stehen hier nicht zur Debatte. Dennoch sei eine Bemerkung dazu hier erlaubt. Auf der Basis eines fünfjährigen Experimentierens mit computergestützten, automatisch ausgewerteten Klausuren hat es sich als vorteilhaft erwiesen, Klausuren so zu konstruieren, dass die Hälfte der erreichbaren Punkte in strukturierten Aufgaben und die andere Hälfte in Freitextaufgaben vergeben werden. Da die verwendete Klausursoftware die differenzierten Klausurergebnisse in einer EXCEL-Tabelle ausgibt, ist es ein Leichtes, die Ergebnisse der strukturierten Aufgaben mit denen der Freitextaufgaben zu korrelieren. In den meisten Fällen traten hohe Korrelationen auf, ebenso statistische Signifikanzen (5%-Niveau). In einer Klausur mit ca. 120 Teilnehmern konnte durch direkten Vergleich der beiden Teilnoten festgestellt werden, dass bei keinem Studierenden in den beiden Klausurteilen eine größere Notendifferenz als eine halbe Note auftrat. Damit hat man zwar keine saubere Validitätsprüfung, aber man erhält mindestens den Hinweis, dass man mit beiden Klausurteilen das Gleiche misst. In Fällen, bei denen die Korrelationen nicht befriedigend ausfielen, konnte sehr schnell herausgefunden werden, woran dies mit hoher Wahrscheinlichkeit lag. In der ausgegebenen Tabelle waren nämlich auch die Ergebnisse der einzelnen Aufgaben deskriptiv statistisch beschreibbar, und man konnte etwa sehen, dass manche Aufgaben zu leicht oder zu schwer waren, d.h. dass sie keine adäquate Leistungsdifferenzierung bei den Klausurteilnehmern erzeugt haben. Kommt so etwas etwa bei den strukturierten Aufgaben an mehreren Stellen vor, nicht aber bei den Freitextaufgaben, dann kann sich keine hohe Korrelation ergeben.

Das heißt in aller Kürze: Bei automatisch ausgewerteten Klausuren erhält man je nach *Software* Daten, die eine rasche und nicht sehr arbeitsaufwändige Analyse ermöglichen. Die Analyse kann beispielsweise Hinweise darauf geben, dass die Aufgaben nicht gut genug konstruiert waren, dass die Themen der Freitextaufgaben zu schwer waren usw. Das kann bis hin zur Konsequenz führen, für gewisse Themen mehr Zeit in den Lehrveranstaltungen zu veranschlagen. Natürlich kann man solche Auswertungen auch auf der Basis von herkömmlichen Papierklausuren machen. Computerunterstützte Klausurengestaltung und -auswertung erleichtert die Arbeit jedoch deutlich.⁹

⁸ Auf die Möglichkeit, bearbeitete standardisierte Klausur-Fragebögen durch einen geeigneten *Scanner* einzulesen, kann an dieser Stelle nur verwiesen werden. In diesem Fall können die Fragebögen im Papier- und Bleistift-Verfahren bearbeitet werden; die Eingabe der einzelnen Antworten per Hand ist jedoch überflüssig.

⁹ Vgl. dazu auch Wollersheim, H.-W.: eTesting in Lehramtsstudiengängen – mehr als ein technisches Problem? Foliensatz eines Vortrags im Rahmen des Workshops „Computerunterstütztes Prüfen (e-assessment)“ an der Universität Hamburg, 6.12.2007.

4.1 Lern- und Testaufgaben

Man kann auf der Basis der Literatur und dem professionellen Wissen der Didaktiker die verschiedenen Aufgabentypen als eine Teilklasse des Wissens verstehen, das sich in Wissensarten differenziert.¹⁰ Aufgaben sind interaktive Wissensarten, die zwischen der reinen Vorgabe von rezeptiv aufzunehmenden Informationen und den noch komplexeren Szenarien kooperativer Wissensgenerierung im Lernen stehen. Aufgaben geben nämlich einerseits Informationen, lassen aber anderes aus, was der Lernende zu ergänzen hat. Aufgaben produzieren gezielt Leerstellen, die der Lernende ausfüllen muss. Die Produktion der Leerstelle in Aufgaben kann die Ausfüllung der Leerstelle eindeutig bestimmen oder eher offen gestalten. Diese Variation erreicht man durch Art und Ausmaß der Informationsvorgabe. Flechsig hat vor diesem Hintergrund Aufgaben der Struktur nach als *Formulare* bestimmt.¹¹

Wissen hat in didaktischen Kontexten immer mindestens drei Funktionen: Informationsfunktion, Instruktionsfunktion und Kommunikationsfunktion. Diese Dreiteilung deckt sich mit der Dreiteilung medialer Strukturierung in (rezeptive) Darstellungsmedien, Interaktionsmedien und Kommunikationsmedien. Interaktionsmedien sind solche, bei denen die Testperson mit einem Medium interagiert. Das mag ein Aufgabenblatt oder ein Computerbildschirm sein. Kommunikationsmedien sind solche, bei denen die Testperson mit anderen Testpersonen interagiert – sei es per Brief, *Chat*, Forum oder in einer virtuellen Welt (z.B. *Second Life*).

Aufgaben haben im didaktischen Kontext des Lernens generell die Funktion, das Können der Lernenden zu aktualisieren. Bei dieser Aktualisierung können drei Unterfunktionen unterschieden werden.

- *Lernaufgaben* haben die Funktion, das Können zu initiieren, das Können ein erstes Mal hervorzubringen.
- *Selbsttestaufgaben* haben die Funktion, die Lernenden über den Stand in ihrem Lern- und Bildungsprozess aufzuklären.
- *Testaufgaben* haben die Funktion zu zertifizieren, welches Können im Bildungsprozess erworben wurde.

Aufgaben sind also funktional dreifach zu differenzieren. Strukturell unterscheiden sie sich trotz ihrer funktionalen Differenzierung jedoch nicht. Diese funktionale Differenzierung wird herausgestellt, weil alle drei Funktionen mit geeigneter *Aufgabensoftware* eingelöst werden können. Wenn also z.B. eine Aufgabe durch Selbsttestaufgaben im Internet begleitet wird, dann wird damit den Studierenden die Möglichkeit geboten, ihr Verständnis zu testen und gegebenenfalls in der nächsten Vorlesungsstunde nachzufragen. Ein anderes Szenario könnte sein, eine Vorlesung zum Thema Forschungsmethoden zu halten und parallel dazu Übungen machen zu lassen. Dazu werden Lernaufgaben und das zu ihrer Lö-

¹⁰ Vgl. Meder, N. u.a. (2006): Web-Didaktik. Eine neue Didaktik webbasierten, vernetzten Lernens. (Wissen und Bildung im Internet Bd. 2). Bielefeld, S. 119-174, insbes. S. 127.

¹¹ Haller, H.-D./Flechsig, K.-H./List, J.: CEDID – Software und Kurse zum computergestützten didaktischen Design; Haller, H.-D. (1995): Wissensorganisation mit CEWID, einem wissensorientierten und tätigkeitsunterstützenden System. In: Meder, N./Jaenecke, P./Schmitz-Esser, W. (Hrsg.): Konstruktion und Retrieval von Wissen. Frankfurt, S. 14-21.

sung notwendige Wissen ins Internet gestellt. Die Studierenden müssen diese Lernaufgaben von Woche zu Woche lösen. Die Aufgabensoftware registriert die Studierenden und kumuliert ihre jeweils erreichte Punktzahl bis zum Ende des Semesters. Wer eine vorab bestimmte Punktzahl am Ende erreicht hat, hat die Lehrveranstaltung erfolgreich absolviert. Mit einem solchen technologisch gestützten Studienarrangement schafft man *workload* und *credits* auf Seiten der Studierenden, ohne auf Präsenzveranstaltungen und entsprechendes Personal angewiesen zu sein. Auch hier also liegt das Hauptaugenmerk auf der Arbeitseffizienz, ohne jedoch zu übersehen, dass man mit einem solchen Studienszenario einen hohen Grad an Standardisierung und Objektivität erreicht. Das mag nicht für jede Lehrveranstaltung bzw. jede Thematik möglich oder auch nur sinnvoll sein, ist aber gleichwohl eine Option.

4.2 Aufgaben: Struktur, Funktion, Sequenzen und Transfer

Aufgaben sind *interaktive Wissensarten*. In der Interaktion können zwei Konstruktionsfaktoren der Aufgabenstellung unterschieden werden:

- erstens wird Information gegeben, und zwar die Rahmeninformation für die Aufgabenstellung,
- zweitens werden offene Stellen produziert, so genannte Leerstellen, die vom Lernenden durch seine Aktivität auszufüllen sind.

