

**Tobias HALBHERR¹, Nora DITTMANN-DOMENICHINI,
Thomas PIENDL & Claudia SCHLIENGER (Zürich)**

Authentische, kompetenzorientierte Online-Prüfungen an der ETH Zürich

Zusammenfassung

Dieser Artikel erörtert Möglichkeiten der Gestaltung authentischer, kompetenzorientierter Online-Prüfungen jenseits bisher etablierter Aufgabenformate. Es werden Lösungen für die Gestaltung von Prüfungsumgebungen vorgestellt, welche einen selektiven und gegen Betrug abgesicherten Zugriff auf Ressourcen wie Software, Dateien, Netzwerke oder Internet ermöglichen. Basierend auf einer Diskussion prüfungsdidaktischer Grundlagen und anhand von Beispielen aus der Praxis an der ETH Zürich wird dargelegt, wie die Verfügbarkeit solcher Ressourcen die unmittelbare und authentische Abbildung von Kompetenzen in Prüfungsaufgaben deutlich vereinfacht – oder überhaupt erst ermöglicht – und damit grundlegend neue Möglichkeiten der Prüfungsgestaltung eröffnet.

Schlüsselwörter

Online-Prüfung, Kompetenzorientierung, E-Learning, summative Prüfungen, E-Klausuren, E-Prüfungen

¹ E-Mail: tobias.halbherr@let.ethz.ch



Authentic, competence-oriented online examinations at ETH Zurich

Abstract

This paper discusses the design of authentic, competence-oriented online examinations that go beyond established examination task paradigms. Online examination environments that enable selective access to digital resources, such as professional software, files, networks or the Internet, are presented. Based on didactical considerations and examples from online examination practice at ETH Zurich, this paper shows how the availability of such computer-based resources can facilitate the design of authentic, competence-oriented examination tasks – or make them possible in the first place – and thus open doors to entirely new opportunities in examination design.

Keywords

Online examination, competence orientation, e-learning, e-assessment, summative assessment, authentic assessment

1 Einführung

1.1 Zielsetzung

Prüfungen sind ein zentraler Bestandteil unseres Hochschulwesens: Zum einen dienen sie als Hauptselektionsinstrument und sollen sicherstellen, dass Studierende den Leistungsanforderungen der jeweiligen Hochschule genügen. Zum anderen wird studentisches Lernverhalten nicht durch die Unterrichtsgestaltung allein, sondern auch in hohem Maß durch die erwartete Prüfung bestimmt. Eine zielführende Diskussion der Qualität von Lehre muss deshalb die Qualität von Prüfungen zwingend mitberücksichtigen, wobei uns insbesondere zwei Aspekte wichtig sind: Validität und Lernförderlichkeit.

Validität: Prüfungen sollen als Messinstrumente aussagekräftige, d. h. valide Einschätzungen studentischer Leistungen ermöglichen (APA & NCME, 1999; HASSELHORN & GOLD, 2006; AEA, 2012). Direkte Voraussetzungen für Validität – und damit Teilmengen derselben – sind die Objektivität – die Unabhängigkeit einer Messung von Messumständen und messenden Personen – sowie die Reliabilität – die Genauigkeit und Reproduzierbarkeit einer Messung. Darüber hinaus muss zusätzlich zwingend sichergestellt werden, dass auch tatsächlich gemessen wird, was man zu messen beabsichtigt (NEWTON & SHAW, 2014; RUCH, 1924, S. 13). Dies bedingt entsprechend gestaltete Prüfungsaufgaben.

Lernförderlichkeit: Das Hauptziel einer Hochschule ist die Ausbildung qualifizierter, gut vorbereiteter Studienabgänger/innen. Prüfungen sollen dieses Ziel unterstützen oder zumindest nicht kompromittieren. Seneca bemerkte schon im 1. Jahrhundert nach Christi treffend: „Nicht für das Leben, sondern für die Schule lernen wir“ („non vitae, sed scholae discimus“, epistulae morales ad Lucilium 106, 11-12). Prüfungen beeinflussen studentisches Lernverhalten wie kaum ein anderer einzelner Faktor (CILLIERS, SCHUWIRTH, ADENDORFF, HERMAN & VAN DER VLEUTEN, 2010; JONES, 1923). Es liegt in der Natur der Sache, dass Studierende sich gezwungen sehen, ihr Lernverhalten auf ihre Prüfungen auszurichten: Wollen sie ihre Ausbildung erfolgreich fortsetzen beziehungsweise abschließen, müssen sie ihre Prüfungen bestehen. Handelt es sich zusätzlich um Prüfungen mit hohen Durchfallquoten und starkem Leistungsdruck, bleibt den Studierenden letztlich sehr wenig Freiraum, ihr Lernen unabhängig von ihren Erwartungen bezüglich Form und Inhalt der Prüfung zu gestalten. Doch auch bei Prüfungen mit niedrigen Durchfallquoten bestimmen Art und Inhalt der erwarteten Prüfungsaufgaben maßgeblich, wie Studierende lernen (MARCHANT, 2002).

In diesem Beitrag konzentrieren wir uns auf eine Diskussion neuer Möglichkeiten der *Gestaltung authentischer, kompetenzorientierter Prüfungen mittels computer-gestützter Prüfungsumgebungen*. Grundlage für diese Diskussion bilden die im Folgenden besprochenen prüfungsdidaktischen Ansätze sowie Erfahrungen und

Beispiele aus der Online-Prüfungspraxis² an der ETH Zürich. Besondere Aufmerksamkeit liegt dabei auf der Verbesserung der Qualität von Prüfungen im Hinblick auf deren Validität und Lernförderlichkeit. Die technische Service-Infrastruktur, welche solche Prüfungen ermöglicht, wird ebenfalls kurz vorgestellt. Für eine allgemeinere Diskussion von Online-Prüfungen aus didaktischer, organisatorischer und technischer Sicht verweisen wir auf HALBHERR, REUTER, SCHNEIDER, SCHLIENGER & PIENDL (2014), HILLIER & FLUCK (2013) oder WEBB, GIBSON & FORKOSH-BARUCH (2013).

1.2 Didaktische Ansätze

Im Zuge der kognitiven Wende der zweiten Hälfte des 20. Jahrhunderts hat sich ein neuer Konsens gebildet bezüglich der Frage: Wie funktioniert Lernen? Die behavioristische, auf Vermittlung und Abspeichern von Wissen fokussierte Perspektive wurde ersetzt durch ein kognitiv-konstruktivistisches Paradigma. Dieses betrachtet Lernen als aktiven Prozess des Lernenden in der Auseinandersetzung mit seiner Umwelt und unter Einbezug individueller Erfahrungen sowie erworbenen Vorwissens. Neueren Befunden aus den Neurowissenschaften hält diese Betrachtungsweise stand (z. B. GAZZANIGA, IVRY & MANGUN, 2009). In der Tat bleibt sie für das Verständnis der Wirksamkeit von Lehr- und Lernprozessen weiterhin zentral und wird durch neurowissenschaftliche Perspektiven zwar ergänzt, keinesfalls aber ersetzt (SCHUMACHER & STERN, 2012). Die kognitive Wende führte zur Entwicklung zahlreicher pädagogischer und didaktischer Ansätze, welche helfen sollen, Lehren und Lernen konkreter, aktiver und authentischer zu gestalten. Im Kontrast dazu sind summative Prüfungen leider bis heute häufig in der Überprüfung erworbenen, abstrakten Wissens verhaftet geblieben.

Das Konzept des Constructive Alignment nach BIGGS & TANG (2011) sowie die Lernzieltaxonomie nach BLOOM (ANDERSON, KRATHWOHL & BLOOM,

² „Online-Prüfungen“ sind Prüfungen, die an einem Computer durchgeführt werden, welcher mit einem Netzwerk (Intra- oder Internet) verbunden ist.

2001) sind zwei der einflussreichsten didaktischen Ansätze, welche sich im Zuge der kognitiven Wende in der didaktischen Praxis etabliert haben. Constructive Alignment postuliert, dass gute Didaktik eine kohärente Abstimmung von Lehren, Lernen und Prüfen voraussetzt. Die Lernzieltaxonomie wiederum schlägt eine Klassifikation der kognitiven Fähigkeiten und Fertigkeiten vor, die im Anschluss eines Lernprozess bei den Lernenden erwartet werden. Die Reihenfolge der sechs definierten Kategorien (Wissen, Verstehen, Anwendung, Analyse, Synthese und Evaluation) folgt einer Idee zunehmender Komplexität. Ziel ist es, mit Hilfe der Taxonomie die Kongruenz der drei Komponenten Lernziele, Lernaktivitäten und Prüfungen einfacher bestimmen zu können, indem diese aufgrund der vorgeschlagenen Kategorien klassifiziert werden. Dabei zeigt sich, dass die postulierten Lernziele oft auf komplexere, über die Wissenskategorie hinausgehende Kategorien zielen. Die Entwicklung entsprechender Prüfungsfragen scheitert jedoch häufig und diese zielen – trotz Bemühungen und entsprechender Schulungen – letztendlich „nur“ auf die Kategorie Wissen (KRATHWOHL, 2002). Erstens muss aufgrund dieser Beobachtung davon ausgegangen werden, dass in der Praxis ein gutes Alignment zwischen Prüfung, Lehren und Lernen häufig nicht gegeben ist, sobald Lernziele jenseits der Wissenskategorie geprüft werden sollen. Ein derartiges Misalignment ist problematisch sowohl aus Sicht der Validität als auch der Lernförderlichkeit. Zum einen wird etwas anderes („Wissen“) gemessen als im Lernziel vorgegeben (z. B. „Anwendung“). Zum anderen werden damit fragwürdige Lernanreize gesetzt, welche – da in der Prüfung begründet – höchst verhaltenswirksam sind. Statt „die Anwendung zu üben“ wird „Wissen gepaukt“. Diese Diskrepanz wird häufig von den Studierenden bemerkt und kann starke negative Auswirkungen auf die Lernmotivation haben (z. B. STRUYVEN, DOCHY & JANSSENS, 2005; DOCHY & MCDOWELL, 1998; BIGGS 1996). Zweitens kann man folgern, dass die Lernzieltaxonomie die selbst gesetzten Ziele – zumindest im Hinblick auf die Gestaltung kongruenter Prüfungsaufgaben – in der didaktischen Praxis nur unzureichend einzulösen vermag. In der prüfungsdidaktischen Beratung von Examinatoren setzen wir deshalb zunehmend auf die im Folgenden besprochenen Heuristiken.

Während die Lernzieltaxonomie versucht, gute Prüfungsdidaktik durch eine heuristische Kategorisierung der (vermeintlichen) kognitiven Anforderung von Prüfungsaufgaben zu vermitteln, verfolgen die Konzepte „Kompetenzorientierung“ und „Authentizität“ einen anderen Ansatz. WEINERT (2001) definiert Kompetenzen als „die bei Individuen verfügbaren oder durch sie erlernbaren *kognitiven Fähigkeiten und Fertigkeiten*, um *bestimmte Probleme zu lösen*, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen *in variablen Situationen* erfolgreich und verantwortungsvoll nutzen zu können“³. Bei einem kompetenzorientierten Ansatz müssen die in Lernzielen formulierten Kompetenzen die Studierenden also befähigen, bestimmte Probleme in variablen Situationen tatsächlich erfolgreich lösen zu können. In einer kompetenzorientierten Betrachtung von Prüfungen verlagert sich die Aufmerksamkeit von der Überprüfung abstrakter, nicht direkt beobachtbarer kognitiver Prozesse weg (z. B. nach ANDERSON, KRATHWOHL & BLOOM, 2001), hin zur Beurteilung beobachtbaren Verhaltens: Die Einschätzung der Fähigkeit, mithilfe vermittelter Kompetenzen, bestimmte konkrete – bekannte oder neuartige – Probleme erfolgreich lösen zu können. Es wird also nach geeigneten Möglichkeiten gesucht, Kompetenzen im Prüfungssetting zur Anwendung zu bringen.

