

Allgemeine Konzeption eines formativen eAssessments für die Studienvorbereitung im Bereich Mathematik

TP 3.2 Formatives eAssessment

Katja Derr

DHBW Mannheim

Januar 2016



Dieses Werk wurde im Rahmen des Projekts optes erstellt und steht unter der Lizenz
Creative Commons 'Namensnennung - Weitergabe unter gleichen Bedingungen 3.0 Deutschland'.

Inhaltsverzeichnis

1	Einleitung: Formatives eAssessment im Selbststudium	2
2	Eigenschaften von Aufgaben	4
2.1	Aufgaben zur Übung/Reproduktion	5
2.2	Aufgaben auf dem Niveau Anwenden/Verstehen	5
2.3	Praxisbezug als Motivation in den Ingenieurwissenschaften?	6
2.4	Aufgabenschwierigkeit	7
3	Aufgabentypen Online	8
3.1	Pro und contra geschlossene/offene Aufgabenformate	8
3.2	Geschlossene Online-Fragetypen	10
3.3	Halb offene Online-Fragetypen	12
3.4	Offene Online-Fragetypen	15
3.5	Fazit zu Online-Aufgabentypen	16
4	Feedback	17
4.1	Verifikationsfeedback (Richtig/Falsch)	17
4.2	Elaborationsfeedback (Erweitertes Feedback)	17
4.3	Fazit zu Feedback	18
5	Testkonzept im Projekt optes	20
5.1	Curriculum	20
5.2	Lernprozessinitiiierende Leistungserfassung: Diagnostischer Einstiegstest	21
5.3	Lernprozessbegleitende Leistungserfassung: Üben und Anwenden	22
5.4	Lernerfolgskontrolle	24
6	Hinweise zur Gestaltung von Aufgaben	25
6.1	Test-Startseite	25
6.2	Test-Einstellungen	25
6.3	Aufgabengestaltung	26
6.4	Fragetypen	30
6.5	Feedback	33
7	Glossar	39
8	Quellen	40
9	Literatur	41

1 Einleitung: Formatives eAssessment im Selbststudium

Im Teilprojekt 3.2 entsteht ein Fragenpool der Grundlagenmathematik, der in unterschiedlichen Lehr-Lernszenarien im optes-Vorkurs zum Einsatz kommt und für dessen Entwicklung das vorliegende Dokument als theoretische Grundlage dient. Nach einer Begriffsklärung werden die unterschiedlichen Einsatzgebiete von Aufgaben und Tests im Lernprozess diskutiert und zu einer Handreichung für die Projektpartner zusammengefasst.

In der angelsächsischen Literatur findet sich der Begriff Assessment im Zusammenhang mit allen Formen der Leistungserhebung, wobei zwischen *summativen* Assessments, die Eingang in eine Gesamtbewertung (Note) finden, und *formativen* Assessments in Form von Übungen, Selbsttests oder Gruppenarbeiten unterschieden wird (Sadler, 1989; Black & Wiliam, 1998; Black & Wiliam, 2009). Formatives Assessment hat dabei die Aufgabe, Lehrenden und Lernenden einen Überblick über den aktuellen Lernstand einer Gruppe oder eines Individuums zu verschaffen. Es lässt sich als Prozess beschreiben, der dem summativen Assessment vorgeschaltet ist und „alle Formen von Leistungsbeurteilung“ umfasst, „die Informationen über die Diskrepanz zwischen Lernzielen und aktuellem Lernstand liefern und dadurch den Lehrenden und/oder den Lernenden selbst helfen, den weiteren Lernprozess zu gestalten“ (Klieme et al., 2010, S. 64f). Der Abgleich der eigenen Kenntnisse mit dem erwarteten Wissensstand sollte möglichst häufig erfolgen, damit Missverständnisse und Fehlinterpretationen sich nicht verfestigen (und dann im summativen Assessment zu Tage treten).

Neben der Leistungserfassung und der Rückmeldung über den Wissensstand von Lernenden ist der Einsatz von formativem Assessment immer auch ein Mittel zur Aktivierung. Die Bearbeitung von Aufgaben kann dabei unterschiedliche didaktische Ziele haben, vom Memorieren einer Formel über das Einüben eines mathematischen Verfahrens bis zur Förderung des Verständnisses durch selbständige Erarbeitung eines Lösungswegs (Renkl, 1991). Physikalische oder technische Fragestellungen können dabei zur Veranschaulichung mathematischer Lerninhalte dienen und zur Diskussion anregen (Mazur, 1997).