Auf letztere Weise – durch Lösen der Aufgaben - werden Informationen in aktives Wissen transferiert, was nichts anderes heißt, als dass gelernt oder Wissen aktualisiert wird. Aufgaben haben in erster Linie die Funktion, Wissen zu generieren: Sie sind Lernaufgaben. In zweiter Linie besteht ihre Funktion darin, Wissen und Können zu überprüfen. In dieser zweiten Funktion sind Aufgaben Diagnosewerkzeuge. So können sie im Selbsttest dem Studierenden eine Rückmeldung über seinen Lernerfolg geben, in einem Abschlusstest feststellen, welche Lernziele erreicht wurden, oder sie können in einem Einstufungstest den Lernenden so weit informieren, dass er das für ihn richtige Kursniveau, das es zu wählen gilt, einschätzen kann.

Aufgaben sind die einzige bekannte Form, Kompetenzen zu entwickeln, zu aktualisieren und zu überprüfen. Welche Aufgaben welchen Kompetenzen entsprechen, ist noch nicht ausreichend erforscht. Unser diesbezügliches Wissen beruht auf ungesicherter professioneller Erfahrung. Die zentrale Forschungsaufgabe wird durch die Frage markiert, welche Lernoperationen entlang welcher Regeln und vor welchem Wissenshorizont vollzogen werden müssen, damit die Leerstelle in den Aufgabenformularen richtig gefüllt, die Aufgabe mithin gelöst wird.

4.3 Typologie von Prüfungsaufgaben

Abgesehen von ihrer Struktur, ihrem funktionalen Einsatz und ihrer Sequenzierung können Aufgaben unterteilt werden in *entdeckende Aufgaben*, *Ordnungsaufgaben*, *Antwortaufga-*

ben, Ankreuzaufgaben und Unterscheidungsaufgaben. Diese Unterteilung (vgl. Abb. 1) orientiert sich weitgehend an den kognitiven Operationen, die Lernende bei der Lösung von Aufgaben vollziehen müssen. Die damit vorgenommene Klassifikation der Aufgaben ist zum großen Teil unabhängig von den anderen in 4.2 genannten Unterscheidungsmerkmalen. Im Folgenden werden verschiedenen Aufgabenarten definiert; ebenso wird geprüft, inwieweit sie sich im Hinblick auf computergestützte automatische Auswertung eignen.

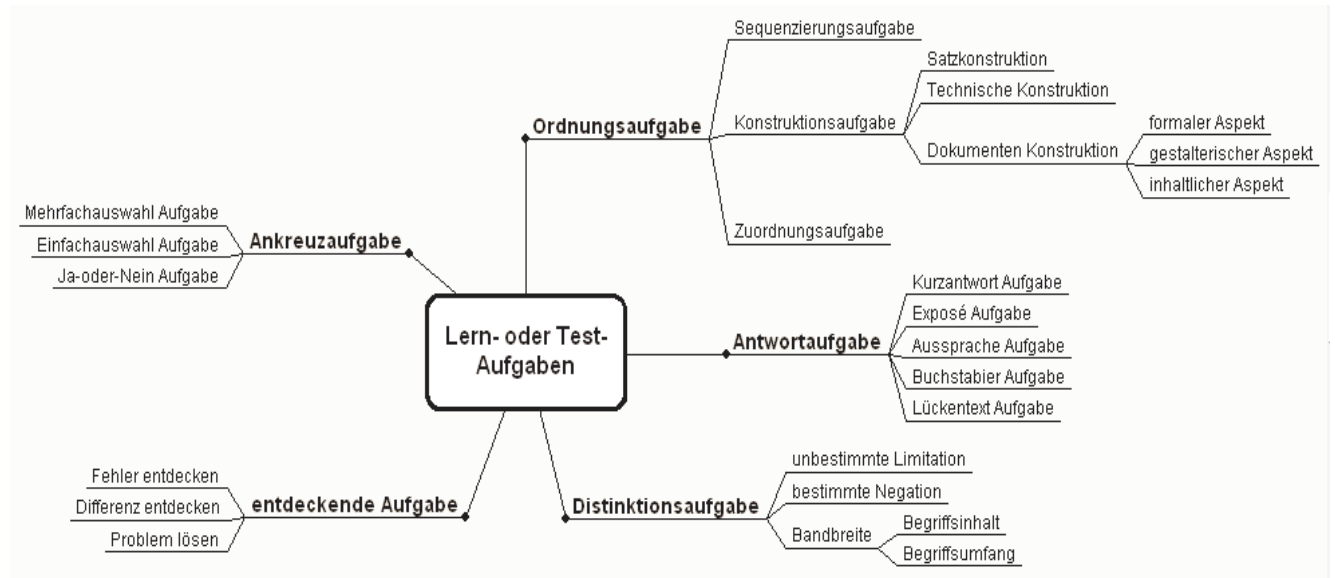


Abb. 1: Typologie der Aufgaben (Meder u.a. 2006, S. 75)

4.3.1 Entdeckende Aufgaben

In Aufgaben der Art *entdeckende Aufgabe* sollen Regeln, Gesetzmäßigkeiten oder Zusammenhänge in einem vorliegenden Material – auch über Fehler und Differenzen – identifiziert werden. Es geht dabei eher um kreative kognitive Prozesse. Die entdeckenden kognitiven Prozesse werden unterschieden in *Fehler entdecken*, *Differenz entdecken*, *Problem lösen*:

- *Fehler entdecken*: Der Lernende soll Fehler in einem Dokument, einem Bild oder einem Text – gelegentlich auch durch Vergleich – entdecken. Insoweit diese Fehler syntaktisch eindeutig zu bezeichnen sind – etwa durch Markierung oder sprachlich eindeutige Charakterisierung, kann mit automatisierter Auswertung gearbeitet werden. Wenn aber der Fehler nur in einer semantisch frei gestalteten Ausführung identifiziert werden kann, dann eignen sich keine derzeit bekannten *Software*prozeduren zur Auswertung, denn die Lösung muss dann in einem Freitext formuliert werden. Die wichtigste kognitive Operation, die mit diesem Aufgabentyp abgeprüft werden kann, ist die der Überprüfung auf Stimmigkeit, auf Passung bzw. auf Brüche und Differenzen. Dabei können Operationen des Vergleichens und Unterscheidens eine große Rolle spielen. Eine Aufgabe solchen Typs kann man z.B. sehr leicht konstruieren, indem man eine charakteristische Textpassage aus einer Theorie nimmt und durch Auswechseln eines *terminus*

technicus einen Fehler produziert, der dann vom Studierenden durch Mausklick auf das Wort identifiziert werden muss.

- *Differenz entdecken*: Unterschiede im vorliegenden Material sollen gefunden werden. Ein typisches Beispiel wird mit *Original und Fälschung* bezeichnet. Insoweit diese Differenzen syntaktisch eindeutig zu bezeichnen sind – etwa durch Markierung oder sprachlich eindeutige Charakterisierung – kann mit automatisierter Auswertung gearbeitet werden. Wenn aber die Differenzen nur in einer semantisch frei gestalteten Ausführung identifiziert werden können, dann eignen sich wiederum keine der derzeit bekannten *Software*prozeduren zur Auswertung. In solchen Aufgaben fordert man die Operationen des Vergleichens und Unterscheidens vor dem Hintergrund von Wissen, das als Distinktionsbasis fungieren kann, indem es die Aufmerksamkeit auf wesentliche Stellen, die den Fehler ausmachen könnten, richtet.
- *Problem lösen*: Mit Hilfe eines gegebenen Materials und gewisser Hilfsmittel soll ein beschriebenes Problem gelöst werden. Dabei können Probleme unterschieden werden, je nachdem ob sie *well-defined* (gut definiert bzw. gut bestimmt) oder *ill-defined* (schlecht definiert bzw. unterbestimmt) sind. *Well-defined* Probleme sind in jedem Falle lösbar, und ihre Lösung ist bei engen Variationen eindeutig. Von daher sind *well-defined* Probleme prinzipiell computerauswertbar. *Ill-defined* Probleme sind offene Probleme, deren Lösung nicht eindeutig und nicht klar ist, weil keine Hilfsmittel angegeben werden können. Von daher sind sie derzeit noch nicht computerunterstützt auswertbar. Ob die Forschung im Bereich Künstliche Intelligenz hier eine Lösung findet, ist nicht absehbar.

4.3.2 Ordnungsaufgaben

Ordnungsaufgaben werden unterschieden in *Sequenzierungsaufgabe*, *Konstruktionsaufgabe* und *Zuordnungsaufgabe*. Sie zielen auf das Verständnis ab, nach welchen Regeln und in welchen Strukturen etwas, das mit dem Lerngegenstand zu tun hat, angeordnet ist.

- *Sequenzierungsaufgaben* sind Aufgaben, in denen gegebene Gegenstände, Sachverhalte oder Begriffe in eine Reihe gebracht werden müssen, z.B.: eine Liste von Worten soll nach dem Alphabet sortiert werden. Anspruchsvoller ist etwa die Aufgabe, aus einer Liste von Wörtern einen hierarchisch geordneten Begriffsbaum zu erstellen oder aus einer Liste von Sätzen eine argumentative Kette zu bilden. Computergestützt wird das in der Regel so realisiert, dass die gegebenen Elemente mit der Maus an bestimmte Stellen des Bildschirms verschoben werden müssen. Diese Stellen können dann vom Programm ausgelesen und auf Richtigkeit geprüft werden. Es ist klar, dass die Lösung der *Software* vorgegeben werden muss.

Bei manchen Sequenzierungsaufgaben geht es um die Anwendung von Faktenwissen. Bei anderen anspruchsvolleren Aufgaben dieses Typs kann es auch noch um die Identifizierung bzw. Entdeckung der Regel der Sequenzierung gehen, wie bei der Aufforderung zur Ordnung folgender Elemente im Rahmen der Thematik anthropologischer Grundlegung der Erziehungswissenschaft:

Bauen Sie die folgenden Stichworte sachlich folgerichtig in eine Argumentation ein:

Triebüberschuss

Mangel an Instinkten

Bildsamkeit

Teilweise funktionsloser Körperbau (Seh-Hand-Feld)

Offenheit

oder

Immanuel Kant beschreibt vier Erziehungsziele, die mit Stufen der Vernunftentwicklung verbunden sind. Bauen Sie diese Erziehungsziele sachlich folgerichtig im Sinne Kants in eine Argumentation ein.

Kultivierung

Disziplinierung

Moralisierung

Zivilisierung

- *Konstruktionsaufgaben* sind Aufgaben, in denen aus Teilen das Ganze zusammengesetzt werden soll. Die *Konstruktionsaufgabe* wird üblicherweise weiter in *Satzkonstruktion*, *Technische Konstruktion* und *Dokumenten-Konstruktion* differenziert.
 - Satzkonstruktionsaufgaben nennt man solche, in denen ein Satz z.B. aus den Satzgliedern zusammengesetzt werden soll oder aus einzelnen Worten unter Berücksichtigung der richtigen Beugung der Wortarten ein grammatikalisch richtiger Satz zu bilden ist. Lässt sich aus dem vorgegebenen Wortmaterial nur *ein* sinnvoller Satz bzw. eine endliche Menge von Sätzen bilden, dann kann auch dieser Aufgabentyp automatisch ausgewertet werden. Im Kontext des Schreib-Lese-Lernkomplexes werden solche Aufgaben in der Grundschule gern eingesetzt. Im Studierendenkontext tauchen sie nur beim Fremdsprachenstudium auf – durchaus auch bei gehobenen Stilübungen.
 - Ist die Aufgabe eine technische Konstruktion, so wird die Zusammensetzung einer Maschine, eines Bauwerkes oder ähnlichem aus seinen Einzelteilen, den Bauelementen gefordert. Z.B. könnte eine solche Aufgabe lauten: Aus welchen Teilen setzt sich ein PC wie zusammen? Die Teile – Rechner, Bildschirm, Maus, diverse Kabel etc. – könnten bildlich vorgegeben sein und müssen mit der Maus verschoben und kombiniert werden. Auch dies lässt sich dann automatisiert auswerten. Im Kontext der Erziehungswissenschaft kann dieses Aufgabenformat für die Planung und Konstruktion einer Unterrichts- oder Weiterbildungseinheit eingesetzt werden, aber auch bei der Konstruktion einer Falllösung im Kontext sozialpädagogischer Fallarbeit und Fallkonferenzen.
 - *Dokumenten-Konstruktion* ist eine Aufgabe, die das Erstellen eines Textes unter Berücksichtigung der formalen, gestalterischen und/oder inhaltlichen Vorschriften verlangt: Es soll z. B. ein Stunden- oder Versuchsprotokoll nach einem vor-

gegeben Muster angefertigt werden. Hier ist der Formularcharakter besonders deutlich. Wenn die zu ordnenden Textteile vorgegeben sind, dann kann auch hier automatisiert ausgewertet werden. Müssen die Textteile aber auch erstellt, oder gegebene Textteile noch angepasst werden, dann stößt die computergestützte Auswertung an ihre Grenzen, und die Aufgabe wird in ihren Bestandteilen zur Freitextaufgabe (siehe auch dort). Typische erziehungswissenschaftliche Dokumenten-Konstruktionsaufgaben sind im Bereich der Unterrichts- und Seminar-Planung – allgemein in der Veranstaltungsplanung gegeben, wenn gefordert ist, dass man nach einem vorgegebenen Muster zu verfahren hat. - Mit solchen Aufgaben testet man kognitive Operationen, die komplex ineinander greifen: Operationen in der Teil-Ganzes-Relation sowie in der Träger- (ist-Grundlage-für) sowie Bestandsrelation (besteht-aus).

Zuordnungsaufgabe: Elemente zweier Mengen werden einander nach vorgegebenen Kriterien zugeordnet. Das können z.B. in der einen Liste Bilder sein, in einer zweiten Liste Begriffe. Anspruchsvoller und dem studentischen Niveau angemessen könnte es sein, einer Liste von Thesen Elemente aus einer Liste von Argumenten zuzuordnen oder Elemente einer Liste von Thesen in historische Epochen einzuordnen. Sind die jeweiligen Listen vorgegeben, dann kann automatisiert ausgewertet werden.

Aufgabenbeispiele:

Ordnen Sie folgende Aussagen den drei Wissensformen – Alltagswissen, Professionswissen, wissenschaftliches Wissen – nach Peter Vogel zu:

AW	PW	WW	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es wird erworben durch alltägliche Internalisierung/ Sozialisation
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es dient Angehörigen pädagogischer Berufe zum kompetenten Handeln im Beruf
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es wird begrenzt durch seinen prinzipiell vorläufigen (hypothetischen) Charakter
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es wird erworben durch wissenschaftliche Ausbildung und Erlernen beruflicher Handlungsschemata und -routinen
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es dient dem Wissens- und Erkenntnisfortschritt auf dem Gebiet der Erziehung
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es wird begrenzt durch negative Lösungseffekte und die Zuständigkeit anderer Professionen
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es dient Laien zur Bewältigung alltäglicher pädagogischer Probleme
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es wird erworben durch wissenschaftliches Studium (Lehre und Forschung)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Es wird begrenzt durch negative Lösungseffekte

oder:

Ordnen Sie die nachfolgenden Zitate jeweils einer der in der Vorlesung behandelten historischen Epochen bzw. „Denkformen“ der Pädagogik zu!

- A: Aufklärungspädagogik
- B: Neuhumanismus
- C: Reformpädagogik
- D: NS-Pädagogik
- E: Kritische Erziehungswissenschaft
- F: Konstruktivistische Pädagogik

Es folgen 13 längere, bis zu 10 Zeilen umfassende Originalzitate.

4.3.3 Antwortaufgaben

Antwortaufgaben werden differenziert in *Kurzantwort-Aufgabe*, *Exposé-Aufgabe*, *Aussprache-Aufgabe*, *Buchstabier-Aufgabe* und *Lückentext-Aufgabe*.

- In *Kurzantwort-Aufgaben* werden freie Antworten verlangt, die im Allgemeinen nicht länger als ein bis zwei Sätze sind. Wird beispielsweise eine eindeutige Definition verlangt, wie das etwa im Kontext der Wissenschaftstheorie und der Methodologie empirischer Forschung gegeben sein kann, dann kann auch hier automatisiert ausgewertet werden. Probleme gibt es hier, wenn sprachliche Fehler gemacht werden. Das könnte eine intelligente Auswertungssoftware mit Rechtschreib- und Grammatikkorrektur auffangen. Wenn dies nicht oder nur unvollständig geht, dann müssen solche Aufgaben von Hand nachkorrigiert werden. Ein Vorteil bleibt dabei dennoch insofern, als nur die von der Software als falsch gewerteten Lösungen überprüft und gegebenenfalls in ihrer Wertung geändert werden müssen. Bei diesem Aufgabenformat ergeben sich die zu prüfenden Lernoperationen aus der jeweiligen Fragestellung. Man sieht leicht, dass solche Prüfungsfragen mit ein bis zwei Sätzen zu beantworten sind.

Beispiele einer solchen Aufgabe können sein:

Benennen Sie drei Gütekriterien eines wissenschaftlichen Testes!

oder:

Mit der beginnenden Neuzeit und vor allem im 18. Jahrhundert wurden die Forderungen nach einer systematischen Erziehung der nachwachsenden Generation in Familie und Schule immer zahlreicher. Zählen Sie wichtige gesellschaftliche Veränderungen bzw. Probleme auf, die zu solchen Forderungen führten.

