

Kompetenzorientiert prüfen mittels Bildungsstandards – keine Lösung

Kompetenzorientiert prüfen mittels Bildungsstandards stellt keine Lösung für Prüfungen an Berufsbildenden Schulen dar.



UNIV.-PROF. DR. CHRISTOPH METZGER
 Direktor des Institut für Wirtschaftspädagogik
 Universität St. Gallen
 christoph.metzger@unisg.ch

1 Problemstellung: Bildungsstandards und Prüfungspraxis

An berufsbildenden wie auch allgemeinbildenden Schulen der Sekundarstufe II ist es seit geraumer Zeit üblich, dass Schülerinnen und Schüler im Hinblick auf den Schulabschluss eine umfangreiche Arbeit (z.B. Selbständige Vertiefungsarbeit oder Maturaarbeit genannt) zu verfassen und zu präsentieren haben, wobei die gezeigte Leistung notenrelevant ist. Folgender Ablauf sei angenommen:

Die Arbeit im Umfang von 20 bis 30 Seiten wird in Zweiergruppen verfasst, wofür sich die Teams selbst bilden sollen. Als Bearbeitungszeitraum stehen drei Monate zur Verfügung. Während dieser Zeit haben die Teams mit der betreuenden Lehrkraft in einem ersten Gespräch die Forschungsfrage, in einem weiteren Gespräch die Disposition zu diskutieren, im späteren Verlauf den Entwurf der Arbeit einzureichen, diesen nach Durchsicht durch die Lehrkraft mit ihr zu besprechen und schließlich die überarbeitete Schlussfassung, einschließlich einer Reflexion zu Prozess und Ergebnis der Arbeit (sog. Lernjournal), abzugeben. Es dürfen sämtliche zugänglichen Informationsquellen einschließlich Interviews verwendet werden, sie sollen aber in der für wissenschaftliches Arbeiten üblichen Form nachgewiesen werden. Innerhalb eines Monats nach Abgabe der Arbeit erhalten die Teams die Arbeit zurück, versehen mit einem differenzierten Kommentar zu Form und Inhalt. Im Verlaufe eines weiteren Monats erhalten die Teams schließlich Gelegenheit, ihre Arbeit – in der durch die Lehrkraft didaktisch sinnvoll geplanten Reihenfolge – in einem 15-minütigen Vortrag vor der Klasse und Gästen zu präsentieren und während weiterer 15 Minuten zu diskutieren. Nun erst setzt die Lehrkraft die Note fest, die sich nach einem den Schülern bereits zu Beginn dieses Projektes bekannt gemachten Raster errechnet, der die wesentlichen Beurteilungskriterien sowie das Gewicht der Arbeitsphasen von der Forschungsfrage über die Disposition bis zur Präsentation und Diskussion enthält. Die Lehrkraft bespricht mit jedem Team die Arbeit, die Reflexion sowie die erteilte Note.

Diese Lehrkraft ist überzeugt, dass sie mit dem gewählten Verfahren handlungsorientiert und damit dem aktuellen Prüfungsverständnis entsprechend prüfe. Konfrontiert mit der Diskussion um Bildungsstandards fühlt sie sich – und mit ihr viele Kolleginnen und Kollegen – allerdings verunsichert. Beim Studium der weit

herum beachteten Expertise „Zur Entwicklung nationaler Bildungsstandards“ zu Händen des Deutschen Bundesministeriums für Bildung und Forschung (Klieme et al. 2003) empfindet sie zwar einerseits folgende Aussage beruhigend:

Standards sollen ein Kriterium für die Lernergebnisse jeder Schule, jeder Klasse und konsequenterweise jeder einzelnen Schülerin und jedes einzelnen Schülers sein. Sie dürfen aber nicht mit Benotungs- und Prüfungsregeln oder gar mit Regelungen zum Übergang auf weiterführende Schulen verwechselt werden.

Die Bildungsstandards decken erklärtermaßen nicht das gesamte Curriculum, sondern nur einen Kern in zentralen Domänen des Lernens ab. Zum Zweiten ist die Benotung und die Vergabe von Abschlüssen mit gutem Grund Gegenstand von Entscheidungen, die Lehrerinnen und Lehrer nach professionellen Maßstäben, aber selbstständig, unter Abwägung verschiedenster Aspekte, in ihrer pädagogischen Verantwortung treffen müssen. Beide Argumente sprechen dafür, Bildungsstandards nicht als Kriterien für Notengebung, Zertifizierung und Selektionsentscheidungen zu missbrauchen (S. 39).

Gleichwohl fühlt sich die Lehrperson verunsichert, wenn gleichenorts die Bildungsstandards zusammenfassend charakterisiert werden: „Bildungsstandards . . . greifen allgemeine Bildungsziele auf. Sie benennen die Kompetenzen, welche die Schule ihren Schülerinnen und Schülern vermitteln muss, damit bestimmte zentrale Bildungsziele erreicht werden. . . . Die Kompetenzen werden so konkret beschrieben, dass sie in Aufgabenstellungen umgesetzt und prinzipiell mit Hilfe von Testverfahren erfasst werden können“ (S. 13). Denn aus Erfahrung weiß unsere Lehrperson sehr wohl, dass das von ihr und ihren Kolleginnen praktizierte Prüfungsverfahren zwar durchaus wichtige, komplex miteinander verbundene Kompetenzen erfassen soll, aber wohl nicht den Ansprüchen an Präzision und Objektivität genügt, die sie mit dem Begriff „Testverfahren“ – nicht zuletzt im Nachgang zu den PISA-Untersuchungen – assoziiert.

Vom Beispiel sich lösend, stellt sich verallgemeinernd die Frage, wie weit Bildungsstandards vereinbar sind mit dem aktuellen Verständnis kompetenzorientierten Prüfens – oder allgemeiner der Schülerbeurteilung –, wie es nicht nur in der Prüfungstheorie aussieht, sondern auch in der Schul- und Prüfungspraxis, wenn

auch unterschiedlich elaboriert, verwirklicht wird. Der Vergleich des Konzepts von Bildungsstandards mit einem dem aktuellen prüfungstheoretischen Verständnis entsprechenden Prüfungsmodell wird folgende These untermauern: Das Konzept von Bildungsstandards kann den Anforderungen an ein kompetenzorientiertes Prüfen und der damit verbundenen Schülerbeurteilung nicht standhalten, ja es birgt sogar die Gefahr eines Rückfalls in Zeiten objektivierte psychometrisch orientierten Prüfens.