Beim Einsatz von formativem Assessment zur Unterstützung von Lernprozessen ist zu unterscheiden, welche und wie viele Inhalte das Assessment umfasst, ob nur ein oder mehrere Lernziele adressiert werden und wann es stattfindet - zu Beginn, während oder am Ende des Lernprozesses. Shavelson et al. (2008) unterscheiden zwischen

- „On-the-Fly Formative Assessment“, das sich auf das mehr oder weniger spontane Bearbeiten von Übungsaufgaben im Unterricht bezieht,
- „Planned-for-Interaction Formative Assessment“, das als Impuls oder Aktivierungsmaßnahme im Unterricht eingesetzt werden kann sowie
- „Curriculum-Embedded Formative Assessment“, das sich auf die vorgeschriebenen Lehrinhalte bezieht und damit auch die Vorbereitung auf das summative Assessment beinhalten kann (Shavelson et al., 2008, S. 300f).

Im Präsenzunterricht reagiert die Lehrperson auf den Lernfortschritt und wählt Übungsaufgaben und Tests aus, die (idealerweise) genau auf die Lücke zwischen Lernziel und aktuellem Wissensstand der Lernenden zielen, „this just right gap“ (Heritage, 2007, S. 141).

Im Online-Selbststudium wird der Einsatz von Selbsttests und Übungsaufgaben durch das didaktische Design vorgegeben; Entscheidungen über Zeitpunkt, Dauer und Häufigkeit der Lernaktivitäten liegen allerdings bei den Lernenden. Die größere Verantwortung für Organisation und Ablauf des Lernprozesses erfordert einen höheren Grad der Selbständigkeit, sowie die Fähigkeit zum „selbstregulierten Lernen“ (Zimmerman, 1989; Pintrich & de Groot, 1990; Boekaerts et al., 2000).

Brückenkurs-Teilnehmer haben unterschiedliche schulische und berufliche Hintergründe und ihre Vorkenntnisse in Mathematik unterscheiden sich teilweise erheblich (z.B. Schwenk & Berger, 2006; Fischer & Biehler, 2011; Krumke et al., 2012; Bausch et al., 2014). Auch die Fähigkeit zur Selbsteinschätzung ist unterschiedlich ausgeprägt und oft besteht Unsicherheit, welche Kenntnisse an der Hochschule erwartet werden.

Zusätzlich zu den genannten Arten des formativen Assessment ist darum die „Kalibrierung“ zu nennen, ein Abgleich des eigenen Vorwissens mit den Lernzielen des Kurses (Winne, 2004). Formative eAssessments zu Beginn des Lernprozesses machen „die relevanten Lernziele in operationalisierender Form transparent“ (Jacobs, 2008, S.100) und geben den Lernenden ein Werkzeug zur Selbstevaluation bzw. zum „Self-monitoring“ an die Hand (Pressley et al., 1989; Nicol & Macfarlane-Dick, 2006). Als Auslöser für weitere Lernhandlungen sind lernprozessinitiiierende Assessments mit dem „Planned-for-Interaction Formative Assessment“ vergleichbar, es gibt aber auch Verbindungen zum „Curriculum-Embedded Formative Assessment“, da Informationen über die an der Hochschule erforderlichen mathematischen Grundkenntnisse vermittelt werden.

Auf Basis der ersten diagnostischen Leistungserfassung kann dann die Auswahl der Lerninhalte und die Bearbeitung der dazugehörigen Lernmodule erfolgen (in optes: Lernzielorientierte Kurse, LoKs). Hier kommt formatives eAssessment in erster Linie in Form von Übungsaufgaben zum Einsatz („On-the-Fly Formative Assessment“). Je nach Lerninhalt ist außerdem der Einsatz von Aufgaben zur Einführung eines Themas zu Beginn einer Lerneinheit möglich („Planned-for-Interaction Formative Assessment“).

Da der Besuch des Vorkurses freiwillig ist, ist auch die Lernerfolgskontrolle ein formativer Selbsttest. Die Lernenden entscheiden sich zu einem selbst bestimmten Zeitpunkt zur Durchführung einer abschließenden Leistungserfassung, die ggf. wiederholt werden kann.

Basierend auf diesen Vorüberlegungen werden im weiteren Verlauf des Dokuments unterschiedliche Aspekte der Aufgabengestaltung beleuchtet und in Bezug auf ihre Relevanz für den Einsatz im Online-Vorkurs diskutiert. Als Basis für das Konzept, das in Kapitel 5 zusammengefasst wird, dient die folgende Differenzierung formativen und summativen eAssessments im optes-Projekt:

Formatives eAssessment

- Lernprozessinitiiierende Leistungserfassung zu Beginn des Lernprozesses (Diagnose: Information und Kalibrierung)
- Lernprozessbegleitende Leistungserfassung während des Lernprozesses (Steuerung, Aktivierung, Motivierung)
- