

Patrick Schubert

**E-Assessments**

Möglichkeiten und Anforderungen  
im Kontext kompetenzorientierter Hochschuldidaktik

eingereicht als

DIPLOMARBEIT

an der



HOCHSCHULE MITTWEIDA  

---

UNIVERSITY OF APPLIED SCIENCES

Fakultät Elektro- und Informationstechnik

Mittweida, 2011

Erstprüfer: Prof. Dr.-Ing. Robert J. Wierzbicki

Zweitprüfer: Dipl.-Ing. Sieglinde Klimant

Die vorgelegte Arbeit wurde verteidigt am:

### **Bibliographische Beschreibung:**

Schubert, Patrick:

E-Assessments - Möglichkeiten und Anforderungen im Kontext kompetenzorientierter Hochschuldidaktik. - 2011. - 70 S.

Mittweida, Hochschule Mittweida, Fakultät Elektro- und Informationstechnik, Diplomarbeit, 2011

### **Referat:**

In der vorliegenden Arbeit wird der zeitgemäße Einsatz von E-Assessments als Alternative zu klassischen Prüfungsverfahren und effektives Mittel zur Individualisierung von Lernprozessen, speziell im Hochschulbereich, untersucht. Dazu wird die bisherige Anwendungspraxis klassischer Prüfungsverfahren kritisch betrachtet und mit den Möglichkeiten und Anforderungen innovativer E-Assessments verglichen. Anhand dieser Erkenntnisse und lerntheoretischer Aspekte werden daraufhin alternative Lösungsansätze für den zeitgemäßen Einsatz elektronischer Test- und Bewertungsverfahren aufgezeigt. Abschließend wird die praktische Umsetzbarkeit der erarbeiteten Ansätze mit heutigen Lernplattformen geprüft.

# Inhaltsverzeichnis

I	Abbildungsverzeichnis.....	4
II	Abkürzungsverzeichnis .....	5
1.	Einleitung .....	6
1.1	Problemstellung .....	6
1.2	Zielsetzung .....	7
1.3	Methodische Vorgehensweise.....	8
2.	Theoretische Grundlagen des E-Assessment .....	10
2.1	Begriffserklärung und Definition von Assessment und E-Assessment .....	10
2.2	E-Learning als Anwendungsbereich des E-Assessment .....	11
2.3	Technische Hintergründe zum Einsatz von E-Assessments.....	12
2.3.1	Entwicklung von E-Assessments .....	12
2.3.2	Klassifikationen von E-Assessments .....	12
2.3.3	Sicherheits- und rechtliche Aspekte .....	16
2.4	Pädagogische Grundlagen .....	20
2.4.1	Behaviorismus.....	20
2.4.2	Kognitivismus .....	21
2.4.3	Konstruktivismus.....	22
3.	E-Assessment als Nachfolger klassischer Prüfungen - Motivation für den Einsatz.....	24
3.1	Aktuelle Prüfungssituation an deutschen Hochschulen .....	25
3.1.1	Prüfungsalltag .....	25
3.1.2	Prüfungsplanung .....	28
3.1.3	Prüfungsformen.....	29
3.1.4	Bisheriger Einsatz von E-Assessments .....	34
3.2	Anforderungen an E-Assessments als zeitgemäße Leistungsnachweise aus heutiger Sicht.....	36
3.2.1	Überprüfung kompetenzorientierter Lernziele .....	36
3.2.2	Gestaltung eines schülerzentrierten Unterrichts.....	40
3.2.3	Verfügbarkeit/ Durchführbarkeit von E-Assessments.....	43
3.3	Rahmenbedingungen des E-Assessment - Interface-Design und Assessment-Umgebung .....	45
3.3.1	Intuition (Intuition).....	46

3.3.2	Consistency (Beständigkeit) .....	46
3.3.3	Optimization (Optimierung).....	47
3.4	Empirische Untersuchung - Instant-Feedback-Experiment.....	48
3.4.1	Versuchsbeschreibung und -durchführung des Experiments .....	49
3.4.2	Versuchsergebnisse des Experiments .....	49
3.4.3	Auswertung des Experiments .....	52
4.	Moodle 2.0 - Assessment-Möglichkeiten und -Grenzen der populären Lernplattform .....	53
4.1	Kurzbeschreibung der Lernplattform Moodle und deren Notwendigkeit im E-Learning.....	53
4.2	Vorhandene Testarten und Erweiterungsmöglichkeiten in Moodle im Vergleich mit anderen Lernplattformen.....	54
4.3	Möglichkeiten der Umsetzung entwickelter E-Assessment-Ansätze in Moodle.....	58
5.	Zusammenfassung und Ausblick .....	61
III	Anhang .....	64
IV	Literaturverzeichnis .....	66
V	Erklärung zur selbständigen Anfertigung der Arbeit .....	70

# I     **Abbildungsverzeichnis**

Abbildung 1: Klassifikation von E-Assessments nach BULL & MCKENNA (2004) .....	14
Abbildung 2: Sicherheitstechnische Anforderungen an E-Assessments im Vergleich mit klassischen Prüfungen .....	16
Abbildung 3: Reiz-Reaktion-Schema mit Gehirn als „Black-Box“ .....	20
Abbildung 4: Abbildung des Lernprozesses im Kognitivismus nach Baumgartner (1997) .....	21
Abbildung 5: Schematisches Lernmodell des Konstruktivismus .....	23
Abbildung 6: Abhängigkeit von Validität, Reliabilität und Objektivität .....	24
Abbildung 7: Appendix- und Herz-Kreislauf-Modell nach EUGSTER (2009/2010) .....	29
Abbildung 8: Kontinuierliche Erfassung von Lern- und Lösungsprozessen im Vergleich zur punktuellen Leistungsmessung.....	37
Abbildung 9: Beispielschema für einen semi-automatischen E-Assessment-Workflow .....	39
Abbildung 10: Individuelle Lernpfad-Bestimmung durch E-Assessments .....	40
Abbildung 11: Beispielschema für einen semi-automatischen Feedback-Workflow.....	43
Abbildung 12: Beispielfrage aus dem Experiment .....	49
Abbildung 13: Antwortverteilung der Kontroll- und Feedback-Gruppe im Instant-Feedback-Experiment.....	50
Abbildung 14: Beispiel-Benutzeroberfläche der Lernplattform „Moodle“ .....	53
Abbildung 15: Kurzantwort-Frage in „Moodle“ .....	55
Abbildung 16: OpenMeetings-Plugin in „Moodle“ .....	57

## II Abkürzungsverzeichnis

<b>ARPAnet</b>	Advanced Research Projects Agency Network
<b>CAA</b>	Computer-Assisted-Assessments
<b>CBA</b>	Computer-Based-Assessments
<b>E-Form</b>	elektronische Form
<b>ICT</b>	Information and Communication Technology
<b>iGRAF</b>	Instant Graphical Feedback
<b>JISC</b>	Joint Information Systems Committee
<b>LAN</b>	Local Area Network
<b>MC</b>	Multiple-Choice
<b>OMR</b>	Optical Mark Recognition
<b>PCs</b>	Personalcomputer
<b>SCORM</b>	Sharable Content Object Reference Model
<b>TASA</b>	Text-Guided Automated Self-Assessment

# 1 Einleitung

## 1.1 Problemstellung

Die Präsenz von elektronischen Hilfsmitteln ist in heutigen Lernprozessen allgegenwärtig. Galten bis vor einiger Zeit noch Bücher und somit das Medium Papier als „persistente“ Informationsquelle Nummer eins, bieten mittlerweile computergestützte Lösungen gleichwertige Qualitäten und zudem noch viele Ergänzungen. Generell werden Informationen heute vorwiegend elektronisch festgehalten, verarbeitet und verteilt. Der Umgang mit entsprechenden Geräten, zum Beispiel PCs, Notebooks oder Smartphones ist vertraut und Voraussetzung für viele alltägliche Tätigkeiten. (JIM, 2008)

Durch dieses Nutzungsverhalten haben sich auch die Einstellung zum Lernen und die Lernmethoden verändert. E-Learning, also das wörtlich übersetzte „elektronische Lernen“ ist heute ein wichtiger Bestandteil des Lernprozesses. Gemeint ist damit nicht nur der gezielte Einsatz von E-Learning als Ergänzung oder Alternative zur Präsenzlehre sondern auch als Mittel zur akuten Wissensbeschaffung im täglichen Leben. Unbekannte Wörter werden „gegoogelt“, Definitionen in der „Wikipedia“ nachgeschlagen und Computerprobleme oft in Foren diskutiert. (Fischer & Schwendel, 2009) All diese alltäglichen Vorgänge lassen sich als intentionaler (absichtlicher) oder inzidenteller (beiläufiger) Wissenserwerb, also Lernen, bezeichnen.

Entsprechend naheliegend scheint somit auch eine Erweiterung des Einsatzes von elektronischen Hilfsmitteln auf den Bereich der Leistungsüberprüfung, der aktuell zumeist nach wie vor durch schriftliche oder mündliche Prüfungen abgedeckt wird. Durch den Einsatz von elektronischen Test- und Bewertungsverfahren, also „E-Assessments“ könnte das Prüfungssystem revolutioniert und zeitgemäßer gestaltet werden. Einige Hochschulen haben bereits das Potential von E-Assessments erkannt und setzen elektronische Prüfungen und Leistungsbewertungen erfolgreich ein. Viele weitere Hochschulen prüfen aktuell den Einsatz entsprechender Methoden. Allerdings beschränkt sich der Einsatz zumeist noch auf eine direkte Umsetzung konventioneller Prüfungen in die „E-Form“, beispielsweise bei Multiple-Choice-Prüfungen. Der Anspruch an ein, aus dieser Sicht, gutes E-Assessment lautet demnach: So gut zu sein, wie die äquivalente klassische Prüfung.

Die Sinnhaftigkeit für die Einführung von E-Assessments begründet sich aber viel mehr in den vorhandenen Schwächen konventioneller Prüfungen, unter methodischen, didaktischen und ökonomischen Gesichtspunkten, die durch den Einsatz von elektronischen Test- und

Bewertungsverfahren beseitigt, beziehungsweise zumindest vermindert werden sollen. Der Anspruch an ein zeitgemäßes E-Assessment muss demnach lauten: besser zu sein, als die äquivalente klassische Prüfung. Dieser Anspruch lässt sich nicht mehr durch bloßes adaptieren von bekannten Vorgehensweisen in ein neues Medium erfüllen, sondern bedarf umfassender, tiefgreifender Änderungen und der Bereitschaft zu Neuem. Ein erfolgreicher Einsatz schließt nicht zuletzt den Kreislauf der Nutzung von elektronischen Medien im Lernprozess, der bisher meist durch den Einsatz von „Paper-and-Pencil-Prüfungen“, auch wenn zuvor alternative Lernmethoden zum Einsatz kamen, unterbrochen wurde.

## 1.2 Zielsetzung

Die eingangs beschriebene Problematik lässt erkennen, dass der Einsatz von E-Assessments ein Thema ist, dem aktuell viel Beachtung geschenkt wird. Gleichzeitig besteht durch den bisher nur geringfügigen qualitativen Einsatz entsprechender Verfahren, speziell im Hochschulbereich, ein großer Informationsbedarf an methodischen, praktischen und rechtlichen Fakten. Letzteres soll allerdings im Rahmen dieser Arbeit nur beiläufig betrachtet werden.

Hauptziel der vorliegenden Arbeit ist die Untersuchung, inwiefern sich durch den Einsatz von E-Assessments klassische Prüfungsverfahren, speziell im Hochschulbereich, didaktisch und methodisch verbessern lassen und welche Voraussetzungen entsprechende E-Assessment erfüllen müssen um dieses Ziel zu erreichen. Diese Untersuchung erfolgt im Kontext vorhandener und zukünftiger Lernmodelle und Lernparadigmen.

Dazu wird einerseits die aktuell angewandte Prüfungspraxis, die häufig aus einmaligen Leistungsfeststellungen am Ende eines Kurses oder Moduls besteht, generell in Frage gestellt und deshalb auf ihre Validität, speziell bei einem Einsatz von schriftlichen und mündlichen Prüfungen, im aktuellen Kontext geprüft. Andererseits werden die Möglichkeiten und Grenzen von E-Assessments sowie ihrer Umsetzbarkeit untersucht. An Hand dieser Erkenntnisse werden anschließend alternative Lösungsansätze zur zeitgemäßen, individuellen Leistungsbewertung mithilfe von E-Assessments erarbeitet. E-Assessments werden dabei nicht nur als Alternative oder Ersatz konventioneller Prüfungen zur Leistungsbewertung (summative E-Assessments) betrachtet, sondern auch als Medium zur Eignungstestung (diagnostische E-Assessments) und vorwiegend zur kontinuierlichen Messung und Überwachung des Lernfortschritts (formative E-Assessments). Somit werden E-Assessments nicht nur als Prüfungsersatz, sondern auch als aktives Element zur Gestaltung eines schülerzentrierten Unterrichts gesehen.



Eine Betrachtung der E-Assessment-Möglichkeiten der Lernplattform „Moodle“ im Vergleich mit anderen Lernplattformen wird zeigen, ob und in welchem Umfang sich die zuvor erarbeiteten Strategien mit bestehenden Systemen bereits umsetzen lassen.

Im Rahmen dieser Arbeit werden detaillierte technische oder rechtliche Aspekte der Umsetzung von E-Assessments nicht näher erläutert. Außerdem enthält die vorliegende Arbeit keine Anleitung zur Erstellung von E-Assessments mit der Lernplattform „Moodle“.

Diese Arbeit möchte einen Beitrag zur aktuellen akademischen Debatte über elektronische Assessments leisten und richtet sich an Verantwortliche an Hochschulen und anderen Bildungseinrichtungen, die sich über zeitgemäße E-Assessments informieren oder diese einsetzen möchten und dabei alternative Lern- und Prüfungsmethoden zur Gestaltung eines virtuellen, schülerzentrierten Unterrichts nutzen wollen. Detaillierte Vorkenntnisse im Bereich elektronischer Test- und Bewertungsverfahren sowie tiefgreifendes technisches, pädagogisches oder psychologisches Hintergrundwissen sind zum Verständnis dieser Arbeit nicht notwendig.

### **1.3 Methodische Vorgehensweise**

Um die oben beschriebene Zielsetzung der vorliegenden Arbeit, unter der Prämisse einem großen Leserspektrum einen umfassenden Einblick in die Thematik zeitgemäßer E-Assessments zu geben, möglichst optimal umsetzen zu können, werden eingangs sowohl begriffliche als auch theoretische Hintergründe zum Thema E-Learning und E-Assessment erläutert. Diesbezüglich werden auch Grundlagen zur technischen Entwicklung und zu den allgemeinen Anforderungen an elektronische Assessments, vorerst ungeachtet ihrer didaktischen Aufgaben, dargestellt.

Eine anschließende Erläuterung der wichtigsten lerntheoretischen Grundsätze schafft dann eine Basis für die spätere kombinierte Betrachtung der Möglichkeiten und Anforderungen von E-Assessments im Kontext zur aktuellen Lern- und Prüfungssituation im Hochschulbereich und deren Spezifikationen. Dazu werden zunächst verschiedene klassische Prüfungsformen vorgestellt und hinsichtlich ihrer Validität, Objektivität und Reliabilität untersucht um daraufhin alternative Lösungsansätze vergleichen zu können. Die Betrachtung der Möglichkeiten sowie der Vor- und Nachteile von E-Assessments erfolgt anhand aktueller internationaler Literatur und neuer Erkenntnisse aus Forschung und Technik. Um einen möglichst umfassenden Einblick in die komplexe Thematik der elektronischen Assessments im aktuellen lerntheoretischen Kontext geben zu können werden vor allem Lösungsansätze präsentiert und konkrete Beispiele bewusst kurz gehalten, da

jede Anwendung individuell betrachtet werden muss und sich wenig verallgemeinern lässt. In der abschließenden Betrachtung der theoretischen Grundlagen und erarbeiteten Lösungsansätze in Verbindung mit der populären Lernplattform „Moodle“, werden die Erkenntnisse beispielhaft in die Praxis übertragen und hinsichtlich ihrer heutigen Umsetzbarkeit geprüft.

## 2 Theoretische Grundlagen des E-Assessment

### 2.1 Begriffserklärung und Definition von Assessment und E-Assessment

Für den Begriff E-Assessment existiert gegenwärtig kein einheitliches Begriffsverständnis, sondern vielmehr eine Ansammlung verschiedener Definitions- und Begriffserklärungen sowie diverser, teilweise synonym verwendeter Bezeichnungen. Erstmals international publiziert wurde der Begriff „E-Assessment“ laut RUEDEL im Jahr 2000 (Ruedel, 2010). Seitdem wird er auch im deutschen Sprachraum zunehmend als Sammelbezeichnung für Prüfungen- und Testszenarien verwendet, bei denen elektronische Hilfsmittel, vorwiegend Computer, zum Einsatz kommen. Es existieren allerdings auch andere Bezeichnungen wie: „Online-Assessment“, „E-Examination“ oder „elektronische Prüfung“ und „Computer-Prüfung“ als eingedeutschte Varianten, welche aber nur eine geringe Trennschärfe aufweisen und langsam durch den Begriff „E-Assessment“ abgelöst werden. Da das Wort „Assessment“ eine große Bandbreite an Bedeutungen und Übersetzungsmöglichkeiten liefert, lassen sich mit ihm viele Test-Szenarien, von einfachen Multiple-Choice-Prüfungen, bis hin zu komplexen, lernprozessintegrierten Bewertungs- und Testverfahren, beschreiben. Dieser Umstand macht auch eine exakte Definition schwierig und führte zu zahlreichen verschiedenen Definitionsversuchen.

Das britische „JISC“ (Joint Information Systems Committee), ein Komitee das britische Universitäts- und Forschungseinrichtungen bei der Nutzung moderner Informations- und Kommunikationstechnologie unterstützt, definiert E-Assessment wie folgt:

**Definition 1:** „e-Assessment is the end-to-end electronic assessment processes where ICT is used for the presentation of assessment activity, and the recording of responses. [...]“ (JISC, 2007)

Übersetzt bedeutet das, dass E-Assessment den gesamten elektronischen Bewertungs-, Test- oder Prüfungsprozess beschreibt, bei dem Informations- und Kommunikationstechnik als Präsentationsmedium und zum Erfassen der Ergebnisse zum Einsatz kommt.

Eine weitere Definition bezeichnet E-Assessments als:

**Definition 2:** „[...] eine Lernfortschrittskontrolle [...], die mit Hilfe elektronischer Medien vorbereitet, durchgeführt und nachbereitet wird. Eine besondere Rolle spielt dabei die (teil-) automatische Durchführung von Korrekturen im Rahmen des technisch Möglichen“ (Eilers et al., 2008)

Beide Definitionen heben unterschiedliche Facetten des E-Assessment hervor und machen nochmals deutlich, wie wichtig eine kontextbezogene Definition des Begriffs sein kann um einen einheitlichen Ausgangspunkt für weitere Betrachtungen zu schaffen. Im Rahmen dieser Arbeit wird unter E-Assessments folgendes verstanden:

**Definition 3:** E-Assessments sind die Summe elektronisch erstellter und durchführbarer Test- oder Bewertungsverfahren, die in einem Lernprozess integriert sind und diesen beeinflussen können. Ein einzelnes E-Assessment kann sowohl diagnostisch, formativ oder summativ sein.

## 2.2 E-Learning als Anwendungsbereich des E-Assessment

Um einen effektiven Einsatz von E-Assessments im Lernprozess erreichen zu können, ist es notwendig, zunächst einige Grundlagen und Begrifflichkeiten des E-Learning, also des „elektronisch unterstützen Lernens“ als Anwendungsbereich und somit Bezugsrahmen des E-Assessment zu betrachten. Es ist natürlich auch denkbar E-Assessments in die klassische Präsenzlehre zu integrieren, was durchaus in ausgewählten Anwendungsbereichen auch sinnvoll ist, doch hauptsächlich ist ein Einsatz dann von Vorteil, wenn der gesamte Lehr- und Lernprozess elektronisch unterstützt ist und somit eine gute Interaktion zwischen „Learning“, „Teaching“ und „Assessment“ herrscht. So soll im weiteren Verlauf dieser Arbeit vor allem die symbiotische Verbindung von E-Learning und E-Assessment als zeitgemäße Lehrmethodik vorgestellt werden.

### Begriffserklärung E-Learning und E-Education

Wenn im allgemeinen Sprachgebrauch sowie in der Fachliteratur von „E-Learning“ gesprochen wird, so wird damit häufig der gesamte Lehr- und Lernprozess gemeint, bei dem elektronische Hilfsmittel zur Erstellung, Distribution, Präsentation von Lernmaterialien und Interaktion zwischen Lehrenden und Lernenden zum Einsatz kommen. DIETINGER weist aber darauf hin, dass der Begriff „E-Learning“ genau genommen nur reine Lernaspekte betrachtet und nicht die Lehrperspektive beschreibt (Dietinger, 2003). Trotzdem hat sich der Begriff E-Learning weitestgehend als Oberbegriff für alle Aktivitäten des elektronischen Lernens und Lehrens durchgesetzt.

Eine differenziertere Betrachtung beschreibt die elektronisch unterstützte Lehre hingegen als „E-Education“, die sich wiederum aus den Bestandteilen „E-Learning“ und „E-Teaching“ zusammensetzt (Baumgartner et al., 2004). Diese Darstellung kann dann durch „E-Assessment“ als weiterer essentieller Bestandteil noch erweitert werden (Gruttmann, 2009). Betrachtet man

also die elektronisch unterstützte Lehre als komplexes Gebilde mit drei Hauptbestandteilen, so ließe sich E-Learning wie folgt definieren:

**Definition 1:** „E-learning is learning which takes place as a result of experiences and interaction in an Internet environment. It is not restricted to a regular school day and can take place in a variety of locations including home, school and community locations e.g. libraries, cafes etc.“ (Campbell, 2001)

E-Learning beschreibt demnach nur den Lernprozess an sich, der aus Erfahrungen und Interaktionen in einer virtuellen (Internet-) Umgebung resultiert, und nicht an eine reguläre Präsenzlehre oder physische Orte gebunden ist.

Generell wird aber der Begriff E-Learning, wie bereits eingangs beschrieben, weniger atomar betrachtet und vielmehr als zusammenfassende Beschreibung aller Tätigkeiten, Aktivitäten und Arbeitsschritte, die größtenteils elektronisch stattfinden und eine Vermittlung von Wissen oder Kompetenzen zum Ziel haben angesehen - den eigentlichen Prozess des Lernens mit eingeschlossen.

**Definition 2:** „E-Learning kann begriffen werden als Lernen, das mit Informations- und Kommunikationstechnologien unterstützt bzw. ermöglicht wird. Wichtig ist, dass diese Technologien mit dem Lernprozess selbst unmittelbar verbunden sind und nicht nur rudimentäre Hilfsmittel darstellen.“ (Seufert et al., 2001)

An dieser Stelle wird abermals deutlich, dass das unter Punkt 2.2 bereits aufgezeigte Vorhandensein mehrerer Definitionsvarianten für viele Bezeichnung und Schlagworte aus dem Bereich der „elektronischen Lehre“ gilt, und auch hier kein einheitliches Begriffsverständnis existiert. Es ist also notwendig, eigene Vorgehensweisen präzise zu beschreiben und Lösungsansätze zu konkretisieren. Im Rahmen dieser Arbeit bezieht sich der Begriff E-Learning immer auf einen komplexen Lernprozess, von der Entwicklung der Lehrinhalte, über didaktische Methoden und Organisationsformen, bis hin zum Lernen und Bewerten von Lernleistungen selbst. Wenn einzelne Bereiche gesondert betrachtet werden sollen, wird von „Lehren“ oder „Lernen“ gesprochen.

## 2.3 Technische Hintergründe zum Einsatz von E-Assessments

### 2.3.1 Entwicklung von E-Assessments

Die Entstehung des E-Assessment ist eng mit der Entwicklung der Informations- und Kommunikationstechnik, speziell des Computers und später auch des Internets, verbunden. Sie ist aber keineswegs so neu, wie die Bezeichnung es vermuten lässt. Als eine der ersten Anwendungen lässt sich der „Automatic Grader“ aus dem Jahr 1959, also drei Jahre vor Entwicklungsbeginn des legendären ARPAnet - dem Vorgänger des heutigen Internet, nennen. Er ermöglichte es die Programmierfähigkeiten von vielen Studierenden gleichzeitig zu überprüfen und sie außerdem bei ihrer Verbesserung zu unterstützen (Hollingsworth, 1960). Für damalige Verhältnisse kann dieses Programm bereits als sehr fortschrittlich angesehen werden, da es neben einer summativen Bewertungsmöglichkeit der Programmierleistungen den Studenten gleichzeitig gute und vor allem schnelle Feedbackmöglichkeiten bot.

Im Laufe der letzten 15 bis 20 Jahre, erfolgte durch verschiedene technische Innovationen auch immer eine Weiterentwicklung der E-Assessment Systeme, die durch den Einsatz neuer Technologien stetig in ihrer Komplexität, Benutzerfreundlichkeit und letztendlich auch in ihrer didaktischen Qualität verbessert werden konnten. Dies führte zu einer großen Popularitätssteigerung von E-Assessments und damit letztendlich auch zu einer wachsenden Akzeptanz. Es entstanden daraufhin erste Klassifizierungen von E-Assessment Systemen anhand ihrer technischen Grundlagen, von denen auf die wichtigsten nachfolgend kurz eingegangen werden soll.

### 2.3.2 Klassifikationen von E-Assessments

Ein erste Einteilung von E-Assessments stammt von BULL & MCKENNA aus dem Jahr 2004 und unterscheidet diese in:

- Stand-alone-,
- Closed-Network- und
- Internet Systeme,

wobei auch Kombinationen der einzelnen Merkmal möglich sind (Bull & McKenna, 2004). Das Unterscheidungskriterium ist demnach das genutzte Kommunikationsmedium der Systeme.