Ein wichtiger Leitgedanke kompetenzorientierter Prüfungsgestaltung ist die Authentizität (GIELEN, DOCHY & DIERICK, 2003). Mit Authentizität ist der Grad der unmittelbaren Übereinstimmung von Prüfungen, beziehungsweise Prüfungsaufgaben, mit entsprechenden kriterienrelevanten ‚realen‘ Problemstellungen gemeint. Solche ‚realen‘ Problemstellungen können sich sowohl auf Zielkompetenzen im Berufsfeld beziehen als auch auf notwendige Zwischenschritte im Kompe-

³ Die nicht hervorgehobenen „*motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten*“ sowie die „*verantwortungsvolle Nutzung*“ – Aspekte welche häufig unter den sogenannten überfachlichen Kompetenzen subsummiert werden – werden hier bewusst ausgeklammert. Prüfungssituationen sind aus psychometrischer Sicht wie auch aus praktischen und ethischen Überlegungen (vgl. z. B. DARLEY & BATSON, 1973) unserer Einschätzung nach klar nicht geeignet für die Erhebung solcher Aspekte.

tenzerwerb innerhalb des aktuellen Bildungskontexts. In authentischen Prüfungen werden Kompetenzen mittels realistischer Problemstellungen so direkt, d. h. ‚*authentisch*‘ wie möglich erhoben. Diese Vorgehensweise soll die Gestaltung valider und im Sinne des Alignments kongruenter Prüfungsaufgaben deutlich erleichtern – insbesondere für Prüfende ohne vertiefte Fachkenntnisse in Testentwicklung oder kognitiver Leistungsmessung. Darüber hinaus werden authentische Prüfungen von Studierenden als bedeutungsvoller und relevanter wahrgenommen und können damit einen positiven Einfluss auf Lernmotivation und Engagement entfalten (GULIKERS, BASTIAENS & KIRSCHNER, 2004; HERRINGTON & HERRINGTON, 1998; SAMBELL, MCDOWELL & BROWN, 1997).

2 Neue Möglichkeiten der Aufgabengestaltung mit Computern

Wie lässt sich die Gestaltung authentischer, kompetenzorientierter Prüfungsaufgaben mit einer hohen Kohärenz zu Lernzielen und Lehr- und Lernaktivitäten (Alignment) erleichtern? Traditionelle papierbasierte Prüfungen bieten dazu vergleichsweise begrenzte Möglichkeiten, weil Kandidatinnen und Kandidaten in der Produktion ihrer Antworten auf handschriftliche Methoden wie Texte, Skizzen oder mathematische Herleitungen beschränkt sind. Für zahlreiche Lernziele lassen sich auf diese Weise ‚reale‘ Problemstellungen im Sinne der Authentizität und Kompetenzorientierung heutzutage kaum mehr nachbilden. Einerseits sind Papier und Stift schon lange nicht mehr das Standardwerkzeug am zeitgenössischen Arbeitsplatz, sondern Computer mit Tastatur, Maus, fachspezifischer Software sowie Datei- und Netzzugang. Andererseits bieten Computer im Vergleich zu Papier und Stift deutlich vielfältigere Möglichkeiten, ‚reale‘ Problemstellungen nachzubilden, auch dort wo diese nicht computergebunden sind.

2.1 Der Computer als Werkzeug

Ein Großteil der Tätigkeiten, auf die ein Studium vorbereitet, beinhalten das Arbeiten an und mit Computern. Im Kontext von Alignment sowie dem Anliegen authentischer, kompetenzorientierter Prüfungen ist es deshalb naheliegend, entsprechende Kompetenzen auch direkt am Computer prüfen zu lassen. Maschinenbaustudierende zu einer Problemstellung Skizzen von Maschinenelementen anfertigen und mit Texten erläutern zu lassen, ist eine Sache – sie Bauteile am Computer tatsächlich entwerfen zu lassen, etwas vollkommen anderes. An der ETH Zürich werden solche Online-Prüfungen zurzeit in Fächern wie Statistik, Numerik, Informatik, Finanzmathematik, rechnergestützte Wissenschaften oder Simulationsmethoden durchgeführt. Weitere Prüfungen sind geplant in den Studiengängen Maschinenbau, Bauingenieurwesen, Management, Gesundheitswissenschaften oder Erdwissenschaften. Neben der Arbeit mit reinen Softwarewerkzeugen ist es auch möglich, via peripherer Verbindungen Hardware-basierte Werkzeuge in entsprechende Online-Prüfungen zu integrieren. So wird an der ETH Zürich zum Beispiel eine Prüfung in pflanzenpathologischer Diagnostik mit Hilfe von Mikroskopen mit USB-Anschluss durchgeführt. Bei all diesen Prüfungen geht es dabei *nicht* per se um ein Mitprüfen der viel diskutierten „IT-Kompetenzen“, sondern darum, ob fachspezifische Kompetenzen beherrscht werden, welche auf der Verwendung *konkreter* IT-Werkzeuge basieren.

Solche Prüfungen mit dem Computer als Werkzeug eröffnen zahlreiche Möglichkeiten der Aufgabengestaltung. Studierende können aufgefordert werden, etwas von Grund auf selbst zu erstellen, z. B. ein Programm, einen Algorithmus oder ein CAD-Modell, oder etwas durchzuführen, z. B. eine Berechnung oder eine Mikroskopierprobe. Es können unvollständige Arbeitsvorlagen verteilt werden, welche vervollständigt, erweitert oder verbessert werden sollen, oder es müssen darin enthaltene Fehler identifiziert oder behoben werden. Aufgaben am Computer können auch ohne weiteres um traditionelle Formate ergänzt werden, zum Beispiel indem Studierende in einem Aufsatz ihre Vorgehensweise schriftlich begründen oder identifizierte Fehler analysieren müssen. Es ist sogar möglich, entsprechende ge-

schlossene Aufgaben zu gestalten. So kann z. B. eine Arbeitsvorlage auch als Fallbeispiel für eine Mehrfachwahlaufgabe dienen.

Da Kompetenzen unmittelbar überprüft werden können, vereinfacht sich die Gestaltung authentischer, kompetenzorientierter Prüfungsaufgaben enorm. Die zuverlässige und sinnvolle Unterscheidung besserer und schlechterer Leistungen kann ebenfalls erleichtert werden, da mit den Aufgaben auch die Leistungskriterien der zu erhebenden Kompetenzen authentischer abgebildet werden – statt dass aus abstrahierten Korrelaten – wie Umschreibungen in Textform, Entwurfsskizzen oder Mehrfachwahlaufgaben – indirekt auf diese rückgeschlossen werden müsste. So kann man direkt überprüfen, ob ein programmierter Algorithmus tatsächlich rechnet, was er rechnen soll, ob eine Fragestellung eigenständig und mittels geeigneter statistischer Methoden sinnvoll beantwortet wird oder ob das Durchführen einer Mikroskopierprobe zu gewünschten Ergebnissen führt. In der Praxis zeigt sich allerdings, dass Studierenden aufgrund der höheren Komplexität solcher Aufgaben häufig mehr Zeit für das Bearbeiten gewährt werden muss als bei herkömmlichen Aufgabenformaten. Zudem kann sich die Erhebung von Teilwissen beziehungsweise die Unterscheidung unterschiedlich fortgeschrittener, aber noch nicht abgeschlossener Kompetenzentwicklungen als schwieriger erweisen (FÄSSLER, HINTERBERGER, BOSIA & DAHINDEN, 2005). Es kann deshalb Sinn machen, die Prüfungszeit zu verlängern und ergänzend herkömmliche textbasierte Aufgaben sowie Aufgaben, welche Faktenwissen prüfen, einzusetzen.

2.2 Simulation ‚realer‘ Problemstellungen im Computer

Auch heute findet nicht alle Arbeit am Computer statt. Dennoch bietet der Computer auch in ICT-fernen Bereichen mehr Möglichkeiten, reale Problemstellungen authentisch und kompetenzorientiert nachzubilden als Papier. So können Bild, Ton und Video (einfacher) in die Prüfung eingebunden werden und der Computer bietet durch Interaktion die Möglichkeit einer komplexen Repräsentation ‚realer‘ Problemstellungen sowie Problemlösungen. Ein einfaches Beispiel ist eine Prüfung in

systematischer Botanik, welche die schweizerische Flora mittels eines frei navigierbaren virtuellen Herbariums in die Prüfungsräume bringt.⁴ Mittels photographisch detailliert und einheitlich dokumentierter Pflanzen (Habitus, Habitat, Querschnitte etc.) müssen die Kandidatinnen und Kandidaten Bestimmungen vornehmen und weiterführende Wissensfragen beantworten. Ein ähnliches Beispiel ist die Bestimmung von Vogelarten anhand ihres Gesangs. Solche „Simulationen“ können aber auch weit darüber hinausgehen. Beispiele wären ein virtueller Patient für Prüfungen in der Medizin, eine virtuelle Zelle für die Biologie oder simulierte Umwelten für Prüfungen in Ökonomie, Statik, Kernphysik, Klimaforschung, Soziologie, Chemie, Elektrotechnik und zahllose Fächer mehr. In der Aufgabenstellung hat man hier fast alle Freiheiten.

Dabei können die im Computer nachgebildeten Problemstellungen einerseits im Rahmen klassischer Aufgabenformate, am Computer oder auf Papier, bearbeitet werden. Andererseits ist es aber auch möglich, Aufgaben innerhalb der simulierten Umgebungen selbst oder durch Manipulation derselben bearbeiten zu lassen. Die relevanten Leistungsdaten wiederum können durch die simulierende Software selbst oder durch eine Form von Datei- oder Datenexport in einer Prüfungssoftware erfasst werden (z. B. mittels Screenshots, Videoaufzeichnungen oder Simulationskennzahlen).

2.3 Der Computer als Informationsquelle

Computer sind heutzutage auch das wichtigste Werkzeug zur Beschaffung von Informationen. Das erfolgreiche Navigieren, Auffinden und Beurteilen relevanter Information ist in vielen Disziplinen von überragender Bedeutung. Umgekehrt bleibt die Bedeutung von verinnerlichtem, bewusst abrufbarem Faktenwissen sowie dem Verständnis komplexer Zusammenhänge auch in einer hoch vernetzten Welt mit ubiquitärem Zugriff auf Informationen bestehen. Erstens ist deklaratives semantisches Faktenwissen innert Sekundenbruchteilen abrufbar, während selbst

⁴ Vgl. <http://www.balti.ethz.ch/tiki-index.php?page=Pruefungsbilder>

eine gezielte Internetsuche Sekunden, eher aber Minuten oder gar Stunden dauert. Zweitens ist semantisches Faktenwissen eine Voraussetzung für den Ausbau neuronaler semantischer Netzwerke sowie für die damit interagierenden kognitiven Prozesse (GAZZANIGA, IVRY & MANGUN, 2009). Je komplexer kognitive Prozesse und je stärker sich diese bewusster Kontrolle entziehen, wie z. B. Kreativität oder das Lösen neuartiger Probleme, desto größer die Bedeutung entsprechender gut ausgeprägter neuronaler semantischer Netzwerke. Sowohl Open-book-Prüfungen, welche den Zugang zu externen Informationsquellen erlauben, als auch Closed-book-Prüfungen, welche aus dem Gedächtnis abrufbares Wissen verlangen, haben somit ihre Daseinsberechtigung. Im Sinne der Kompetenzorientierung und Authentizität muss dabei für jede Prüfung (oder Prüfungsaufgabe) einzeln abgewogen werden, ob sich der Open- oder Closed-book-Modus besser zur Erhebung der zu messenden Zielkompetenzen eignet. Beide Arten von Prüfungen können am Computer umgesetzt werden.

In elektronischen Open-book-Prüfungen muss ein selektiver Zugang auf erlaubte Ressourcen gewährleistet werden, während der Zugriff auf nicht erlaubte Hilfsmittel und Kommunikationskanäle verwehrt bleibt. Ressourcen können dabei entweder in Form lokal gespeicherter Dateien zur Verfügung gestellt werden oder es wird der selektive Zugang zu Ressourcen über Intranet und/oder Internet ermöglicht. Ein uneingeschränkter Netzzugriff ist dabei grundsätzlich nicht erwünscht, da jedes Netzwerk, insbesondere aber das Internet, immer auch ein Kommunikationskanal ist und somit nicht mehr sichergestellt wäre, dass Kandidatinnen und Kandidaten ihre Prüfungsleistung auch tatsächlich selbst erbracht haben. Ressourcen in Dateiform können Vorlesungsunterlagen, Übungen oder begleitende Literatur sein, aber auch studentische Notizen und Arbeiten. So kann Studierenden einer Informatikveranstaltung zum Beispiel der Zugriff auf während des Semesters in Übungen entwickelte Prozeduren und Code-Fragmente gewährt werden, damit diese als Bausteine in der Bearbeitung von Prüfungsaufgaben verwendet werden können. Eine Maßnahme, die in diesem konkreten Fall sowohl aus Sicht der Authentizität interessant ist – bei ‚realen‘ Problemstellungen in der Programmierung wird nur selten der gesamte Code von Grund auf neu geschrieben –, als auch der Kohärenz zwi-

schen Lehr- und Lernaktivitäten und Prüfung (Alignment) dient. Das Internet ermöglicht den Zugriff auf eine schier unerschöpfliche Vielfalt an Ressourcen wie Help-Files, FAQs, Bibliotheken, Repositories, öffentliche Datensammlungen oder Forschungsdatenbanken. Einschränkungen sind lediglich dadurch gegeben, dass der Zugriff auf erwünschte Ressourcen technisch bisweilen nur ermöglicht werden kann, indem gleichzeitig der Zugriff auf andere, unerwünschte Ressourcen oder Kommunikationskanäle ermöglicht wird. So verwenden z. B. zahlreiche Webseiten für die interne Suche externe Suchmaschinen wie Google, so dass eine sinnvolle Nutzung derselben nur möglich ist, wenn gleichzeitig der Zugriff auf Google – inklusive aller im Google-Cache gespeicherter Inhalte – ebenfalls freigeschaltet würde.

Der Zugriff auf Ressourcen in elektronischen Open-book-Prüfungen kann dabei zwei unterschiedliche Funktionen erfüllen. Erstens als unterstützende oder erleichternde Rahmenbedingung für Prüfungen, bei welchen ein Abspeichern der entsprechenden Wissensdomänen im Gedächtnis kein zentrales Element des zu messenden Kompetenzkonstrukts ist. Zweitens in Prüfungen, bei welchen der erfolgreiche Umgang mit entsprechenden Ressourcen oder die zielführende Informationssuche an sich Teil des zu messenden Kompetenzkonstrukts sind. So kann es im Sinne authentischer, kompetenzorientierter Prüfungen zum Beispiel durchaus Sinn machen, Studierende Fragestellungen zu ihnen noch unbekannten Fachthemen mittels Zugriff auf Forschungsdatenbanken und -Bibliotheken beantworten zu lassen. Schließlich ist das schnelle und zuverlässige Auffinden relevanter Informationen und das kompetente Einarbeiten in neue Fachthemen eine Schlüsselkompetenz in zahlreichen Disziplinen.

3 Service-Infrastruktur

Wir haben versucht aufzuzeigen, wie Prüfungen durch den Einsatz von Computern kompetenzorientierter und authentischer gestaltet werden können. Um Prüfungen am Computer aber überhaupt erst möglich zu machen, müssen zunächst die relevanten organisatorischen, technischen und rechtlichen Fragen im Bereich Online-

Prüfungen geklärt und entsprechende Hürden im Auf- und Ausbau einer Dienstleistung „Online-Prüfungen“ überwunden werden. Eine ausführliche Diskussion dieser Themen findet sich z. B. in PIENDL, HALBHERR & SCHNEIDER (2014), NIELSEN et al. (2015) oder SCHULZ & APOSTOLOPOULOS (2014). Das Finden geeigneter Wege zur Absicherung der Prüfungscomputer ist dabei eine der größeren Herausforderungen: Der Zugriff auf installierte Software, auf Funktionen des Betriebssystems sowie Internet und Intranet muss verhindert oder selektiv gesteuert werden können. Betrugsversuche durch den Zugriff auf unerlaubte Hilfsmittel, z. B. mittels Web-Recherche, spezialisierter Software, E-Mail, Chat, Dateiserver oder elektronischen Spickzetteln, wären sonst sehr einfach. Eine technische Absicherung der Prüfungscomputer kann die menschliche Prüfungsaufsicht allerdings nicht ersetzen, denn auch ein vollkommen abgesicherter Computer kann herkömmliche Betrugsversuche mittels Zuflüstern, SMS oder Papiernotizen nicht verhindern. Im Folgenden stellen wir als Beispiele kurz die technischen Setups für abgesicherte Online-Prüfungen vor, die zurzeit an der ETH Zürich im Einsatz sind. Für eine detailliertere Diskussion verweisen wir auf HALBHERR et al. (2014) sowie auf REUTER & HALBHERR (2015).

3.1 Lokale Rechner mit Safe Exam Browser (SEB)

Safe Exam Browser⁵ ist eine frei verfügbare Open-Source-Software für Microsoft Windows und MAC OS X, die einen Rechner in eine abgesicherte Prüfungsstation verwandelt (SCHNEIDER, VOLK, LEHRE, BAUER & PIENDL, 2012). SEB versetzt den Rechner in einen sogenannten Kiosk-Modus, der den Zugriff auf weitere Software sowie auf unerwünschte Systemfunktionen verhindert. Gleichzeitig verbindet sich eine integrierte Browserkomponente über das Internet oder ein LAN mit der eigentlichen Prüfungssoftware, zum Beispiel das Prüfungsmodul eines Learning Management Systems (LMS). Die LMS Moodle und ILIAS stellen Funktionen zur Verfügung, welche sicherstellen, dass nur mit SEB auf die Prüfung zu-

⁵ www.safeexambrowser.org

gegriffen werden kann. Solche Prüfungssetups sind vergleichsweise einfach in Aufbau und Betrieb. Man bleibt aber in der Prüfungsgestaltung auf den Funktionsumfang der gewählten Prüfungssoftware beschränkt. In der Regel bieten solche Online-Prüfungsumgebungen eine Auswahl klassischer Aufgabenformate sowie die Möglichkeit der einfachen Einbettung von Multimedia Inhalten.

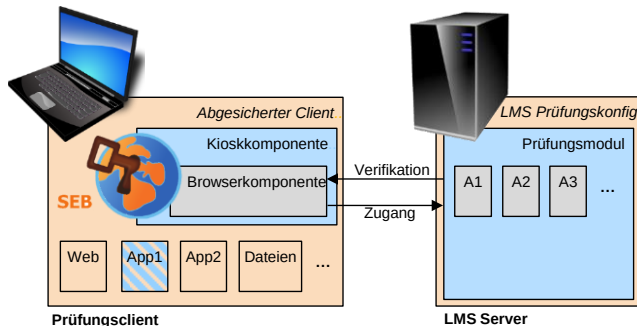


Abb. 1: Setup mit lokalen Prüfungsclients und Safe Exam Browser (SEB).
Auf blaue Bereiche kann zugegriffen werden, auf rote nicht.

3.2 Virtuelle Desktops mit SEB und Drittapplikationen

Nach Bedarf kann die Absicherung durch die Kioskkomponente von SEB gelockert und der Zugriff auf Drittapplikationen (z. B. Matlab, SPSS, PDF-Reader etc.) ermöglicht werden. Zahlreiche Applikationen führen allerdings durch ihren großen Funktionsumfang Hintertüren in die durch SEB abgesicherten Prüfungsumgebungen ein. So verfügen zum Beispiel zahlreiche Programme über einen integrierten Hilfe-Browser, mit welchem mit etwas Einfallsreichtum auch auf das Internet zugegriffen werden kann. Um solche Hintertüren zu schließen, können zusätzlich zu SEB tiefgreifende Änderungen in der Systemkonfiguration nötig werden. Im Prinzip ist es möglich, dass diese auf dem physischen Prüfungsclient vorgenommen werden. Ein solches Vorgehen zeigte sich allerdings in der Praxis bei einer großen Anzahl an Prüfungscomputern sowie dicht aufeinanderfolgenden Prüfungen

als weder ausreichend flexibel noch ausreichend zuverlässig. Eine gute Lösung dieses Problems sind virtuelle Prüfungsclients (REUTER & HALBHERR, 2015). Mit einer sogenannten Virtuellen Desktop Infrastruktur (VDI) können vollständig virtualisierte Desktop-Rechner zur Verfügung gestellt werden. Für jede Prüfung kann nun ein virtueller Rechner als Masterkopie des benötigten Prüfungsclients konfiguriert werden, welcher den gewünschten Zugriff auf Drittapplikationen, Dateien, Netzwerkbereiche und/oder ein LMS-Prüfungsmodul gewährt. Vor der Prüfung kann die erforderliche Anzahl virtueller Prüfungsclients einfach ausgerollt werden, indem diese entsprechend häufig vom Master kopiert werden. Solche Setups mit SEB, VDI und einer Prüfungssoftware sind in Aufbau und Betrieb deutlich anspruchsvoller und aufwändiger. Dafür ermöglichen sie die einfache und flexible Bereitstellung der komplexen – authentischen, kompetenzorientierten – Prüfungssetups, wie sie in Kapitel 2 besprochen wurden.

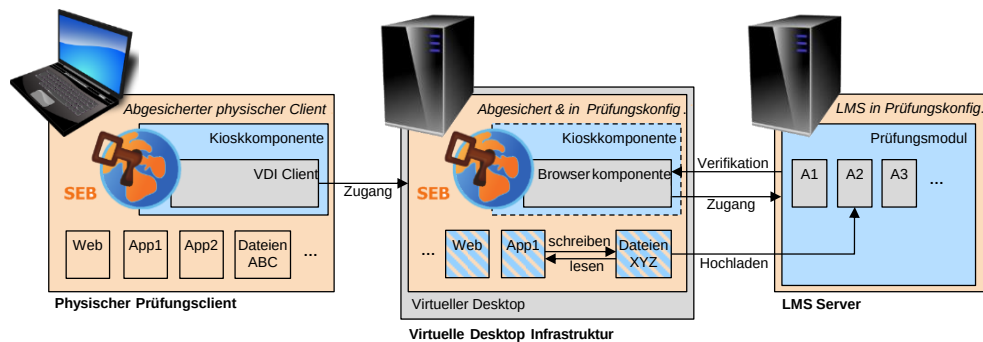


Abb. 2: Setup mit virtuellen Prüfungsclients, SEB und Drittapplikationen.
Auf blaue Bereiche kann zugegriffen werden, auf rote nicht.

3.3 Linux-Setup

Bei Lehrveranstaltungen in informatiknahen Bereichen werden Programmierkenntnisse häufig in Linux-Umgebungen vermittelt und geübt. Entsprechende Prüfungen finden deshalb im Sinne des Alignment ebenfalls in Linux-Umgebungen statt. Da-

bei werden die Prüfungsclients in einem speziellen Prüfungsmodus gestartet, welcher alle prüfungsrelevanten Sicherheitsfunktionen aktiviert. Die Prüfungsrechner sind untereinander nicht erreichbar und die Kommunikation nach außen beschränkt sich auf für die Prüfungsdurchführung unerlässliche Funktionen. Es besteht keine Anbindung an ein LMS. Prüfungsaufgaben und allfällige prüfungsrelevante Dateien werden lokal auf den Prüfungsclients zur Verfügung gestellt oder ausgedruckt. Auch dieses Setup ermöglicht authentische, kompetenzorientierte Online-Prüfungen, insbesondere wie sie in Kapitel 2.1 besprochen wurden.

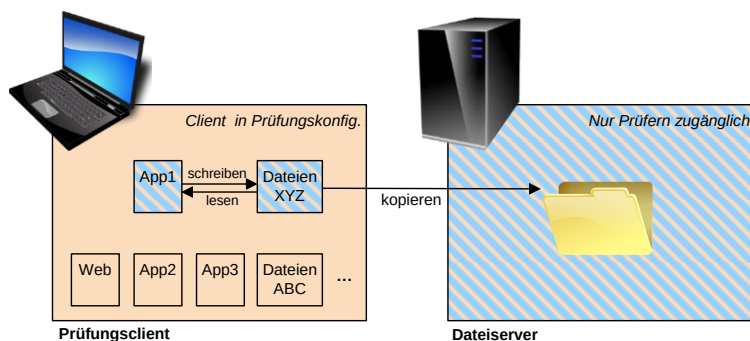


Abb. 3: Linux Setup. Auf blaue Bereiche kann zugegriffen werden, auf rote nicht.

4 Diskussion

Der Computer eröffnet zahlreiche neue Wege zur Gestaltung authentischer, kompetenzorientierter Prüfungen. Aufgrund der besprochenen prüfungsdidaktischen Fachliteratur darf man hiervon sowohl Verbesserungen der Validität entsprechender Kompetenzmessungen als auch substantielle positive Effekte auf Lernverhalten und Lernmotivation von Studierenden erwarten. Anekdotische Berichte von Examinatorinnen und Examinatoren der ETH Zürich bestätigen diese Vorhersagen. Eine empirische Überprüfung derselben steht allerdings noch aus. In diesem Artikel haben wir uns auf die Möglichkeiten zur Gestaltung und Durchführung authentischer, kompetenzorientierter Prüfungen insbesondere im Kontext selektiver,

summativer Prüfungen im Hochschulbereich konzentriert. Basierend auf solchen Prüfungen wird entschieden, welche Art von Fachleuten mit welchen Kompetenzen, Talenten, Stärken und Schwächen in Zukunft in einer Disziplin wirken werden – und welche nicht. Deshalb ist eine sinnvolle Wahl von Selektionskriterien sowie eine valide Erhebung derselben von weitreichender Bedeutung. Aufgrund des hohen Selektionsdrucks ist hier zudem der extrinsisch motivierende Einfluss von Prüfungen auf studentisches Lernverhalten am größten. Aber auch in Bereichen mit geringerem Selektionsdruck hat durch Bologna die Bedeutung und Zahl summativer Leistungskontrollen in Form von Klausuren zugenommen. Die Bedeutung anderer, weniger stark standardisierter Formen von Leistungskontrollen, wie Seminararbeiten oder Vorträgen, hat im Zuge dieses Wechsels gleichzeitig de facto abgenommen. In Klausuren mit traditionellen Aufgabenformaten ist die Wahrscheinlichkeit groß, dass Faktenwissen und Auswendiglernen ein zu großes Gewicht erhalten im Vergleich zu vertieftem Verständnis und dem erfolgreichen Umgang mit komplexen Problemstellungen. Aufgrund dieser Überlegungen sind wir überzeugt, dass die hier diskutierten Ansätze für authentische, kompetenzorientierte Prüfungen am Computer einen wesentlichen Beitrag zur Verbesserung der Qualität von Hochschullehre leisten können. In der Tat laden die neuen Möglichkeiten computerbasierter Prüfungen zu einer grundsätzlichen Überprüfung und allfälligen Neugestaltung der derzeitigen Prüfungspraxis ein. Dabei sehen wir die in diesem Artikel besprochenen Ansätze – wie auch Online-Prüfungen allgemein – keinesfalls als allein gültige Lösung, sondern als wertvolle Ergänzung des Werkzeugrepertoires von Examinatorinnen/Examinatoren. Authentische, kompetenzorientierte Online-Prüfungen können helfen, die Lücke zwischen hoch standardisierten aber unzureichend authentischen Klausuren auf der einen und minimal vorstrukturierten praxisnahen, betreuten Arbeiten auf der anderen, zu schließen. Insbesondere vermögen sie Praxisnähe und Authentizität und damit Validität und eine positivere Lernmotivation dort zu schaffen, wo folgenschwere Entscheidungen in der Ausbildung Studierender getroffen werden müssen.

Voraussetzung hierfür ist allerdings, dass eine entsprechende neue Prüfungspraxis überhaupt erst stattfinden kann. Dies bedingt erhebliche Investitionen in Entwick-

lung und Betrieb entsprechender Infrastrukturen und Dienstleistungen. Darüber hinaus bringt die Technologisierung von Prüfungen auch neuartige Fragestellungen und Risiken mit sich – zum Beispiel der Umgang mit technischen Problemen während einer Prüfung, Fragen der Datenhaltung und des Datenschutzes oder die neu notwendige Koordination zwischen Didaktik-Spezialistinnen und -Spezialisten, IT-Dienstleistenden und Examinatorinnen und Examinatoren der Hochschule. In diesem Zusammenhang ist zu beachten, dass der Betrieb von Infrastrukturen für die diskutierten authentischen, kompetenzorientierten Online-Prüfungen deutlich aufwändiger und anspruchsvoller ist als bei Infrastrukturen, welche sich auf klassische Aufgabenformate beschränken. Es macht deshalb durchaus Sinn, sich in einem ersten Schritt auf Online-Prüfungen mit herkömmlichen Aufgabenformaten zu konzentrieren. Nachdem erste Erfahrungen gesammelt und Betriebserfahrung aufgebaut wurden, kann anschließend das Serviceangebot schrittweise und mit dosiertem Risiko erweitert werden. Ein Mangel an Nachfrage hingegen ist gemäß unserer Erfahrung kaum zu befürchten. Zu groß scheinen das Interesse für sowie die Vorteile von Online-Prüfungen. Neu geschaffene Kapazitäten waren in unserem Fall immer innerhalb weniger Semester vollständig ausgelastet. Derzeit werden um die zehn Prozent aller schriftlichen Prüfungen der ETH Zürich als Online-Prüfungen abgelegt – ca. 100 Prüfungen mit insgesamt ca. 10.000 Kandidatinnen und Kandidaten pro Jahr. Bei rund der Hälfte der bestehenden Online-Prüfungen handelt es sich um Varianten authentischer, kompetenzorientierter Online-Prüfungen, welche alle in ihrer jetzigen Form auf Papier nicht möglich wären. Der Flaschenhals für die Verbreitung von Online-Prüfungen an der ETH Zürich war und ist dabei die Geschwindigkeit, mit welcher die Hochschule im Stande ist, neue Kapazitäten zur Verfügung zu stellen – und nicht die Bereitschaft der Examinatorinnen und Examinatoren, sich auf diese neue Art Prüfungen einzulassen.

Die in diesem Artikel beschriebenen Ansätze sind das vorläufige Ergebnis von über sechs Jahren Online-Prüfungspraxis an der ETH Zürich sowie der begleitenden Auseinandersetzung mit prüfungsdidaktischen Themen. Bedauerlicherweise werden Online-Prüfungen häufig noch immer ausschließlich im Rahmen althergebrachter Paradigmen und Methoden diskutiert und Vorteile vor allem in vermeint-

lich realisierbaren ökonomischen Einsparungen ausgemacht. In der Tat sind Online-Prüfungen in der Vorstellung vieler gleichzusetzen mit Multiple-Choice-Prüfungen am Computer. Eine solch enge Perspektive birgt die Gefahr, dass die Chancen verpasst werden, die mit den vielfältigen neuen und fundamental andersartigen Möglichkeiten des Computers einhergehen. Computer haben fast sämtliche Arbeitsprozesse von Grund auf revolutioniert und völlig neue überhaupt erst möglich gemacht. Warum sollte es im Prüfungsbereich anders sein?

5 Literaturverzeichnis

American Psychological Association & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. American Educational Research Assn.

Anderson, L. W., Krathwohl, D. R. & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Boston, MA: Allyn & Bacon.

Association of Educational Assessment – Europe (2012). *European Framework of Standards for Educational Assessment 1.0*.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.

Biggs, J. & Tang, C. (2011). *Teaching for quality learning at university*. Maidenhead: McGraw-Hill International.

Cilliers, F. J., Schuwirth, L. W., Adendorff, H. J., Herman, N. & van der Vleuten, C. P. (2010). The mechanism of impact of summative assessment on medical students' learning. *Advances in health sciences education*, 15(5), 695-715.

Darley, J. M. & Batson, C. D. (1973). "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27(1), 100.

Dochy, F. J. R. C. & McDowell, L. (1998). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279-298.

Faessler, L., Hinterberger, H., Bosia, L., & Dahinden, M. (2005). Assessment as an instrument to evaluate quality of instruction. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Vol. 2005, No. 1, S. 3555-3562).

Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (2009). *Cognitive Neuroscience. The Biology of the Mind* (S. 312-363). New York: Norton.

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 (Hrsg.), *Optimising new modes of assessment: In search of qualities and standards* (S. 37-54). Dordrecht: Kluwer Academic Publishers.

Gulikers, J. T., Bastiaens, T. J. & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52(3), 67-86.

Halbherr, T., Reuter, K., Schneider, D., Schlienger, C. & Piendl, T. (2014). Making Examinations more Valid, Meaningful and Motivating: The Online Exams Service at ETH Zurich. *EUNIS Journal of Higher Education IT*, 2014-1.

Hasselhorn, M. & Gold, A. (2006). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren* (S. 354-363). Stuttgart: W. Kohlhammer Verlag.

Herrington, J. & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Education Research & Development*, 17(3), 305-322.

Hillier, M. & Fluck, A. (2013). Arguing again for e-exams in high stakes examinations. In *2013 Australian Society for Computers in Learning and Tertiary Education Conference* (S. 1-11).

James, H. W. (1927). The effect of handwriting upon grading. *The English Journal*, 16(3), 180-185.

Jones, H. E. (1923). Experimental studies of college teaching: The effect of examination on the permanence of learning. *Archives of Psychology*, 10, 1-70.

- Klein, J. & Taub, D.** (2005). The effect of variations in handwriting and print on evaluation of student essays. *Assessing Writing*, 10(2), 134-148.
- Krathwohl, D. R.** (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212-218.
- Marchant, G. J.** (2002). Student reading of assigned articles: Will this be on the test? *Teaching of Psychology*, 29(1), 49-51.
- Marshall, J. C. & Powers, J. M.** (1969). Writing neatness, composition errors, and essay grades. *Journal of Educational Measurement*, 6(2), 97-101.
- Newton, P. & Shaw, S.** (2014). *Validity in Educational and Psychological Assessment*. London: SAGE.
- Nielsen, K. G., Petersen, L., Wallstedt, B., Basse, P., Hansen, P. S., Hansen, S. S. & Sørensen, S. M.** (2015). Evaluation of Digital Assessment. *EUNIS Journal of Higher Education IT*, 2015-1.
- Piendl, T., Halbherr, T. & Schneider, D.** (2014). Online-Prüfungen an der ETH Zürich: Vom Projekt zum Service. In T. Škerlak, H. Kaufmann & G. Bachmann (Hrsg.), *Lernumgebungen an der Hochschule. Auf dem Weg zum Campus von morgen*. Münster, New York, München, Berlin: Waxmann.
- Reuter, K. & Halbherr, T.** (2015). Kompetent und sicher: Online-Prüfungen mit Virtueller Desktop Infrastruktur und Safe Exam Browser an der ETH Zürich. In *Tagungsband GML² 2015: Grundfragen Multimedialen Lehrens und Lernens. E-Examinations: Chances and Challenges*.
- Ruch, G. M.** (1924). *The Improvement of the Written Examination*. Chicago, IL: Scott, Foresman and Company.
- Sambell, K., McDowell, L. & Brown, S.** (1997). But is it fair? An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23(4), 349-371.
- Schneider, D. R., Volk, B., Lehre, M., Bauer, D. & Piendl, T.** (2012). Der Safe Exam Browser. Innovative Software zur Umsetzung von Online-Prüfungen an der ETH Zürich. In *Digitale Medien-Werkzeuge für exzellente Forschung und Lehre* (S. 431-441).

Schulz, A. & Apostolopoulos, N. (2011). eExaminations Put To Test: Potenziale computergestützter Prüfungen. In *Hamburger eLMAGAZIN #07: eAssessment auf dem Prüfstand* (S. 38-40).

Schulz, A. & Apostolopoulos, N. (2014). E-Examinations at a Glance – Die Computerisierung des Prüfungswesens an der Freien Universität Berlin. In T. Škerlak, H. Kaufmann & G. Bachmann (Hrsg.), *Lernumgebungen an der Hochschule. Auf dem Weg zum Campus von morgen*. Münster, New York, München, Berlin: Waxmann.

Schumacher, R. & Stern, E. (2012). Neurowissenschaften und Lehr-Lern-Forschung: Weiches Wissen trägt zu lernwirksamem Unterricht bei? / Neuroscience and Research on Learning and Instruction: What Kind of Knowledge Contributes to Educational Outcome? *Die Deutsche Schule*, 104(4), 383.

Struyven, K., Dochy, F. & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review 1. *Assessment & Evaluation in Higher Education*, 30(4), 325-341.

Webb, M., Gibson, D. & Forkosh-Baruch, A. (2013). Challenges for information technology supporting educational assessment. *Journal of Computer Assisted Learning*, 29(5), 451-462.

Weinert, F. E. (2001). *Leistungsmessungen in der Schule* (S. 27). Weinheim, Basel: Beltz.

Autorinnen und Autoren



Lic. phil. Tobias HALBHERR || ETH Zürich, Lehrentwicklung und -technologie || Haldenbachstrasse 44, CH-8092 Zürich

www.let.ethz.ch/people/tobiahal

tobias.halbherr@let.ethz.ch



Dr. Nora DITTMANN-DOMENICHINI || ETH Zürich, Lehrentwicklung und -technologie || Haldenbachstrasse 44, CH-8092 Zürich

www.let.ethz.ch/people/norad

nora.dittmann@let.ethz.ch



Dr. Thomas PIENDL || ETH Zürich, Lehrentwicklung und -technologie || Haldenbachstrasse 44, CH-8092 Zürich

www.let.ethz.ch/people/piendlt

thomas.piendl@let.ethz.ch



Dr. Claudia SCHLIENGER || ETH Zürich, Lehrentwicklung und -technologie || Haldenbachstrasse 44, CH-8092 Zürich

www.let.ethz.ch/people/schlclau

claudia.schlienger@let.ethz.ch