(1) _____

(2) _____

(3) _____

(4) _____

- In *Exposé-Aufgaben* wird ein Kurzaufsatz von ein bis drei Seiten verlangt, der die Problem- oder Aufgabenlösung knapp skizziert. Persönliche praktische Erfahrung hat gezeigt, dass hier die automatisierte Auswertung scheitert. Es soll zwar schon in der Linguistik entwickelte Programme geben, die die Qualität solcher Texte ziemlich valide einschätzen können, es ist aber keine Prüfungssoftware bekannt, in der ein solches Programm implementiert ist. Norbert Meder hat in einer Eigenentwicklung¹² den Versuch einer Teilauswertung gemacht. Bei der Essay-Aufgabe „Explizieren Sie das Konzept der Bildung bei Schiller!“ kann man in der Software hinterlegen, dass folgende Wörter vorkommen müssen: Stofftrieb, Formtrieb, Spieltrieb, Schönheit, Freiheit, Erscheinung, Ewigkeit, lebendige Zeit. Die Software prüft, ob diese Wörter im Text vorkommen und vergibt dabei proportional maximal 50% der erreichbaren Punkte. Die Nachkorrektur von Hand vergibt die restlichen 50% der möglichen Punkte nach Maßgabe der sprachlichen und logisch-argumentativen Form, der Gedankenführung usw. Neben dieser Teilerleichterung, die zugleich auch eine gewisse Objektivierung ausmacht, erspart die Tatsache, dass der Text getippt vorliegt, etwa 50-75 Prozent Korrekturzeit gegenüber handschriftlichen Texten. Das wurde im Selbsttest geprüft. Zu diesem Aufgabentypus können auch solche dreiteiligen Aufgabenformate gerechnet werden, in denen zunächst (a) ein komplexes Problem oder ein Fall dargestellt wird, dann (b) Fragen und Aufgabenstellungen zu diesem Fall formuliert werden und dann (c) die als Freitext formulierte Aufgabenlösung teils per Computer, teils durch den Lehrenden selbst beurteilt wird.

¹² Siehe www.exam2go.de. Vgl. auch Homeister, D./ Ziegler, S.: Verständnisfrage. Freitextantworten mit neuronalen Netzen auswerten. In: Linus-Magazin 4 (2008), S. 104-112.

Erläutern Sie (auf zwei handschriftlichen Seiten) das Vorgehen der Unterrichtsforschung im Rahmen des Prozess-Produkt-Paradigmas – und verdeutlichen Sie demgegenüber die Besonderheiten des Angebot-Nutzungs-Modells.

Vorab konstruierter Erwartungshorizont: Im Rahmen des Prozess-Produkt-Paradigma (PPP) versucht die empirische Unterrichtsforschung, den Zusammenhang zwischen Schul- und Unterrichtsbedingungen einerseits und den kurz- und längerfristigen Lernergebnissen der Schüler andererseits zu klären. Dabei werden Lehrer-, Schüler und Schulvariablen sowie die konkreten Abläufe des Unterrichts empirisch möglichst breit erfasst (= Prozess). Zugleich werden bestimmte Lernergebnisse bzw. Auswirkungen des Unterrichts erfasst (=Produkte). Auf dieser Basis wird eine korrelative Beziehung (nicht: Kausalität) zwischen Bedingungen bzw. Prozessen und den Produkten hergestellt. Entscheidend ist, dass messbare Lernergebnisse und z.T. auch Sozialisationswirkungen das Erfolgskriterium von Schule und Unterricht sind. Was nicht messbar ist, kann nicht in dieses Modell eingehen. Alles in allem geht das PPP von einer mehr oder weniger starken Determinationskraft des Prozesses auf die Produkte aus, wenngleich nur Korrelationen errechnet, nicht aber Kausalitäten experimentell geprüft werden.

Das Angebots-Nutzungs-Modell basiert auf dem PPP. Es unterscheidet sich aber in einem wesentlichen Punkt: Dem Verständnis des Lernens der Schüler sowie dem Verhältnis von Lehren des Lehrers und Lernen der Schüler. Im Rahmen dieses Modells ist Unterricht eine vom Lehrer bereit gestellte Gelegenheit zum Lernen. Es ist ein Angebot, das Schule/Lehrer machen. Entscheidend ist nun, wie Schüler darauf eingehen, es also nutzen. Lernen ist ein selbsttätiger Akt, er kann nicht ‚erzwungen‘ werden. Schüler nutzen das Unterrichtsangebot u in individueller und je unterschiedlicher Weise; sie sind auf diese Weise auch mitverantwortlich für das Zustandekommen von Wirkung des Unterrichts. Das bedeutet nicht, dass der Lehrer nunmehr keine Verantwortung mehr hat für das Zustandekommen von Wirkungen: Sein Angebot muss so gestaltet sein, dass es auch möglichst angenommen wird.

Wichtige Elemente aus der Musterantwort - bis zu 6 - müssen genannt werden; hierfür gibt es maximal 6 Punkte.

4.3.4 Ankreuzaufgaben

Ankreuzaufgaben, wie sie in klassischen *Multiple-choice*-Tests vorkommen, werden in *Mehrfachauswahl*-, *Einfachauswahl*-, und *ja/ nein-Aufgabe* unterschieden. In allen Fällen hat der Lernende aus Antwortvorgaben auszuwählen.

- Im ersten Fall der Mehrfachauswahl sind mehrere richtige Antworten aus vorgegebenen Antworten auszuwählen.
- Im zweiten Fall der Einfachauswahl ist nur eine unter den vorgegebenen Antworten richtig.
- Im dritten Fall muss entschieden werden, ob eine oder mehrere Aussagen richtig oder falsch sind.

Alle drei Fälle prüfen letztlich Urteilskraft ab – also ein sehr anwendungsbezogenes Wissen. Der Studierende muss in allen drei Formaten Wahr-Falsch-Entscheidungen treffen. Jenseits von einfachem Raten ist dies nur möglich, wenn der oder die Studierende Kontextwissen als Distinktionsbasis für die Wahr-Falsch-Entscheidungen aktualisiert. Zur

Vermeidung von Raten muss insbesondere bei der Mehrfachauswahl darauf geachtet werden, dass alle Antwortalternativen – auch die falschen – nahe an der Wahrheit liegen.

Beispiele:

Wodurch ist nach Giesecke *pädagogisches Handeln* gegenüber anderen sozialen Handlungsformen (medizinisches, politisches, administratives Handeln) bestimmt?

- Weil es gesellschaftliche Probleme löst.
- Weil es das Ziel verfolgt, Gesundheit langfristig zu erhalten.
- Weil unter der Perspektive des lebenslangen Lernens schnelle Veränderungen auf dem Arbeitsmarkt möglich sind.
- Weil es das menschliche Zusammenleben vereinfacht.
- Weil es unter der Maßgabe betrieben wird, Lernen zu ermöglichen.

In der Hermeneutik...

- ... sind Verstehen und Erklären eins.
- ... geht es um Naturgesetze.
- ... geht es um Bedeutung.
- ... betreibt man die „Kunst der Auslegung“.

Die hermeneutische Differenz ...

- ... pendelt zwischen Verstehendem und dem vom Autor Gesagten.
- ... gehört nicht notwendig zur hermeneutischen Situation.
- ... ist umso größer, je komplexer der Text ist.
- ... bewegt sich zwischen Vorverständnis und Textverständnis.

In der empirischen Lehrerforschung unterscheidet man mindestens drei Paradigmen: das Persönlichkeitsparadigma, das Prozess-Produkt-Paradigma und das Expertenparadigma. Welche der folgenden Aussagen zu diesen Paradigmen treffen zu?

- Das Persönlichkeits-Paradigma umfasst die Suche nach allgemein gültigen Merkmalen der positiven Lehrerpersönlichkeit.
- Das Persönlichkeitsparadigma untersucht insbesondere die Interaktion zwischen Lehrer und Schülern.
- Im Prozess-Produkt-Paradigma gelten Schülerleistungen als primärer Indikator für den Erfolg der Arbeit des Lehrers, denn gleiche Unterrichtsbedingungen führen zu immer gleichen Ergebnissen.
- Im Rahmen des Prozess-Produkt-Paradigmas versucht man, die Wirkungen einzelner Verhaltensmuster und Fertigkeiten der Lehrenden auf wesentliche Aspekte des Schülerverhaltens zu analysieren.
- Beim Expertenparadigma wird der Blick ausschließlich auf das handlungsleitende Wissen und Können des Lehrers gerichtet.
- Das Expertenparadigma folgt der Einsicht, dass dem Handeln erfolgreicher Lehrer professionseigenes Wissen zugrunde liegt.
- Nur das Expertenparadigma geht von der Trainierbarkeit positiven Lehrerverhaltens aus.

4.3.5 Distinktions-/Diskriminationsaufgaben

Eine Aufgabe der Art *Distinktionsaufgabe* – oft auch *Diskriminationsaufgabe* genannt – bezieht sich im Wesentlichen auf Begriffslernen. Differenziert wird danach,

- ob die Unterscheidung über eine *unbestimmte Limitation* des Begriffes erfolgt. Man grenzt den Gegenstand dann von seinem Kontext und von anderen (materiellen oder symbolischen) Gegenständen durch unbestimmte Negation ab, d.h. ohne seine bestimmenden Merkmale zu nennen. Beispiel: Die Seele ist nichts Materielles.
- ob die Unterscheidung über eine *bestimmte Negation* des Begriffes erfolgt. Hier untersucht man die Eigenschaften (Merkmale) der Gegenstände, anhand derer sie identifiziert und von ähnlichen, aber andersartigen Gegenständen abgegrenzt werden. Die natürlichen Zahlen (die Null und die positiven ganzen Zahlen) sind nicht die negativen ganzen Zahlen (die nicht-negativen Zahlen).
- ob die Unterscheidung über die explizite Bestimmung des Begriffes erfolgt. Die explizite Bestimmung kann erfolgen
 - durch die Angabe des *Begriffsumfangs*, indem die Gegenstände benannt und aufgezählt werden, die unter den Begriff fallen, und
 - durch die Angabe des *Begriffsinhalts*, indem die Merkmale des Begriffes, das sind die gemeinsamen Eigenschaften der Gegenstände, die unter ihn fallen, aufgelistet werden.

Die Distinktionsaufgabe wird hier nur der angestrebten Vollständigkeit halber aufgeführt. Dem Format nach kann sie in den bisher beschriebenen Formen gestellt werden. Gesonderte Erwähnung findet sie hier auch deshalb, weil sie Begriffsverständnis überprüft, was für wissenschaftliches Wissen zentral ist.

Alle hier in einer Typenhierarchie angeordneten Aufgaben lassen sich ins Internet stellen und auch weitgehend automatisch auswerten. Ausgenommen von einer automatischen Auswertung sind die Kurzantwort- und die Exposé-Aufgabe. In diesen beiden Fällen muss die versuchte Aufgabenlösung von einem Tutor oder Korrektor beurteilt werden.

4.4 Die automatisierte Auswertung von Klausuren

Die meisten Internet-Lernplattformen unterstützen die Konstruktion von Aufgaben. Sie unterscheiden sich darin, welche und wie viele Aufgabenformate man nutzen und wie benutzerfreundlich man Aufgaben anlegen kann. Das ist wichtig, weil es beim Aufbau einer Klausur auf die Variation und die Vielfalt der Aufgaben ankommt. Denn die verschiedenen Aufgabenformate verlangen auch verschiedene kognitive Operationen, wie weiter oben gezeigt wurde. Aber noch wichtiger als dies ist die Auswertung der Aufgaben und der Klausur. Hier gibt es gravierende Unterschiede bei den in Frage kommenden *Softwarepaketen*. Das fängt bei der Speicherung der Eingabe der Studierenden an: Werden beispielsweise Aufgaben nur zum Selbsttest unterstützt, dann bleibt häufig eine dauerhafte Speicherung aus. Man kann dann zwar eine Aufgabe wiederholen, aber nicht mehr auf frühere Ergebnisse zurückgreifen. Bei Lernaufgaben verzichtet man manchmal darauf, die Lösungen

auf dem *Server* zu speichern, sondern speichert nur auf dem lokalen Rechner. Das scheint in diesem Fall zu genügen, hat aber Nachteile: Man kommt nicht von jedem Rechner aus an seine Lösungen. Das schränkt gerade studentisches Arbeiten erheblich ein.

Für den Klausurbetrieb ist die sofortige Speicherung auf einem *Server* unerlässlich. Wenn beispielsweise der lokale Arbeitsrechner eines Klausurteilnehmers abstürzt, dann darf nichts verloren sein, und der Teilnehmer muss an einem anderen Rechner sofort weitermachen können. Das zeigt zugleich ein anderes praktisches Problem: Klausurräume müssen immer für einen solchen Fall auch mit Reserve-Rechnern ausgestattet sein.

Am meisten muss man bei der Wahl einer Klausur*software* darauf achten, was sie mit Bezug auf die Auswertung leistet, denn in solchen Leistungen liegt die Unterstützung der Arbeit des Hochschullehrers begründet. Auf der Basis langjähriger praktischer Erfahrungen ist es unerlässlich, dass ausnahmslos immer die Möglichkeit der Nachkorrektur einer Klausur durch den Veranstalter, als von Hand, gegeben sein muss. Dies schließt eine Änderung der automatisch vergebenen Punkte ein. Das scheint trivial, wird aber nicht immer unterstützt. Eine weitere Forderung besteht darin, die Bepunktung nachträglich ändern zu können. Zeigt sich beispielsweise, dass eine Aufgabe zu schwer oder schlecht gestellt war, dann muss man sie aus der Bewertung herausnehmen oder ihren Punktwert – und damit ihr Gewicht – verringern können.

Die endgültig korrigierte und bewertete Klausur eines Teilnehmers muss in einem vernünftigen, Papier sparenden Format auszudrucken sein – sei es zur Archivierung oder zur Einsicht für den Studierenden. Das gilt auch für die Punktlisten, die der Hochschullehrer als Übersichten braucht.

5. Bewertungsinstrument Korrekturraster – ein Beispiel

Entsprechend dem Grundprinzip modularisierter Studiengänge werden Lehren und das Prüfen an Kompetenzen orientiert, die vermittelt werden sollen bzw. deren Vorliegen in Rahmen von Prüfungen ermittelt werden soll. Die eingeforderten Kompetenzen sind in der Regel in den Modulhandbüchern ausgewiesen. Sie können in Raster übertragen werden. Kompetenzraster ermöglichen eine praktikable Bewertung der zu vermittelnden und zu prüfenden Kompetenzen anhand und entlang von Inhalten. Das folgende Beispiel bezieht sich auf den Bachelor-Studiengang „Erziehungswissenschaft“ am Institut für Berufs- und Weiterbildung im Fachbereich Bildungswissenschaften der Universität Duisburg-Essen. Im engeren Sinne wird auf das Einführungs-Modul 1 im ersten Semester Bezug genommen. In seinem Kontext soll das Grundprinzip einer Arbeit mit Kompetenzrastern erläutert werden.¹³ Das Modulhandbuch sieht für das Modul 1 drei aufeinander bezogene Veranstaltungen vor:

- Eine Vorlesung, die in die Studiengrundlagen der Erziehungswissenschaft einführt.

¹³ URL der Essener Bachelor-Studienstruktur: http://biwi.uni-duisburg-essen.de/sites/biwi/files/UDE_20-20Studienverlaufsplan_20Bachelor_20EW.pdf; heruntergeladen am 25.1.2008, 23.12h.

- Ein auf der Arbeit an Texten gegründetes Seminar, das zu erläuternde und zu problematisierende Inhalte der Studiengrundlagen vertieft, und ein
- Tutorium, in dem Aufgaben aus dem Proseminar sowie Studientechniken (zitieren, recherchieren, argumentieren, Gruppenarbeit, Kurzreferate etc.) an den Inhalten eingeübt werden.¹⁴

Die für das Einführungsmodul zu erwerbenden Kompetenzen sind Grundkenntnisse, Orientierungswissen und Studientechniken. In der Vorlesung wird Orientierungswissen vermittelt, im Seminar werden Grundbegriffe erörtert und vertieft, und im Tutorium werden Studientechniken einübt. Mit „*Orientierungswissen*“ ist ein Wissen gemeint, das man braucht, um sich in der Welt bzw. auf einem bestimmten Gebiet zurechtzufinden, ohne es schon in spezifischer Weise ganz erklären zu können und in ihm tätig werden zu können. Wer Orientierungswissen hat, weiß mehr oder weniger, dass es den betreffenden Sachverhalt gibt, er weiß aber nicht unbedingt, etwas damit anzufangen.¹⁵ Es geht wie bei den beiden anderen Qualifikationszielen (Grundkenntnisse, Studientechniken) dabei um echtes Einführungswissen. Problembestände der Erziehungswissenschaft sollen als solche wahrgenommen werden, theoretische Lösungen deuten sich vielleicht schon an, es besteht aber im ersten Semester noch kein breites Zusammenhangswissen. „Grundkenntnisse“ meint demgemäß ein basales Wissen um Gegenstände und Inhalte der Erziehungswissenschaft. Es fasst „Erklärungswissen (Know-Why)“, „Quellenwissen (Know-Where)“ und eine bestimmten Art von „Handlungswissen (Know-How)“ zusammen. Wenn dann in anderen Studienkontexten pädagogisches Handeln, also z. B. Beraten, angesprochen wird, liegen im Rahmen eines ersten Zugreifens Grundkenntnisse und ein Wissen um Zusammenhänge vor. „Studientechniken“ meinen eher technische Handlungsformen wissenschaftlichen Handelns.