2 Charakterisierung von Bildungsstandards

Bildungsstandards drücken aus, über welche Kompetenzen Schüler verfügen müssen, sodass daraus geschlossen werden kann, dass sie die Bildungsziele erreicht haben. Der Focus liegt damit auf dem Output oder der Performance, im Gegensatz zu Opportunity-to-learn-Standards, verstanden als Ressourcen für einen wirksamen Lehr-Lern-Prozess, oder Content Standards als bloße Definition der Lerninhalte (s. Aff in dieser Nummer von wissenplus).

Wie im einleitenden Zitat bereits wiedergegeben, bedarf es dreier Komponenten, damit Bildungsstandards ihre nachfolgend beschriebenen Funktionen erfüllen können (Klieme et al. 2003). Sie basieren erstens auf allgemeinen und weitherum akzeptierten Bildungszielen. Diese konkretisieren sie zweitens in Form von Kompetenzanforderungen, basierend auf einem jeweils auszuwählenden mehrdimensionalen Kompetenzmodell. Anhand dessen ist mindestens zu definieren, über welche Teilkompetenzen auf welchem Schwierigkeits- oder Qualitätsniveau (Kompetenzstufe) in einem bestimmten Bereich (Fach, Domäne) Schüler zu einem bestimmten Zeitpunkt (z.B. Schulstufe, Schultyp) verfügen müssen. Mit anderen Worten beantworten Bildungsstandards die Frage: Was müssen Schüler wie gut können, um den jeweiligen „grundlegenden Handlungsanforderungen“ (S. 15) zu genügen? Drittens werden die Bildungsstandards in „Aufgabenstellungen bzw. Testverfahren“ (S. 16) konkretisiert, mittels derer outputorientiert ermittelt werden kann, welche Kompetenzen sie auf welchem Kompetenzniveau erreicht haben. Für Beispiele sei auf die Expertise von Klieme et al. und auf Dubs (2004) verwiesen.

Bildungsstandards sollen zunächst eine Orientierungsfunktion erfüllen (Klieme et al., 2003 S. 45), indem sich die Betroffenen, besonders Lehrende und Lernende auf verbindliche Ziele beziehen können bzw. – so meine Interpretation – müssen. Auch wenn dabei Klieme et al. – und auch etwa Dubs (2006) – betonen, dass damit für die innerschulische Curriculumentwicklung ein grosser Freiraum im Sinne der teilautonomen Schule bestehe, ist meines Erachtens der Verpflichtungscharakter der Orientierungsfunktion nicht zu übersehen. Zweitens sollen Bildungsstandards ermöglichen, die Lernergebnisse zu erfassen und zu bewerten, dies auf der Ebene des betreffenden Bildungssystems oder eines Teils davon (z.B. berufsbildende Schulen), auf der Ebene einzelner Schulen sowie beim einzelnen Schüler. Bezüglich des Schülers von „bewerten“ (Klieme et al. 2003, S. 4) zu sprechen ist allerdings insofern irrelevant, als die Autoren davon abraten, standardbezogene Tests für die Notengebung beizuziehen. Insgesamt dienen Bildungsstandards also der Qualitätssicherung und -verbesserung einschließlich der auf den einzelnen Schüler bezogenen Diagnostik und Förderung, jedoch auch – wie Oser und Renold (2005, S. 127), bezogen auf Kompetenzen von Lehrpersonen, deutlich machen – der Normierung und Erhöhung der Vergleichbarkeit. Sie sind deutlich outputorientiert.

Im Hinblick auf eine Prüfungskonzeption ist es wesentlich, ei-

nen Blick auf die Merkmale zu werfen, wie sie Klieme et al. (2003, S. 17–27) für gute Bildungsstandards benennen:

Bildungsstandards sollen zunächst verständlich formuliert und mit realistischem Aufwand realisierbar sein. Im Weiteren sollen sie der maßgeblichen Struktur eines Faches bzw. Domäne oder Disziplin entsprechend formuliert sein und den Erwerb von Kompetenzen als einen bis zu einem bestimmten Zeitpunkt kumulativ und systematisch vernetzten Lernprozess auffassen. Schließlich sind zwei Merkmale für das Prüfen besonders relevant. Zum einen sollen sich Bildungsstandards auf den Kernbereich eines Faches fokussieren, also nicht die ganze Breite eines Bereiches umfassen, sondern, wie Dubs (2006, S. 19) es nennt, für die Zukunft unstrittige, bedeutsame Lernbereiche und Lerninhalte. Diesbezüglich wird auch vom Kerncurriculum gesprochen. Zum anderen drücken Bildungsstandards einerseits Mindestanforderungen an alle Schüler aus – Klieme et al. verstehen diese Mindeststandards sogar als schulformübergreifend –, andererseits können aber Standards auch situationsgerecht, etwa nach Schultypus differenziert, d.h. höher oder tiefer angesetzt werden.

3 Bildungsstandards auf dem Prüfstand eines Prüfungsmodells

3.1 Ein Rahmenmodell als Basis des Prüfens

Die Gestaltung von schulischen Prüfungen (s. Abb. 1), die letztlich der Beurteilung von Schülerinnen und Schülern, oder vorsichtiger ausgedrückt der Beurteilung ihrer Kompetenzen, dient, bedarf im Kern der Entscheidungen über die Inhalte (WAS prüfen) und Formen des Prüfens (WIE prüfen) sowie die Formen des Auswertens, besonders der Bewertung und der Kommunikation. Diese Entscheidungen hängen maßgeblich von zwei übergeordneten Fragen ab: Erstens ist zu klären, wozu überhaupt geprüft werden soll, zweitens ist zu entscheiden, in welchem Maße den üblichen Anforderungen an eine Prüfung (besonders Gültigkeit und Zuverlässigkeit) entsprochen werden soll, d.h. wie gut geprüft werden soll. Schließlich sind Funktionen des Prüfens, Anforderungen an das Prüfen sowie der eigentliche Prüfungsprozess nicht als geschlossenes System zu betrachten, sondern eingebettet in den gesellschaftlich-curricularen Rahmen des fraglichen Bildungssystems, hier der beruflichen Grundbildung im Besonderen, sowie in das begleitende bzw. vorherrschende Lehr-Lern-Verständnis.

Im Folgenden wird das Konzept der Bildungsstandards dahingehend beurteilt, wieweit es den einzelnen Entscheidungselementen dieses Prüfungsmodells entspricht.

3.2 Bildungsstandards im Lichte des gesellschaftlich-curricularen Rahmens

Bildungsstandards spiegeln den gesellschaftlich-curricularen Auftrag nur zu einem geringen Teil wieder und sind deshalb eine unzureichende Grundlage für ein umfassendes, integratives Prüfungskonzept.

Charakteristisch für die berufsbildenden Schulen – so wie ich sie von außen betrachtet mit Blick auf Lehrpläne sowie Unterrichts- und Prüfungsmaterialien verstehe – ist, dass sie, vorwiegend vollzeitschulisch angelegt, sowohl berufsorientierte als auch nicht-berufsorientierte, schwergewichtig handlungsorientierte Kompetenzen anstreben und über schulinterne Prüfungen zur Berufsmatura führen, welche zum Hochschulstudium berechtigt, aber auch den Einstieg in verschiedene Berufsfelder ermöglichen. Sie haben also sog. allgemein- und berufsbildende Funktionen auf einem ■

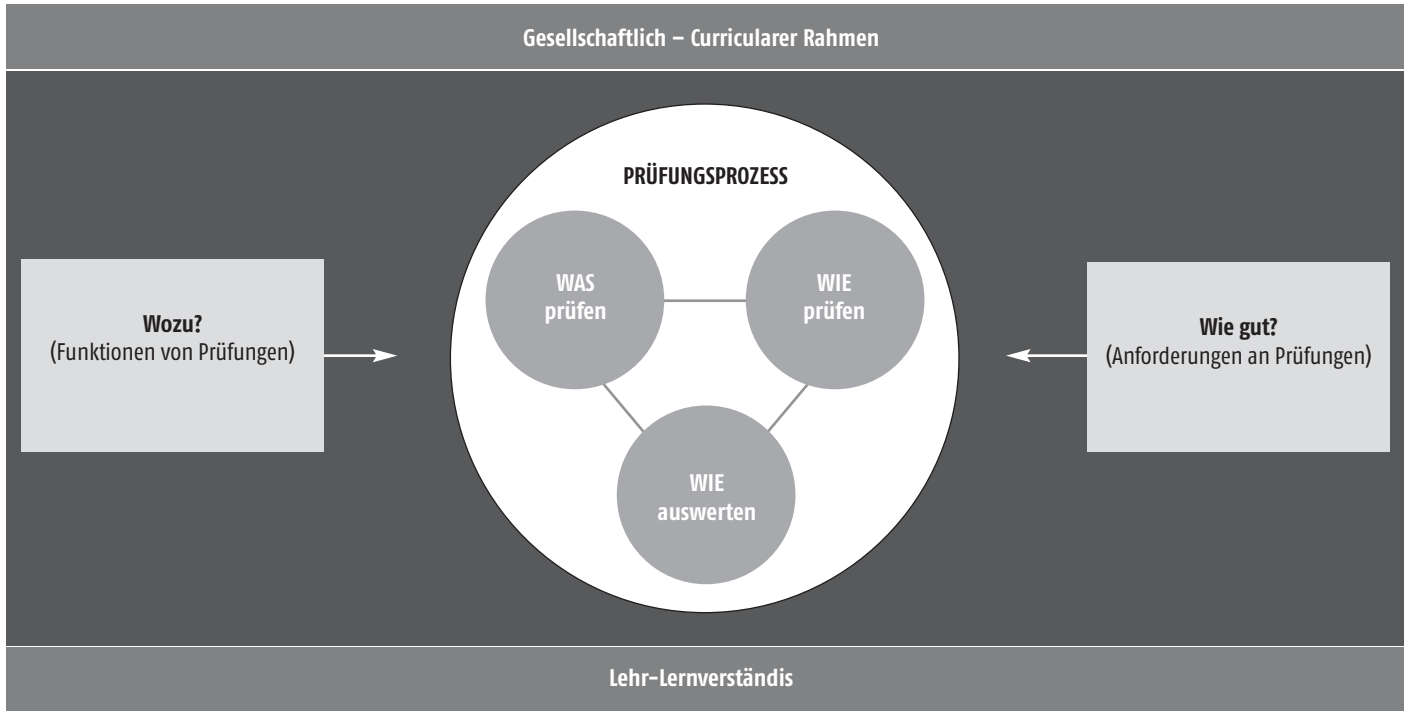


Abb. 1: Rahmenmodell des Prüfens (Metzger & Nüesch, 2004)

durchaus anspruchsvollen Niveau zu erfüllen. Wie Aff in dieser Ausgabe von wissenplus ausführt, ist mit der Einführung von nationalen Bildungsstandards die Hoffnung verbunden, dass die Vergleichbarkeit der schulischen Abschlüsse erhöht wird. Wenn, wie oben beschrieben, Bildungsstandards allerdings nur ein Kerncurriculum sowie Mindestanforderungen ausdrücken sollen, ist jedoch zu erwarten, dass diese Vergleichbarkeit gering bleiben wird, weil sie die genannte umfassende Zielsetzung nicht abbilden können.

3.3 Bildungsstandards im Lichte des Lehr-Lern-Verständnisses
 Bildungsstandards nehmen das aktuelle Lehr-Lern-Verständnis, mindestens bezüglich des Prüfens, nicht genügend auf, indem sie die Prozesskomponente des Lernens ausblenden.

Es darf aktuell davon ausgegangen werden, dass auch an berufsbildenden Schulen versucht wird, ein kognitivistisch-konstruktivistisches Lehr-Lern-Verständnis umzusetzen. Gemäß diesem konstruieren Lernende ihr Wissen und Können mittels verschiedener Denkoperationen und -strategien und möglichst auch praktischem Handeln aktiv, zielorientiert, mit dem Vorwissen verknüpfend, dieses erweiternd und differenzierend, eingebettet in einen für sie bedeutsamen inhaltlichen wie auch sozialen Kontext, auf Transfer ausgerichtet und in hohem Maße interaktiv (Achtenhagen 1992; Reinmann-Rothmeier & Mandl 1998; Shepard 2001; Shuell & Moran 1994). Dazu müssen die Lernenden ihr Lernen selber regulieren können, indem sie sich für ihr Lernen selbst verantwortlich fühlen und strategisch verhalten (Zimmerman 1994; Segers, Dochy & Cascallar 2003, S. 2), was aber durchaus vereinbar damit ist, dass die Lehrperson die Wissenskonstruktion fördert, indem sie unterstützt, berät, anleitet, darbietet und erklärt (Reinmann-Rothmeier & Mandl 1998). Wie oben charakterisiert, sind Bildungsstandards klar ein Instrument der Output-Steuerung, blenden also den Prozess des Lernens und Lehrens als einen in sich immer wieder einmaligen Vorgang aus. Nur indirekt kann aus dem

geforderten Kompetenzmodell und entsprechenden Testaufgaben auf notwendige Lehr-Lern-Prozesse geschlossen werden. Ein umfassendes Prüfungskonzept sollte jedoch die Prozesskomponente sichtbar aufnehmen – wie im Folgenden noch differenzierter auszuführen ist.

3.4 Bildungsstandards im Lichte der Funktionen von Prüfungen

Bildungsstandards können die formative Funktion von Prüfungen gar nicht und die summative Funktion nur unzureichend erfüllen.

Prüfungen können bzw. sollen grundsätzlich zwei unterschiedliche Funktionen erfüllen, die formative und die summative (Bloom, Hastings & Madaus 1971). Summative Prüfungen (neuerdings auch „assessment of learning“ genannt (Birenbaum 2003, S. 3) informieren Betroffene und Interessierte über nachgewiesene Kompetenzen zum Zwecke der Zulassung, Zuweisung und Wegweisung, haben also für die Lernenden weiter reichende, letztlich selektive Konsequenzen (Metzger & Nüesch 2004, S. 5). Sie werden deshalb zu einem für das Gros einer Lerngruppe sinnvollen Zeitpunkt (z.B. sog. Hausarbeiten oder Klausuren), aber auch über einen längeren Zeitraum hinweg (z.B. eine größere Abschlussarbeit) gegen Ende einer Lerneinheit oder eines längeren Ausbildungsabschnittes an einer bedeutsamen Schnittstelle (z.B. Maturitätsprüfung) durchgeführt und schließen die entsprechende Lernphase – im günstigeren Fall – ab.

Gerade weil im aktuellen Lehrverständnis der Prozesscharakter besonders betont wird, ist die Aufgabe des Prüfens aber besonders darin zu sehen, die Lernenden in ihrer Konstruktion des Wissens und Könnens zu unterstützen. Dazu dienen die formativen Prüfungen (neuerdings auch „assessment for learning“ genannt [Birenbaum 2003, S. 3]). Sie sollen das Lernen in Richtung des geforderten bzw. vom Lehrenden oder Lernenden gewünschten konstruktiven Prozesses und Ziels unterstützen und frei von unnötigem Zeitdruck lenken (Gielen, Dochy & Dierick 2003, S. 48), wobei diese Lenkung möglichst unmittelbar, d. h. kurzfristig greifen soll (Shepard, 2001). So

sollen qualitative, d.h. besonders verstehens- und auch motivationsorientierte Einsichten in den Prozess des Lernens möglich werden (Shepard 2001, S. 1091; OECD 2005). All diese Wirkungen sollten sich aber nicht nur bezogen auf die Bewältigung einer spezifischen Lernaufgabe einstellen. Vielmehr sollte der Prozess formativen Prüfens transferorientiert als Element strategischen Lernens die Lernkompetenz im Sinne der Fähigkeit und des Willens, selbstständig zu lernen, unterstützen (OECD 2005, S. 2). Verlauf und Ergebnisse formativer Prüfungen sollen damit aber auch Rückschlüsse auf den Lehrprozess ermöglichen, sowohl in Richtung des Lehrerfolgs als auch notwendiger didaktisch-methodischer Anpassungen des Lehrprozesses (Metzger 1986, S. 95–96; OECD 2005b).

Da im Konzept der Bildungsstandards die Testaufgaben outputorientiert und auf zeitliche Schnittstellen ausgerichtet sind, können sie die formative, also unmittelbare Lenkung des Lehr- und Lernprozesses nicht wahrnehmen. Selbst wenn, wie ausgeführt, solchen Testaufgaben auch die Möglichkeit der Diagnostik und Förderung der einzelnen Lernenden zugeordnet wird, erfolgt diese Kontrolle zu spät und ohne angemessene Lernhilfen. Es ist auch mit Blick auf die Praxis zu bezweifeln, dass am Ende eines langen Ausbildungsabschnittes – also z.B. am Ende eines Schuljahres – festgestellte Lerndefizite und -schwierigkeiten von Lernenden und Lehrenden nur annähernd so effizient und motiviert angegangen werden, wie wenn die Rückmeldungen sowie Lernhilfen integriert in den Lehr-Lern-Prozess erfolgen. Sowohl für den formativen wie für den summativen Fall ist wiederum einzuwenden, dass die Fokussierung auf Kerncurriculum und Mindeststandards kein repräsentatives Bild des Lehr-Lern-Prozesses widerspiegelt und im summativen Fall damit auch keinen gültigen, d.h. genügend umfassenden Nachweis bezüglich der für eine berufsbildende Schule relevanten Kompetenzen ermöglicht.

3.5 Bildungsstandards im Lichte der Anforderungen an Prüfungen

Den Bildungsstandards entsprechende Testaufgaben können den Anforderungen an Prüfungen nicht genügen. Ganzheitlichere Prüfungen müssen vielfältigere Kompetenzdimensionen und diese auch auf komplexeren Anforderungsstufen erfassen, was mittels bildungsstandardorientierten Aufgaben nicht geschehen wird bzw. kann.

Formative wie auch summative Prüfungen haben den vier Anforderungen Gültigkeit, Chancengerechtigkeit, Ökonomie und Zuverlässigkeit in einem genügenden Maße zu entsprechen (AERA, APA & NCME 2002; Linn & Gronlund 1995; Messick 1989; Metzger, Dörig & Waibel 1998). (1) Gültigkeit bedeutet, dass Prüfungen inhaltlich sowie von ihrer Form und den Beurteilungsverfahren her wirklich das überprüfen, was die Schülerinnen und Schüler gemäß Kompetenzanforderungen können sollen. Das heißt besonders: Erstens müssen die Prüfungen angemessen über die curricular relevanten Kompetenzen einschließlich der Inhalte hinweg streuen, zweitens sollen die Aufgaben die für die Prüfung ausgewählten relevanten Kompetenzen auch wirklich, d.h. in jener Komplexität erfassen, wie sie gemäß Curriculum und vorangegangenen Lehr-Lernprozessen angestrebt wurden. (2) Eng verknüpft mit der Gültigkeit ist die Anforderung der Chancengerechtigkeit. Danach sollen Prüfungen in Form und Inhalt den jeweiligen schulischen Lernbedingungen (z.B. Lehrkräfte, eingesetzte Unterrichtszeit, Lehrmittel) und externen Lernbedingungen (z.B. Zugang zu Medien, sprachliches Umfeld) entsprechen. (3) Die Prüfungen sollen das,

was sie erfassen sollen, möglichst zuverlässig erfassen, d.h., es dürfen keine Messfehler auftreten, die das Prüfungsergebnis und letztlich dessen Bewertung und Interpretation wesentlich verfälschen. Dies ist mittels möglichst hoher Objektivität, d.h. für alle betroffenen Prüflinge gleichartiger und präziser Bedingungen in der Durchführung, Auswertung und Interpretation anzustreben. (4) Schließlich sollen Prüfungen ökonomisch sein. Sie sollen den Nutzen, den sie aufgrund ihrer Funktion zu erbringen haben, mit einem für die Konstruktion, Durchführung und Auswertung vertretbaren Aufwand erbringen.

Im kognitivistisch-konstruktivistischen Modell des Prüfens wird nun der Gültigkeit höchste Bedeutung zugemessen (GIPPS 1995; MOSS 1994). Ein differenzierter, aussagekräftiger Kompetenznachweis ist demnach nur möglich, wenn Prüfungen die geforderten Kompetenzen wirklich in der inhaltlich nötigen Breite und Tiefe verlangen und das relevante Können der Lernenden in jenen Kontexten überprüfen, in denen das Wissen und die Fähigkeiten eine Bedeutung und einen Anwendungsbezug haben (Resnick & Resnick, 1992). In diesem Prüfungsverständnis werden einerseits die Lernprozesse stärker betont, andererseits wird das Wissen als weniger objektivierbar betrachtet. Überdies kommt der Chancengerechtigkeit von Prüfungen eine hohe Bedeutung zu (Darling-Hammond, Aness & Falk 1995). Einschränkungen werden im kognitivistisch-konstruktivistischen Modell bei der Zuverlässigkeit und Ökonomie in Kauf genommen. So wird akzeptiert, dass hochgültige Prüfungsaufgaben an Zuverlässigkeit einbüßen, weil sie oft umfangreicher, komplexer, kontextgebundener und individualisierter sind und deshalb schon in der Konstruktion und Durchführung, besonders aber in der Auswertung nicht denselben Objektivitätsansprüchen genügen können wie reproduktive Prüfungsaufgaben. Ausgehend von der Maxime hoher Gültigkeit folgt schliesslich, dass die Ökonomie erst dann in Betracht zu ziehen ist, wenn zwei gleichermaßen gültig scheinende Prüfungsmethoden unterschiedlich aufwändig sind.

Die Kurzbeschreibung der Anforderungen lässt erkennen, dass in aller erster Linie verlässlich zu erfassen ist, was die Lernenden wie gut können. Unabdingbare Basis dafür ist eine curricular und im für die Lernenden aktuell erlebbaren Lehr-Lern-Prozess abgesicherte Bestimmung der Kompetenzanforderungen. So steht denn folgerichtig auch im Konzept von Bildungsstandards die Komponente des Kompetenzmodells mit Kompetenzarten und -anforderungen im Mittelpunkt. Dass ein solches Modell mehrdimensional ist, zeigt schon die von Klieme et al. (2003) und in der Diskussion über Bildungsstandards allgemein immer wieder beigezogene Definition von Weinert (2001), wonach unter Kompetenzen „die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme lösen zu können, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (S. 27–28) zu verstehen sind. Wie auch Aff, Fortmüller sowie Hauer und Stock in ihren Aufsätzen dieser Nummer von wissenplus deutlich herausarbeiten, können basierend auf dieser Definition in einem Kompetenzmodell verschiedenste Dimensionen verortet werden. So lassen sich – bei aller Überlappung – zunächst kognitive und nicht-kognitive (also z. B. Einstellungen) Kompetenzen unterscheiden. Selbst wenn man sich primär auf die kognitiven Kompetenzen konzentriert, lässt sich in verschiedene Richtungen differenzieren, so etwa: ■

- (a) nach Komplexität und Funktion variierende generelle Arten des Wissens (z.B. deklaratives, prozedurales und metakognitives Wissen) sowie unterschiedlich intensive und komplexe kognitive Prozesse (so z.B. Erinnern, Verstehen, Anwenden, Analysieren, Evaluieren, Kreieren) in ihren zahlreich denkbaren Verbindungen einschliesslich Präsentationsmodi wie schriftlich, mündlich oder praktisch (siehe dazu die weiterentwickelte Bloom'sche Taxonomie: Anderson & Krathwohl 2001);
- (b) nach Komplexität von Problemstellung, Lösungshorizont und Kontext (Dörner 1976; für die Volkswirtschaftslehre siehe den Vorschlag von Dubs [2004]);
- (c) nach dem Grad geforderter Selbständigkeit (z. B. bezüglich Reproduktion bzw. Transfer);
- (d) nach der Differenziertheit der zugrunde gelegten Struktur eines Fachgebietes, und nicht zuletzt
- (e) nach dem Grad der Fachspezifität, was etwa in der Unterscheidung in Sach-, Sozial- und Selbstkompetenz oder in der Unterscheidung in fachliche und überfachliche Kompetenzen zum Ausdruck kommt (z.B. Euler & Hahn 2004), ohne damit gleich in den kaum mehr griffigen Bereich völlig genereller Schlüsselkompetenzen zu gelangen.

Die einzelne Lehrkraft oder auch eine Prüfungskommission einer einzelnen Schule mag sich nun annähernd außer Stande fühlen, Prüfungen zu gestalten, welche alle oder auch nur wenige dieser Dimensionen von Kompetenzen erfassen. Sie werden es aber zweifellos besser können, als es über national dekretierte Bildungsstandards bzw. Kompetenzmodelle und Testaufgaben möglich ist. Denn die Lehrpersonen werden bereits den Unterricht entlang solcher Kompetenzen entwickelt und gestaltet haben, basierend auf intensiven fachdidaktischen Überlegungen und mit Rücksicht auf die jeweilige ganz konkrete Lernsituation der Schülerinnen und Schüler. Demgegenüber bringt die Orientierung an einem Kerncurriculum, das nur einen Teil der Unterrichtszeit (z.B. 50%) und an Mindeststandards (z.B. entsprechend der Formel „80% der Schülerinnen und Schüler müssen 80% der Ziele erreichen“ [Dubs 2006, S. 19]) mit hoher Wahrscheinlichkeit eine Reduktion auf wenige der aufgeführten Dimensionen und darin auf die jeweils tiefste Komplexitätsstufe mit sich. Im Besonderen ist zweierlei zu befürchten. Erstens wird die Themenvielfalt auf wenige Themen reduziert, womit die größeren fachlichen Zusammenhänge verloren gehen. Zweitens werden nur vergleichsweise isoliert solche kognitive Prozesse geprüft, die für eindeutig lösbare Probleme notwendig sind, wenn denn überhaupt bis auf die Stufe des Problemlösens vorgedrungen wird. Differenziertere Denk- und Arbeitsweisen sind nicht erforderlich. Durch beide Verkürzungen werden Gültigkeit und Chancengerechtigkeit verringert – dies zugunsten der Zuverlässigkeit, die durch den Standardisierungsgrad der Testaufgaben bezüglich Entwicklung, besonders aber auch Auswertung erreicht wird. Im Rahmen der schweizerischen kaufmännischen, zentralen Lehrabschlussprüfungen liess sich dies wiederholt feststellen (Metzger & Waibel 1993; Amos et al. 2004). Die ökonomischen Effekte des Einsatzes auf Bildungsstandards beruhender Testaufgaben müssten wohl erst noch untersucht werden.

3.6 Bildungsstandards im Lichte des Prüfungsprozesses

Aus den bisher betrachteten Elementen des Prüfungsmodells ergibt sich zwangsläufig das Erfordernis einer variantenreichen, integrativen Gestaltung der Prüfungsprozesse. Die an Bildungsstan-

dards orientierten Testaufgaben können gerade aufgrund ihrer Fokussierung auf Kerncurriculum und Mindestanforderungen diese Komplexität nicht bewältigen.

Der kognitivistisch-konstruktivistisch orientierten Optimierung der Prüfungsanforderungen folgend, ergeben sich folgende praktische Konsequenzen für den eigentlichen Prüfungsprozess.

Prüfungen müssen inhaltlich repräsentativ sein, d.h. Kompetenzen in jener Breite und Tiefe sowie jenen Kontexten erfassen, wie sie in der schulischen Ausbildung von den Lernenden zu erwerben waren. Dies erfordert den Einsatz produkt- und prozessorientierter Prüfungsmethoden. Als Durchführungsmodi sind schriftliche, mündliche und praktische Prüfungen, sinnvoll aufeinander abgestimmt, einzusetzen. Aufgaben zur freien ausführlichen Bearbeitung (z.B. Fallaufgaben, Berichte, Konstruktionen, Projekte, Präsentationen, selbstständige Vertiefungsarbeiten, Maturaarbeiten) sind gegenüber Kurzantwort- und Auswahlaufgaben und Aufgaben zur stark eingeschränkten Bearbeitung weitestgehend vorzuziehen. Die Prüfungen sollen sich im Weiteren nicht allein auf das Ende der Ausbildung und dies in der Praxis oft fälschlicherweise nur mittels kurz zu bearbeitender und sog. objektiv lösbarer Aufgaben konzentrieren, sondern zeitraumbezogen einerseits Ergebnisse von Zwischenprüfungen oder -zeugnissen, andererseits auch begleitend zur Ausbildung verschiedene über eine längere Zeitspanne zu erbringende Leistungen (z.B. genannte umfangreiche Arbeiten, Portfolio) einbeziehen. Überdies sollen neben individuell zu erbringenden Leistungen auch mehrere Lernende gemeinsam eine Leistung zu erbringen haben (kooperative Formen des Prüfens). Konsequenterweise sollte dabei das Gruppenergebnis beurteilt werden, womit zwangsläufig nicht nur die Qualität individueller Leistungen, sondern auch jene der Gruppenprozesse bedeutsam werden. Komplexere Durchführungsmodi, Bearbeitungsformen und Prüfungsinhalte machen es nötig, dass verschiedene, miteinander zusammenhängende Aspekte einer komplexeren Leistung analytisch anhand verschiedener, angemessen zu gewichtender Kriterien und Indikatoren für unterschiedliche Qualitätsstufen beurteilt werden. Solche Korrekturraster oder „Rubrics“ (z.B. Herman, Aschbacher & Winters 1992) basieren oft nicht auf eindeutigen Musterlösungen, sondern verlangen nach einer Einschätzung anhand formalerer Standards, werden allerdings wenn immer möglich aufgabenspezifisch entwickelt oder enthalten mindestens auch solche Elemente (für ein Beispiel zur Beurteilung betriebswirtschaftlicher Falllösungen s. Metzger, Dörig & Waibel 1998, S. 122). Damit für die Prüflinge die Nähe zu den Lernbedingungen gewahrt bleibt, sollten Abschlussprüfungen möglichst dezentral, d.h. von den Lehrenden selbst konstruiert und durchgeführt werden. Für eine primär an der Gültigkeit orientierte Prüfung gilt schließlich, dass die Leistungen eines Individuums nicht im Vergleich zu den übrigen Mitgliedern derselben oder einer außenstehenden Gruppe, sondern entsprechend den aufgabenbezogenen Anforderungen bewertet werden.