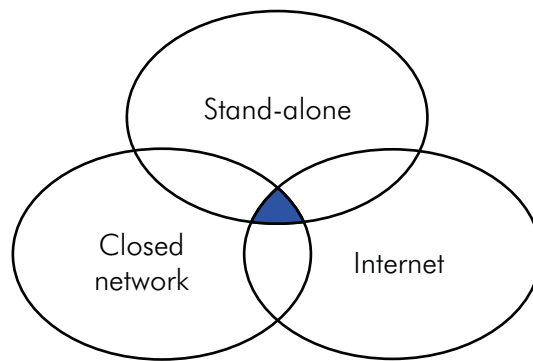


Abb. 1: Klassifikation von E-Assessments nach BULL & MCKENNA (2004)

„Stand-alone“ Systeme verfügen über keine direkte Möglichkeit der Kommunikation mit anderen Systemen oder einem Server. Die Assessment-Software und die Benutzereingaben werden direkt auf dem System gespeichert und gegebenenfalls auch dort ausgewertet. Alle Reaktionen des Systems auf Benutzereingaben sind intern vorgegeben und können nicht von außen, beispielsweise durch die Interaktion mit einem Server, beeinflusst werden.

„Closed-Network“ Systeme zeichnen sich durch eine Interaktionsmöglichkeit der Assessment-Software mit anderen Komponenten im gleichen geschlossenen Netzwerk aus. In einem praxisnahen Kontext betrachtet, könnte das bedeuten, dass mehrere Assessment-Clients auf unterschiedlichen Computern in einem „local area network“ (LAN) mit einem Assessment-Server verbunden sind, der eine zentrale Bearbeitung bestimmter Administrations-Aufgaben übernimmt. Solche administrative Aufgaben können beispielsweise die Nutzerauthentifizierung, die Bereitstellung der Prüfungsaufgaben, die Entgegennahme der jeweiligen Prüfungsantworten und die Auswertung oder Speicherung dieser Informationen sein.

„Internet“ Systeme entsprechen in ihrem Grundaufbau den eben vorgestellten Closed-Network Systemen. Sie nutzen lediglich anstatt eines geschlossenen, ortsgebundenen Netzwerkes das Internet als Kommunikationsmedium. Dadurch sind diese Systeme sehr flexibel hinsichtlich ihrer räumlichen Gestaltung, aber auch sicherheitstechnisch am Anfälligsten und somit wesentlich komplexer.

Aufbauend auf diese Einteilung erweitern später CONOLE & WARBURTON diese drei E-Assessment Typen, die sie als „Computer-Based-Assessments“ (CBA) zusammenfassen um sogenannte „Computer-Assisted-Assessments“ (CAA). Dieser Kategorie ordnen Sie „OMR“ (Optical Mark Recognition) und „E-Portfolios“ zu (Conole & Warburton, 2005).

Eine eigenständige Klassifizierung von E-Assessments aus dem Jahr 2006 stammt von STEINBERG und gliedert diese in folgende vier Bereiche:

- „Computer-Based-Assessment“,
- „Computer-Assisted/Mediated-Assessment“,
- „Web-Based-Assessment“ und
- „Web-Assisted-Assessment“.

Dabei werden als „Assisted“ alle die Systeme bezeichnet, bei denen die Aus- und Bewertung der Assessment-Ergebnisse nicht allein durch den Computer vorgenommen wird. Systeme der anderen Kategorien müssen die Ermittlung der Assessment-Ergebnisse eigenständig vornehmen, bei „Web-Based“ sogar in „Echtzeit“ (Steinberg, 2006). Hier liegt also das Unterscheidungskriterium nicht mehr im Kommunikationsmedium, sondern in der Verarbeitung und Auswertung der Assessment-Ergebnisse.

Obwohl die beschriebenen Klassifizierungen erst einige Jahre alt sind, sind nach heutigen Gesichtspunkten betrachtet, viele der vorgestellten Szenarien nur noch wenig praxisrelevant. Die rasante technische Weiterentwicklung führt zu einem Angleichen der technischen Möglichkeiten der einzelnen Systeme und bewirkt beispielsweise, dass heute prinzipiell jedes zeitgemäße E-Assessment-System als internetbasiert angesehen werden kann, was sich, dem allgemeinen Trend folgend, zukünftig auch nicht ändern wird. Auch die automatisierte Auswertbarkeit wird heute weniger, wenn auch immer noch spürbar, durch technische Grenzen eingeschränkt, als durch didaktische Überlegungen. Generell muss der Fokus von einem zu technologisch orientierten hin zu einem pädagogisch geprägten Assessment-Verständnis gelenkt werden, um die Qualität zukünftiger Systeme weiter zu verbessern.



### 2.3.3 Sicherheits- und rechtliche Aspekte

Auch wenn für die Entwicklung künftiger E-Assessments pädagogische Ziele und Anforderungen im Vordergrund stehen werden, ist die Betrachtung von Sicherheitsaspekten ein wichtiger Punkt und formales Kriterium für die Einsetzbarkeit eines jeden Systems. Ohne die Einhaltung bestimmter Restriktionen wird ein E-Assessment keine rechtssichere Alternative zu klassischen „Paper-and-Pencil-Prüfungen“ darstellen können und somit, speziell im Hochschulbereich, nur von geringer praktischer Relevanz sein.

Die Anforderungen lassen sich in folgende vier Hauptziele gliedern:

- eindeutige Zuordnung der Assessmentleistung zu einer Person,
- Verhinderung unberechtigter Datenzugriffe und -manipulation,
- Ausfallsicherheit der Assessment-Komponenten sowie
- dauerhafte und unveränderliche Speicherung jeder einzelnen Assessmentleistung (gegebenenfalls in Form eines Schriftstückes) (Reepmeyer, 2008)

Diese Eckpunkte müssen natürlich auch in klassischen Prüfungsverfahren eingehalten werden und sind keinesfalls eine Neuerfindung elektronischer Prüfungsszenarien. Lediglich in der Methodik der Umsetzung sind große Unterschiede vorhanden, welche nachfolgend tabellarisch gegenüber gestellt werden sollen.

Anforderungen	Umsetzung in klassischen Prüfungen	mögliche Umsetzung in E-Assessments
eindeutige Zuordnung der Assessmentleistung zu einer Person	<ul style="list-style-type: none"> <li>• Authentifizierung mittels Ausweis</li> <li>• Beaufsichtigung der Leistungserbringung</li> <li>• Verhinderung von Kommunikation unter den Prüfungsteilnehmern</li> <li>• Bestätigung der eigenen Prüfungsleistung mittels Unterschrift</li> <li>• Zufallsfragen, Zufallsreihenfolge, verschiedene Gruppen</li> </ul>	<ul style="list-style-type: none"> <li>• Authentifizierung mittels Benutzerkennwort und Assessment-Kennwort</li> <li>• Beaufsichtigung der Leistungserbringung</li> <li>• Nutzung spezieller Software (SaveExamBrowser, etc.)</li> <li>• Bestätigung der eigenen Prüfungsleistung mittels elektronischer Signatur oder Unterschrift</li> </ul>
Verhinderung unberechtigter Datenzugriffe und -manipulation	<ul style="list-style-type: none"> <li>• Geheimhaltung der Prüfungsaufgaben bis zum Prüfungsbeginn</li> <li>• Versiegelung der Prüfungen nach Abgabe</li> </ul>	<ul style="list-style-type: none"> <li>• Verschlüsselung der Assessment-Aufgaben und der Datenübertragung zwischen Client und Server</li> <li>• Ausdruck und Unterschrift der eigenen Prüfungsleistung nach Abgabe</li> </ul>
Ausfallsicherheit der Assessment-Komponenten	<ul style="list-style-type: none"> <li>• Ersatzstift, Ersatzblätter (einfach realisierbar, liegt primär in der Pflicht der Prüfungsteilnehmer)</li> </ul>	<ul style="list-style-type: none"> <li>• Verwendung redundanter Systeme</li> <li>• kontinuierliche Speicherung der Benutzereingaben auf dem Server</li> <li>• Bereithaltung von Ersatz-Hardware</li> <li>• Administrierbarkeit des Systems während des Assessment</li> </ul>
dauerhafte und unveränderliche Speicherung jeder einzelnen Assessmentleistung	<ul style="list-style-type: none"> <li>• Versiegelung der Prüfung nach Abgabe</li> <li>• sichere Lagerung im Safe</li> </ul>	<ul style="list-style-type: none"> <li>• redundante, dezentrale, verschlüsselte Speicherung der Daten im Read-Only Modus</li> <li>• sichere Lagerung der ausgedruckten Assessment-Leistungen im Safe</li> </ul>

Abb. 2: Sicherheitstechnische Anforderungen an E-Assessments im Vergleich mit klassischen Prüfungen

Es wird deutlich, dass viele einfache und vertraute organisatorische Abläufe im elektronischen Prüfungskontext plötzlich komplexer und aufwendiger werden. Dem gegenüber kann natürlich auch an anderen Stellen wieder Aufwand eingespart werden, sodass organisatorisch betrachtet eigentlich weniger eine Mehrbelastung seitens der Lehrenden auftritt, sondern vielmehr eine Verschiebung der Aufgabenfelder und Neugewichtung der Prioritäten.

Im Folgenden wird noch einmal detaillierter auf die vier Hauptkriterien einer technisch und rechtlich sicheren Prüfungsleistung eingegangen.

### **Eindeutige Zuordnung der Assessmentleistung zu einer Person**

Das die Gewährleistung der eindeutigen Zuordnung einer Prüfungsleistung zu einer bestimmten Person ein entscheidendes Kriterium für die Aussagekraft einer Prüfung an sich darstellt ist offensichtlich. Durch den Einsatz von E-Assessments ergeben sich allerdings vielfältige Angriffsmöglichkeiten dieser Eindeutigkeit, die in einem klassischen Prüfungsszenario nicht gegeben sind. Physisch anwesende und authentifizierte Personen in einem Prüfungsraum können nicht einfach durch andere ersetzt oder manipuliert werden. Im elektronischen Umfeld ist dies aber durchaus möglich.

Die Gewährleistung der eindeutigen Authentifizierung eines E-Assessment-Teilnehmers hängt stark von den technischen Gegebenheiten des Assessments ab. Findet dieses in einem speziellen Prüfungsraum („in-classroom“-Lösung) unter Beaufsichtigung statt, können mittels klassischen Ausweiskontrollen und Passwörtern sowie verschlüsselten Verbindungen auf digitaler Ebene weitgehend gleiche Sicherheitsbedingungen wie in Paper-and-Pencil-Prüfungen geschaffen werden. Ist das E-Assessment hingegen nicht an einen festen Raum gebunden („at-home“-Lösung) sondern beispielsweise auch von zuhause durchführbar ist praktisch keine Kontrolle der Authentizität einer Person mehr möglich.

### **Verhinderung unberechtigter Datenzugriffe und -Manipulation**

In einer klassischen Prüfungssituation finden alle relevanten Arbeitsschritte der Leistungserbringung des Prüfungsteilnehmers in seinem unmittelbaren physischen Umfeld statt. Das Prüfungsdokument wird ihm aus einer vertrauenswürdigen Quelle überreicht, anschließend bearbeitet und zurückgegeben. Eine vom Prüfungsteilnehmer ungewollte Manipulation seiner „Daten“ ist weitgehend ausgeschlossen und auch eine gewollte Manipulation kann durch ausreichende Beaufsichtigung während der Prüfung nahezu vermieden werden.

In den meisten E-Assessment-Szenarien wird hingegen das „Prüfungsdokument“, also sowohl der Aufgaben- als auch der Lösungsteil, mehrere Male während des Assessments den Prüfungsraum verlassen und wieder „betreten“. (Dahinden & Hinterberger, 2010) Das liegt in der Client/Server Struktur der meisten zeitgemäßen Assessment-Systeme und dem damit verbundenen kontinuierlichen Datenaustausch begründet. Dies ist aber keineswegs als Nachteil zu verstehen, denn nur so ist eine effektive Wartung der Hard- und Software sowie deren Ausfallsicherheit überhaupt möglich. Entscheidend ist, dass die Kommunikation zwischen allen beteiligten Instanzen vor, während und nach des Assessments ausreichend geschützt und verschlüsselt stattfindet. Eine Manipulation ist dann nicht wahrscheinlicher als in einer klassischen Prüfungssituation.

### **Ausfallsicherheit der Assessment-Komponenten**

Die eben erwähnte Ausfallsicherheit ist ein entscheidender Punkt für die Qualität eines E-Assessment Systems. Das gilt sowohl hardware- als auch softwareseitig und stellt nach wie vor den größten empfundenen Unsicherheitsfaktor sowohl auf Seite der Prüfenden als auch der Geprüften im Bezug auf den Einsatz von E-Assessments dar. Diese Skepsis ist allerdings unbegründet, denn bezogen auf seine Ausfallsicherheit ist die Durchführung eines E-Assessment mittlerweile als Standardsituation zu betrachten. Viele technische Systeme in kritischen Einsatzbereichen erfordern ein Maximum an Zuverlässigkeit und entsprechend gefundenen Lösungen können einfach übernommen werden. So lassen sich Serversysteme durch redundante Komponenten und Notstromaggregate sehr ausfallsicher betreiben und Assessment-Clients einfach austauschen, wenn während des Assessment ein Problem auftritt. Durch die kontinuierliche Kommunikation zwischen Client und Server, kann jeder Assessment-Teilnehmer problemlos sein Eingabegerät wechseln, sollten technische Probleme auftreten, ohne dass Eingaben verlorengehen oder zeitliche Nachteile entstehen.

Trotzdem findet an dieser Stelle ein entscheidender Wechsel im Verantwortungsbewusstsein seitens der Prüfer und Geprüften statt. Lag bei klassischen Paper-and-Pencil Prüfungen die Sorgfaltspflicht für die Funktionsfähigkeit der Arbeitsmittel (Papier, Stift, etc.) zumindest primär, eindeutig beim Geprüften, ist jetzt die prüfende Instanz für die Gewährleistung problemlos funktionierender Eingabegeräte verantwortlich. Auch ist an dieser Stelle der finanzielle Aufwand für das Bereithalten redundanter Systeme im E-Assessment erheblich gestiegen. Konnte man Ersatzstifte, Zusatz-Papier, etc. nahezu unbegrenzt, finanziell vernachlässigbar vorhalten sowie problemlos austauschen erfordert dies jetzt vor allem in finanzieller Hinsicht einen höheren Aufwand.

## **Dauerhafte und unveränderliche Speicherung der Assessmentleistungen**

Für die Durchführung eines rechtssicheren E-Assessment nimmt dieser Punkt eine entscheidende Rolle ein, denn kommt es zu einem Rekurs, ist die Bildungseinrichtung in der Beweispflicht (Dahinden & Hinterberger, 2010). Das setzt natürlich eine ausreichend lange und sichere Aufbewahrung der Prüfungsunterlagen voraus.

Zum Einsatz von E-Assessments als Ersatz schriftlicher Prüfungen liegt folgende Entscheidung des Verwaltungsgerichts Hannover aus dem Jahr 2008 vor:

„Eine Prüfung, bei der die auf einem Bildschirm angezeigten Prüfungsfragen ausschließlich durch das Markieren der vom Anwendungsprogramm vorgegebenen Antwortfelder mit einem Eingabegerät beantwortet werden und die Fragen und Antworten ausschließlich als digitale Informationen auf einem Speichermedium verbleiben, stellt keine schriftliche Prüfung dar. Eine schriftliche Prüfung setzt voraus, dass das Prüfungsergebnis von dem Prüfling in Schriftform verfasst wird und als in dieser Form verkörperte Sprache auf einem Dokument (Schriftstück) für jedermann lesbar bleibt.“ (VG Hannover, 2008)

Zwar bezieht sich diese Entscheidung im konkreten Fall auf eine Multiple-Choice Prüfung, doch hat sich die daraus resultierende E-Assessment-Praxis weitgehend durchgesetzt. Die Prüfungsunterlagen werden dabei direkt nach Beendigung des Assessments ausgedruckt und vom Assessment-Teilnehmer noch einmal überprüft und unterzeichnet. Die Unterlagen liegen dann in Schriftform vor und genügen somit den rechtlichen Anforderungen und können sicher „gespeichert“ werden.

## 2.4 Pädagogische Grundlagen

Da der E-Assessment Begriff in dieser Arbeit immer auch im Kontext seines pädagogischen Anwendungsbereiches und somit als Bestandteil des E-Learnings betrachtet wird, ist es notwendig auf einige grundlegende Lerntheorien einzugehen. E-Learning selbst stellt dabei keine neue Lerntheorie dar, sondern baut vielmehr auf bekannte pädagogische Konzepte auf und setzt diese auf die Nutzung neuer Medien um. Die nachfolgenden Erläuterungen sollen auch pädagogischen Laien die Grundprinzipien der wichtigsten Lerntheorien nahebringen.

### 2.4.1 Behaviorismus

Der Behaviorismus stellt eine der ersten wissenschaftlichen Lerntheorien dar und entstand Anfang des 20. Jahrhunderts. Ihm liegt ein einfaches Basismodell zu Grunde, dass Lernen als eine extern steuerbare Verhaltensänderung beschreibt, die als Reaktion des Lerner auf einen äußeren Reiz erfolgt (Baumgart, 2007). Betrachtet werden dabei nur der Reiz (Input) und die ausgelöste Reaktion (Output), die interne Verarbeitung wird ignoriert – das Gehirn als „Black-Box“ betrachtet (Frohberg, 2008).

Eine wesentliche Aussage dieser Theorie besteht darin, dass eine Verhaltensänderung bewirkt werden kann, indem der Lerner wiederholt dem gleichen Reiz ausgesetzt und seine Reaktion darauf entweder als „richtig“ belohnt oder „Fehlverhalten“ bestraft wird. Dieser Prozess geht maßgeblich auf den russischen Psychologen Iwan P. Pavlov zurück und wird als „Konditionierung“ bezeichnet (Pavlov, 1928).

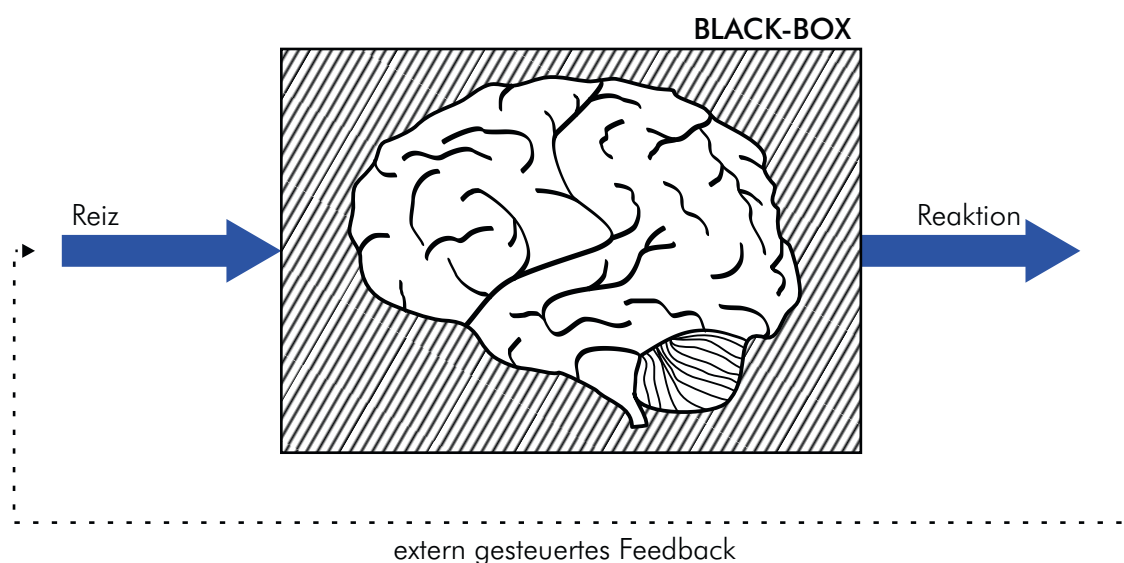


Abb. 3: Reiz-Reaktion-Schema mit Gehirn als „Black-Box“

Ein Großteil des allgemeinen Lernverständnisses geht auf diese Grundlagen zurück und auch heute noch beruhen viele Lernaktivitäten auf behavioristischen Prinzipien. Ein sehr simples aber dennoch treffendes Beispiel dafür ist das Auswendiglernen von Vokabeln zum Erlernen einer Fremdsprache (Frohberg, 2008). Die Aussprache und Schreibweise einzelner Wörter wird immer wieder geübt und abgefragt, bis sie schließlich in den Wortschatz des Lernenden übergehen und er sie fortan benutzen kann. Generell setzt diese Praxis aber voraus, dass ein fixes, allgemeingültiges Wissen existiert und jede Reaktion als eindeutig richtig oder falsch gewertet werden kann (Jank & Meyer, 2002). Der Lehrer nimmt nach diesem Verständnis eine zentrale Rolle ein und bestimmt als „Experte“ den Lernprozess, die Lernziele, überprüft den Lernerfolg und belohnt (gute Note) oder bestraft (schlechte Note) entsprechend. Da der Lernende in diesem Prozess eine eher passive Rolle einnimmt, spricht man von rezeptivem Lernen (Frohberg, 2008).

### 2.4.2 Kognitivismus

In der Mitte des 20. Jahrhunderts begann man die lerntheoretischen Thesen des Behaviorismus genauer zu hinterfragen und sich verstärkt für die Prozesse innerhalb der „Black Box“, dem Gehirn, zu interessieren. Zwar war es nach wie vor nicht möglich, die genauen Vorgänge innerhalb des Gehirns beim Lernen zu beobachten, doch ging man davon aus, dass diese individuell verschieden sind. Diese internen, psycho-mentalenen Prozesse bezeichnete man als Kognition, den sich daraus ableitenden lerntheoretischen Ansatz als Kognitivismus (Flindt, 2005).

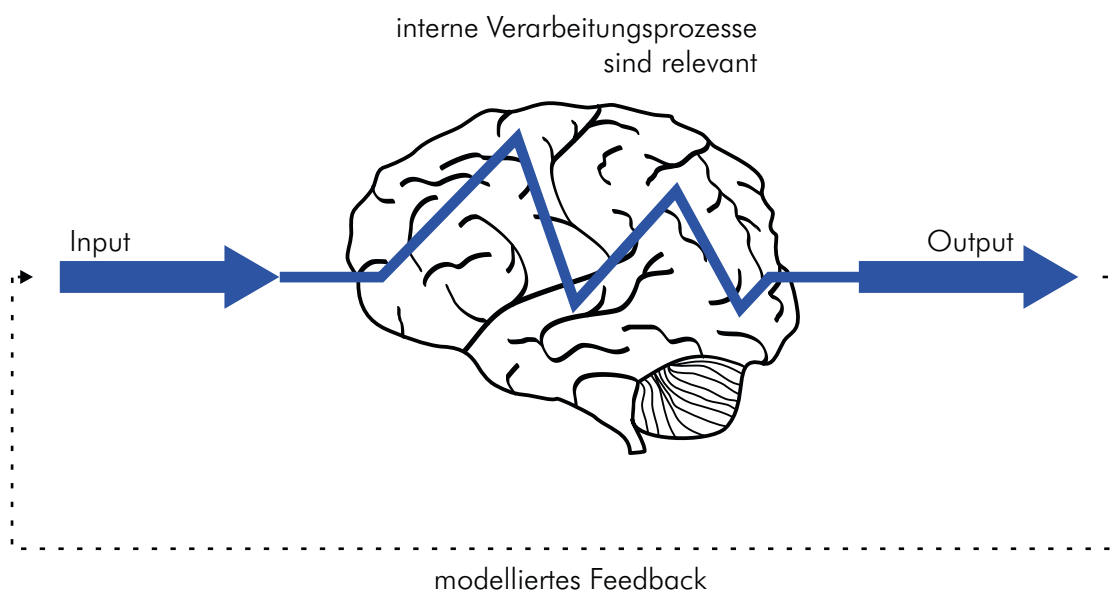


Abb. 4: Abbildung des Lernprozesses im Kognitivismus nach Baumgartner (1997)

Ähnlich wie bereits im Behaviorismus stellt der Kognitivismus allerdings keine einheitliche Strömung dar, sondern besteht vielmehr aus unterschiedlichen Theorien und Ansätzen. Eine entscheidende Gemeinsamkeit ist aber die Abkehr davon, den Lerner als weitgehend passives Element im Lernprozess zu betrachten, das durch äußere Reize steuerbar ist. Man beginnt stattdessen die individuellen Denkprozesse des Lerners in den Mittelpunkt zu stellen und Lernen als Informationsverarbeitung zu verstehen, die abhängig von eigenen Erfahrungen, vorhandenem Vorwissen und bestimmten kognitiven Ressourcen ist (Kerres, 2001). Der eigentliche Lernprozess läuft demnach bei jedem Individuum anders ab und zielt auf eine eigenständige Problemlösung. Es wird akzeptiert, dass es für ein bestimmtes Problem mehrere „richtige“ Lösungen geben kann. Der Lehrer wird dabei nicht mehr als „Experte“ sondern als „Tutor“ betrachtet und nimmt eine beobachtende und helfende Rolle im Lernprozess ein (Baumgartner & Payr, 1997).

### 2.4.3 Konstruktivismus

Als letzte der drei Hauptlerntheorien ist abschließend der Konstruktivismus zu nennen, welcher Ende des 20. Jahrhunderts populär wurde. Seine Ansätze beruhen vor allem auf neuen Erkenntnissen der Neurowissenschaften und weiterentwickelten kognitivistischen Theorien, speziell von Piaget. Deshalb handelt es sich beim Konstruktivismus genau genommen nicht um eine reine Lern-, sondern eine Erkenntnistheorie (Flindt, 2005), deren zentrales Element der „Wissensbegriff“ darstellt. Da auch der Konstruktivismus, ähnlich zu Behaviorismus und Kognitivismus, viele einzelne Theorien und Ansätze in sich vereint, welche in ihren Ansichten teils weit auseinandergehen, sollen im Folgenden nur allgemeine, grundlegende Aspekte betrachtet werden.

Generell beschreiben die Konstruktivisten „Wissen“ nicht als „Abbild der externen Realität, sondern als eine Funktion des Erkenntnisprozesses“ (Schulmeister, 2007), die von jedem Individuum eigenständig erzeugt werden muss. Für Vertreter des „Radikalen Konstruktivismus“ bedeutet dies, dass alles Wissen subjektiv ist und keine objektive Wahrheit existieren kann (Glaserfeld, 2007). Auf den Lernprozess angewendet, stellt sich dann allerdings die Frage, ob Unterricht unter diesen Gesichtspunkten überhaupt sinnvoll ist, denn was soll gelehrt werden, wenn alles Wissen subjektiv ist? Aus diesem Grund hat sich für die Anwendung des Konstruktivismus im Lernprozess eine gemäßigte Sichtweise durchgesetzt.

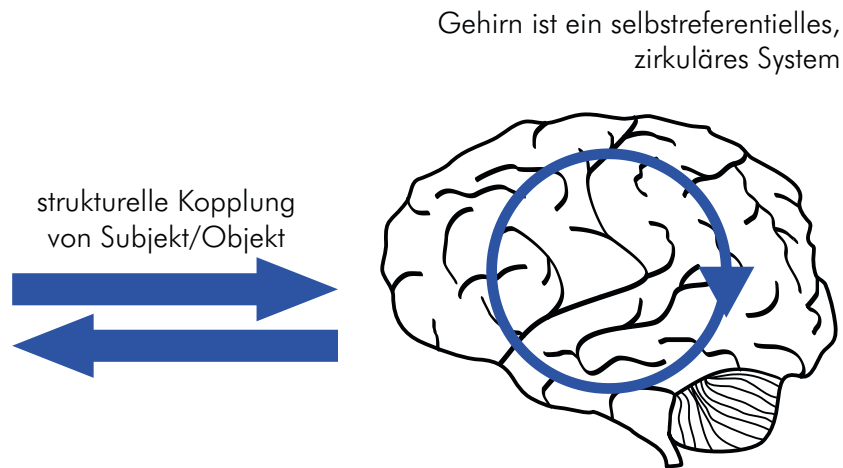


Abb. 5: Schematisches Lernmodell des Konstruktivismus

Entscheidend ist dabei die Akzeptanz des Lerners als aktives und zentrales Element des Lernprozesses sowie die Erkenntnis, dass Wissen von jedem Lerner selbst aus verschiedenen angebotenen Informationen konstruiert werden muss und sich individuell unterscheidet. „Lernen ist demnach nicht länger eine Folge des Lehrens, sondern eine eigenständige Konstruktionsleistung des Lernenden“ (Frohberg, 2008 nach Jank und Meyer, 2002), die am effektivsten in praxisnahen und authentischen Situationen erfolgt. Auch wird die Zusammenarbeit der Lernenden untereinander als positiv für den Lernerfolg eingeschätzt (Kerres, 2001).

Der Lehrer nimmt im Konstruktivismus vor allem eine lernbegleitende, beratende Rolle ein und steht dem Lernenden als „Coach“ zur Seite. Der Lernende wiederum sollte die Möglichkeit haben den Lernprozess individuell beeinflussen zu können um bestmögliche Lernerfolge zu erzielen. Ein reiner Frontalunterricht wird deshalb aus konstruktivistischer Sicht abgelehnt und sollte stattdessen durch eine abwechslungsreiche, anspruchsvolle Lernumgebung ersetzt werden, die den Lernenden aktiv ins Unterrichtsgeschehen einbindet (Flindt, 2005).



### 3 E-Assessment als Nachfolger klassischer Prüfungen - Motivation für den Einsatz

Betrachtet man E-Assessments als eigenständige, alternative Prüfungsmethoden, stellt sich zuerst einmal die Frage nach der Motivation für deren Einsatz, speziell auch in der Hochschullehre. Denn auf bekannte und lang erprobte Vorgehensweisen zu verzichten, bedarf natürlich gewichtiger Gründe. Diese können ganz allgemein entweder in erkannten Schwächen oder Problemen des „alten“ Systems oder aber in überzeugenden Vorteilen und Innovationen des „neuen“ Systems liegen. In der Praxis wird meist eine Kombination aus beiden Varianten vorliegen, die schließlich zu einem schrittweisen Ablösen des „Alten“ durch das „Neue“ führt.

Im Folgenden wird die Motivation für den Einsatz von E-Assessments als Alternative zu klassischen Prüfungsverfahren untersucht, in dem Schwächen und Probleme sowie Vorteile beider Systeme kritisch gegenübergestellt und vor allem aus didaktischer Sicht bewertet werden.

Dazu ist es notwendig zuerst einige Grundbegriffe der Qualität von Leistungsnachweisen im Allgemeinen zu betrachten. In der Fachliteratur wird häufig von drei Hauptanforderungen an qualitative Leistungsnachweise gesprochen: Validität, Reliabilität und Objektivität. Im Rahmen dieser Arbeit soll eine Wichtung der Begriffe vorgenommen werden, welche die Validität als Resultat der Reliabilität und Objektivität ansieht.

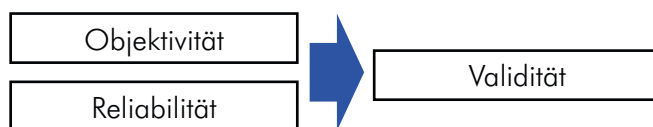


Abb. 6: Abhängigkeit von Validität, Reliabilität und Objektivität

**Validität:** Die Validität eines Leistungsnachweises beschreibt in welchem Maße tatsächlich die Kompetenzen und Lernziele überprüft werden, die auch überprüft werden sollen. Die Validität ist somit allgemein ein Maß für die Qualität eines Leistungsnachweises.

**Reliabilität:** Unter Reliabilität oder „Zuverlässigkeit“ versteht man die Güte der Wiederholbarkeit eines Leistungsnachweises. Nur wenn diese in einem ausreichend hohen Maß gewährleistet ist, sind Prüfungsergebnisse vergleichbar. Diese Anforderung richtet sich vor allem an die Prüfungsaufgaben.

**Objektivität:** Bei der Durchführung und Auswertung von Leistungsnachweisen ist die Objektivität von großer Bedeutung, denn sie beschreibt wie unabhängig eine Bewertung vorgenommen werden kann. Bei maximaler Objektivität erzielt die Auswertung derselben Leistung in mehreren Instanzen stets dieselben Ergebnisse. (Roloff, 2003)

### 3.1 Aktuelle Prüfungssituation an deutschen Hochschulen

Der Einsatz von elektronischen Hilfsmitteln zur Durchführung von computergestützten Assessments ist im Prinzip nichts Neues, denn im Gegensatz zum Menschen sind Computersysteme wesentlich zuverlässiger in der Konstanz ihrer Leistungen (vgl. Objektivität), vorausgesetzt sie wurden entsprechend solide programmiert (vgl. Reliabilität) und eignen sich damit gut zur Überprüfung bestimmter Fähigkeiten ihrer Nutzer. So gehörten einfache „Testprogramme“ mit zu den ersten entwickelten Computerprogrammen überhaupt. Auch sind Computerspiele prinzipiell nichts Anderes als Testprogramme, welche einerseits bestimmte „Skills“ erfordern, diese andererseits aber auch, im Laufe ihrer Nutzung, erzeugen. Trotz dieser eigentlich elementaren Verbindung von Computern und ihrer Fähigkeit uns zu testen spielten E-Assessments bis vor kurzem nur eine unbedeutende Rolle in der Lehre.

Interessant wurden E-Assessments jüngst im hochschulischen Bereich im Zuge der Bologna-Reformen und der mit ihnen gestiegenen Anzahl an Leistungsnachweisen, die sich mit klassischen Prüfungsverfahren nur schwer realisieren lassen. Hinzu kommen steigende Studierendenzahlen, bei einer weitestgehend gleichbleibenden Anzahl an Lehrkräften und eine verstärkte Ausrichtung der Lernziele auf die Vermittlung von Kompetenzen. (Ruedel, 2010) So suchte man nach Alternativen um diese Mehrbelastungen bewältigen zu können und die Qualität der einzelnen Lehrveranstaltungen zu sichern. Dieser Schritt ist sowohl positiv als auch negativ zu bewerten, denn er stellt zwar einerseits das Medium E-Assessment generell als Alternative zu klassischen Prüfungsverfahren vor, bedient sich andererseits dabei viel zu oft ausschließlich an den rationalisierenden Möglichkeiten dieser Prüfungsform, ungeachtet ihrer didaktischen Qualitäten (Reinmann, 2007).

#### 3.1.1 Prüfungsalltag

Spricht man von Leistungsnachweisen im Bereich der Hochschullehre assoziiert man dies meist immer noch automatisch mit dem Begriff der „Prüfung“ und meint damit auch eine ganz bestimmte Prüfungspraxis, welche über viele Jahrzehnte nahezu unverändert Anwendung fand und auch

heute meist noch findet. Gemeint ist damit die Durchführung einer, klassischerweise 90 - 135 minütigen schriftlichen Prüfung als Abschluss eines Kurses/Moduls am Semesterende. Außerdem typisch für Hochschulprüfungen ist das geballte Stattfinden mehrerer dieser Leistungsnachweise in einem eng begrenzten Prüfungszeitraum von zwei bis vier Wochen mit mehreren Prüfungen wöchentlich. Selbstverständlich finden auch andere Formen von Leistungsnachweisen Anwendung im Hochschulbereich, so zum Beispiel mündliche Prüfungen, Hausarbeiten, Referate, Praktika oder auch Übungsaufgaben, doch der Großteil folgt nach wie vor dem beschriebenen Muster (Ruedel, 2010).

Bereits aus dieser Prüfungspraxis, ungeachtet der angewandten jeweiligen Prüfungsmethode, lassen sich nachteilige, beziehungsweise zumindest unvorteilhafte Aspekte ableiten.

### **Einmaligkeit**

Findet pro Kurs/Modul nur eine einzige Abschlussprüfung als Leistungsnachweis statt, ergeben sich daraus zwei Nachteile. Einerseits wird dadurch der Effekt gefördert, dass ausschließlich für die Prüfung gelernt wird, was den eigentlichen konzentrierten Lernprozess auf wenige Wochen unmittelbar vor der Prüfung beschränkt und somit die vorherige Integrationsbereitschaft ins Unterrichtsgeschehen verringert. Andererseits entsteht natürlich vor und während der Prüfung eine enorme Stressbelastung der Lernenden, was sich deutlich und stark verfälschend auf die Prüfungsergebnisse auswirken kann.

### **Dauer**

Die Zeitdauer einer klassischen Prüfung ist meist exakt festgelegt und erlaubt somit keine Anpassung an die individuellen Zeitbedürfnisse der Geprüften. Diese Vorgehensweise erzeugt zwar auf den ersten Blick einen objektiven Prüfungsrahmen, unterstellt aber damit auch, dass jeder Lerner sein Wissen und seine Kompetenzen gleich schnell wiedergeben oder anwenden kann. Praktisch ist dies allerdings nicht der Fall, was wiederum zu einer erheblichen Stressbelastung der Geprüften führt. Anhand dieser Sichtweise wird deutlich, wie behavioristisch nach wie vor die Grundlagen unseres Schul- und Lehrsystems geprägt sind.

### **Position der Prüfung im Lernprozess**

Werden Lernleistungen erst am Kursende erfasst liegt dieser Zeitraum im Hochschulbereich logischer Weise unmittelbar vor den Semesterferien. Für den Lernenden ist damit der Kurs nach dem „Ende“ der Prüfung und bereits vor Erhalt der Prüfungsergebnisse weitestgehend abgeschlossen. Die Wirkung des Leistungsnachweises auf den Lernprozess wird dadurch minimiert und somit viel Potential verschenkt.

## Prüfungszeitraum

Nachteilig für die Qualität einer Prüfung kann auch der zeitliche Abstand zu anderen Leistungsnachweisen sein, insbesondere wenn vornehmlich die Anwendung von Gelerntem (Problemlösung) im Vordergrund steht (Problem: „Black-Out“). Besonders kritisch ist dabei im Bereich der Hochschullehre zu bewerten, dass nahezu alle Leistungsnachweise in einem sehr engen Zeitraum am Semesterende stattfinden, während die Restzeit des Semesters im klassischen Sinne weitgehend prüfungsfrei ist. Ein Resultat dieser Prüfungspraxis ist, dass Lernende oft die Auswahl ihrer Lehrveranstaltungen primär nach Art der Prüfungsleistung und nur sekundär anhand ihrer persönlichen Interessen, Stärken oder Kompetenzen auswählen.

## Korrekturzeiten

Die Korrekturzeiten schriftlicher Hochschulprüfungen belaufen sich meist auf mehrere Wochen, je nach Umfang und Auswertungsaufwand des Leistungsnachweises sowie der Teilnehmerzahl. Je länger dieser Zeitraum wird, desto geringer wird die Wirkung des Leistungsnachweises auf den eigentlichen Lernprozess (Ridgway et al., 2006). Bei einer Dauer von mehreren Wochen und vielen absolvierten Prüfungen sind oft bereits die genauen Inhalte und eigenen Lösungen des Leistungsnachweises längst vergessen. Die Bekanntgabe der Note hat dann nur noch einen „Rating-Charakter“ und so wird auch die Möglichkeit der Prüfungseinsicht meist nur noch von Prüflingen wahrgenommen, welche unzureichend abgeschlossen, sprich die Prüfung nicht bestanden haben.

Viele dieser Punkte verdeutlichen, wie stark behavioristisch auch heute noch zahlreiche Leistungsnachweise, gerade im Hochschulbereich, geprägt sind. Dies soll im Umkehrschluss nicht bedeuten, dass alle behavioristischen Ansätze generell schlecht oder unzeitgemäß sind, nur sollten auch neue Erkenntnisse praktisch im Prüfungsalltag umgesetzt und angewendet werden. Es zeigt sich auch, dass derzeitige Prüfungen oft nicht kompetenzorientiert genug sind, sondern nach wie vor primär lexikalisches Wissen abfragen. Außerdem werden neben den eigentlichen gewollten Prüfungszielen zahlreiche Zusatzkompetenzen geprüft, so zum Beispiel die Belastbarkeit in Stresssituationen oder das Zeitmanagement der Geprüften. Diese Kompetenzen sind zweifelsohne generell bedeutsam aber nicht in jeder Situation wirklich praxisrelevant. Die Verhältnismäßigkeit zwischen geprüften Kompetenzen und Zusatzkompetenzen ist somit ein erstes entscheidendes Kriterium für die Validität von Leistungsnachweisen, welches sich im aktuellen Prüfungseinsatz noch stark verbessern lässt.

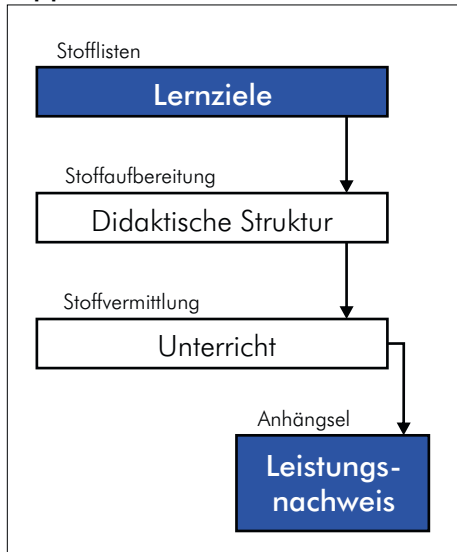
Des Weiteren kann, insbesondere bei Massenprüfungen mit mehreren Korrektoren, die Objektivität dieser Prüfungspraxis in Frage gestellt werden, da Musterlösungen oft nur unzureichend formuliert und spezifiziert sind und somit für ein und dieselbe Prüfungsleistung unterschiedliche Ergebnisse, abhängig von den jeweiligen Korrektoren entstehen können (Roloff, 2002). Selbst die Auswertbarkeit handschriftlicher Prüfungsunterlagen im Allgemeinen kann dabei zu sehr subjektiven Resultaten führen, denn was ein Korrektor noch „dechiffrieren“ kann, ist für andere möglicherweise unleserlich und somit falsch. So verstärkt sich neben der Gefahr inkorrekten Prüfungsergebnissen zu generieren meist auch noch eine zusätzliche Verlängerung der Korrekturzeiten.

Im Allgemeinen lässt sich die, bereits besprochene, mitunter sehr lange Korrekturzeit von Leistungsnachweisen als ein kritischer Punkt der aktuellen Prüfungspraxis im Hochschulbereich ansehen. Wie bereits eingangs beschrieben wurde, geht damit ein entscheidender Nutzen des Leistungsnachweises teilweise oder gänzlich verloren. Gemeint ist die Feedbackfunktion, die es dem Lernenden erst ermöglicht die geforderten Leistungen mit seinen erbrachten zu vergleichen und somit seine eigenen Fähigkeiten detaillierter einschätzen zu können. Nur so können konkrete Defizite aufgearbeitet und somit optimale Lernerfolge erzielt werden. Erfolgt dieses Feedback nicht, oder erst mit großen zeitlichen Differenzen zum eigentlichen Assessment ist dies nicht mehr gewährleistet (Merrill, 2002). Wie groß die zeitliche Differenz maximal sein darf um optimale Ergebnisse zu erzielen hängt wiederum von der Komplexität der geprüften Kompetenzen und den dafür erforderlichen kognitiven Fähigkeiten ab.

### 3.1.2 Prüfungsplanung

Leistungsnachweise spielen für den Lernenden in der Hochschullehre eine immens wichtige Rolle, denn über ihr Bestehen entscheiden sich nicht nur die Fortführbarkeit des Studiums, sondern damit verbunden auch spätere berufliche Perspektiven und Chancen. Man kann also behaupten, dass diese Leistungsnachweise den späteren Lebensweg des Lernenden maßgeblich beeinflussen. Umso entscheidender ist die Frage, welcher Stellenwert diesen Leistungsnachweisen seitens der Lehrenden beigemessen wird. Nach EUGSTER betrachten nach wie vor viele Lehrende im Hochschulbereich Prüfungen als „Curriculares Anhängsel“ ihrer Lehrveranstaltung und stellen deren Planung und Konzeption zu weit in der Hintergrund („Appendix-Modell“). In der Kursplanung werden oft primär nur die Lerninhalte und eine entsprechende didaktische Struktur für deren Vermittlung berücksichtigt und somit die Prüfungsplanung auf einen späteren Zeitpunkt verlegt. (Eugster, 2009/2010) Diese Praxis erschwert ein harmonisches Zusammenspiel zwischen „Lehre“ und „Assessment“ und stellt somit ein weiteres Kriterium für die fragliche Validität vieler derzeitiger Leistungsnachweise dar.

### Appendix-Modell



### Herz-Kreislauf-Modell

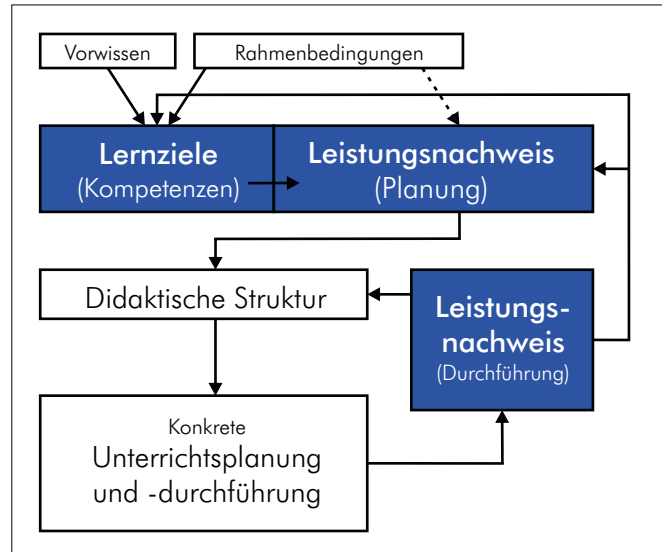


Abb. 7: Appendix- und Herz-Kreislauf-Modell nach EUGSTER (2009/2010)

EUGSTER empfiehlt die Planung der Leistungsnachweise bereits konkret in die Lernzielformulierung zu integrieren und beschreibt dieses Modell als „Herz-Kreislauf-Modell“, da Leistungsnachweise nicht mehr als abschließendes Anhängsel eines Curriculums verstanden werden, sondern aktiv in dieses integriert sind (Eugster, 2009/2010). Durch dieses Prinzip ist eine kontinuierliche Verbindung zwischen, einerseits den zu vermittelnden Kompetenzen, und andererseits ihrer Überprüfbarkeit gewährleistet.

### 3.1.3 Prüfungsformen

Nachdem im Kapitel 3.1.1 Prüfungsalltag bereits der Einsatz und Umgang mit Prüfungen, als hauptsächliche Form der Assessments im Hochschulbereich, näher betrachtet wurde, wird jetzt eine genauere Auseinandersetzung mit den jeweils eingesetzten Prüfungsformen vorgenommen. Anstelle des „Wann“ und „Wie“ steht jetzt das „Was“ und „Womit“ der aktuellen Prüfungspraxis im Mittelpunkt, denn viele heutige Prüfungsformen haben sich in ihrer Anwendungspraxis in den letzten Jahrzehnten nur wenig verändert. Im Gegensatz dazu steigen die Anforderungen an ein modernes Bildungssystem aber stetig an, was zu neuen und verbesserten Lehr- und Lernmethoden und einer generell facettenreicheren Lehrkultur geführt hat. Weshalb sich aber das Prüfungssystem nur verhältnismäßig langsam weiterentwickelt hat, lässt sich unter anderem auf die strikten rechtlichen Anforderungen zurückführen, welche seitens der Lehrmethoden mehr Freiräume bieten. In jedem Fall beschränkt sich auch heute noch der hauptsächlich Einsatz von Assessments im Hochschulbereich auf summative Prüfungen (Reinmann, 2007), die meist schriftlich, seltener mündlich und noch seltener in Form von Referaten oder Projektarbeiten stattfinden. Tatsächlich

scheint auch die Wahl der jeweiligen Prüfungsform von der Gewichtung der Prüfungsleistung im Gesamtkontext des Studiums abzuhängen, denn häufig werden die wertigsten Prüfungen in schriftlicher Form und weniger wertige Prüfungen in mündlicher oder alternativer Form durchgeführt. Dies müsste im Umkehrschluss logischer Weise bedeuten, dass schriftliche Prüfungen die höchste Validität aufweisen und andere Prüfungsformen qualitativ schlechter sind, andererseits wäre die aktuelle Einsatzpraxis, zumindest aus didaktischer Sicht, ungerechtfertigt.

Im Folgenden werden zunächst die am häufigsten eingesetzten Prüfungsformen im Hochschulbereich im Hinblick auf ihren Einsatz als summative Leistungsnachweise betrachtet. Dabei stehen deren Besonderheiten hinsichtlich ihrer Objektivität, Reliabilität sowie Validität im Vordergrund.

### **Schriftliche Prüfung**

Die schriftliche Prüfung stellt, wie bereits eingangs beschrieben, die gebräuchlichste Art des Hochschul-Assessments dar und kann demnach als Standard-Prüfungsform angesehen werden, was wiederum bedeutet, dass sie höchsten qualitativen Anforderungen gerecht werden sollte. Tatsächlich bietet diese Prüfungsform eine hohe Durchführungsobjektivität, das heißt, der Einfluss der prüfungsdurchführenden Person auf das Prüfungsergebnis ist gering. Auch bietet diese Prüfungsform eine hohe Interpretationsobjektivität, was besagt, dass die finale Notenvergabe aus den erreichten Punkten sehr objektiv erfolgt. Kritik kann allerdings seitens der Auswertungsobjektivität angebracht werden, denn gerade bei Multi-Korrektor-Szenarien, wie sie bei teilnehmerstarken Prüfungen oft zu finden sind, setzt eine objektive Bewertung (Punktevergabe) perfekte Musterlösungen und einen absolut kontinuierlichen Auswertungsprozess voraus. Dass dies nur in den seltensten Fällen erreicht werden kann ist absolut menschlich und nahezu unvermeidbar. So werden beispielweise schwer lesbare Handschriften oder Lösungen mit zahlreichen Grammatikfehlern oft schlechter bewertet, wogegen bei gut oder sympathisch empfundenen Personen Fehler eher übersehen werden. Auch spielen Kontrasteffekte - auf eine schwache Prüfungsleistung folgt eine bessere oder umgekehrt - sowie Ermüdungserscheinungen des Korrektors eine entscheidende Rolle und führen oft zu milderer oder härteren Bewertungen. (Roloff, 2003)

Hinsichtlich der Reliabilität lassen sich nur schwer allgemeingültige Aussagen treffen, denn diese hängt vor allem von der Qualität der Prüfungsfragen ab und ist somit von Prüfung zu Prüfung verschieden. Allerdings spielen auch die Rahmenbedingungen der Prüfungsdurchführung eine Rolle, welche sich bei schriftlichen Prüfungen gut vereinheitlichen lassen und somit, nach

klassischer Sichtweise, eine gute Ausgangsposition für eine hohe Reliabilität darstellen. Da die Objektivität ein notwendiges Kriterium für die Reliabilität und diese wiederum mitentscheidend für die Validität eines Assessment, einer Prüfung ist, ergeben sich aus den aktuellen Betrachtungen bereits erste Rückschlüsse auf die Validität schriftlicher Prüfungen. So kann man prinzipiell von guten Validitäts-Voraussetzungen sprechen. Abgesehen vom generellen Prüfungsparadoxon, welches besagt, dass sich jede Prüfung im Grunde selbst im Weg steht und damit eine optimale Leistungserbringung seitens des Geprüften verhindert, fehlt schriftlichen Prüfungen oft die notwendige praktische Nähe. So können viele schriftliche Prüfungsszenarien nicht die eigentlichen Handlungskompetenzen prüfen, sondern eigentlich nur das vorhandene Wissen, also Kenntnisse über diese Handlungskompetenzen testen. Die Gültigkeit oder Validität dieser Prüfungen unter aktuellen Anforderungen ist dann fraglich.

### **Mündliche Prüfung**

Eine weitere, relativ häufig Anwendung findende Prüfungsform im Bereich der Hochschullehre stellt die mündliche Prüfung dar. Im Gegensatz zur schriftlichen Prüfung stehen einerseits der Medienwechsel von Schrift zu Sprache und andererseits die Plastizität des Prüfungsablaufs, sowie der Prüfungsaufgaben an sich. Insgesamt kann die Prüfung wesentlich schülerzentrierter gestaltet werden, was aber gleichzeitig hohe Anforderungen an die Prüfer stellt, denn die Individualisierung führt schnell zu einem Objektivitäts- beziehungsweise Reliabilitätsverlust. Generell sind die Rahmenbedingungen dieser Prüfungsform weitaus flexibler als die der schriftlichen Prüfung und erfordern so vor allem seitens der Prüfer eine kontinuierliche Anpassung an verschiedenste Prüfungssituationen.

Was bedeutet diese Charakteristika für die Objektivität dieser Prüfungsform? Der entscheidendste Punkt ist sicherlich der Einsatz der menschlichen Komponente im gesamten Prüfungsprozess, also von der Präsentation der Aufgaben, über die Rezeption der Antworten, bis hin zur Auswertung der Prüfungsleistung und oft sogar der unmittelbaren Übermittlung des Prüfungsergebnisses an den Geprüften. Diese Faktoren reduzieren zwar erst einmal die Durchführungsobjektivität, trotzdem können versierte Prüfer diesen scheinbaren Qualitätsverlust durch gezielte Fragestellungen und genaue Beobachtungen des Antwortverhaltens des Geprüften wieder ausgleichen. Dennoch unterliegt die Auswertung starken subjektiven Tendenzen, welche durch Sympathie oder Antipathie zwischen Prüfer und Geprüfem dem Verhalten des Prüfungsteilnehmers sowie zahlreichen anderen Faktoren, wie dem ersten Eindruck, Vorleistungen oder Kontrasteffekten hervorgerufen werden können. Für die Interpretationsobjektivität setzen sich diese Faktoren entsprechend fort.



Zum Thema mündliche Prüfungen und Reliabilität trifft ROLOFF folgende Aussage: „Die Reliabilität bezieht sich [...] auf die Präzision, mit der in der mündlichen Prüfung benotet wird, unabhängig davon, welche Leistung gemessen wird.“ (Roloff, 2002) Das bedeutet, dass unterschiedliche Prüfer unabhängig voneinander die Leistung eines Geprüften gleich bewerten würden. Auch hier hängt die Reliabilität wieder stark von den gestellten Prüfungsaufgaben ab, hinzu kommen aber diesmal noch weitere, den Prüfungsrahmen betreffende Umstände, die im Gegensatz zu schriftlichen Prüfungen nicht weitgehend gleich, sondern oft stark verschieden sind. Auch spielt wieder die Verfassung der Prüfenden eine entscheidende Rolle und anders als bei schriftlichen Prüfungen liegt jetzt der Hauptzeitaufwand nicht mehr in der Prüfungskorrektur, sondern in der Prüfungsdurchführung. Im Laufe eines langen Prüfungstages werden sich also nicht nur die Leistungsbewertung, sondern auch die Aufmerksamkeit, Geduld, Formulierungen, Antwortakzeptanz etc. seitens der Prüfenden ändern.

Zusammenfassend kann man also festhalten, dass schriftliche Prüfungen vor allem hinsichtlich ihrer Objektivität mündlichen Prüfungen überlegen sind. Mündliche Prüfungen können hingegen teilweise praxisnäher gestaltet werden und sind individueller auf einzelne Prüflinge oder Prüfungssituationen anpassbar. Eine entscheidende Rolle für die Qualität einer mündlichen Prüfungen spielen die Kompetenzen der Prüfungskommission.

### **Referat/ Präsentation**

Das Referat oder die Präsentation kann eine sehr praxisnahe Form der Leistungsbewertung darstellen, denn im Gegensatz zu den anderen bereits genannten Prüfungsformen, befindet sich der Geprüfte hier in einer, auch im späteren Berufsleben auftretenden Assessment-Situation (Pitch, Meeting, Projektpräsentation, Vorstellungsgespräch, etc.). So spielen neben der Präsentation der eigentlichen Inhalte auch die Art und Weise des Vortrags und die inhaltliche Strukturierung eine entscheidende Rolle. Es werden also neben dem Vorhandensein von Fachkompetenz auch in hohem Maße Zusatzkompetenzen, wie Methoden- und Sozialkompetenzen geprüft. Da diese aber nicht im Gegensatz zu den eigentlichen Prüfungszielen stehen, sondern diese im speziellen Fall mit einschließen, kann hier äußerst praxisnah geprüft werden. Dabei wird vorausgesetzt, dass die Referats- oder die Präsentationinhalte auch dementsprechend ausgelegt und die Rahmenbedingung möglichst authentisch sind. Bezugnehmend auf den problemorientierten Lerngedanken ist es sinnvoll, dass der Geprüfte über ein, im Vorhinein eigenständig bearbeitetes Thema referiert und seine eigene Problemlösungs-Strategie präsentiert. Dieser kognitivistische Ansatz kann die Validität eines Referates, einer Präsentation entscheidend erhöhen, da so einerseits die Motivation

des Geprüften gesteigert und außerdem ein eigener Wissenskontext erzeugt werden. Nur eine Stoffsammlung aus fremden Quellen zusammenzustellen und diese anschließend wiederzugeben ist wesentlich uneffektiver.

### **Projektarbeit**

Eine Projektarbeit ist eine sowohl praxis-, als auch wissenschaftsnahe Form des Assessment und kann beispielsweise auch als gerade beschriebene problemlösungsorientierte Vorarbeit eines Referates oder einer Präsentation dienen. Im Gegensatz zu einer klassischen Prüfung ist eine Projektarbeit weniger ergebnisorientiert, denn im Mittelpunkt steht die Bearbeitung eines bestimmten Problems, einer definierten Aufgabe und der entsprechend angewendete individuelle Lösungsprozess. Die Projektarbeit besteht nicht nur aus einer Präsentation der Resultate, sondern auch aus einer ausführlichen Dokumentation der Vorgehensweise und Teilergebnisse. Außerdem entsteht durch die Bearbeitung der Projektaufgabe häufig ein reelles physisches Produkt, was allgemein motivationsfördernd wirkt und die kontinuierliche Lernbereitschaft erhöht. Die, im Vergleich zu klassischen Prüfungen, meist längere Projektdauer ermöglicht eine Selbstorganisation des Bearbeitungsablaufs und bei einer Teamarbeit auch eine gewisse Freiheit der Aufgabenwahl entsprechend seiner individuellen Stärken und Schwächen. Es kann also auch hier wieder eine hohe Validität erreicht werden, ohne andere Faktoren, wie zum Beispiel Aufwand, Sicherheit oder Personaleinsatz immens steigern zu müssen. Auch hinsichtlich ihrer Objektivität und Reliabilität sind Projektarbeiten mit anderen Prüfungsformen vergleichbar und erfüllen zu dem sehr gut die Forderung nach gesteigerter Kompetenzorientierung.

Generell wird deutlich, dass jede Prüfungsform erwartungsgemäß spezifischen Stärken und Schwächen aufweist und deshalb für bestimmte Anwendungsbereiche besser geeignet ist als für andere. Keine Prüfungsform erreicht dabei perfekte Objektivität, Reliabilität oder Validität. Es ist also entscheidend immer konkret zu prüfen, welche Kompetenzen tatsächlich examiniert werden sollen und wie diese Überprüfung möglichst praxisnah gestaltet werden kann. Diese Vorgehensweise wird nach der Meinung des Autors bis jetzt noch nicht ausreichend umgesetzt, denn oft werden bestehende Verfahren als gegeben hingenommen und nicht kontinuierlich kritisch genug hinterfragt. Vor allem Zusatzpotentiale von Assessments, wie Motivationsförderung, Lernlenkung, Gestaltung eines schülerzentrierten Unterrichts und Selbsteinschätzungsfunktionen werden weitgehend vernachlässigt, was dazu beiträgt, dass lehren und lernen immer noch stark getrennt von der Leistungsbewertung betrachtet wird. Auch werden dadurch Assessments, aus Sicht der Lerner, oftmals als etwas vorwiegend Negatives angesehen..

### 3.1.4 Bisheriger Einsatz von E-Assessments

In vielen Hochschulen werden bereits elektronisch unterstützte Leistungsnachweise durchgeführt, weshalb deren bisherige Einsatzpraxis an dieser Stelle betrachtet wird. Es werden dabei allerdings keine konkreten Projekte vorgestellt, sondern vielmehr häufig angewandte Prinzipien geschildert. Auch wenn diese nicht auf jedes einzelne durchgeführte E-Assessment zutreffen, so ist es doch möglich einen guten Querschnitt durch die aktuelle Anwendungspraxis zu erzielen.

Generell lassen sich die aktuellen E-Assessment-Projekte nach Meinung des Autors in zwei Hauptgruppen unterteilen, die den Großteil dieser Assessment-Praxis abbilden. Die erste Gruppe hat einen eher experimentellen Charakter und geht meist von Einzelnen oder kleinen Forschungsgruppen aus. Ihr Ziel ist häufig die Erprobung der neuen Möglichkeiten, oft vor allem aus technischer Sicht, von elektronischen Assessments. Bei Projekten der anderen, „zielorientierten“ Gruppe stehen primär ökonomische Gesichtspunkte im Vordergrund, mit dem Ziel den gestiegenen Bedarf an Leistungsnachweisen überhaupt bewältigen zu können. Beide Gruppen stellen verschiedenen Anforderungen und Erwartungen an E-Assessments und richten ihre Forschung und Entwicklung entsprechend darauf aus. Während die experimentelle Gruppe dem natürlichen menschlichen Interesse an Neuem und dem damit verbundenen Bestreben, dieses Neue zu entdecken und erkunden nachgeht, versucht die zweite Gruppe hauptsächlich bereits bekannte Eigenschaften des Neuen möglichst effizient für sich zu nutzen. Beide Ansätze sind bereits ein guter Schritt in eine richtige Richtung, jedoch ergeben sich auch jeweils gewisse problematische Effekte.

Betrachtet man ganz allgemein die Nutzung einer neuen Technologie für einen bestimmten Anwendungsbereich, ergeben sich zwei Möglichkeiten. Entweder existiert ein konkretes Ziel, dessen Umsetzbarkeit mit Hilfe der neuen Technologie geprüft werden soll um abschließend entscheiden zu können, welches Potential vorhanden ist und wie effizient sich damit das gesetzte Ziel verwirklichen lässt (Ansatz der „zielorientierten Gruppe“). Oder, man nähert sich dem Neuen ohne feste Ziele und analysiert alle Eigenschaften und Möglichkeiten und entscheidet daraufhin in welchen Anwendungsbereichen diese neue Technologie angewendet werden kann um Bestehendes zu optimieren (Ansatz der „experimentellen Gruppe“).

Bei der experimentellen Gruppe besteht dabei im Bereich des E-Assessments das generelle Problem, dass die entwickelten Produkte oft völlig überladen sind und verschiedenste Techniken nur um des Nutzens Willen genutzt werden, wobei oft keine klaren Ziele vorhanden sind. Konkret nutzen

derartige Assessment-Systeme oft unzählige verschiedene Fragetypen und Präsentationsmedien, die sich aber oft auf wenige reduzieren ließen, ohne die didaktische Qualität des Assessment zu vermindern. Im Gegenteil, der Geprüfte muss sich so ständig auf neue Situationen und wechselnde Bedingungen einstellen und ist damit oft völlig überfordert (Ricketts & Wilks, 2002).

Andererseits ist eben gerade dieses experimentieren wichtig für die Entwicklung einer Technologie, solange die Experimentierphase auch als solche angesehen und nicht als fertiges Produkt akzeptiert wird. Die zielorientierte Gruppe macht sich die Ergebnisse des Experimentierens zunutze und versucht diese möglichst effektiv einzusetzen um wiederum ihre Ziele zu erreichen. Entscheidend ist dabei vor allem die richtige Zielformulierung, die derzeit oft ausschließlich in der Rationalisierung von Prüfungen mittels Computerunterstützung liegt, um schlicht die „Bologna-Konsequenzen“ damit auffangen zu können. Das didaktische Weiterentwicklungspotential bleibt dabei meist auf der Strecke, die Folge sind oft elektronische Multiple-Choice-Prüfungen oder ähnliche relativ triviale Lösungen, die durch ihre weitestgehend automatische Auswertbarkeit zwar große Zeiteinsparungen bewirken, dem Gedanken der Kompetenzorientierung genauso wenig entsprechen wie ihre konventionellen Pendanten (Reinmann, 2007).

Viele der unter Punkt 3.1.1 und 3.1.2 aufgezeigten Schwachpunkte der klassischen Prüfungspraxis werden also auch durch die aktuelle Anwendungsweise von E-Assessments nicht verbessert sondern weitestgehend übernommen, denn an den prinzipiellen Abläufen hat sich nur wenig geändert, lediglich das Prüfungsmedium wurde ausgetauscht.

Auch soll an dieser Stelle noch einmal darauf hingewiesen werden, dass MC oder geschlossene Fragetypen generell nicht als negativ oder unzeitgemäß angesehen werden, nur sollte ihr Einsatz dort stattfinden, wo er auch didaktisch sinnvoll und nicht nur zeiteffizient ist. Außerdem muss natürlich angemerkt werden, dass nach wie vor kein detaillierter rechtlicher Rahmen für die Durchführung elektronischer Leistungsnachweise in Deutschland existiert, beziehungsweise geltende Gesetze die Möglichkeiten von E-Assessments stark einschränken, was deren Weiterentwicklung logischerweise eminent behindert.

## 3.2 Anforderungen an E-Assessments als zeitgemäße Leistungsnachweise aus heutiger Sicht

Im nachfolgenden Kapitel werden ausgewählte, aus Sicht des Autors wichtige, Anforderungen an heutige Leistungsnachweise, speziell in Verbindung mit elektronisch unterstützten Assessment-Verfahren vorgestellt und diskutiert. Eine entscheidende Rolle spielt dabei der Ansatz, Leistungsnachweise nicht losgelöst von anderen Komponenten im Lern- und Lehrprozess zu betrachten (Eugster, 2009/2010), sondern diese als einen essentiellen Teil heutiger Lernkultur zu verstehen. Dabei werden Leistungsnachweise nicht ausschließlich als summative „Prüfungen“ angesehen, sondern vordergründig als Mittel zur Lernfortschrittskontrolle, individuellen Lernlenkung und Selbsteinschätzung. E-Assessments sind dabei kein „Allheilmittel“ oder Garant für eine automatische Verbesserung der Lehre oder Erhöhung der Validität von Leistungsnachweisen. Entscheidend ist der sinnvolle Einsatz von adäquaten Assessment-Szenarios an den richtigen Stellen im Lernprozess und unter den richtigen Voraussetzungen.

Ein wichtiges Ziel zeitgemäßer Assessments, ist aus Sicht des Autors, nicht nur die Fähigkeit Lernerfolge zuverlässig messen zu können, sondern auch die Fähigkeit diese in entscheidendem Maße positiv zu beeinflussen und mitzubestimmen. Eine Grundlage dafür liefert ROTH mit seiner Aussage, dass Bildungserfolg neben der Intelligenz, zu etwa gleichen Teilen auch von Motivation und Fleiß abhängt. Diese Faktoren sind zwar einerseits abhängig von der Persönlichkeit und elterlichen Prägung des Lerners, andererseits aber auch signifikant von seiner Umwelt (Roth, 2010). Lässt sich also durch Assessments, ähnlich zu anderen „Konkurrenzsituationen“ die Motivation und der Fleiß der Lerner erhöhen, wirkt sich das in großem Maße positiv auf deren Lernerfolge aus. Speziell im E-Learning-Bereich kann dies von großer Bedeutung sein, da andere motivierende Faktoren, wie beispielsweise entsprechende Fähigkeiten des Lehrers im „Face-to-Face-Teaching“ oder der direkte Kontakt zu Mitlernern erheblich eingeschränkt, beziehungsweise zumindest virtualisiert werden.

### 3.2.1 Überprüfung kompetenzorientierter Lernziele

Die, im Zuge der Bologna-Reform häufig angesprochenen Kompetenzorientierungen, beschreiben eine Anpassung der Lernzielformulierungen an die heutigen Bedürfnisse eines modernen Bildungssystems. Auch wenn keine einheitliche Definition zum Kompetenzbegriff existiert, kann man zusammenfassend, folgende Aussage treffen: Unter Kompetenzen sind individuell verfügbare oder erworbene Fähigkeiten und Fertigkeiten zu verstehen, die es ermöglichen bekannte Sachverhalte

problemlösungsorientiert auch auf unbekannte Situationen anzuwenden. (Dany et. al, 2008) Vom Lernenden wird folglich erwartet, dass konkrete Erkenntnisse aus verschiedenen Einzelleistungen koordiniert auf ein praxisnahes, unbekanntes Problem angewendet werden können. Um dies zu realisieren müssen auch im Lehrprozess vielfältige, pragmatische Szenarien inszeniert werden, die dem Lernenden ein eigenständiges Arbeiten und individuelles Problemlösen ermöglichen. Gleichzeitig bedeutet Kompetenzorientierung aber auch eine Fokussierung auf das Ergebnis eines Lernprozesses und die zu erreichenden Kompetenzen (Reinmann, 2007).

Die damit verbundenen Anforderungen an kompetenzorientierte Assessments sind entsprechend hoch und führen auch E-Assessments, unter der Maßgabe eines praxistauglichen Aufwand-Nutzen Verhältnisses, schnell an die Grenzen des momentan technisch Möglichen. Dabei ist es zunächst unentscheidend, wann, wie und in welcher Häufigkeit die Assessments konkret im Lernprozess positioniert werden (diese Problematik wird unter Punkt 3.2.2 näher betrachtet). Viel entscheidender ist die Frage, wie man überhaupt adäquat Kompetenzen prüfen kann, dies gilt dann sowohl im Einzelnen wie auch in der Masse. Die meisten bisherigen Prüfungsszenarien kommen dafür sicherlich nicht in Frage, denn sie dienen vorwiegend der Überprüfung von sogenanntem „Lexikalischen“, also dem Faktenwissen. Überhaupt scheint eine direkte Erfassung von Kompetenzen nicht möglich, lediglich bestimmte Leistungen lassen sich erfassen und bewerten um daraufhin auf die vorhandenen Kompetenzen zu schließen (Klieme et al., 2003).

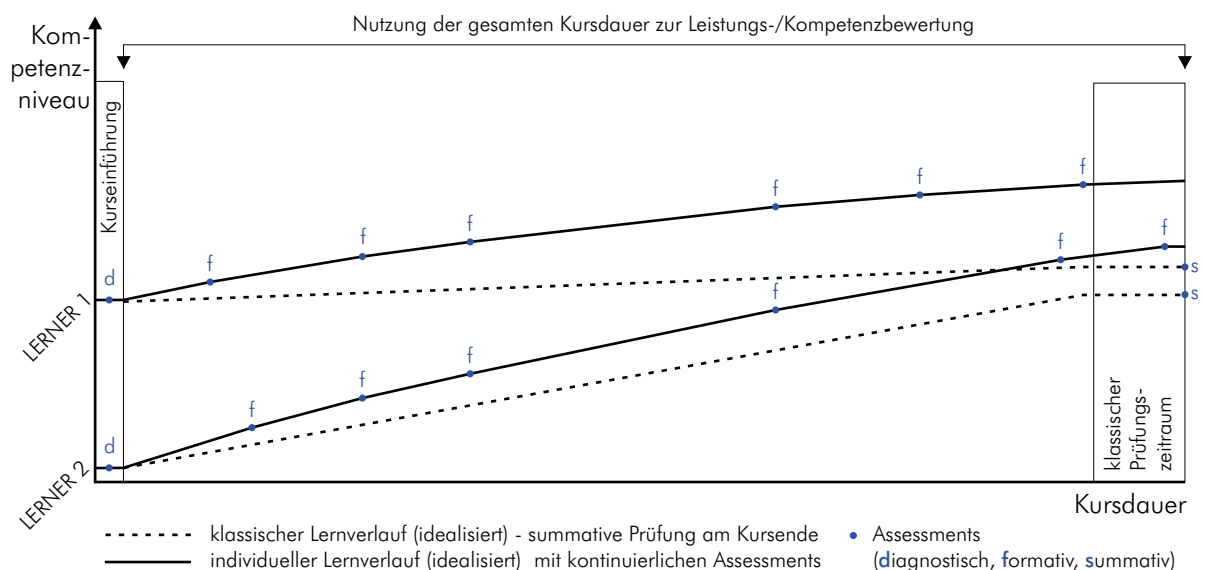


Abb. 8: Kontinuierliche Erfassung von Lern- und Lösungsprozessen im Vergleich zur punktuellen Leistungsmessung

Ein Ansatz dennoch eine Kompetenzbewertung mittels Assessments vornehmen zu können, beruht in der Fokussierung des Lösungsprozesses anstelle der Lösung selbst (Reinmann, 2007), denn die kombinierte Betrachtung des Lösungsansatzes, der Vorgehensweise und des resultierenden Ergebnisses unter Berücksichtigung des Wissensstandes des Lerners und anderer spezifischer Faktoren liefern weit mehr Daten als das solitäre Ergebnis selbst.

Dieser Ansatz ist auch gut mit konstruktivistischen Grundlagen vereinbar, denn er berücksichtigt individuelle Denkmodelle und muss folglich auch verschiedene Lösungen akzeptieren. Für eine Umsetzung als E-Assessment bedeutet das natürlich eine nahezu unbegrenzte Anzahl an Lösungsmöglichkeiten und den weitestgehenden Verzicht auf geschlossene Fragenformate, wie Multiple-Choice (MC), etc. Generell lässt sich eine konsequent individuell bearbeitbare Problemfrage auch nicht hinsichtlich ihrer Antwortformate beschränken, was bedeutet, dass dem Geprüften freigestellt werden muss, welche Präsentationsmittel (Formel, Text, Zeichnung, etc.) er zur Darstellung seines Lösungsprozesses verwenden möchte. Die technische Umsetzung gestaltet sich dabei relativ einfach, denn neben der eigentlichen Aufgabenstellung könnte dem Geprüften eine Art „Whiteboard“ zur Verfügung gestellt werden, das sich durch modular angelegte Antwortformate individuell füllen lässt. Schwieriger gestaltet sich dann allerdings die Auswertung eines solchen Assessments. Durch die individuelle Gestaltbarkeit gibt es keine festes, maschinenlesbares Auswertungsschema, wie dies beispielsweise bei MC-Tests der Fall ist. Der Workload für die prüfenden Instanzen würde sich dadurch immens erhöhen.

Eine Lösung dieses Problems bietet aber der Ansatz eines semi-automatischen Assessments, das mittels „Data-Mining“ oder ähnlichen Verfahren Informationen aus, für die automatische Verarbeitung sonst unlesbaren Daten extrahiert und auswerten kann. (Bescherer et. al, 2012) Dabei können bestimmte Musterlösungsfragmente, die in besonders häufigen Lösungsstrategien enthalten sind, hinterlegt werden und nur Sonderfälle an einen menschlichen Prüfer weitergeleitet werden. Die manuelle Bewertung dieser Fälle kann dann wiederum modularisiert und automatisch in den Musterlösungskatalog übernommen werden um den Datenbestand zu ergänzen. Da ein solch komplexes System mit völlig offenen Antwortformaten, nach aktuellem Forschungsstand des Autors noch nicht existiert und noch entsprechend umfangreiche Entwicklungsarbeit und -zeit nötig ist, kann das Prinzip des semi-automatischen Assessment natürlich auch auf definiertere Strukturen angewendet werden. So können Korrekturzeiten verkürzt und effizienter gestalten werden ohne dabei die Qualität des Ergebnisses, durch die Verwendung ausschließlich maschinenauswertbarer Fragentypen, zu beeinträchtigen.

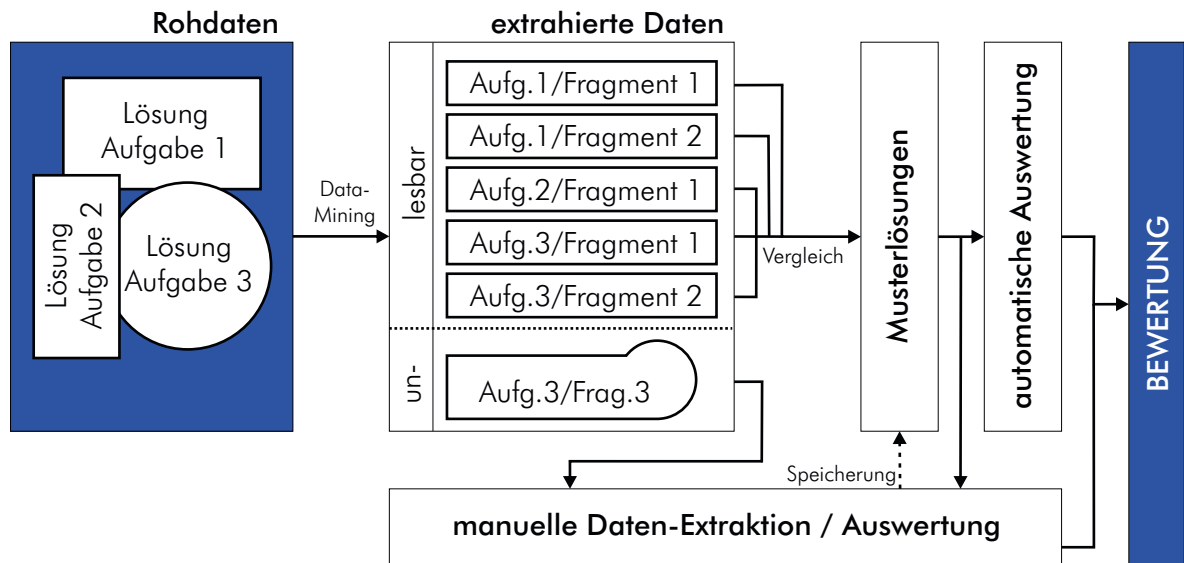


Abb. 9: Beispielschema für einen semi-automatischen E-Assessment-Workflow

Eine weitere Option der summativen Leistungsbewertung ist die Nutzung eines E-Portfolios, welches eine elektronische Sammlung mehrerer Arbeitsleistungen darstellt und aus einer Vielzahl verschiedener Komponenten bestehen kann (Stratmann et. al, 2009). Ein Portfolio ist also ein Art Referenzmappe, die meist personenbezogen darstellt an welchen Projekten die betreffende Person mitgewirkt, welche spezifischen Probleme sie wie gelöst oder welche Aufgaben sie wie bearbeitet hat. Diese Datensammlung erfolgt meist über einen längeren Zeitraum, beispielsweise über ein Semester, und ist damit in der Lage, neben einer finalen summativen Bewertung, auch den Lernfortschritt zu dokumentieren und diesen als weiteren Benchmark einzusetzen. So kann ein Portfolio als eine Art „Container“ oder gemeinsamer Rahmen gesehen werden, der verschiedenste Leistungsnachweise enthalten kann und eine hohe Durchführungsobjektivität aufweist. Für den Lernenden hat diese Form des Assessments zudem den Vorteil, dass seine Leistungen kontinuierlich dokumentiert und nicht nur, wie in einer klassischen Prüfungssituation, in einem sehr kurzen Zeitraum punktuell erfasst werden.

Auch wenn diese Form der leistungsresultierenden Kompetenzbewertung sicherlich nicht für Massenveranstaltungen geeignet ist, abgesehen von der auch hier in geringem Maße anwendbaren Semi-Automatik, bietet sie trotzdem vielfältige Vorzüge im Bezug auf die Qualität der Lehre. So wird einerseits der Lerner durch eine kontinuierliche Dokumentation seiner Leistungen gezwungen sich dauerhaft mit dem Lernstoff zu beschäftigen, andererseits besteht für die bewertenden Instanzen eine umfassende Datensammlung zur finalen Kompetenzbewertung. Die Bewertung kann unter Umständen natürlich auch ausgelagert werden um, einerseits



den Workload auch mit Massenveranstaltungen kompatibel zu machen und andererseits nachfolgenden Bewertungsgremien (andere Bildungseinrichtung, Personalabteilung, etc.) eine objektive Beurteilung, der für sie relevanten Leistungen, zu ermöglichen.

### 3.2.2 Gestaltung eines schülerzentrierten Unterrichts

Neben der Fähigkeit, Kompetenzen valide summativ bewerten zu können, sollten zeitgemäße E-Assessments auch als elementares Mittel zur Unterrichtsgestaltung und Lernlenkung gesehen werden. Moderne Assessments sind essentielle Bestandteile des Lernprozesses und keine obligatorischen, aber separierten „Anhängsel“. Gerade im Bereich des E-Learning kann so auf relativ einfache Art und Weise eine Individualisierung des Lernprozesses erreicht werden, denn im Gegensatz zum „Face-to-Face-Teaching“ kann jedem Lerner unproblematisch eigener „learning-content“ zur Verfügung gestellt und bedürfnisorientiert präsentiert werden. In der klassischen Präsenzlehre wäre dies nur mit sehr hohem Aufwand realisierbar und würde in einer konventionellen „in-classroom“-Situation sicherlich zu erheblichem Chaos führen. In einer virtuellen Umgebung ist dies nicht der Fall und so kann für jeden Lerner ein individueller Pfad durch eine Ausgangsmasse an „learning-content“ generiert werden, an dessen Knotenpunkten E-Assessments stattfinden um den nächsten effektivsten Pfadabschnitt zu ermitteln. Diese Assessments sind nicht zwingend als umfangreiche Tests zu verstehen, sondern können je nach Lernabschnitt beispielsweise auch nur aus einer einfachen „ja/nein“-Frage bestehen.

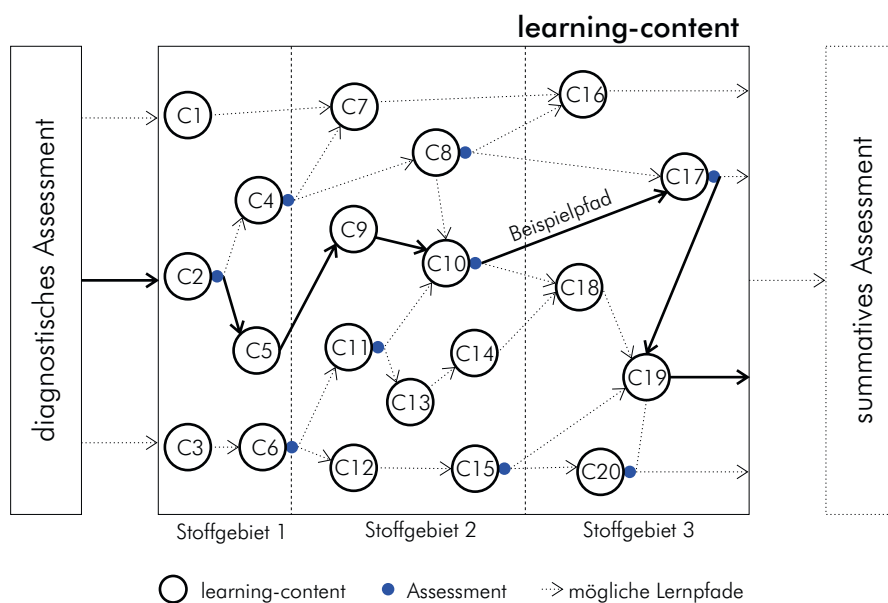


Abb. 10: Individuelle Lernpfad-Bestimmung durch E-Assessments

Welche konkreten Vorteile bietet das für den Lerner? Einerseits ist so eine individuelle Berücksichtigung der Vorkenntnisse und des Wissensstandes jedes Einzelnen möglich, es entsteht also keine Unter- oder Überforderung. Andererseits kann jeder Lerner trotz einer eigenständigen Mitbestimmung seines Lernpfades innerhalb fester Grenzen gezielt gefördert werden. Ein späterer summativer Leistungsvergleich ist also trotz aller Individualität möglich.

Ein weiterer entscheidender Punkt für die Maximierung von Lernerfolgen ist die Förderung eigenständigen Lernens und das Schaffen eines eigenen Lernbewusstseins. (Zhou, 2012) Dies kann aus Sicht des Autors durch eine Kombination von drei Hauptmerkmalen erreicht werden: (a) Ausrichtung der Lerninhalte an praxisrelevanten, hinreichend schwierigen Problemstellungen innerhalb des Verstehenshorizontes des Lerners, (b) Präsentation von zeitnahe, individuellen Feedback im Anschluss an Assessments und „on-demand“ und (c) Integration von „Community-Tasks“ in den Lernprozess.

Die Ausrichtung der Lerninhalte auf Problemlösungsszenarien bringt den Vorteil, dass Wissen und Kompetenzen in einem praktischen Rahmen im direkten Umfeld des Lerners erworben werden können. So kann eine eigenständige Filterung und Umwandlung von kontextbezogenem in abstraktes Wissen vorgenommen werden, was als sehr effektiv gilt (vgl. konstruktivistischer Ansatz). Außerdem wird durch den realen Bezug des Problems zur Umwelt des Lerners die Motivation und Lernbereitschaft in erheblichem Maße erhöht was wiederum gute Lernerfolge verspricht (Merrill, 2002). Das auch die spätere Überprüfung der Lernleistungen in einem ähnlich praxisnahen Kontext geschehen sollte ist ja bereits im vorherigen Abschnitt 3.2.2 erwähnt wurden.

Wie kann gezieltes Feedback die Lernleistung und Motivation eines Lerners positiv beeinflussen? Unter der Maxime, dass die Abbildung eines bestimmten Sachverhaltes im Gehirn jedes Lerners immer subjektiv über ein spezifisches „mentales Modell“ erfolgt, welches sich immer von den Modellen anderer Lerner unterscheiden wird, ist es wichtig, das Feedback individuell auf den Lerner anzupassen. Nur dann können größtmögliche Erfolge erzielt werden. (Ilfenthaler, 2012) Das entsprechende Methoden bereits technisch umsetzbar sind beschreibt ILFENTHALER und nennt Beispielsoftware wie: TASA (Text-Guided Automated Self-Assessment) oder iGRAF (Instant Graphical Feedback). (Ilfenthaler, 2012) Es muss dabei allerdings immer in verschiedene Arten der Feedbacks unterschieden werden, die hinsichtlich ihrer Komplexität und dem damit verbundenen Informationsgehalt für den Lernenden variieren. Die trivialste Form des Feedbacks beinhaltet lediglich eine Aussage über die Richtig- oder Falschbeantwortung einer bestimmten

Frage und bietet damit wenig Parallelen zu konstruktivistischen Ansätzen. Eine Komplexitätsstufe höher wird auch noch die richtige Lösung hervorgehoben, was vor allem bei mehr als zwei Antwortmöglichkeiten bereits erste Lerneffekte hervorrufen kann. Allerdings beschränkt sich auch dieses Feedback auf geschlossene Fragenformate, deren Antwort sich als eindeutig richtig oder falsch deklarieren lässt (vgl. MC-Test).

Wesentlich informativer wird das Feedback, wenn neben der richtig/falsch Bewertung ein statischer Hinweis zur gegebenen Antwort oder zum richtigen Ergebnis präsentiert wird. Der Lerner kann dann anhand der Musterlösung eigene Fehler nachvollziehen und diese gegebenenfalls korrigieren. Wie bereits eingangs beschrieben wurde ist diese Art der Abstraktion allerdings eher ungeeignet um effektive Lernerfolge zu generieren, weshalb ein individuelles Feedback, zugeschnitten auf den Lösungsweg des Lerners, unter Beachtung seiner Kenntnisse und Vorleistungen erfolgen sollte. Entscheidend ist, dass dem Lerner immer nur so viel Informationen zur Verfügung gestellt werden, dass dieser zum nächsten Schritt seines Problemlösungsprozesses geführt wird. (Kong, 2012) Auch hier kann, abhängig vom jeweiligen „learning-content“, wieder eine semi-automatische Feedback-Logik zum Einsatz kommen, die bestimmte Standardfälle und Muster erkennt und selbstständig bearbeitet. Das Feedback kann so in vielen Fällen unmittelbar nach einer Nutzerinteraktion erfolgen (= instant Feedback).

Sonderfälle können zur manuellen Bearbeitung weitergeleitet werden, was eine gezielte Hilfestellung oder Förderung ermöglicht. Der Workload an „Standardproblemen“ für die Dozenten kann dabei, je nach Güte des Feedback-Algorithmus, auf ein Minimum reduziert werden, was genügend Freiraum für eine individuelle Einzelbetreuung gestattet. In einer Ausbaustufe diese Systems könnte zwischen dem maschinellen und dem Dozenten-Feedback noch eine Community-Feedback-Ebene eingeführt werden.

Nicht standardkonforme Lösungen könnten dann vorerst an andere Lerner, deren Lösungswege für das gleiche oder ein ähnliches Problem bereits als valide eingestuft wurden, weitergeleitet werden. Hilft das Feedback dem Lerner ist die Anfrage beendet und die Community-Lösung kann später, nach einer eventuellen manuellen Prüfung, in den Datenbestand übernommen werden. Hilft das Feedback dem Lerner nicht, wird die Anfrage zur „Dozenten-Ebene“ weitergereicht. So ließe sich der Feedback-Workload für die Dozenten nochmals verringern, ohne spürbare Konsequenzen für den Lerner. Eine Belohnung der Hilfesteller durch Zusatzcredits, etc. wäre denkbar.

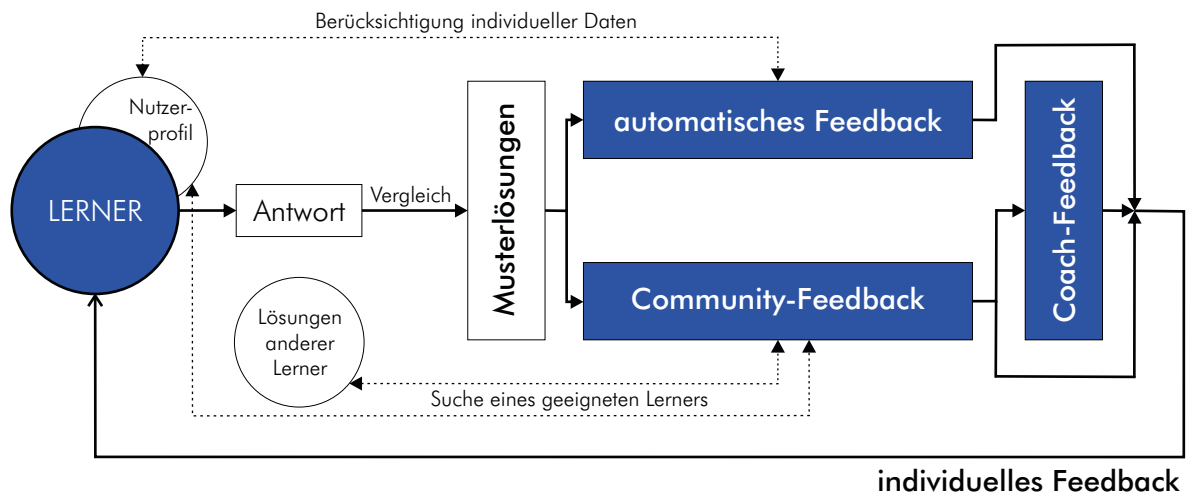


Abb. 11: Beispielschema für einen semi-automatischen Feedback-Workflow

Das letzte Hauptmerkmal für die Gestaltung eines schülerzentrierten Unterrichts mittels innovativer E-Assessments beschäftigt sich mit der Förderung von Teamwork und Teamfähigkeit (Kerres, 2001). Entsprechende Kompetenzen sind heute unverzichtbar und gängige Praxis bei der Bearbeitung komplexer Sachverhalte im späteren Berufsalltag. Eine realistische Vorbereitung sollte also bereits im Hochschulstudium geleistet werden. Assessments können dabei einmal als Grundlage für die Zusammenstellung der Teams zum Einsatz kommen um Lerner mit verschiedenen Charakteren, Lerntypen und Kompetenzniveaus auszuwählen. Andererseits können später auch die Leistungen der einzelnen Gruppen mittels E-Assessments erfasst werden. Unabhängig davon, ob jedes Teammitglied einen eigenständigen Teil der Gesamtaufgabe bearbeitet, oder mehrere Personen ein und dieselbe Problematik behandeln. Ein Vorteil im Einsatz von E-Assessments gegenüber klassischem Face-to-Face-Teamassessment besteht vor allem in der örtlichen Ungebundenheit der einzelnen Teammitglieder voneinander. Vorhandene Zeitressourcen können so effektiver genutzt und zur tatsächlichen Bearbeitung der Projektaufgabe genutzt werden, anstatt in Fahrzeit, Raumsuche, etc. investiert werden zu müssen.

### 3.2.3 Verfügbarkeit/ Durchführbarkeit von E-Assessments

Generell ist die Verfügbarkeit von E-Learning-Content und E-Assessments, aus Sicht des Autors, ein entscheidender Punkt für die Akzeptanz und Effektivität moderner Lernumgebungen und gleichzeitig auch ein großer Unterschied zum klassischen Unterricht. Zeitliche Ressourcen sind heute eines der hochwertigsten Güter überhaupt und in ihrer Existenz stark begrenzt. Viele Dinge passieren deshalb „to go“ um auch vorhandene Leerlaufzeiten noch effizient nutzen zu können.

Entsprechend rasant verlief die technische Entwicklung der Anwenderhard- und -software in diesem Bereich in den letzten Jahren, mit dem Ziel immer leistungsfähigere und benutzerfreundlichere Geräte auf den Markt zu bringen. Viele ehemals „stationäre“ Tätigkeiten lassen sich heute „mobil“ ausführen und sorgen für eine entsprechend zunehmende Erwartungshaltung nach Mobilität in vielen Anwendungsbereichen. Diese Tendenz ist demnach auch im Bereich der Bildung allgemein und somit auch im Hochschulbereich zu erwarten, beziehungsweise bereits zu verzeichnen.

Zeitgemäße E-Assessments sollten sich an diesen Trends und heutigen Einstellungen zum Lernen orientieren und so mobil wie möglich gestaltet werden. Dabei ist zu beachten, dass sich mit steigender Mobilität natürlich auch die Aussage des Assessment ändert, da beeinflussende Umweltfaktoren nicht mehr kontrolliert werden können. Gemeint ist damit weniger die Konzentrationsfähigkeit oder das Ablenkungspotential des Assessment-Teilnehmers, als die Hilfestellung Dritter oder das Benutzen von Suchmaschinen, Nachschlagewerken, etc. An dieser Stelle müssen innovative pädagogische Konzepte ansetzen um Assessments zu entwickeln, die trotz der nahezu unbegrenzten Anzahl an unbekanntem Einflussfaktoren auf die Assessment-Situation valide Ergebnisse liefern. Es wird dann sicherlich weniger die konkrete Einzelleistung eine ausschlaggebende Rolle spielen, als die Leistungs-Entwicklung sowie die spezifische Vorgehensweise und Methodik des Lerner, beim Lösen einer Vielzahl unterschiedlicher Assessments.

Auch ist davon auszugehen, dass sich durch die Gestaltung des Assessments an sich, die Bereitschaft des Lerner zur bewussten Täuschung verändern lässt und in dem Moment minimiert wird, wo der leistungsfördernde Gedanke des Assessment maximal und der selektierende Gedanke minimal gehalten wird. Solche Assessments könnten dann vornehmlich formativ eingesetzt werden um den individuellen Lernprozess zu beeinflussen und den spezifischen Lernpfad des Einzelnen zu bestimmen, bei dem primär nur seine eigenen Leistungen im Vordergrund stehen. Diese Assessments sollten mobil verfügbar sein, damit der Lerner jederzeit die Möglichkeit hat sie durchzuführen, unabhängig seines Aufenthaltsortes oder seines Zeitbudgets. Die Assessments müssen schnell verfügbar, immer speicherbar und in der Lage sein auf plötzliche Abstürze, Kommunikationsstörungen, etc. zu reagieren. Außerdem sollten die Assessments abwechslungsreich und fordernd in den „learning-content“ integriert werden und den Lerner in passenden Abständen an ihre Existenz erinnern. Anderenfalls wird ihre Nutzung schnell als kompliziert, unpraktisch, anstrengend, uninteressant oder langweilig empfunden, was den Effekt des Assessments auf den Lernerfolg weitgehend eliminieren würde.

Möchte man ergänzend an der klassischen Prüfungsform, zur summativen Kompetenz- und Leistungsbewertung festhalten, können selbstverständlich auch dort E-Assessments eingesetzt werden. Um die Eigenständigkeit oder Nachvollziehbarkeit der Leistungserbringung dann allerdings gewährleisten zu können ist eine „in-classroom“-Lösung zu empfehlen. Ein mögliches Setup eines solchen „prüfungsnahe“ Assessment wird im nächsten Kapitel beispielhaft beschrieben, muss aber im Detail selbstverständlich auf jede Lern- und Prüfungssituationen individuell angepasst werden.

### **3.3 Rahmenbedingungen des E-Assessment - Interface-Design und Assessment-Umgebung**

Für einen erfolgreichen Einsatz von E-Assessments ist es neben der detaillierten Auseinandersetzung mit den theoretischen Hintergründen ebenso wichtig bestimmte Rahmenbedingungen zu beachten und diese bei der Umsetzung des Assessment-Konzeptes einzuhalten. Eine entscheidende Grundvoraussetzung dafür ist es natürlich, ein solches Konzept tatsächlich zielgerichtet zu entwickeln und dessen Validität kontinuierlich zu hinterfragen und auch auf dessen technische Umsetzbarkeit zu prüfen. Auch sollten bereits die späteren Einsatzbedingungen bekannt sein um gegebenenfalls schon bei der Entwicklung berücksichtigt werden zu können.

Dass das Interface-Design generell ein sehr elementarer Aspekt jeder Mensch-Maschine-Schnittstelle ist, ist inzwischen allgemein bekannt und nicht umsonst wird vor allem dann, wenn der Benutzer ein Kunde ist und somit relativ frei zwischen verschiedenen Angeboten wählen kann, auf ein durchdachtes Interface-Design Wert gelegt. Auch wenn dabei das primäre Ziel im Verkauf einer Ware oder Dienstleistung und somit in der Absatzsteigerung liegt, können die angewendeten Prinzipien nahezu eins zu eins auch auf E-Assessments übertragen werden. Ein zeitgemäßes E-Assessment unterscheidet sich in seinen Anforderungen bezüglich seiner Usability nicht wesentlich von anderen Internetangeboten - es ist de facto eine „Website“ und somit nichts grundlegend neues.

Was aber charakterisiert kundenorientierte Internetangebote wie Webshops, Social-Networks oder Business-Websites und welche Verbindung besteht zu E-Assessments? Das charakteristische eines online-Angebots ist, dass dieses zumeist freiwillig genutzt wird und zuvor vom Nutzer aus einer Vielzahl von Angeboten ausgewählt werden muss. Es herrscht also ein enormer Konkurrenzdruck zwischen den verschiedenen Angeboten um die Gewinnung von Nutzern und somit potentiellen

Kunden. Dies sorgt für eine Anpassung der „User-Interfaces“ an die Ansprüche der Nutzer und an technische Gegebenheiten um eine intuitive Benutzbarkeit gewährleisten zu können und eine positive „User-Experience“ zu erzeugen.

Dieser „Weg-des-geringsten-Widerstandes“ ist auch für E-Assessments essentiell, denn er beinhaltet nicht nur eine gute und instruktionsarme Benutzbarkeit der Assessment-Oberfläche, sondern beschreibt auch gleichzeitig den Verzicht auf Unnützes und die Fokussierung auf das Entscheidende eines jeden qualitativen Assessment - dessen Validität. Diese Validität ist nur dann gewährleistet, wenn es dem Geprüften ermöglicht wird, das Maximum seiner kognitiven Fähigkeiten tatsächlich für die Erbringung der beabsichtigten Prüfungsleistung zu nutzen und nicht unnötig Kapazitäten für das Verstehen und Bedienen der Assessment-Software investiert werden müssen. (vgl. Chandler & Sweller, 1991)

### **3.3.1 Intuition (Intuition)**

Unter „Intuition“ ist in diesem Zusammenhang die intuitive, also die gewohnte oder erwartete, Möglichkeit der Benutzung der Assessment-Software zu verstehen. Das bedeutet einerseits, dass E-Assessments generell nicht dazu geeignet sind an ihnen neue, ungewohnte Nutzungskonzepte in der Benutzer- oder Menüführung auszuprobieren, auch wenn diese scheinbar bestimmte Arbeitsabläufe vereinfachen oder verkürzen können. Durch die Unvertrautheit der Nutzer mit diesen Abläufen müssen zu viele geistige Kapazitäten für deren Verständnis und Speicherung aufgewendet werden, sodass die eigentliche Prüfungsleistung vermindert wird, was wiederum dem ursprünglichen Prüfungsgedanken im Rahmen dieser Arbeit entgegensteht. Andererseits ist es auch hier nach wie vor nötig die zukünftigen Nutzer der Assessment-Software mittels Instruktionen und Probedurchläufen mit dem System und seiner Bedienung vertraut zu machen, denn es können trotz aller Benutzerfreundlichkeit nicht alle individuellen Vorgehensweisen berücksichtigt werden. Ohne das an dieser Stelle auf konkrete Praxissituationen eingegangen werden soll, lässt sich generell sagen, dass innovative Assessments nicht durch ein ungewöhnliches Interface-Design innovativ werden, sondern durch ein durchdachtes Gesamtkonzept und somit durch ihre Validität.

### **3.3.2 Consistency (Beständigkeit)**

Ein weiterer Punkt, der im Rahmen dieser Arbeit als entscheidend für den Erfolg eines E-Assessments angesehen wird ist die „Consistency“ des Interface-Design, sowohl im „Mikro-“ als auch im „Makro-Umfeld“ einzelner Assessments.

Was ist mit „Consistency“ (übersetzt: Beständigkeit, Stetigkeit aber auch Einheitlichkeit) gemeint? Im Mikro-Umfeld, also sich auf ein einzelnes Assessment beziehend, verdeutlicht es, dass Funktionen, Menüs oder Abläufe, welche im Assessment einmal eingeführt wurden in ihrer Form und Funktionalität beständig sein sollten. Auch diese Erkenntnis ist Prinzipien kommerzieller Websites entlehnt und ermöglicht dem Nutzer eine gute und sichere Orientierung innerhalb des Systems. Dadurch verringert sich einerseits wieder die kognitive Last zur Nutzung der Software, andererseits wird dadurch auch ein „Vertrauensverhältnis“ zwischen Nutzer und Software aufgebaut, was zur Akzeptanz des Systems und zur Stressreduktion vor und während des Assessments beitragen kann. Für das Makro-Umfeld, also Assessment-übergreifend, lassen sich diese Erkenntnisse auch anwenden, denn ein Nutzer (Schüler, Student) wird in der Regel mehrere Assessments innerhalb eines Hochschulstudiums zu bewältigen haben und je einheitlicher diese Systeme aufgebaut und je vertrauter die Bedienungsabläufe sind, desto größer wird die Akzeptanz und Leistungsfähigkeit des Nutzers innerhalb eines bestimmten Systems werden. Diese vereinheitlichten Rahmenbedingungen können demnach auch zur Erhöhung der Validität der einzelnen Assessments führen.

### 3.3.3 Optimization (Optimierung)

Auch wenn sich Einheitlichkeit und Optimierung, im Sinne von Spezialisierung, eigentlich widersprechen, kann es im Bezug auf E-Assessments durchaus sinnvoll sein, beide Möglichkeiten bei der Konzeption zu berücksichtigen und später umzusetzen. Denn neben den beschriebenen Vereinheitlichungen gibt es auch Parameter die gezielt auf die praktischen Rahmenbedingungen der E-Assessments allgemein, beispielsweise fakultätsintern sowie für jedes Einzel-Assessment angepasst werden sollten. Gemeint sind dabei vor allem Anpassungen an vorhandene Hardware, Prüfungsräumlichkeiten oder andere allgemeingültige Prüfungsmodalitäten im Makro-Umfeld sowie spezielle Anpassungen im Mikro-Umfeld an individuelle Prüfungsanforderungen.

Verdeutlichen lässt sich die „Optimization“ am Besten in Form eines kurzen Beispiels:

Angenommen es stehen an einer Hochschule drei spezielle Hörsäle zu Verfügung, welche zur Durchführung der meisten E-Assessments genutzt werden und dementsprechend mit einer geeigneten Infrastruktur ausgestattet sind. Aus Kosten- und Platzgründen werden diese Räume allerdings auch für normale Lehrveranstaltungen genutzt, weshalb sich für Tablet-PCs entschieden wurde, die nur bedarfsweise verteilt werden können. Hochschulweit wird eine einheitliche Assessment-Software verwendet (Consistency), welche an diese Gegebenheiten angepasst ist (Optimierung im Makro-Umfeld). Diese Anpassungen



könnten sein: spezielle Optimierung der Software für die vergleichsweise geringe Auflösung der Tablet PCs, Verwendung kontrastreicher Farben zur Erleichterung der Lesbarkeit, Optimierung der Navigation für „Touch-Bedienung“, Nutzung der Kamerafunktionalität zum Einfügen handschriftlicher Aufgaben (z.B. Formeln, Rechnungen, etc.) und so weiter.

Diese Grundfunktionalitäten sind dann allen Nutzern vertraut und können, dank einheitlicher Verwendung aufwands- und kostenarm vermittelt werden. Trotzdem ist es sicherlich notwendig und deshalb auch sinnvoll gewisse Modalitäten für eine Mathematikprüfung anders zu gestalten als beispielsweise für ein Assessment im Bereich Programmierung, bei dem Unterlagen und Aufzeichnungen zur Prüfung genutzt werden dürfen („Open-Book“). Diese Anpassung im Mikro-Umfeld eines Assessments können dann zur Vereinfachung oder praxisnäheren Gestaltung beitragen, denn die Grundfunktionen sind bekannt und erprobt und werden lediglich geringfügig erweitert beziehungsweise vermindert.

### **3.4 Empirische Untersuchung - Instant-Feedback-Experiment**

Feedback wird im Rahmen dieser Arbeit als eminent wichtiger Bestandteil des effektiven Lernens angesehen und dementsprechend oft thematisiert. Aus Sicht des Autors ist ein Feedback dabei umso wirkungsvoller, je schneller es erfolgt und je individueller es auf den Lerner angepasst ist. Unter zeitlichen Gesichtspunkten nimmt somit das „Instant-Feedback“ eine Sonderform ein, denn es erfolgt unmittelbar auf eine Aktion des Lerners und ist demnach praktisch nur computergestützt generierbar. Die Anforderungen an die Feedback-Software steigen dabei mit zunehmender Individualität des Feedbacks an. Wie wirkungsvoll dieses Feedback in einer realen Lern- und Testsituation ist, wird im Rahmen dieser Arbeit anhand eines Online-Surveys mit 171 Probanden untersucht. Die detaillierten Versuchsergebnisse sind im Anhang aufgelistet.

#### **3.4.1 Versuchsbeschreibung und -durchführung des Experiments**

Ziel des Surveys ist es, die Wirkung von statischem „Instant-Feedback“ auf bestimmte Aktionen einer Testperson zu beobachten und das daraus resultierende (Lern)-Verhalten dieser Person zu dokumentieren. Um ein möglichst effektives Aufwand-Nutzen-Verhältnis zu erreichen wird dabei neben der Feedback-Wirkung auch die „Multi-Tasking-Fähigkeit“ der Testpersonen im Survey untersucht. Die Ergebnisse der „Multi-Tasking-Tests“ werden allerdings im Rahmen dieser Arbeit nicht betrachtet.

Die Aufgabe der Versuchsteilnehmer besteht darin, fünf zufällig ausgewählte Logikfragen, wie sie aus klassischen IQ-Tests bekannt sind, zu lösen. Es werden dabei jeweils drei Gruppen geometrischer Figuren gezeigt, welche durch eine weitere Gruppe logisch ergänzt werden sollen. Es stehen pro Frage fünf mögliche Lösungsvarianten zur Verfügung, von denen jeweils nur eine korrekt ist. Am Anfang des Experiments wird die Vorgehensweise anhand eines Beispiels erklärt.

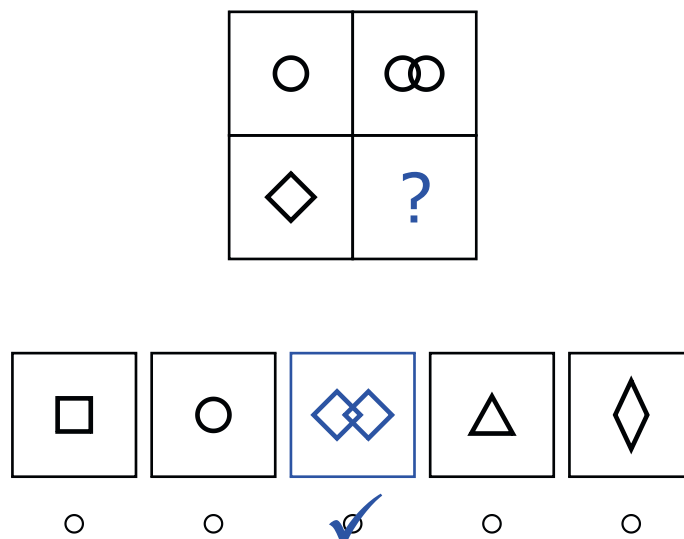


Abb. 12: Beispielfrage aus dem Experiment

Da, wie bereits eingangs beschrieben, auch die „Multi-Tasking-Fähigkeit“ der Probanden im gleichen Experiment untersucht werden soll, läuft im Hintergrund des Logik-Tests ein Kurzfilm ab, dessen Handlung und Details sich die Versuchsteilnehmer ebenfalls einprägen sollen. Vor Beginn des eigentlichen Tests werden die Probanden in zwei Gruppen eingeteilt - eine Feedback-Gruppe und eine Kontrollgruppe ohne Feedback. Diese Einteilung erfolgt zufällig und ohne das Wissen der Versuchsteilnehmer. Die Testpersonen der Feedback-Gruppe erhalten anschließend pro Frage 20 Sekunden Zeit, diese zu beantworten. Im Anschluss an jede Frage bekommen sie ein Feedback über die Richtigkeit ihrer Antwort und gegebenenfalls einen Lösungshinweis zur Frage, mit der Möglichkeit diese in gleicher Zeit anschließend noch einmal zu beantworten. Wird die Frage beim ersten Versuch direkt richtig beantwortet, folgt zeitgleich zum „RICHTIG!-Feedback“ die nächste Logikfrage. Jeder Teilnehmer der Feedback-Gruppe hat also die Möglichkeit eine falsch beantwortete Frage in einem zweiten Versuch mit Hilfe eines zusätzlichen Hinweises noch korrekt zu lösen. Ist dies nicht der Fall folgt auch hier die nächste Frage. Der Versuchsablauf für die Gruppe ohne Feedback gestaltet sich ähnlich, mit dem Unterschied, dass nach Beantwortung einer Frage kein Feedback zur Richtigkeit der Lösung oder ein Lösungshinweis erfolgt.

Um die Wirkung des statischen „Instant-Feedback“ auf die Lösungsfähigkeit der Probanden hinsichtlich der Testfragen zu untersuchen, sind jeweils die erste und vierte sowie die dritte und fünfte Frage des Tests identisch. Die jeweils abgegebenen Antworten der Versuchsteilnehmer und die Anzahl der benötigten Versuche (maximal zwei) zur Richtigbeantwortung der Logik-Fragen werden dabei gespeichert.

### 3.4.2 Versuchsergebnisse des Experiments

Am Online-Survey haben insgesamt 171 Testpersonen, verteilt auf einen vierwöchigen Zeitraum, teilgenommen. Dabei konnten 121 gültige Versuchsdurchläufe gespeichert werden, welche sich wiederum auf 50 Feedback und 71 Non-Feedback-Probanden aufteilen. Um den Effekt des Feedback bewerten zu können, ist es notwendig zunächst die Kontrollgruppe zu betrachten und ihr Abschneiden bei den Testfragen zu untersuchen.

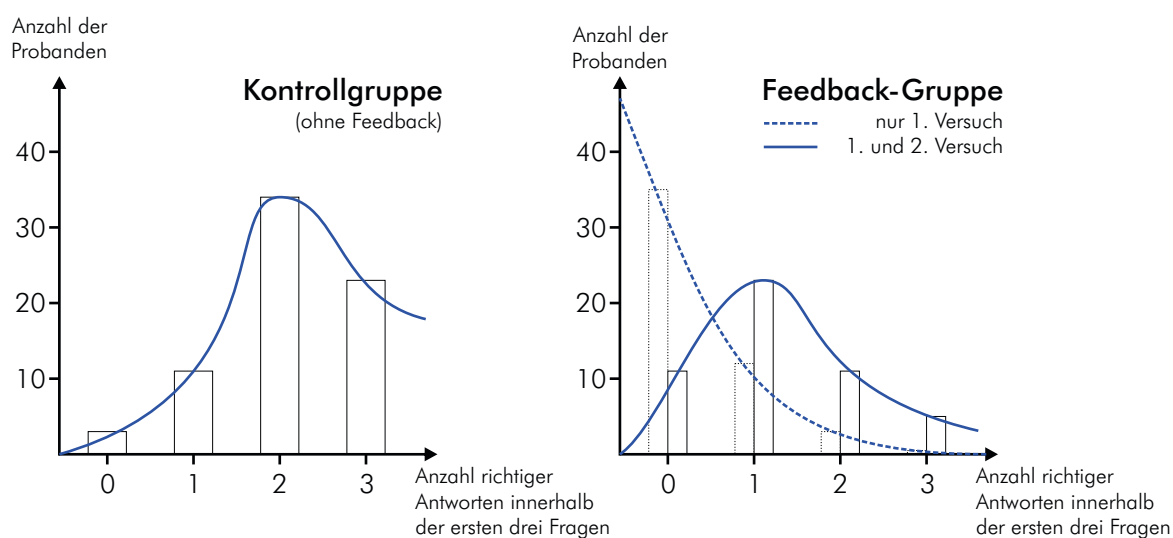


Abb. 13: Antwortverteilung der Kontroll- und Feedback-Gruppe im Instant-Feedback-Experiment

Betrachtet man das linke, obige Diagramm und die zugehörige Funktionskurve, fällt zunächst auf, dass die Anzahl der richtigen Antworten unter den drei ersten Fragen, pro Teilnehmer, weitgehend normalverteilt ist. Das heißt, die Mehrheit der Versuchspersonen hat den Test durchschnittlich gut absolviert und im Mittel 2,08 Fragen richtig beantwortet. Dies entspricht den erwarteten Ergebnissen und ist ein guter Ausgangspunkt für alle folgenden Betrachtungen. Es wurden dabei vorerst nur die ersten drei Fragen betrachtet, weil Frage vier und fünf jeweils identisch mit Frage eins beziehungsweise drei sind.

Weiter lässt sich zum Ergebnis der Kontrollgruppe sagen, dass 31 Probanden mindestens eine der Doppelfragen erst falsch und beim zweiten Auftreten dann richtig beantwortet haben ohne zwischendurch ein Feedback zu ihrer Lösung zu erhalten. Dieses Phänomen tritt allerdings in 24 der genannten 31 Fälle im Zusammenhang mit Frage eins und vier auf, wohingegen bei Frage drei und fünf kein spürbarer Anstieg der Richtigbeantwortung festzustellen ist. Es ist daher davon auszugehen, dass die Versuchsteilnehmer die erste Frage benötigt haben um das Prinzip des Tests zu verstehen und deshalb hier so oft falsch geantwortet haben. Bei einem zweiten Auftreten der Frage wurde sich also nicht an die vorherige Lösung erinnert, sondern an dieser Stelle war einfach das prinzipielle Vorgehen beim Lösen solcher Aufgaben verstanden. Diese Annahme wird auch durch die „Rückfallquote“ bestätigt, welche beschreibt wie viele Fragen nach eingangs erfolgreicher Lösung beim zweiten Auftreten trotzdem falsch beantwortet wurden. Hier liegt die Anzahl bei sieben, was exakt der Differenz aus 31 und 24 entspricht, was wiederum der bereinigten Anzahl der Richtigbeantwortungen beim zweiten Auftreten einer ehemals falsch beantworteten Doppelfrage gleichkommt. Zusammenfassend betrachtet bedeutet das, dass die Wahrscheinlichkeit der Falschbeantwortung einer Doppelfrage, innerhalb der Kontrollgruppe, beim erstmaligen Auftreten gleichhoch ist wie bei zweiten Mal. Es erfolgt also keine Verbesserung der Lösungsfähigkeit der Probanden hinsichtlich der Logik-Fragen innerhalb des Tests.

Anders verhalten sich die Ergebnisse der Feedback-Gruppe. Hier beginnen die allgemeinen Testergebnisse der Versuchspersonen wesentlich schlechter, denn gerade einmal drei der 50 Probanden konnten zwei Fragen innerhalb des Tests beim ersten Versuch richtig beantworten. Dabei ist zu beachten, dass pro Frage jeweils maximal zwei Versuche möglich sind und die Testpersonen nach jedem Versuch ein Feedback erhalten. Unabhängig davon sind wieder die Fragen eins und vier sowie drei und fünf identisch. Die Mehrheit der Probanden absolvierte also den Test mit durchschnittlich weniger als einer richtigen Antwort, betrachtet man jeweils nur den ersten Versuch (Durchschnitt: 0,36). Dafür konnten viele Fragen mit Hilfe des zweiten Versuchs erfolgreich gelöst werden, was sich einerseits auf das Feedback andererseits natürlich auch auf die quasi Verdopplung der Zeit zurückführen lässt. Bezieht man den zweiten Versuch in die Auswertung mit ein, hat jeder Teilnehmer im Schnitt 1,2 Fragen korrekt beantwortet, was einen erheblichen Anstieg im Vergleich zum ersten Versuch bedeutet.

Neben dem direkten Einfluss des „Instant-Feedback“ auf die Lösungsfähigkeit einer Frage im ersten beziehungsweise zweiten Versuch, sollte aber auch untersucht werden, wie die Probanden beim erneuten Auftreten einer bekannten Frage reagieren. Dabei fiel auf, dass 15 Doppelfragen,

welche beim ersten Auftreten trotz Feedback gar nicht gelöst werden konnten, beim erneuten Auftreten sofort richtig beantwortet wurden. Doppelfragen, die bereits beim zweiten Versuch gelöst wurden, konnten beim erneuten Auftreten mit noch höherer Wahrscheinlichkeit richtig beantwortet werden. Hier lag die Anzahl der Fragen bei 20 Stück. Die „Rückfallquote“ innerhalb der Feedback-Gruppe betrug allerdings 10%, das heißt 10 der 100 relevanten Fragen (eine Frage je Fragenpaar x Teilnehmerzahl) wurden nach erfolgreichem Lösen beim zweiten Auftreten trotzdem wieder falsch beantwortet. Eine Sonderstellung der ersten Frage, ähnlich wie bei der Kontrollgruppe, konnte hier nicht festgestellt werden, was auf die Möglichkeit des „2. Versuchs“ und die damit verbundene Zeitverdopplung zurückzuführen ist.

### 3.4.3 Auswertung des Experiments

Vergleicht man die Daten beider Gruppen miteinander, fällt auf, dass die Feedback-Gruppe jeweils im zweiten Versuch deutlich besser abschnitt als im Ersten, wohingegen bei der Kontrollgruppe keine Veränderung zu verzeichnen war. Dies kann aber nur zum Teil auf das Feedback zurückgeführt werden, denn einerseits spielt hier die Verdopplung der Zeit eine entscheidende Rolle, andererseits verringern sich durch den Wegfall einer falschen Lösung die Auswahlmöglichkeiten um 20%, was wiederum die Wahrscheinlichkeit, die richtige Lösung zu erraten, erhöht.

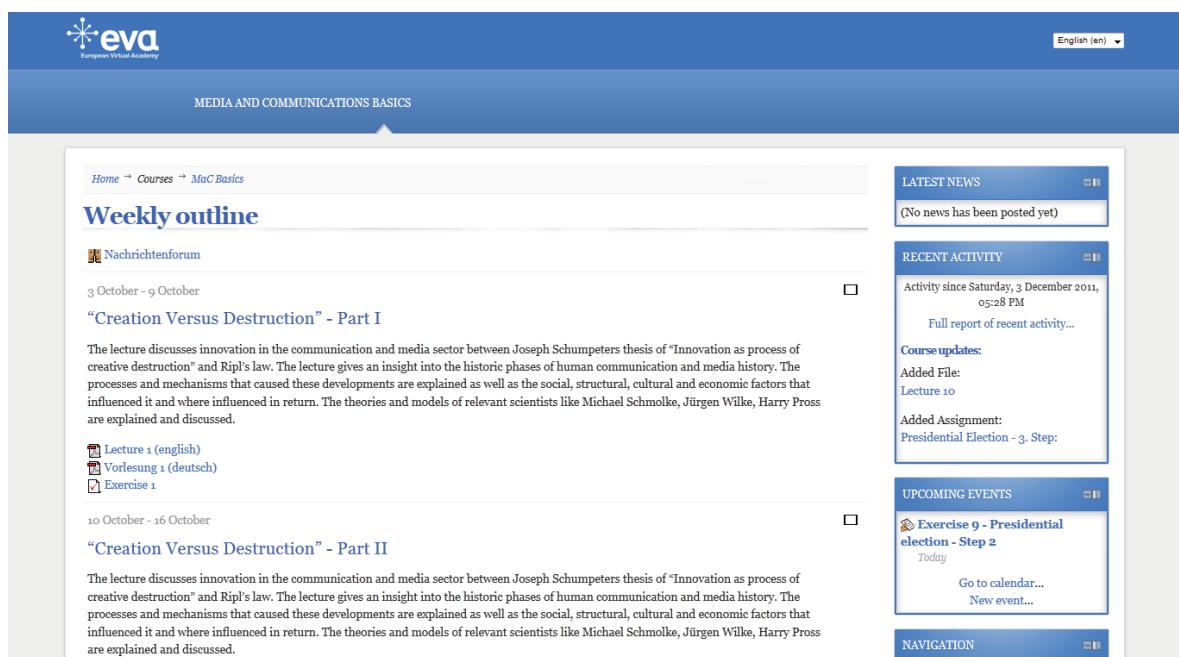
Kritisch müssen auch die „Rückfallquote“ und das generell schlechtere Abschneiden der Feedback-Gruppe betrachtet werden, denn diese Daten sollten sich erwartungsgemäß nicht so gravierend von denen der Kontrollgruppe unterscheiden. Es ist deshalb zu vermuten, dass hier die Rahmenbedingungen die Ergebnisse des Versuchs beeinträchtigt haben. Als mögliche Ursache kann dabei der Hintergrundfilm gesehen werden, welcher bei der Feedback-Gruppe länger als bei der Kontrollgruppe ist. Bei einer unzureichend schnellen Internetverbindung kann es so zu wiederholten Unterbrechungen des Video-Streamings kommen, was die Aufmerksamkeit der Probanden erheblich stören kann. Auch ist es denkbar, dass dadurch einige Testfragen nur langsam geladen und somit erst verspätet angezeigt wurden.

Viele der eingangs beschriebenen und erwarteten positiven Effekte des „Instant-Feedback“ konnten im Experiment nicht nachgewiesen werden. Dies stellt aber nicht das generelle Potential von „Instant-Feedback“ oder Feedback im Allgemeinen in Frage, sondern verdeutlicht nur, dass die eingesetzte Versuchsumgebung und der Versuchsablauf nicht ideal gewählt wurden. Dabei ist vor allem das bereits beschriebene Hintergrundvideo als kritisches Element zu sehen.

## 4 Moodle 2.0 - Assessment-Möglichkeiten und -Grenzen der populären Lernplattform

### 4.1 Kurzbeschreibung der Lernplattform Moodle und deren Notwendigkeit im E-Learning

Die Open-Source Software „Moodle“ ist eine online-orientierte Lernplattform, mit deren Hilfe sich bekannte physische und organisatorische Strukturen aus dem klassischen Lernalltag virtualisieren lassen. Man kann „Moodle“ deshalb auch als eine virtuelle Lernumgebung bezeichnen, welche die notwendigen Rahmenbedingungen für E-Learning- und E-Assessment-Aktivitäten bereitstellt. Dazu lassen sich einerseits verschiedene Kursbereiche mit jeweils mehreren thematisch zusammengehörigen Kursen anlegen innerhalb dieser dann spezifische Arbeitsmaterialien und Aktivitäten angeboten werden. Andererseits lassen sich auch personelle Strukturen abbilden, wie zum Beispiel: Administratoren, Dozenten, Coaches, Studenten, Gasthörer, etc., damit ein realistischer Studienablauf gewährleistet ist. Auch sonstige Aktivitäten aus dem klassischen Studienalltag, wie Einschreibungen, termingebundene Aktivitäten, Bewertungen, Chats und Foren, etc. sind realisierbar. Eine Lernplattform wie „Moodle“ ist also ein elementares Konstrukt, ähnlich den Gebäuden, Hörsälen und Personalstrukturen einer Hochschule, für ein virtuelles, computergestütztes Lernen. (Moodle, 2011)



The screenshot displays the Moodle interface for a course titled "MEDIA AND COMMUNICATIONS BASICS". The main content area shows a "Weekly outline" with two entries:

- 3 October - 9 October**: "Creation Versus Destruction" - Part I. The description states: "The lecture discusses innovation in the communication and media sector between Joseph Schumpeters thesis of 'Innovation as process of creative destruction' and Ripl's law. The lecture gives an insight into the historic phases of human communication and media history. The processes and mechanisms that caused these developments are explained as well as the social, structural, cultural and economic factors that influenced it and where influenced in return. The theories and models of relevant scientists like Michael Schmolke, Jürgen Wilke, Harry Pross are explained and discussed." Below this entry are links for "Lecture 1 (english)", "Vorlesung 1 (deutsch)", and "Exercise 1".
- 10 October - 16 October**: "Creation Versus Destruction" - Part II. The description is identical to the first entry.

The right sidebar contains several widgets:

- LATEST NEWS**: (No news has been posted yet)
- RECENT ACTIVITY**: Activity since Saturday, 3 December 2011, 05:28 PM. Full report of recent activity... Course updates: Added File: Lecture 10. Added Assignment: Presidential Election - 3. Step.
- UPCOMING EVENTS**: Exercise 9 - Presidential election - Step 2. Today. Go to calendar... New event...
- NAVIGATION**

Abb. 14: Beispiel-Benutzeroberfläche der Lernplattform „Moodle“

## 4.2 Vorhandene Testarten und Erweiterungsmöglichkeiten in Moodle im Vergleich mit anderen Lernplattformen

Standardmäßig stehen in „Moodle“ vier verschiedene Grundtestarten zur Verfügung: Tests, Aufgaben, Lektionen und Workshops (MoodleDocs, 2011). Unterschieden werden diese Testarten anhand ihrer Anforderungen hinsichtlich der Auswertbarkeit. So werden unter der Aktivität „Tests“ alle Assessments zusammengefasst, welche sich, bis auf die Ausnahme „Freitextfrage“, automatisch im System auswerten lassen und dem Lerner sofort ein Feedback seiner Leistungen liefern. Es sind dabei folgende Fragenformate vorhanden: Berechnet, Berechnet Multiple Choice, Einfachberechnet, Freitext, Kurzantwort, Lückentextfrage, Multiple-Choice, Numerisch, Wahr/Falsch, Zufällige Kurzantwort Zuordnung, Zuordnung (MoodleDocs, 2011). Alle Fragenformate haben, unabhängig ihrer genauen Funktionsweise, gemeinsam, dass sich nur vordefinierte Antwortmöglichkeiten wirklich sinnvoll auswerten lassen und mit einem spezifischen Feedback versehen werden können. Diese Form der Auswertung lässt sich natürlich softwareseitig am einfachsten umsetzen, stößt aber beispielsweise beim Frageformat „Kurzantwort“ schnell an ihre Grenzen, wo der User eine Frage in Form einer kurzen Texteingabe beantworten soll.


Verdeutlichen lässt sich dies am besten anhand eines einfachen Beispiels:

In einem „Moodle“-Assessment zum Thema Geografie soll der Teilnehmer folgende Frage beantworten: „Wie heißt die Hauptstadt des asiatischen Staates Brunei?“. Die richtige Antwort „Bandar Seri Begawan“ ist dabei im System hinterlegt und mit einem kurzen Feedback zur Bedeutung des Namens versehen um das Wissen des Lerners zu ergänzen und zu festigen. Außerdem sind die häufigsten Falschantworten auf diese Frage „Kuala Lumpur“ und „Jakarta“ hinterlegt und geben dem Teilnehmer das Feedback, dass es sich hierbei um die Hauptstädte benachbarter Staaten handelt, nicht aber um die Bruneis. Was sich aber nur schwer abfangen lässt sind Rechtschreib-, beziehungsweise Tippfehler, welche in einer Prüfungssituation zweifelsohne häufig passieren werden. Die Antwort „Bander Seri Begawan“ würde vom System als falsch bewertet, wobei die Stadt eigentlich eindeutig richtig vom User erkannt wurde. Auch ein kleiner Tippfehler, wie „Bandar Seri Begawam“, der nicht sofort auffällt, führt zu einer Fehlbewertung, denn hier kommt es im Wesentlichen ganz klar auf die richtige Zuordnung der Stadt und nicht auf die einhundertprozentig korrekte Schreibweise an.

1 Wie heißt die Hauptstadt des asiatischen Staates Brunei?

Punkte: 1,00/1,00

Antwort:  ✓

 Das ist richtig!

**Bandar Seri Begawan** ist malaiisch und bedeutet: *Hafen des verehrten Herrschers*. Die Hauptstadt von Brunei hat etwa 33.000 Einwohner und liegt am Brunei River, der unweit in das Südchinesische Meer mündet. 2007 wurde die Stadtfläche von ursprünglich 12,87 km<sup>2</sup> auf ihre heutigen 100,36 km<sup>2</sup> vergrößert.

Richtig  
Punkte: 1,00/1,00.

Abb. 15: Kurzantwort-Frage in „Moodle“

Anhand dieses Beispiels lässt sich verdeutlichen, dass oft ähnliche Kleinigkeiten, die aus der programmierorientierten Entwicklung der Frageformate stammen, problematisch für die Validität des E-Assessment werden können und so häufig didaktische Möglichkeiten ungenutzt bleiben, da ihr Einsatz nicht fehlerresistent genug ist. Würde „Moodle“ die Eingaben und Lösungen auch phonetisch miteinander vergleichen könnten tipp- und schreibfehlerbehaftete Antworten trotzdem korrekt ausgewertet werden.

Eine weitere Grundtestart in „Moodle“ stellt die „Aufgabe“ dar, welche nicht automatisch ausgewertet, sondern dem Trainer oder Dozent zur Korrektur vorgelegt wird. Es kann dabei zwischen drei Formaten gewählt werden, die es erstens ermöglichen eine Aufgabe direkt im System zu beantworten und abzuschicken oder zweitens, eine oder mehrere Dateien hochzuladen und drittens, die Aufgabe unabhängig von „Moodle“, per E-Mail, Post oder persönlich in einem Face-to-Face-Meeting abzugeben. (MoodleDocs, 2011) Die Vor- und Nachteile dieser Interaktionsart sind bekannt und lassen sich für jeden konkreten Fall leicht gegeneinander abwägen um festzustellen, ob diese Testart geeignet ist oder nicht.

Eine entscheidende Gemeinsamkeit der ersten beiden Grundtestarten ist die gesonderte Präsentation der Assessment-Elemente vom eigentlichen „learning-content“. Im Gegensatz dazu steht die Grundtestart „Lektion“, welche die Content-Vermittlung und -Überprüfung miteinander in einer Aktivität kombiniert und eine individuellere Anpassung an den Lerner ermöglicht. Der Inhalt wird dabei seitenweise präsentiert und mit Verzweigungen verknüpft. Diese Verzweigungen führen den Lerner, entweder anhand seiner eigenen Auswahl oder aufgrund eines automatisch ausgewerteten Testergebnisses, zu einer neuen Inhalts- oder Testseite. Testelemente können mit Feedbackinformationen, innerhalb der eingangs beschriebenen Restriktionen, versehen werden und ermöglichen eine zusätzliche Verbesserung der Selbsteinschätzung.



Die letzte Grundtestart, welche „Moodle“ standardmäßig bereitstellt, ist der sogenannte „Workshop“. Er unterteilt sich in vier Phasen: Vorbereitungs-, Einreichungs-, Beurteilungs- sowie Bewertungsphase, und ermöglicht neben einer besseren Eigenbewertung seiner Leistungen ein Community- oder Peer-Feedback (MoodleDocs, 2011). Dabei können alle Workshop-Teilnehmer, nach einer Selbsteinschätzung auch ausgewählte (zufällig oder manuell) Fremdleistungen anderer Lerner bewerten. Diese Erkenntnisse der Beurteilungsphase sollen später dem Trainer oder Dozent in der Bewertungsphase behilflich sein, eine objektivere Einschätzung zu finden und die Meinung anderer Teilnehmer dabei zu berücksichtigen.

An dieser Stelle wird bereits deutlich, dass das Standardangebot an Testarten und Fragetypen in „Moodle“ stark begrenzt und auf relativ triviale Anwendungsbereiche beschränkt ist. Andere Lernplattformen wie „ILIAS“ oder „Blackboard“ stellen da ein komplexeres Angebot zur Verfügung. So bieten beide Alternativen zum Beispiel sogenannte „Hot Spot Questions“ oder „Image-Map Fragen“, bei denen der User eine bestimmte Stelle in einer Grafik, oder einem Foto markieren muss um die Aufgabe zu lösen. Außerdem ermöglicht es „ILIAS“ dem Assessment-Autor durch spezielle „Flash-“ und „Java-Applet-“ Fragen nahezu unbegrenzte Aufgabenformate zu generieren, welche, wenn sie eine eigenständige Bewertung der Nutzereingaben vornehmen, ihre Ergebnisse über eine Schnittstelle zurück an „ILIAS“ übergeben können um direkt in die automatische Bewertung einzufließen. (IliasDocs, 2011) Ähnliche Möglichkeiten scheinen auch in der Lernplattform „Blackboard“ vorhanden zu sein, deren genaue Funktionalität konnte allerdings im Rahmen dieser Arbeit, aufgrund fehlender Demozugänge, nicht untersucht werden (Blackboard, 2011). In „Moodle“ lässt sich ein Flash-Assessment nur mittels sogenannter SCORM-Lernpakete anlegen, welche mit einem externen Autorentool erstellt und dann als Aktivität in „Moodle“ importiert werden können (MoodleDocs, 2011). Dabei ist allerdings zu beachten, dass der SCORM-Standard nicht so streng spezifiziert ist, wie beispielsweise DIN-Standards und somit nicht alle SCORM-Anwendungen externer Editoren zuverlässig mit „Moodle“ kommunizieren. Abgesehen von diesen Unterschieden, in einigen spezifischen Anwendungsbereichen, sind allerdings die Fragenformate und Assessment-Bausteine in den meisten Lernplattformen weitgehend gleich und zumindest Standardfragen, wie Multiple Choice, Kurztextfragen, Berechnungen, Wahr/Falsch oder Zuordnungen sind eigentlich in allen Systemen vorhanden. Open-Source Lösungen wie „Moodle“ oder „ILIAS“ bieten zusätzlich den Vorteil, zahlreicher vorhandener Erweiterungsmöglichkeiten, welche von einer großen Online-Community oder kommerziellen Anbietern ständig entwickelt beziehungsweise weiterentwickelt werden. So lässt sich zum Beispiel „Moodle“ mit den E-Portfolio-Plattformen „Mahara“ oder „Chalk &

Wire“ (Mahara, 2011; Chalk & Wire, 2011) verbinden oder um die Team-Meeting Software „OpenMeetings“ erweitern, um die vorhandenen Assessment-Funktionalitäten auszubauen. Der notwendige Implementierungsvorgang ist allerdings meist anspruchsvoll und die Benutzung der Tools gewöhnungsbedürftig, da, im Gegensatz zu Erweiterungen kommerzieller Produkte, die einheitliche Handhabung fehlt. Auch sind nicht bei allen verfügbaren Erweiterungen die gleichen Sicherheitsstandards vorhanden, was gerade bei einem Einsatz in summativen Assessments natürlich wünschenswert wäre.

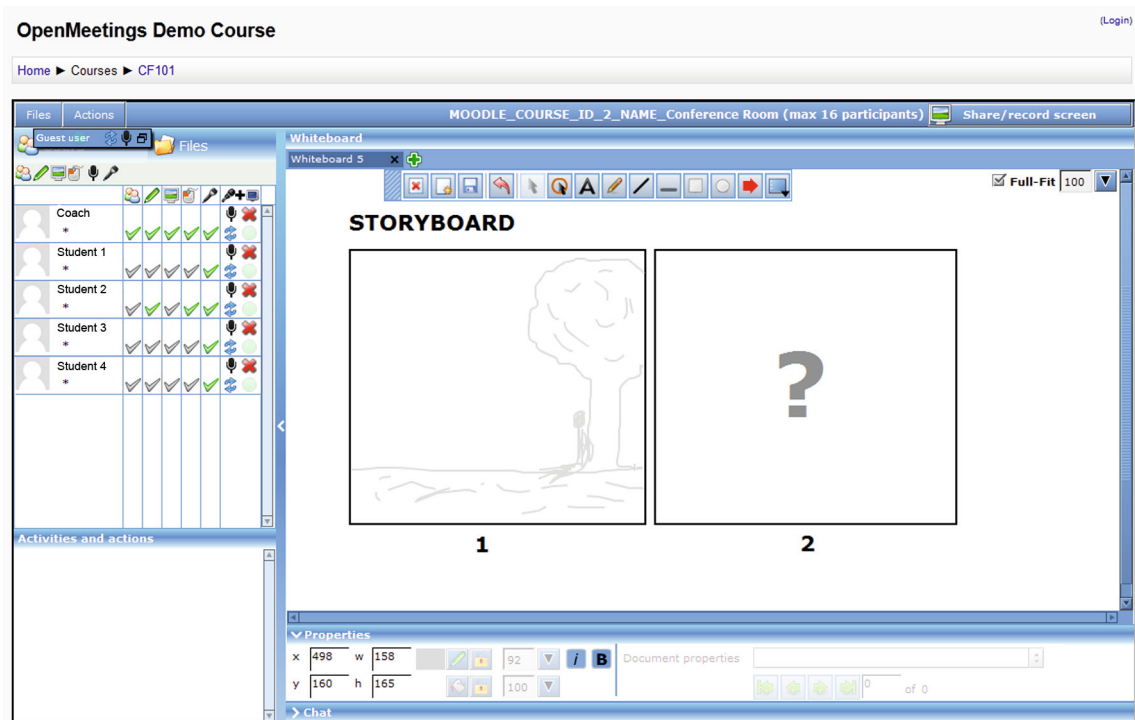


Abb. 16: OpenMeetings-Plugin in „Moodle“

Für welchen Typ Lernplattform und welches konkrete Produkt man sich also letztendlich entscheidet, hängt stark von den spezifischen Anforderungen ab. Bezogen auf Assessments sollte man dabei auf die individuelle Anpassbarkeit vorhandener Bausteine achten, beziehungsweise auf die Möglichkeit der Implementierung eigener Module, denn die standardmäßig vorhandenen Tests und Quizes gängiger Lernplattformen sind für anspruchsvolle Assessment-Szenarien in der Regel nicht geeignet.

### 4.3 Möglichkeiten der Umsetzung entwickelter E-Assessment-Ansätze in Moodle

In den vorangegangenen Kapiteln wurde bereits erläutert, dass sich ein zeitgemäßes Assessment aus heutiger Sicht nicht unbedingt durch die Verwendung bestimmter Frageformate oder Aufgabentypen auszeichnet, sondern vor allem der sinnvolle kombinatorische Einsatz von „Learning-Content“ und jeweils korrespondierenden Assessments entscheidend ist. Mit der Komplexität der Kurs- oder Lernziele steigen also auch die Anforderungen an die summativen, formativen und diagnostischen Assessments. Möchte man alle aufgezeigten Einsatzmöglichkeiten von E-Assessments nutzen und jedem Lerner einen individuellen Lernprozess ermöglichen, ist die Lernplattform „Moodle“ dafür eigentlich nicht geeignet.

Möchte man hingegen nur klassische lineare Unterrichtsstrukturen und konventionelle Prüfungen, welche vornehmlich lexikalisches Wissen prüfen, in eine elektronische Form bringen, ist „Moodle“ durchaus ein geeignetes System. Vor allem die automatische Auswertbarkeit der meisten Fragenformate und die Schnelligkeit, mit der einfachste Kursstrukturen und Tests erzeugt werden können, bringen dann eine Zeit- und Aufwandsersparnis gegenüber Paper-and-Pencil Prüfungen für die Dozenten. Eine hinreichend sichere Testumgebung kann mit Hilfe eines Secure-Browser, wie zum Beispiel dem „Save Exam Browser“ oder dem „Student Test Viewer“ erzeugt werden, um andere Ressourcen außerhalb „Moodles“ auf den Testcomputern zu sperren. Archiviert werden können die Prüfungsfragen und zugehörigen Antworten mittels eines Ausdrucks nach Prüfungsende, deren Richtigkeit der Geprüfte noch durch eine Unterschrift bestätigt. Auch für den Lerner bringt diese Form des Assessments, oder besser der elektronischen Prüfung bereits einige Vorteile. So erhält er erstens direkt nach Prüfungsende, oder bereits während der Prüfung ein Feedback bezüglich seiner Leistungen und eines etwaigen Bestehens oder Nichtbestehens. Andererseits können durch den Computereinsatz auch andersartige Aufgabenformate, beispielsweise mit Multimediainhalten oder Internetzugriff realisiert werden, was in einer klassischen Papierprüfung nur umständlich möglich wäre.

Dem entgegen stehen allerdings auch negative Aspekte, so zum Beispiel auf formaler Ebene, die eher gewöhnungsbedürftige Benutzerführung und die teilweise unübersichtliche Inhaltsanordnung. So sind zum Beispiel „Moodle“-Seiten generell scroll-orientiert aufgebaut, was allgemein als kritisch eingeschätzt wird. Auch fehlen geeignete Übersichtselemente, welche eine schnelle Navigation zwischen den einzelnen Aufgaben und ein markieren bereits gelöster oder teilgelöster Aufgaben ermöglichen, was gerade in Prüfungssituationen sehr wichtig ist.

Obwohl also „Moodle“ eigentlich nicht für einen Assessment-Einsatz, unter im Rahmen dieser Arbeit als zeitgemäß erachteten Bedingungen, in der Hochschullehre geeignet ist, können aus Sicht des Autors dieser Arbeit trotzdem moderne E-Assessment-Szenarien mit „Moodle“ erstellt werden. Es sollten allerdings ausreichend gewichtige Gründe für eine Benutzung von „Moodle“ unter anderen Gesichtspunkten vorliegen, denn sonst ist die Verwendung einer alternativen Software durchaus vorzuziehen. Geeignete Voraussetzungen, die „Moodle“ für zeitgemäße Assessments liefert sind zum Beispiel:

- die Möglichkeit Aufgaben zu terminieren,
- zahlreiche Community-Aktivitäten, wie Foren, Text-, Voice- und Videochats,
- eine Kombination aus Instant-Feedback- und Manuell-Feedback-Aufgaben (semi-automatisches Feedback)
- sowie das bereits beschriebene Aufgabenformat „Lektionen“.

Mit Hilfe einer „Lektion“ können verschiedene Lern- und Testelemente in einer Aktivität verbunden werden und ermöglichen so eine individuelle „Learning-Content“-Präsentation.

So lassen sich jenseits der eigentlichen „Moodle“-Testelemente, mit Hilfe von ausreichend komplex gestalteten Aufgaben, problemlösungsorientierte formative und summative Assessments erstellen. Diese können auch in Gruppen gelöst werden. Am Anfang eines Kurses oder einer Lerneinheit wird das zu bearbeitende Projekt in „Moodle“ vorgestellt, welches später gelöst werden soll. Dabei kommen Lektionen zum Einsatz, die relevante Lerninhalte und formative Assessments zur Lernpfadbestimmung miteinander verbinden. Erarbeitete Ergebnisse lassen sich dann über Chats und Foren mit Teammitgliedern diskutieren und in eine finale Form bringen, welche dann in einer virtuellen Präsenzveranstaltung anderen Kursteilnehmern vorgestellt oder per Mail oder Upload abgegeben werden kann.

Das direkte automatische Feedback auf Assessment-Leistungen innerhalb der Lektionen ermöglicht es dem Lerner eigene Stärken und noch vorhandene Schwächen einerseits überhaupt frühzeitig zu erkennen und andererseits durch die Individualisierung des Lernprozesses zu fördern oder auszugleichen. Der Dozent, Trainer oder Coach kann dabei die Lernpfade und Navigationsmuster überwachen und in besonderen Fällen, welche vom System unzureichend behandelt wurden manuell eingreifen und zusätzlichen Content, Extra-Assessments oder Hilfestellung zur Verfügung stellen. Diese Vorgehensweise erhöht zwar den Workload der Lehrenden, führt aber gleichzeitig

zu einer Verbesserung der Lernsituation und ist insbesondere in „Moodle“ nötig, da das System selbst nicht über ausreichend Intelligente Strukturen verfügt um den tatsächlichen Lernfortschritt des Lerners bewerten zu können.

Die bereits angesprochene Schnittstelle zwischen „Moodle“ und „Mahara“ ermöglicht außerdem ein Erstellen von E-Portfolios aus „Moodle“ heraus. Auch wenn dies, durch die Notwendigkeit der Nutzung eines zweiten Programmes nicht ideales ist, lässt sich dadurch die Funktionalität von „Moodle“ entscheidend erweitern. So können Arbeitsleistungen und Lösungswege dokumentiert und alternativ zu klassischen Prüfungssituationen bewertet werden. Auch die erwähnte Erweiterung „OpenMeetings“, welche neben einem Text-, Voice und Videochat ein gemeinsames Whiteboard zur Verfügung stellt, kann als Assessment-Modul genutzt werden. Vor allem kreative Aufgaben lassen sich so live bearbeiten, Lösungswege gemeinsam erörtern und anschließend von einem beteiligten Coach und allen anderen Teilnehmern bewerten.

Zusammenfassend lässt sich beurteilen, dass „Moodle“ zwar kein optimales Mittel zu Durchführung zeitgemäßer E-Assessment-Szenarien ist, andere Systeme aber auch nur bedingt besser sind. Wenn „Moodle“ trotzdem im Hochschulbereich zum Einsatz kommen soll, kann es vor allem durch seine zahlreichen Erweiterungsmöglichkeiten punkten. Für einen Einsatz als innovative Lernumgebung muss es allerdings erst aufwändig angepasst und umgestaltet werden. Auch dann ist „Moodle“, primär nur als Organisationsrahmen und zur Präsentation der Lerninhalte und Assessments geeignet. Die eigentliche Steuerung der Inhalte und die Bewertung der Lernleistung und des Lernerfolgs müssen manuell, durch einen Dozent, Trainer oder Coach erfolgen, da hier die entsprechend notwendige „künstliche Intelligenz“ softwareseitig nicht vorhanden ist.

## 5 Zusammenfassung und Ausblick

An Assessments im Hochschulbereich werden heute hohe Anforderungen gestellt. Diese liegen einerseits in europäischen Vorgaben begründet, welche im Rahmen des Bologna-Prozesses festgelegt wurden und sich beispielweise im Rahmen der Umstellung auf Bachelor-/Masterstudiengänge in einer Forderung nach Kompetenzorientierung und der Einführung eines Creditsystems äußern. Andererseits stehen sich aber auch die Forderung nach einer individuelleren Ausrichtung der Lehre auf den Einzelnen und gleichzeitig steigende Studierendenzahlen sowie ein größerer Bedarf an Leistungsnachweisen bei einer weitgehend gleichbleibender Anzahl an Lehrkräften gegenüber. Eine wieder aufflammende Bildungsdiskussion war die Folge, welche nicht selten in einer überspitzten Debatte um den Einsatz von E-Assessments, entweder als „Allheilmittel“, oder als „Todesstoß“ deutscher Bildungseinrichtungen gipfelte.

In der hier vorliegenden Arbeit wurden deshalb zunächst die bisher gängige Prüfungspraxis und die am häufigsten dabei zum Einsatz kommenden Prüfungsformen an deutschen Hochschulen untersucht und hinsichtlich ihrer Fähigkeiten betrachtet, heutigen Anforderungen gerecht werden zu können. Dabei wurde festgestellt, dass es hinsichtlich der Validität, Objektivität und Reliabilität klassischer Prüfungsformen und dem zu bewältigenden Mehrbedarf an Leistungsnachweisen entscheidende Defizite gibt und konventionelle Verfahren nur wenig geeignet sind zuverlässig Kompetenzen zu prüfen. Dass aber trotzdem ein vorschnelles Handeln bei der Einführung von E-Assessments als summativer Prüfungsersatz nicht zu empfehlen ist, wurde daraufhin ebenfalls dargelegt. Viele E-Assessment- oder E-Learning-Angebote im Allgemeinen sind didaktisch zu unausgereift und mehr auf eine Rationalisierung des Lern- und Prüfungsprozesses bedacht, als eine Modernisierung und Qualitätssteigerung der Lehre. Bei einer einfachen Übertragung von bekannten Vorgehensweisen in eine elektronische Assessments ist keine Verbesserung der Kompetenzerfassung zu erwarten.

In den darauf folgenden Kapiteln wurde beschrieben, wie unter Beachtung verschiedener Lerntheorien, innovative Lösungsansätze zum Einsatz von zeitgemäßen E-Assessments im Bereich der Hochschullehre aussehen könnten, und welche Anforderungen diese aus Sicht des Autors erfüllen müssen, um erfolgreich zu sein. Dabei werden Assessments einerseits als möglicher Ersatz summativer Prüfungen angesehen, andererseits aber auch ihr Einsatzpotential zur Lernlenkung, Motivationssteigerung und Erschaffung eines schülerzentrierten Unterrichts beschrieben. Dabei wird vor allem deutlich, dass der sinnvolle Einsatz von Assessments an den richtigen Stellen im

Lernprozess und deren konsequente Verknüpfung mit den Lerninhalten entscheidend für die Assessment- und Lernqualität ist.

Abschließend wurde anhand der Lernplattform „Moodle“ beispielhaft untersucht, inwiefern sich die entwickelten Lösungsansätze bereits heute umsetzen lassen und wie sich der vorhandene Workload dabei in der Praxis für die Lehrenden und Lernenden ändert. Dabei wurde ebenfalls deutlich, dass gerade bestehende Lernplattformen noch nicht die notwendigen Mittel bereitstellen um mit vertretbarem Aufwand wirklich zeitgemäße E-Learning/E-Assessment-Kombinationen zu erstellen. Die Assessment-Fähigkeiten heutiger Lernplattformen orientieren sich oft noch zu sehr an der Überprüfung lexikalischen Wissens unter der Verwendung trivialer Testformen, welche sich „einfach“ erstellen und computergestützt auswerten lassen.

Damit zukünftige E-Assessments ein fester Bestandteil, der bereits heute akzeptierten E-Learning-Kultur werden und klassische „Paper-and-Pencil-Prüfungen“ qualitativ höherwertiger ersetzen können müssen sowohl auf technischer, als auch auf rechtlicher Seite noch zahlreiche Entwicklungsschritte vollzogen werden. Viele innovative Technologien und Softwarelösungen, wie beispielsweise Data-Mining-Algorithmen, Mobile-Learning oder „Augmented-Reality“-Ansätze könnten bereits heute eine Vielzahl der Detailbereiche des E-Assessment-Spektrums abdecken, sind aber in aktuellen Lernplattformen oder virtuellen Lernumgebung noch nicht ausreichend implementiert. Erst wenn sich die Qualität der Lernplattformen dahingehend verbessert, lassen sich auch speziell ausgerichtete didaktische Strategien wirklich umsetzen und führen zu einer Verbesserung der Lernsituation, einer Verringerung des Workloades und schließlich sicherlich auch zu einer Steigerung der Qualität der Lehre insgesamt.

Für den heutigen Einsatz von E-Assessments ist es somit erforderlich, nach einer konkreten Formulierung der angestrebten Lernziele, detailliert abzuwägen welche Überprüfungs- und Bewertungsmechanismen eingesetzt werden sollen und können, welche Parameter diese schlussendlich wirklich erfassen und wie zuverlässig dies geschieht. Erst dann kann entschieden werden, ob hier ein E-Assessment wirklich eine, allen Anforderungen entsprechende, Verbesserung darstellt oder ob nicht doch eine konventionelle Prüfungs- oder Testform vorzuziehen ist.

Für viele Anwendungsbereiche, gerade im diagnostischen oder formativen Einsatz ist dies bereits heute der Fall und das E-Assessment stellt eine würdige Alternative zu konventionellen Verfahren und ein zusätzliches Instrument zu Individualisierung und Modernisierung der Lehre

dar. Für eine überzeugende, kompetenzorientierte Anwendung im summativen Bereich fehlen allerdings vor allem noch rechtliche Sicherheit und intelligente Softwarelösungen mit ausgereiften Bewertungsalgorithmen.



### III Anhang

#### Anhang A: „Instant-Feedback-Experiment“ - Allgemeine Nutzerdaten

Teilnehmer	Feed-back	Endpos
1314868186	1	8
1314865746	1	4
1314865419	0	8
1314864037	0	8
1314691076	0	8
1314815491	0	8
1314815248	1	8
1314873502	0	2
1314874180	0	8
1314877706	1	8
1314881977	1	2
1314882635	1	8
1314885556	0	2
1314885577	0	8
1314885610	0	8
1314885645	1	8
1314885792	0	8
1314885827	0	8
1314886065	1	8
1314886525	1	8
1314886613	1	8
1314886685	0	8
1314886857	1	2
1314887132	1	8
1314887551	1	8
1314888123	0	8
1314888530	1	8
1314889118	1	2
1314889375	0	8
1314889386	0	8
1314889595	0	7
1314889697	1	8
1314889722	0	7
1314890416	0	8
1314890715	1	8
1314890720	1	8
1314890819	0	4
1314890989	0	8
1314891458	0	8
1314891891	1	2
1314892196	0	8
1314892220	0	8
1314892601	0	8
1314892793	1	8
1314892803	1	8
1314892955	1	2
1314893624	1	3
1314893663	0	1
1314893844	0	8
1314894034	0	4
1314894190	0	8
1314894364	1	8
1314894420	1	8
1314894701	1	4
1314894774	0	8
1314895050	1	4
1314895255	0	8

Teilnehmer	Feed-back	Endpos
1314895293	1	4
1314895829	0	8
1314895849	1	8
1314896110	1	8
1314896136	1	8
1314896521	0	8
1314897336	0	3
1314897422	1	2
1314897506	1	8
1314898003	0	8
1314899287	0	8
1314899780	0	8
1314901088	0	8
1314901695	1	8
1314901768	0	2
1314901952	1	8
1314902918	0	2
1314903704	1	8
1314904590	1	8
1314905344	0	2
1314907546	0	1
1314908994	0	8
1314909453	0	8
1314909496	1	8
1314910823	0	8
1314911598	0	8
1314915427	0	8
1314929319	1	8
1314946947	1	2
1314947638	0	8
1314951020	0	8
1314952488	0	8
1314953333	0	8
1314955629	1	4
1314957104	0	1
1314957753	1	2
1314958436	1	4
1314959637	0	8
1314965219	0	8
1314968007	1	2
1314968291	0	2
1314970620	1	1
1314972154	1	8
1314973005	0	2
1314975308	0	8
1314976726	1	8
1314979452	0	8
1314981810	0	2
1314981888	1	8
1314984031	1	2
1314987721	1	2
1314989525	1	8
1314992226	1	2
1315039763	1	8
1315046881	0	8
1315046944	0	4
1315048675	1	8

Teilnehmer	Feed-back	Endpos
1315049075	1	4
1315050636	0	8
1315052955	1	8
1315063358	0	8
1315064018	0	8
1315075366	1	2
1315080429	0	8
1315127855	0	1
1315136312	0	8
1315147064	1	8
1315160846	0	8
1315161918	1	8
1315214248	0	8
1315226821	0	2
1315238731	0	8
1315248872	1	3
1315250481	1	8
1315253883	0	7
1315265422	0	8
1315291108	0	8
1315302393	1	2
1315305848	1	8
1315329186	1	1
1315385730	0	8
1315392235	0	2
1315435354	0	8
1315478604	0	8
1315502389	0	8
1315731577	1	8
1315829121	1	8
1315939283	0	8
1316606097	0	1
1316724571	1	8
1316772797	0	8
1316773259	1	8
1316785185	0	2
1316815851	0	2
1316925828	1	8
1316945586	0	7
1316949597	0	8
1316952651	0	8
1316963765	0	8
1316985699	1	8
1317059294	1	8
1317113641	0	2
1317144882	1	2
1317236241	0	2
1317305109	1	8
1317628933	0	8
1317915205	1	8
1317917145	0	8
1317923217	1	8
1317925505	0	8
1317929109	0	8
1317981869	1	7
1318169868	0	8
1318194544	0	4

Anhang B: „Instant-Feedback-Experiment“- Testergebnisse

sessionId	F1	F2	F3	F4	F5	Fragen
1314868186	0	2	0	0	0	6,4,3,
1314865746	-	-	-	-	-	1,3,6,
1314865419	1	1	1	1	1	2,3,7,
1314864037	1	1	1	1	1	2,3,5,
1314815491	0	1	1	1	0	6,3,1,
1314815248	0	2	0	1	0	3,4,1,
1314691076	0	0	1	1	0	4,2,6,
1314874180	0	1	1	1	1	6,4,1,
1314873502	1	1	0	0	1	3,5,2,
1314882635	0	2	0	1	2	2,1,3,
1314885577	1	1	1	1	1	3,5,4,
1314877706	0	0	2	2	1	3,4,7,
1314885645	0	0	1	0	0	2,1,7,
1314885610	0	1	1	1	1	4,5,2,
1314885792	1	1	1	1	1	3,1,6,
1314886065	2	2	0	1	1	2,4,5,
1314886525	2	0	1	1	1	5,6,7,
1314886685	1	1	1	1	1	1,6,7,
1314886613	0	0	2	2	2	1,3,2,
1314887132	0	2	2	0	1	2,3,4,
1314887551	0	1	0	1	2	5,2,6,
1314888123	1	1	1	1	1	6,2,3,
1314885827	0	1	0	0	0	7,6,1,
1314888530	0	0	2	1	1	3,2,4,
1314889375	0	1	0	1	1	4,6,2,
1314889386	0	0	1	1	1	2,4,1,
1314889722	1	1	1	1	1	6,4,3,
1314889697	0	0	1	0	1	7,6,2,
1314890416	0	1	1	1	1	4,6,1,
1314890715	2	0	0	1	0	1,4,3,
1314890720	0	0	0	1	0	7,1,2,
1314890989	0	1	1	1	1	2,3,6,
1314892196	0	1	1	1	1	7,6,2,
1314892220	1	0	1	1	1	2,7,6,
1314892601	0	0	1	1	1	5,4,3,
1314892803	2	0	1	1	0	5,6,7,
1314892793	2	2	0	1	0	7,2,5,
1314891458	1	1	0	1	1	3,5,4,
1314893844	1	1	1	1	1	1,6,2,
1314894190	1	1	0	1	1	6,2,5,
1314894364	0	1	1	0	1	5,1,6,
1314894420	2	1	2	1	0	7,2,1,
1314894701	-	-	-	-	-	1,7,2,
1314894774	1	1	1	1	1	6,1,7,
1314889595	0	1	0	1	0	6,2,1,
1314895255	1	0	1	1	1	7,1,4,
1314895293	-	-	-	-	-	1,5,3,
1314895829	0	1	1	1	1	3,7,6,
1314895849	2	2	2	1	1	4,7,6,
1314896110	2	0	0	1	1	4,6,2,
1314896136	0	2	0	0	2	6,2,3,
1314896521	0	1	1	1	1	7,2,3,
1314897506	2	0	0	1	1	1,7,6,
1314898003	1	0	1	1	1	3,1,4,
1314899287	1	1	1	1	1	6,1,5,
1314899780	1	1	1	1	1	5,3,1,
1314901088	0	1	1	1	1	2,6,7,
1314901695	0	0	0	2	0	7,6,5,
1314901952	0	1	0	1	2	5,2,7,
1314903704	2	2	0	0	1	5,4,2,
1314904590	2	2	1	1	1	3,6,4,

sessionId	F1	F2	F3	F4	F5	Fragen
1314908994	1	1	1	1	1	5,7,6,
1314909453	1	1	1	1	1	2,7,6,
1314909496	0	0	0	2	1	4,1,7,
1314910823	1	0	1	1	1	5,2,6,
1314911598	1	1	1	1	1	7,4,3,
1314915427	1	0	1	1	1	1,5,6,
1314929319	0	2	0	2	2	6,5,1,
1314947638	1	1	1	1	1	3,2,5,
1314951020	1	1	1	0	1	4,6,2,
1314952488	0	1	1	1	1	7,2,4,
1314953333	1	1	1	1	1	4,6,2,
1314955629	-	-	-	-	-	7,6,2,
1314958436	-	-	-	-	-	1,6,4,
1314959637	1	1	0	1	0	4,1,7,
1314965219	0	1	1	1	1	1,5,4,
1314972154	2	0	2	2	1	4,1,5,
1314975308	0	1	1	0	0	6,7,3,
1314976726	1	1	0	0	1	3,5,2,
1314979452	1	1	0	1	1	2,6,7,
1314981888	1	0	0	0	1	4,1,2,
1314989525	2	0	0	1	0	2,7,5,
1315039763	2	0	2	1	1	5,1,3,
1315046881	1	1	0	1	1	4,3,5,
1315046944	-	-	-	-	-	6,3,5,
1315048675	1	0	1	1	1	2,5,7,
1315049075	-	-	-	-	-	1,5,3,
1315050636	0	0	0	1	0	1,3,6,
1315052955	2	0	0	0	2	3,2,4,
1315063358	1	0	1	1	1	1,6,3,
1315064018	1	1	1	1	1	1,6,2,
1315080429	0	1	1	1	1	4,2,6,
1315136312	0	0	1	1	1	1,4,2,
1315147064	0	0	0	2	0	3,1,6,
1315160846	1	0	0	1	1	2,5,3,
1315161918	0	0	0	0	0	7,3,1,
1315214248	0	1	0	1	0	1,6,3,
1315238731	1	1	0	0	1	2,4,5,
1315250481	0	0	0	2	2	5,2,4,
1315253883	1	0	1	1	1	2,3,4,
1315265422	0	0	1	1	1	3,7,5,
1315291108	0	0	1	1	1	2,5,1,
1315305848	2	2	1	1	1	3,7,2,
1315385730	0	1	1	1	1	6,7,4,
1315435354	1	1	1	1	1	5,2,3,
1315478604	1	1	0	0	1	1,2,3,
1315502389	0	0	0	0	0	3,2,5,
1315731577	0	2	0	0	1	3,2,4,
1315829121	0	0	0	0	0	3,6,5,
1315939283	1	1	0	1	1	1,5,6,
1316724571	0	0	0	2	0	3,4,5,
1316772797	1	0	1	1	1	6,5,7,
1316773259	0	2	0	0	0	6,4,1,
1316925828	1	0	2	1	1	2,6,3,
1316945586	0	0	0	0	0	7,6,3,
1316949597	1	1	1	1	0	3,7,6,
1316952651	0	1	1	0	1	5,6,4,
1316963765	0	1	1	1	1	5,3,2,
1316985699	0	0	0	0	0	2,6,3,
1317059294	2	2	2	2	1	5,7,6,
1317305109	0	0	2	0	2	3,6,1,
1317628933	1	0	1	1	1	5,2,4,

sessionId	F1	F2	F3	F4	F5	Fragen
1317915205	0	0	1	2	1	6,3,4,
1317917145	1	1	0	1	1	7,2,3,
1317923217	0	0	0	0	0	2,6,5,
1317925505	1	1	1	1	1	3,6,4,
1317929109	1	1	1	1	1	2,7,1,
1317981869	0	0	2	1	1	1,2,4,
1318169868	1	1	1	1	1	4,3,6,
1318194544	-	-	-	-	-	5,4,1,

## IV Literaturverzeichnis

### **Baumgart, 2007**

Baumgart, Franzjörg (Hrsg.): Entwicklungs- und Lerntheorien: Erläuterungen-Texte-Arbeitsaufgaben. - 2.Aufl. - Bad Heilbrunn: Klinkhardt, 2007

### **Baumgartner & Payr, 1997**

Baumgartner, Peter & Payr, Sabine: Erfinden lernen. In: Konstruktivismus und Kognitionswissenschaft. Kulturelle Wurzeln und Ergebnisse. Zu Ehren Heinz von Foerstes. K., H. Müller und F. Stadler. Wien-New York: Springer. 8, 1997

### **Baumgartner et. al, 2004**

Baumgartner, Peter & Häfele, Hartmut & Maier-Häfele, Kornelia: Content Management Systeme in e-Education - Auswahl, Potenziale und Einsatzmöglichkeiten. Innsbruck: Studien Verlag, 2004

### **Bescherer et. al, 2012**

Bescherer, Christine & Herding, Daniel & Kortenkamp, Ullrich & Müller, Wolfgang. & Zimmermann, Marc: E-Learning Tools with Intelligent Assessment and Feedback for Mathematics Study. In: Intelligent and Adaptive Learning Systems - Technology Enhanced Support for Learners and Teachers. Hershey: IGI Global, 2012

### **Blackboard, 2011**

Blackboard learn, Assessment, Accreditation, Analytics, Improvement  
(online verfügbar unter: <http://www.blackboard.com/Platforms/Learn/Products/Blackboard-Learn/Assessment-Accreditation-Analytics.aspx>) [letzter Zugriff: 04.12.2011]

### **Bull & McKenna, 2004**

Bull, Joanna & McKenna, Colleen: Blueprint for Computer-assisted Assessment. London: RoutledgeFalmer, 2004

### **Campbell, 2001**

Campbell, Nola: e-teaching, e-learning and e-education - A paper to inform the development of the ICT Strategy in New Zealand for the Ministry of Education. Waikato: School of Education - University of Waikato, 2001

(online verfügbar unter: [http://cms.steo.govt.nz/NR/rdonlyres/C11315DE-804A-4831-AB75-3D4E77393DD8/0/eteaching.htm#\\_Toc386771573](http://cms.steo.govt.nz/NR/rdonlyres/C11315DE-804A-4831-AB75-3D4E77393DD8/0/eteaching.htm#_Toc386771573)) [letzter Zugriff: 04.12.2011]

### **Chalk & Wire, 2011**

Website

(online verfügbar unter: <http://chalkandwire.com/index.php/product/technology>) [letzter Zugriff: 04.12.2011]

### **Chandler & Sweller, 1991**

Chandler, Paul & Sweller, John: Cognitive Load Theory and the Format of Instruction. Wollongong: University of Wollongong, 1991

(online verfügbar unter: <http://ro.uow.edu.au/edupapers/128>) [letzter Zugriff: 04.12.2011]

### **Conole & Warburton, 2005**

Conole, Gráinne & Warburton, Bill: A review of computer-assited assessment. In: ALT-J, Research in Learning Technology Vol. 13, No. 1, March 2005, pp. 17-31. Southampton: University of Southampton, 2005

### **Dahinden & Hinterberger, 2010**

Dahinden, Markus & Hinterberger, Hans: Aspekte technischer und legaler Sicherheit von Online-Prüfungen. In: Ruedel, C., Mandel, S. (Hrsg.): E-Assessment, Einsatzszenarien und Erfahrungen an Hochschulen. Münster: Waxmann, 2010

### **Dany et. al, 2008**

Dany, Sigrid & Szczyrba, Birgit & Wildt, Johannes (Hrsg.): Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG, 2008

### **Dietinger, 2003**

Dietinger, Thomas: Aspects of E-Learning Environments. Graz: Graz University of Technology, 2003  
(online verfügbar unter: [http://know-center.tugraz.at/wp-content/uploads/2010/12/Thesis\\_Thomas\\_Dietinger.pdf](http://know-center.tugraz.at/wp-content/uploads/2010/12/Thesis_Thomas_Dietinger.pdf)) [letzter Zugriff: 04.12.2011]

### **Eilers et. al, 2008**

Eilers, Björn, Gruttmann, Susanne & Kuchen, Herbert: Konzeption eines integrierbaren Systems zur computergestützten Lernfortschrittskontrolle. In: H. L. Grob & J. vom Brocke & C. Buddendick (Hrsg.): E-Learning-Management. München: Vahlen Verlag, 2008

### **Eugster, 2009/2010**

Eugster, Balthasar: Aufbau und Organisation innovativer Prüfungsformen in der Modulprüfung, Neunte Tagung der Bologna-Koordinatoren/ -innen. Ulm: HRK, Universität Ulm, 2009

(online verfügbar unter: <http://www.hrk.de/bologna/de/download/dateien/Eugster.pdf>) [letzter Zugriff: 04.12.2011] und CSPC Tagung 2010

(online verfügbar unter: <http://www.youtube.com/watch?v=gbDA7lduNuQ>) [letzter Zugriff: 04.12.2011]

### **Fischer & Schwendel, 2009**

Fischer, Helge & Schwendel, Jens (Hrsg.): E-Learning an sächsischen Hochschulen - Strukturen - Projekte - Einsatzszenarien. Dresden: TUDpress - Verlag der Wissenschaften GmbH, 2009

(online verfügbar unter: [https://bildungsportal.sachsen.de/e3677/e3700/e3701/e3980/e\\_learning\\_publication\\_bps\\_2009\\_ger.pdf](https://bildungsportal.sachsen.de/e3677/e3700/e3701/e3980/e_learning_publication_bps_2009_ger.pdf)) [letzter Zugriff: 04.12.2011]

### **Flindt, 2005**

Flindt, Nicole: e-learning, Theoriekonzepte und Praxiswirklichkeit. Heidelberg: Ruprecht-Karls-Universität, 2005

### **Frohberg, 2008**

Frohberg, Dirk: Mobile Learning. Zürich: Universität Zürich, 2008

(online verfügbar unter: [http://www.ifi.uzh.ch/pax/uploads/pdf/publication/1230/m-learning\\_frohberg\\_komprimiert.pdf](http://www.ifi.uzh.ch/pax/uploads/pdf/publication/1230/m-learning_frohberg_komprimiert.pdf)) [letzter Zugriff: 04.12.2011]

### **Glasersfeld, 2007**

von Glasersfeld, Ernst: Radikaler Konstruktivismus: Ideen, Ergebnisse, Probleme. - 7. Aufl. - Berlin: Suhrkamp Verlag, 2007

### **Gruttmann, 2009**

Gruttmann, Susanne Johanna: Formatives E-Assessment in der Hochschullehre - Computergestützte Lernfortschrittskontrollen im Informatikstudium. Münster: MV-Verlag, 2010

### **Hollingsworth, 1960**

Hollingsworth Jack: Automatic graders for programming classes. In: Communications of the ACM. New York: ACM, 1960

### **Ifenthaler, 2012**

Ifenthaler, Dirk: Intelligent Model-Based Feedback - Helping Learners to Monitor their Individual Learning Progress. In: Intelligent and Adaptive Learning Systems - Technology Enhanced Support for Learners and Teachers. Hershey: IGI Global, 2012

### **IliasDocs, 2011**

ILIAS 4.1 Benutzerdokumentation

(online verfügbar unter: [http://www.ilias.de/docu/goto\\_docu\\_lm\\_1986.html](http://www.ilias.de/docu/goto_docu_lm_1986.html)) [letzter Zugriff: 04.12.2011]

### **Jank & Meyer, 2002**

Jank, Werner & Meyer, Hilbert: Didaktische Modelle. - 5. völlig überarbeitete Aufl. - Berlin: Cornelsen Scriptor Verlag, 2002

### **JIM, 2008**

Medienpädagogischer Forschungsverbund Südwest (Hrsg.): JIM 2008 - Jugend, Information, (Multi-) Media - Basisstudie zum Medienumgang 12- bis 19-Jähriger in Deutschland. Stuttgart, 2008

(online verfügbar unter: [http://www.mpfs.de/fileadmin/JIM-pdf08/JIM-Studie\\_2008.pdf](http://www.mpfs.de/fileadmin/JIM-pdf08/JIM-Studie_2008.pdf)) [letzter Zugriff: 04.12.2011]

### **JISC, 2007**

JISC: Effective Practice with e-Assessment, An overview of technologies, policies and practice in further and higher education. Bristol: HEFCE, 2007  
(online verfügbar unter: <http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf>)  
[letzter Zugriff: 04.12.2011]

### **Kerres, 2001**

Kerres, Michael: Multimediale und telemediale Lernumgebungen. Konzeption und Entwicklung. - 2. Aufl. - München: Oldenbourg Verlag, 2001

### **Klieme et. al, 2003**

Klieme, Eckhard et al.; Bundesministerium für Bildung und Forschung (Hrsg.): Zur Entwicklung nationaler Bildungsstandards. Bonn, Berlin: BMBF, 2003

### **Kong, 2012**

Kong, Siu Cheung: A Knowledge-Based Approach od Modeling an Internet-Based Intelligent Learning Environment for Comprehending Common Fraction Operations. In: Intelligent and Adaptive Learning Systems - Technology Enhanced Support for Learners and Teacher. Hershey: IGI Global, 2012

### **Mahara, 2011**

Mahoodle://Integrating Mahara with Moodle  
(online verfügbar unter: <https://wiki.mahara.org/images/d/d5/Mahoodle.pdf>) [letzter Zugriff: 04.12.2011]

### **Merrill, 2002**

Merrill, M. David: First Principles of Instruction. Logan: Utah State University, 2002  
(online verfügbar unter: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.22.1979&rep=rep1&type=pdf>) [letzter Zugriff: 04.12.2011]

### **Moodle, 2011**

Website  
(online verfügbar unter: <http://www.moodle.de/>) [letzter Zugriff: 04.12.2011]

### **MoodleDocs, 2011**

Moodle 2.0 documentation  
(online verfügbar unter: [http://docs.moodle.org/20/en/Main\\_page](http://docs.moodle.org/20/en/Main_page))  
[letzter Zugriff: 04.12.2011]

### **Pavlov, 1928**

Pavlov, Ivan Petrovich; Gantt, W. Horsley (Trans): Lectures on Conditioned Reflexes: Twenty-Five Years of Objective Study of the Higher Nervous Activity Behavior of Animals. New York: Liverwright Publishing Corporation, 1928

### **Reepmeyer, 2008**

Reepmeyer, Jan-Armin: Rechtssichere E-Prüfungen. Münster: Westfälische Wilhelms-Universität Münster, 2008  
(online verfügbar unter: [http://www.his.de/publikation/seminar/Workshop\\_E-Pruefung/TOPO3.pdf](http://www.his.de/publikation/seminar/Workshop_E-Pruefung/TOPO3.pdf))  
[letzter Zugriff: 04.12.2011]

### **Reinmann, 2007**

Reinmann, Gabi: Bologna in Zeiten des Web 2.0 - Assessment als Gestaltungsfaktor. Augsburg: Universität Augsburg, Medienpädagogik, 2007  
(online verfügbar unter: [http://www.imb-uni-augsburg.de/files/Arbeitsbericht\\_16.pdf](http://www.imb-uni-augsburg.de/files/Arbeitsbericht_16.pdf)) [letzter Zugriff: 04.12.2011]

### **Ricketts & Wilks, 2002**

Ricketts, C. & Wilks, S. J.: Improving Student Performance Through Computer-based Assessment: insights from recent research. In: Assessment & Evaluation in Higher Education, Vol. 27, No. 5, 2002. Carfax Publishing, 2002  
(online verfügbar unter: [http://www.cblt.soton.ac.uk/principles03/pdf/cba\\_insights.pdf](http://www.cblt.soton.ac.uk/principles03/pdf/cba_insights.pdf)) [letzter Zugriff: 04.12.2011]

**Ridgway et. al, 2006**

Ridgway, Jim & McCusker, Sean & Pead, Daniel: Report 10: Literature Review of E-assessment. Durham-Nottingham: School of Education, University of Durham & School of Education, University of Nottingham, 2006

(online verfügbar unter: <http://hal.archives-ouvertes.fr/docs/00/19/04/40/PDF/ridgway-j-2004-r10.pdf>)  
[letzter Zugriff: 04.12.2011]

**Roloff, 2002**

Roloff, Sighard: Hochschuldidaktisches Seminar - mündliche Prüfungen. Offenburg: Fachhochschule Offenburg, 2002

(online verfügbar unter: [http://www.lehrbeauftragte.net/documents\\_public/MuendlPruef\\_Roloff.pdf](http://www.lehrbeauftragte.net/documents_public/MuendlPruef_Roloff.pdf))  
[letzter Zugriff: 04.12.2011]

**Roloff, 2003**

Roloff, Sighard: Hochschuldidaktisches Seminar - schriftliche Prüfungen. Offenburg: Fachhochschule Offenburg, 2003

(online verfügbar unter: [http://www.lehrbeauftragte.net/documents\\_public/SchriftlPruef\\_Roloff.pdf](http://www.lehrbeauftragte.net/documents_public/SchriftlPruef_Roloff.pdf))  
[letzter Zugriff: 04.12.2011]

**Roth, 2010**

Roth, Gerhard: Intelligenz, Emotionen, Motivation - Wie Lernen gelingt. Bremen: Institut für Hirnforschung, Universität Bremen, 2010

(online verfügbar unter: [http://www.lehrerakademie.uni-bremen.de/Wie\\_Lernen\\_gelingt.pdf](http://www.lehrerakademie.uni-bremen.de/Wie_Lernen_gelingt.pdf)) [letzter Zugriff: 04.12.2011]

**Ruedel, 2010**

Ruedel, Cornelia: Was ist E-Assessment? In: Ruedel, C. & Mandel, S. (Hrsg.): E-Assessment, Einsatzszenarien und Erfahrungen an Hochschulen. Münster: Waxmann, 2010

**Schulmeister, 2007**

Schulmeister, Rolf: Grundlagen hypermedialer Lernsysteme. Theorie – Didaktik – Design. - 4. Aufl. - München: Oldenbourg Wissenschaftsverlag, 2007

**Seufert et. al, 2001**

Seufert, Sabnie & Back, Andrea & Häusler, Martin & Berger, Sonja: E-Learning. Weiterbildung im Internet. Das „Plato-Cookbook“ für internetbasiertes Lernen. Baar: SmartBooks Publishing AG, 2001

**Steinberg, 2006**

Steinberg, Monika: Organisatorisches Konzept für Online-Prüfungsverfahren. Hannover: Leibniz Universität, 2006

(online verfügbar unter: [http://www.sra.uni-hannover.de/fileadmin/uploads/Mitarbeiter/Steinberg/Publikationen/AP7\\_Org\\_Konzept\\_screen\\_v1.pdf](http://www.sra.uni-hannover.de/fileadmin/uploads/Mitarbeiter/Steinberg/Publikationen/AP7_Org_Konzept_screen_v1.pdf)) [letzter Zugriff: 04.12.2011]

**Stratmann et. al, 2009**

Stratmann, Jörg & Preussler, Annabell & Kerres, Michael: Lernerfolg und Kompetenz bewerten - Didaktische Potenziale von Portfolios in Lehr-/Lernkontext. In: Themenheft Nr. 18: Neue Medien und individuelle Leistungsdarstellung - Möglichkeiten und Grenzen von ePortfolios und eAssessments. 2009

**VG Hannover, 2008**

Verwaltungsgericht Hannover: Beschluss vom 10.12.2008, 6 B 5583/08

(online verfügbar unter: <http://www.dbovg.niedersachsen.de/Entscheidung.asp?Ind=0520020080055836%20B>) [letzter Zugriff: 04.12. 2011]

**Zhou, 2012**

Zhou, Mingming: From „Self-Tested“ to „Self-Testing“: A Review of Self-Assessment Systems for Learning. In: Intelligent and Adaptive Learning Systems - Technology Enhanced Support for Learners and Teachers. Hershey: IGI Global, 2012

## V Erklärung zur selbständigen Anfertigung der Arbeit

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Mittweida, 09.12.2011