Die drei Wissensarten – Grundkenntnisse, Orientierungswissen und Studientechniken – dienen der Kategorisierung von zu beurteilenden Kompetenzen. Das heißt, sie haben, um mit Chomsky zu sprechen, den Charakter von Regelwerken. Grundkenntnisse umfassen etwa die Regeln von wissenschaftlichen Erklärungen, Modellen, Ursache-Wirkungs-, Zweck-Mittel-Verhältnissen u. ä. Aber sie betreffen auch Regeln des Umgangs mit verschiedenen Quellen, die Inhalte enthalten, und Regeln zu Praktiken, Techniken und Strategien pädagogischen Handelns. Das Regelwerk der Studientechniken betrifft einerseits z. B. die Regeln des Zitierens, des Recherchierens, des Argumentierens. Andererseits sind auch Regeln verschiedener Arbeitsformen (Einzel-, Partner-, Gruppenarbeit) sowie unterschiedlicher Präsentationsformen (Kurzreferat, Powerpoint-Präsentation, Hausarbeit) gemeint. Das Regelwerk des Orientierungswissens ist das eines Wanderns durch ein assoziatives Netz erziehungswissenschaftlicher Theoriemotive und Theoreme.

¹⁴ Vgl. URL: [http://biwi.uni-duisburg-essen.de/sites/biwi/files/UDE_20-20Bachelor_20EW_20Modulhandbuch_20\(M_C3_A4rz_202007\).pdf](http://biwi.uni-duisburg-essen.de/sites/biwi/files/UDE_20-20Bachelor_20EW_20Modulhandbuch_20(M_C3_A4rz_202007).pdf), S. 1.

¹⁵ Vgl. Meder u.a. (2006): Web-Didaktik...a.a.O.

Als eine Prüfungsleistung ist neben der regelmäßigen Teilnahme und dem als Übung angelegten Verfassen einer Hausarbeit für das erste Semester eine 90minütige *Abschlussklausur* vorgesehen. Sie prüft Inhalte aller drei Veranstaltungen ab. Zur Erklärung des Erstellens und Umgehens mit einem Kompetenzraster soll nun diese Klausur und aus ihr *eine* mögliche Fragestellung herausgegriffen werden. Als Frage sei beispielhaft genannt: *Erörtern Sie den Bildungsbegriff nach Anton Hügli und Hartmut v. Hentig.*¹⁶ Ihre Bewertung kann nach folgendem Raster erfolgen (vgl. Abb. 2; s.u.).

Die drei beschriebenen Kompetenzen in der ersten Spalte sind bezogen auf die Klausurfrage zu konkretisieren. Grundkenntnisse sollen sich hier erweisen, indem Erziehungswissenschaftliche Grundbegriffe erläutert werden. Für die zu Grunde gelegte Klausurfrage und die dahinter stehenden Texte sind das beispielsweise Kontrollpädagogik, Autonomiepädagogik, Bildung und Schule. Orientierungswissen soll sich auf pädagogische Handlungsmodalitäten beziehen, d. h. darauf, dass es Erziehen, (Sich-)Bilden, Zwingen, (Sich-)Befreien, Geschichten-Erzählen, (Gespräche-)Führen, Sprache(n)-Sprechen, Theaterspielen, Naturerfahrungen-Machen, Sich-politisch-Betätigen, Arbeiten, Feste-Feiern, Musik-Machen und Aufbrechen als Möglichkeiten in pädagogischen Zusammenhängen gibt.

Die hier ausgewählten Handlungsmodalitäten sind aus den Texten generiert, genauer: Erziehen, (Sich-)Bilden, Zwingen und (Sich-)Befreien sind aus dem Hügli-Text abgeleitet, Geschichten-Erzählen, (Gespräche-)Führen, Sprache(n)-Sprechen, Theaterspielen, Naturerfahrungen-Machen, sich-politisch-Betätigen, Arbeiten, Feste-Feiern, Musik-Machen und Aufbrechen aus dem v. Hentig-Text. Sie sind im ersten Semester noch nicht Strategien der eigenen professionellen Handlungsweise, aber als Potenzial von Pädagogen bekannt. Außerdem sind Einordnungen in die Wissenschaftsgeschichte bzw. Disziplingeschichte oder Geschichte der pädagogischen Einrichtungen und Institutionen im ersten Semester ebenfalls dem Orientierungswissen zugeordnet. Für die Klausurfrage sind sie hinsichtlich der bildungsgeschichtlichen Bedeutung griechischer und aufklärerischer Pädagogen zu konkretisieren. Immanuel Kant soll im Kontext der Begründung von Pädagogik als Wissenschaft bekannt sein. In gleicher Weise sollen die Studierenden Hartmut von Hentig der Laborschule zuordnen können. Sie sollen wissen, dass die Laborschule im Zusammenhang von Schulreform, also Institutionsreform anzusiedeln ist. Derartiges werden sie in der Vorlesung gehört und im Seminar vertieft haben.

Die Studententechniken sind unterteilt in die Kompetenzen des Bibliographierens, Zitierens, Lesens und Gliederns wissenschaftlicher Texte und des Verfassens von Essays. Mit den bloßen Techniken hat das Tutorium die Studierenden bekannt gemacht. Sie haben sie in einer Hausarbeit mit Übungscharakter im Seminar angewandt und sind dabei im Tutorium unterstützt worden. So wird erwartet, dass Erlerntes auch bei der Bearbeitung der Klausuraufgabe gezeigt werden kann.

¹⁶ Die Textgrundlagen hierfür sind: Hügli, A.: Was die Pädagogik vor ihrer Geburt schon war und warum sie es nicht geblieben ist: Philosophie als „Bildungslehre des Menschen“. In: ders.: Philosophie und Pädagogik. Darmstadt 1999, S. 22-52 und Hentig, H. von: Was bildet den Menschen? In: Kieler Entwürfe für Schule und Kirche. 16/1996. S. 3-13.

In den Spalten drei bis sechs sind vier bzw. fünf Kompetenzstufen aufgeführt. Sie spiegeln den Erwartungshorizont hinsichtlich der Klausurfrage wider. Die Stufung entspricht Graden der Erfüllung zwischen Minimal- und Maximalerwartung. Die Kompetenzstufen steigern sich horizontal und zwar sukzessive von der Note „ausreichend“ (4) bis hin zu der Note „sehr gut“ (1) und darüber hinaus (1+). Dabei bezeichnen die Notenstufen folgende graduelle Unterschiede:

ausgezeichnet	(1+):	spielerische Neukombination mehrerer Theorien, eigener Standpunkt, und komplexe Anwendung vernetzter Theorie auf die Praxis
sehr gut	(1):	reflektierende Bezüge zwischen zwei Theorien, die benannt werden und mit denen argumentativ umgegangen wird und die Bezug auf Praxisprobleme aufweisen
gut	(2)	reflektierende Bezüge innerhalb einer Theorie, die benannt werden und mit denen argumentativ umgegangen wird und die Bezug auf Praxisprobleme aufweisen
befriedigend	(3):	Bezüge innerhalb einer Theorie, die benannt werden, aber weniger reflektiert und nicht argumentativ genutzt werden, die aber Bezug auf Praxisprobleme aufweisen
ausreichend	(4):	reine Reproduktion, d. h. Elemente und Bezüge werden genannt, aber nicht reflektiert

Unterhalb von Ausreichend (5 und 6) liegende Leistungen gelten als nicht bestanden.

Für die Bewertung des konkreten Aufgabenbeispiels lässt sich festhalten: Im Vergleich von Studententechniken und inhaltlichen Kompetenzen haben die Inhalte höheres Gewicht. Deshalb reichen im *technischen Kompetenzbereich* in den Graden korrekte Anwendung (1), wenige/marginale Fehler (2), Fehler (3) und schwerwiegende Fehler (4) letztere noch für ein „Ausreichend“ aus. Verständlichkeit muss allerdings wenigstens gewährleistet sein. Gewichtungen werden sich in diesem Aspekt besonders im direkten Vergleich von Leistungen ergeben. Für die *Wissens-Kompetenzen* ergibt sich: In der ersten Kompetenzstufe sind wenigstens Grundbegriffe, Handlungsmodalitäten und historische Darstellungen zu benennen. Das bedeutet, die Studierenden sollten mindestens wissen, dass es Hügli und v. Hentig etwa um Kontroll- und Autonomiepädagogik, Erziehen und (Sich-)Bilden oder um Schulreform geht. Das sind Minimalkenntnisse, die für ein reines Bestehen vorausgesetzt werden. Für eine befriedigende Leistung sollte z. B. das Verhältnis von Kontroll- und Autonomiepädagogik im Hügli-Text dargestellt werden können. Das wären hier Bezüge innerhalb einer Theorie. Grundständige Reflexion wird im Unterschied zur Kompetenzstufe 1 in der 2. Kompetenzstufe erwartet. Auch soll für eine „3“ auf pädagogische Praxis und ihre Problemlagen rekurriert werden. Diese Erwartung ist der Kompetenzstufe 2 mit den Kompetenzstufen 3 und 4 gemeinsam. Für ein „Gut“ sollte die Reflexions-Komponente deutlich ausgeprägt sein und die Argumentation bestimmen. Um ein „Sehr Gut“ zu erhalten, sollte sich die Reflexion und Argumentation auf beide gefragte Texte bei Gewichtungen gleichmäßig verteilen. Wer weitere Theorien einbezieht und aus deren Neukombinationen einen eigenen Standpunkt gewinnt und wem es in diesem Zusammenhang sogar gelingt,

Theorie und Praxis zu vernetzen, der bewegt sich oberhalb der 4. Kompetenzstufe, also oberhalb der Note „1“.

Eine solche Leistung ist für das Einführungsmodul in höchstem Maße unwahrscheinlich. So erscheint es umgekehrt als wahrscheinlich, dass im ersten BA-Semester durchschnittlich die 2. Kompetenzstufe (Note 3) erreicht wird. Folglich lässt sich diskutieren, ob die Kompetenzstufen noch feiner zu differenzieren sind, um für breite Studierendengruppen sowohl das Bestehen als auch das Erreichen der oberen Notenbereiche zu gewährleisten. Dies heißt flexiblere Notenschemata bei quantitativer horizontaler Differenzierung der Erwartungshorizonte. Das ist die eine Möglichkeit. Die andere besteht darin, mit qualitativen vertikalen Graden zu arbeiten. Etwa: Im Anfangssemester sind Studientechniken noch Lerngegenstand, also Fehler noch akzeptabel. Im weiteren Studienverlauf werden die Techniken aber mehr und mehr vorausgesetzt, demnach Fehler zunehmend inakzeptabel. Oder: In der Studieneingangsphase haben Kenntnisse noch basalen Charakter. Da ist der größte Teil des Wissens noch Orientierungswissen. Mit wachsender Semesterzahl wird ein immer dichter werdendes Netz aus Gewusst-Warum (Erklärungswissen), Gewusst-Wo (Quellenwissen) und Gewusst-Wie (Handlungswissen) erwartet. Das heißt, die Qualität der erwarteten Kompetenzen wird sich verändern. Wer im ersten Semester einführendes Orientierungswissen sammelt und zu vernetzten beginnt, ist gut. Wer im sechsten Semester immer noch auf dem Niveau von Orientierungswissen steht, wird durchfallen. Solches qualitativ vertikales Vorgehen wird für den Studiengangsverlauf vor reiner quantitativ-horizontaler Binnendifferenzierung empfohlen. Dies entspricht ja den Kompetenzen, die in den Modulhandbüchern im Vergleich von erstem und sechstem Semester festgelegt sind, z.B. in Duisburg/Essen im Bachelor-Studiengang Erziehungswissenschaft „Grundkenntnisse“ und „Orientierungswissen“ (Modul 1) gegenüber „zentrale Prinzipien der Beratung“, „Konzepte und Verfahren professioneller Gestaltung von Beratungssituationen“ und „zielgruppenorientierte Anwendung der Methoden“ (Modul 12)¹⁷.

Sind die in den Modulhandbüchern festgeschriebenen Kompetenzen in Kompetenzstufen und -raster übersetzt, vereinfacht das die Korrektur. Inwieweit Studierende die jeweiligen Kompetenzen in ihrem Prüfungsverhalten zeigen, kann im Kompetenzraster angekreuzt werden. Die Note ergibt sich quasi „von selbst“ und wird – bei immer gegebenem Ermessensspielraum – quasi berechenbar. Begründende Kommentare können aus den angekreuzten Feldern übernommen werden. Umgekehrt lassen sich auf dieser Grundlage neue Klausuraufgaben operationalisieren. Ist also einmal ein Kompetenzraster erstellt, ermöglicht dies die schnelle Bewertung vieler Klausuren bei steigendem Prüfungsaufkommen, das ja mit der Einführung von BA-/MA-Studiengängen einhergeht.

Die Arbeit mit Kompetenzrastern bietet sich mit der Modularisierung bereits strukturell an. Prüferinnen und Prüfer werden in der Anwendung solcher Raster sowohl horizontaler als auch vertikaler gradueller Steigerung gerecht: nämlich horizontal innerhalb einer Teilprüfung von der Note „vier“ bis zur Note „eins“ ansteigend; vertikal innerhalb des Studien-

¹⁷ URL: [http://biwi.uni-duisburg-essen.de/sites/biwi/files/UDE_20-20Bachelor_20EW_20Modulhandbuch_20\(M_C3_A4rz_202007\).pdf](http://biwi.uni-duisburg-essen.de/sites/biwi/files/UDE_20-20Bachelor_20EW_20Modulhandbuch_20(M_C3_A4rz_202007).pdf), S. 12.

ganges von der Eingangsphase ansteigend bis zur Abschlussphase. Das hier vorgeschlagene Bewertungsinstrument „Kompetenzraster“ wird zunächst für das Prüfungsinstrument „Papier-Bleistift-Klausur“ vorgeschlagen. Digitale Prüfungsformen bedürfen eigener Bewertungsapparate. Allerdings ist der Einsatz von Bewertungsrastern – bei aller gebotenen Konzentration auf die Klausur – ebenso für andere Prüfungsformate, etwa mündliche Prüfung, Praktikumsbericht, Referat, Hausarbeit denkbar.

Das dargestellte Raster hat Beispielcharakter. Es zeigt, wie eine inhaltliche Frage so operationalisiert werden kann, dass die Antworten auf diese Frage sie mit den in modularisierten Bachelor- und Masterstudiengängen geforderten Kompetenzen (vertikal) zusammenhängen und gemeinsam in Notengraden (horizontal) bewertbar werden. Jeder Lehrende bzw. jeder Prüfer hat zunächst die Aufgabe, seine Lehrinhalte mit der an Kompetenzen orientierten Modularisierung „seines“ BA-/MA-Studienganges vor Ort in Bezug zu setzen. Das erfordert zunächst einigen Aufwand, verheißt danach aber erhebliche Erleichterung. Die für BA/MA neu, komplex und kompliziert erscheinende Prüfungspraxis wird objektiviert und standardisiert. Das Kompetenzraster macht kompetenzorientiertes Beurteilen handhabbar.

Abb. 2: Kompetenzraster Bachelor Erziehungswissenschaft, Studieneingangsphase

	Kompetenzbereiche gemäß der Inhalte von Modul 1	Kompetenzstufe 1 [= Note 4]	Kompetenzstufe 2 [= Note 3]	Kompetenzstufe 3 [= Note 2]	Kompetenzstufe 4 [= Note 1]
Grundkenntnisse	Erziehungswissenschaftliche Grundbegriffe erläutern hier: - Kontrollpädagogik - Autonomiepädagogik - Bildung - Schule	Die Grundbegriffe werden bloß reproduziert , aber nicht reflektiert.	Die Grundbegriffe werden benannt, aber weniger reflektiert und nicht argumentativ genutzt. Innerhalb einer Theorie werden Bezüge hergestellt. Auf Praxisprobleme von Erziehung, Bildung, Unterricht wird dabei ebenfalls Bezug genommen.	Die Grundbegriffe werden reflektiert. Innerhalb einer Theorie werden Bezüge benennend hergestellt. Mit diesen theoretischen Momenten wird argumentiert. Auf Praxisprobleme von Erziehung, Bildung, Unterricht wird dabei ebenfalls Bezug genommen.	Die Grundbegriffe werden reflektiert. Innerhalb beider Theorien werden Bezüge benennend hergestellt. Mit diesen theoretischen Momenten wird argumentiert. Auf Praxisprobleme von Erziehung, Bildung, Unterricht wird dabei ebenfalls Bezug genommen.
	Pädagogische Handlungsmodalitäten erörtern hier: - Erziehen - (Sich)-Bilden - Zwingen - (Sich)-Befreien - Geschichten-Erzählen - Gespräche-Führen - Sprache(n)-Sprechen - Theater-Spielen - Naturerfahrungen-Machen - Sich-politisch-Betätigen - Arbeiten - Feste-Feiern - Musik-Machen - Aufbrechen	Die Handlungsmodalitäten werden bloß reproduziert , aber nicht reflektiert.	Die Handlungsmodalitäten werden benannt, aber weniger reflektiert und nicht argumentativ genutzt. Innerhalb einer Theorie werden Bezüge hergestellt. Auf Praxisprobleme von Erziehung, Bildung, Unterricht wird dabei ebenfalls Bezug genommen.	Die Handlungsmodalitäten werden reflektiert. Innerhalb einer Theorie werden Bezüge benennend hergestellt. Mit diesen theoretischen Momenten wird argumentiert. Auf Praxisprobleme von Erziehung, Bildung, Unterricht wird dabei ebenfalls Bezug genommen.	Die Handlungsmodalitäten werden reflektiert. Innerhalb beider Theorien werden Bezüge benennend hergestellt. Mit diesen theoretischen Momenten wird argumentiert. Auf Praxisprobleme von Erziehung, Bildung, Unterricht wird dabei ebenfalls Bezug genommen.
	Einordnungen in die Wissenschaftsgeschichte / Disziplin-geschichte / Geschichte der pädagogischen Einrichtungen / Institutionen vornehmen hier:	Für historische Darstellungen entsprechend.	entsprechend	entsprechend	Entsprechend

Stuientechniken	<ul style="list-style-type: none"> - bildungsgeschichtliche Bedeutung von Sokrates/Platon und Kant - (Be-)Gründung der Pädagogik als Wissenschaft v.a. durch Kant - Reform der Institution Schule (dreigliedriges Schulsystem) durch die Laborschule 				
	Bibliographieren	Schwerwiegende Fehler beim praktischen und theoretischen Bibliographieren	Fehler beim praktischen und theoretischen Bibliographieren	Wenige, marginale Fehler beim praktischen und theoretischen Bibliographieren	Es wird korrekt bibliographiert; Fragen zur Technik des Bibliographierens werden korrekt beantwortet.
	Zitieren	entsprechend	entsprechend	entsprechend	entsprechend
	Wissenschaftliche Texte lesen und gliedern	entsprechend	entsprechend	entsprechend	entsprechend
	Essays verfassen	Eigene Texte können sachlich / informierend mit deutlichen Strukturierungsproblemen formuliert werden; die Techniken des Paraphrasierens, Zitierens und Bibliographierens finden dabei Anwendung mit schwerwiegenden Fehlern. Verständlichkeit noch gegeben.	Eigene Texte können sachlich / informierend, in formal richtigem Aufbau (Einleitung, Hauptteil, Schluss) mit Strukturierungsproblemen formuliert werden; die Techniken des Paraphrasierens, Zitierens und Bibliographierens finden dabei Anwendung mit Fehlern.	Eigene Texte können sachlich / informierend, in formal richtigem Aufbau (Einleitung, Hauptteil, Schluss) mit leichten Strukturierungsproblemen formuliert werden; die Techniken des Paraphrasierens, Zitierens und Bibliographierens finden dabei Anwendung mit wenigen, marginalen Fehlern.	Eigene Texte können sachlich / informierend, in angemessenem, formal richtigem Aufbau (Einleitung, Hauptteil, Schluss) selbstständig formuliert werden; die Techniken des Paraphrasierens, Zitierens und Bibliographierens finden dabei korrekte Anwendung.

6. Konsequenzen für die Personalstruktur

Weiter oben ist bereits darauf hingewiesen worden, dass sich aus den neuen, modularisierten Studienstrukturen Konsequenzen für die Personalausstattung der Erziehungswissenschaft ergeben. Die gegenüber der traditionellen Praxis sehr deutlich gestiegene Lehrnachfrage sowie der um ein Vielfaches gestiegene Prüfungs- und Zertifizierungsaufwand kann – will man nicht das gesamte Studium in Form von Vorlesungen und standardisierten Abschlussklausuren gestalten – letztlich nur zu folgenden Reaktionen führen: Verstärkung und/oder Umstrukturierung des Personals oder aber Absenkung der Studierendenzahlen.

- Die auf die Studierendenzahlen bezogene *Absenkungsstrategie* ist zum einen bei den der Disziplin genuin zugeordneten Studiengängen möglich (Erziehungswissenschaft im Ein- oder Zwei-Fach-Bachelor; Masterstudiengang Erziehungswissenschaft, spezielle Weiterbildungsstudiengänge, Promotionsstudien). Eine Absenkung der Studierendenzahl ist zum anderen im Bereich der von der Disziplin Erziehungswissenschaft zu erbringenden Anteile innerhalb der erziehungs- und bildungswissenschaftlichen Studien in den zum Lehramt führenden Bachelor- und Masterstudiengängen möglich. Hier muss es das Ziel sein, die Kapazität der an den erziehungs- und bildungswissenschaftlichen Studien beteiligten Disziplinen bzw. Disziplin-Anteilen¹⁸ zur Stellgröße für den Zufluss von Studierenden in die Lehrer-Master zu machen; diese Kapazität liefert zugleich eine ungefähre Größe für die Zahl der Zulassungen zum (voraus liegenden) Zwei-Fach-Bachelor. Die Kapazitätsfrage in den Lehrer-Mastern wird man allerdings faktisch nicht ohne Blick auf den aktuellen und zukünftigen Bedarf an ausgebildeten Lehrkräften erörtern können; hier sind landesweite Absprachen zwischen den beteiligten Ministerien und den Hochschulen zu treffen, wobei eine verlässliche Aufgabenverteilung angestrebt werden sollte (vgl. Zielvereinbarungen). Bei der Festlegung der Kapazitäten (in der Lehre, für die Zulassungszahlen) muss die je ortsspezifische Verteilung der Schwerpunkte der Lehrenden berücksichtigt werden. Es ist nicht möglich, pauschal und normativ-planerisch eine Aufteilung auf *lehramtsbezogen* vs. *hauptfachbezogen* vorzunehmen: dies hängt von den real vertretenen Forschungs- und Lehrschwerpunkten der Lehrenden ab.
- Die Strategie einer *Erweiterung und/oder Umstrukturierung des Personals* mit Blick auf die Lehr- und Prüfungssituation kann bedeuten, dass *zusätzliches*, besonders lehrintensives Personal eingestellt wird (abgeordnete Lehrkräfte, Lehrkräfte für besondere Aufgaben, spezielle Lehr-Professuren etc.). Diese Strategie wird der Erziehungswissenschaft, auch den Fachdidaktiken, von vielen Universitätsleitungen empfohlen. Begleitet wird diese Form der Ausweitung des Personals von dringenden *Umstrukturierungsempfehlungen*: Umwandlung von Stellen mit wenig Lehrdeputat in solche mit höheren Lehrdeputaten – möglichst allerdings in solche Stellen, die zeitlich befristet besetzt werden, damit kurzfristig auf Bedarfsveränderungen reagiert werden kann. Die Besetzung lehrintensiver Stellen auf Zeit, wobei nicht selten noch Aufgaben der wissenschaftlichen Selbstqualifizierung *zusätzlich* definiert werden, ist jedoch faktisch nur sehr schwer möglich. Im Bereich der

¹⁸ Nicht die gesamte Erziehungswissenschaft kann hier verrechnet werden, ebenso nur Teile der Pädagogischen Psychologie, Bildungssoziologie, Politologie und Philosophie. Die letztgenannten Fächer sind in vielen Bundesländern in unterschiedlicher Quantität und unterschiedlichem Verpflichtungsgrad am erziehungswissenschaftlichen Studium innerhalb von Lehramtsstudiengängen beteiligt.

schul- und unterrichtsbezogenen Lehre sind qualifizierte Personen aus dem Schulbereich unter diesen Bedingungen sowie insbesondere angesichts der aktuellen Besoldung von Mitarbeitern nicht zu gewinnen. Außerdem wird durch die Ausweitung des faktisch *nur* noch lehrenden Personals der Zusammenhang von Lehre und Forschung zunehmend gelöst. Für die Qualität der Lehre und für die Disziplin insgesamt hätte dies fatale Folgen. Die Trennung der Personals in Lehrpersonal einerseits und Forschungspersonal andererseits sowie die Spezialisierung von Universitäten in Richtung auf berufsorientierte Ausbildungsgänge einerseits und forschungsintensive Spezialinstitutionen mit reinem Forschungsprofil und hohem Rekrutierungsfaktor für wissenschaftlichen Nachwuchs andererseits wäre die langfristige Folge einer solchen Transformation der Erziehungswissenschaft.

Für eine wissenschaftlich fundierte Ausbildung inklusive einer entsprechenden, die Studien kontinuierlich begleitenden Prüfungspraxis ist die Fundierung durch wissenschaftliche Forschung innerhalb des in der Lehre vertretenen Sachbereichs unabdingbar. Aus diesem Grunde sind der Umstrukturierung und vollständigen funktionellen Teilung des Personals (hier ausschließlich Lehrende, dort ausschließlich Forschende) Grenzen gesetzt. Unabhängig von dieser Frage können und sollten jedoch diejenigen Elemente der Prüfungspraxis, die zu standardisieren sind, angesichts der gegebenen Umstände gleichwohl realisiert werden. Auf diese Weise können Zeit- und sonstige Ressourcen geschont werden, und zwar bei Aufrechterhaltung des inhaltlichen Prüfungsniveaus. Aber trotz Einsatz solcher Formen wird zusätzlicher Personalbedarf bestehen bleiben. Ob dieser zusätzliche Bedarf durch Zuwächse gedeckt wird, ob die Studierendenzahl gesenkt wird, ob der Weg in die Umstrukturierung und Teilung des Personals im oben angesprochenen Sinne verordnet wird, ob am Ende eine Mischung aus allen diesen Elementen zustande kommt – dies alles hängt von dem faktischen Einfluss der Erziehungswissenschaft innerhalb des universitären Verteilungsgeschehens ab und natürlich auch davon, welches Potenzial die Universität der Erziehungswissenschaft in ihrer sachlich nicht zu teilenden Verantwortung für die Weiterentwicklung der eigenen Disziplin *und* für die Ausbildung zu pädagogischen Berufen zuzubilligen bereit ist.