Hier besteht nicht der Raum, die hier aufgezeigten Facetten und Variationsmöglichkeiten von Prüfungsprozessen an Beispielen zu veranschaulichen (s. aber etwa Breuer & Lübke-König 2006; Metzger 2005, 2006; Metzger & Nüesch 2004; Metzger, Dörig & Waibel 1998). Es sollte aber deutlich werden, dass berufsbildende Schulen, basierend auf einem durchaus anspruchsvollen berufs- und allgemeinbildend ausgerichteten curricularen Rahmen und einem kognitivistisch-konstruktivistischen Lernverständnis äußerst vielseitiger, variantenreicher Prüfungsprozesse bedürfen, wenn sie der Gültigkeit als primärer Prüfungsanforderung gerecht werden wollen. Auch auf

dieser konkreten Ebene der Prüfungsprozesse wird deutlich, dass an Bildungsstandards orientierte Testaufgaben, schon gar bezogen auf ein Kerncurriculum und Mindestanforderungen, dieser geforderten Vielfalt nie entsprechen können. Gefährlich wäre auch, wenn sozusagen im Sinne einer Arbeitsteilung der Schule all jene Prüfungsverfahren überlassen würden, die nicht standardisierbar erscheinen (so etwa mündliche Prüfungen im Sinne echter Fach- und Beratungsgespräche oder komplexere, zeitaufwändige Fallaufgaben), denn ein Konzept von Prüfungsprozessen ist kein additives Gebilde, sondern es ist ein integriertes Gebäude, aus welchem – um bildlich zu bleiben – nicht einzelne Bausteine entfernt werden können.

4 Antwort auf die Problemstellung

Der Lehrkraft im einleitenden Beispiel, die ihr Prüfungsverfahren hinterfragt, ist zu raten, ihr sehr prozess- und handlungsorientiertes Verfahren bezüglich einer umfangreichen Schlussarbeit von Schülern als Teil eines „ganzheitlich“ orientierten, integrierten Prüfungskonzeptes beizubehalten. Der Empfehlung von Klieme et al. (2003, S. 39), Bildungsstandards nicht als Kriterien für die Notengebung zu verwenden, ist beizupflichten. Allerdings kann der Begründung, dass diese nämlich „nach professionellen Maßstäben“ (S. 39) der Lehrpersonen erfolgen müsse, nicht gefolgt werden. Zum einen sind professionelle Maßstäbe auch an jene Personen anzulegen, die Bildungsstan-

dards bis hin zu Testaufgaben entwickeln, und diese Maßstäbe sollten nicht wesentlich von jenen professioneller Lehrpersonen abweichen, wenn es um die Diagnose und Förderung von Kompetenzen einzelner Schülerinnen und Schüler geht. Zum anderen darf Prüfen nicht mit Bewerten im Sinne weitreichender Konsequenzen gleichgesetzt werden. Prüfen bedeutet ganz zentral, sich genau dieser Diagnose und Förderung von Schülern anzunehmen, solange es um die formative Funktion von Prüfungen geht. Und schließlich genügen Bildungsstandards den Anforderungen nicht, die an ein kompetenzorientiertes Prüfungskonzept zu stellen sind, wie die Analyse gezeigt haben sollte. Sie sind nicht geeignet, Kompetenzen und Kompetenzentwicklungen der einzelnen Schülerinnen und Schüler funktions- und anforderungsgerecht mittels der erforderlichen Vielseitigkeit an Verfahren zu erfassen, zu fördern und im summarischen Fall auch zu bewerten. Die Orientierung der Schülerbeurteilung an Bildungsstandards und damit verknüpft an Kerncurricula und Mindestanforderungen birgt die Gefahr in sich, schulische Prüfungskonzepte in vielen Dimensionen entgegen aller geforderten und geförderten und damit auch zu prüfenden Kompetenzen auf ein nicht zu rechtfertigendes Minimum zu reduzieren und damit die Gültigkeit als zentrale Anforderung an ein Prüfungskonzept unnötiger- und unberechtigterweise zu verringern. ✘

Literatur

Achtenbagen, F. (1992). Lernen, Denken, Handeln in komplexen ökonomischen Situationen Sechzehn Aussagen zu Ergebnissen des Göttinger Forschungsprojekts, verbunden mit einem Ausblick für eine weitere Forschungs- und Entwicklungsarbeit. In F. Achtenbagen & E. G. John (Hrsg.), *Mehrdimensionale Lehr-Lern-Arrangements. Innovationen in der kaufmännischen Aus- und Weiterbildung* (S. 39–42). Wiesbaden: Gabler.

American Educational Research Association (AERA) & National Council on Measurement in Education (NCME) (2002). *Standards for educational and psychological testing*. Washington, DC: AERA.

Amos, J., Amsler, F., Martin, M. & Metzger, C. (2004). *Evaluation von Abschlussprüfungen der beruflichen Grundausbildung*. Basel: Büro für Kommunikation.

Anderson, L. W. & Krathwohl, D. R. et al. (2001). *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Birenbaum, M. (2003). *New Insights Into Learning and Teaching and Their Implications for Assessment*. In M. Segers, F. Dochy & E. Cascallar (Eds.), *Optimising New Modes of Assessment: Search of Qualities and Standards* (pp. 13–36). Dordrecht: Kluwer Academic Publishers.

Bloom, B. S., Hastings, J. T. & Madans, G. F. (Eds.). (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Breuer, K. & Lübke-König, J. (2006). *Valide Prüfung von Handlungskompetenzen in berufsbezogenen Prüfungen*. In D. Euler (Hrsg.), *Facetten des beruflichen Lernens* (S. 283–314). Bern: h.e.p. Verlag.

Darling-Hammond, L., Ancess, J. & Falk, B. (1995). *Authentic assessment in action*. New York: Teachers College Press.

Dörner, D. (1976). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.

Dubs, R. (2004). *Bildungsstandards – ein erfolversprechender Paradigmawechsel? Ein Umsetzungsversuch als Diskussionsgrundlage im Fach Volkswirtschaftslehre*. In M. Wosnitza, A. Frey & R. S. Jäger (Hrsg.), *Lernprozess, Lernumgebung und Lernidiagnostik. Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert* (S. 38–55). Landau: Verlag Empirische Pädagogik.

Dubs, R. (2006). *Bildungsstandards: Das Problem der schulpraktischen Umsetzung*. Netzwerk, 1, 18–29.

Euler, D. & Halm, A. (2004). *Wirtschaftsdidaktik*. Bern: Haupt Verlag.

Gielen, S., Dochy, F. & Dierick, S. (2003). *Evaluating the Consequential Validity of New Modes of Assessment: The Influence of Assessment on Learning, Including Pre-, Post-, and True Assessment Effects*. In M. Segers, F. Dochy & E. Cascallar (Eds.), *Optimising New Modes of Assessment: Search of Qualities and Standards* (pp. 37–54). Dordrecht: Kluwer Academic Publishers.

Gipps, C. V. (1995). *Beyond testing. Towards a theory of educational assessment*. London: The Falmer Press.

Herman, J. L., Aschbacher, P. R. & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: ASCD.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H. & Prenzel, M. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Berlin: BMBWF.

Linn, R. L. & Gronlund, N. E. (1995). *Measurement and assessment in teaching* (7th ed.). Englewood Cliffs: Prentice Hall.

Messick, S. (1989). *Validity*. In R. E. Linn (Ed.), *Educational measurement* (pp. 13–103). New York: Macmillan.

Metzger, C. & Nüesch, C. (2004). *Fair prüfen. Ein Qualitätsleitfaden für Prüfende an Hochschulen* (Hochschuldidaktische Schriften Bd. 6). St. Gallen: Institut für Wirtschaftspädagogik der Universität St. Gallen.

Metzger, C. (1986). *Formative Prüfungen im Hochschulunterricht*. Aarau/Zürich: Sauerländer/SKV.

Metzger, C. (2005). *Kompetenzorientiert prüfen in der beruflichen Grundbildung der Schweiz: Anspruch und Wirklichkeit gezeigt am Beispiel der kaufmännischen Grundbildung*. *bwp@, Berufs- und Wirtschaftspädagogik online*. www.bwpat.net/ausgabe8/metzger

Metzger, C. (2006). *Lernstrategien funktionsgerecht evaluieren eine didaktische und forschungsmethodische Herausforderung*. Erscheint in *Beiheft der ZBW*.

Metzger, C. & Waibel, R. (1993). *Sind die kaufmännischen Lehrabschlussprüfungen gültig?* *Schweizerische Zeitschrift für kaufmännisches Bildungswesen*, 87, 240–268.

Metzger, C., Dörig, R. & Waibel, R. (1998). *Gültig prüfen. Modell und Empfehlungen für die Sekundarstufe II unter besonderer Berücksichtigung der kaufmännischen Lehrabschluss- und Berufsmaturitätsprüfungen*. St. Gallen: Institut für Wirtschaftspädagogik der Universität St. Gallen.

Moss, P. A. (1994). *Can there be validity without reliability?* *Educational Researcher*, 23 (2), 5–12.

Organisation for Economic Co-operation and Development (OECD). (2005, November). *Formative Assessment: Improving Learning in Secondary Classrooms*. Policy Brief. Author: Paris.

Oser, F. & Renold, U. (2005). *Kompetenzen von Lehrpersonen – über das Auffinden von Standards und ihre Messung*. In I. Gogolin, H. H. Krüger, D. Lenzen & T. Rauschenbach (Hrsg.), *Standards und Standardisierungen in der Erziehungswissenschaft* (S. 119–140). Wiesbaden: Verlag für Sozialwissenschaften.

Reinmann-Rothmeier, G. & Mandl, H. (1998). *Unterrichten und Lernumgebungen gestalten* (Forschungsbericht Nr. 60). München: Ludwig Maximilians Universität.

Resnick, L. B. & Resnick, D. P. (1992). *Assessing the thinking curriculum: New tools for educational reform*. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments. Alternative views of aptitude, achievement and instruction* (pp. 37–75). Boston: Kluwer Academic Publishers.

Segers, M., Dochy, F. & Cascallar, E. (2003). *The Era of Assessment Engineering: Changing Perspectives on Teaching and Learning and the Role of New Modes of Assessment*. In M. Segers, F. Dochy & E. Cascallar (Eds.), *Optimising New Modes of Assessment: Search of Qualities and Standards* (pp. 1–12). Dordrecht: Kluwer Academic Publishers.

Shepard, L. A. (2001). *The Role of Classroom Assessment in Teaching and Learning*. In V. Richardson (Ed.), *Handbook of Research on Teaching* (5th ed., pp. 1066–1101). Washington, DC: American Educational Research Association.

Shuell, T. J. & Moran, K. A. (1994). *Learning theories: Historical overview and trends*. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed., Vol. 6, pp. 3340–3345). Exeter: Pergamon.

Weinert, F. E. (2001). *Vergleichende Leistungsmessung in Schulen – eine unstrittene Selbstverständlichkeit*. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim und Basel: Beltz.

Zimmerman, B. J. (1994). *Dimensions of Academic Self-Regulation: A Conceptual Framework for Education*. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-Regulation of Learning and Performance* (pp. 3–21). Hillsdale, NJ: Lawrence Erlbaum Associates.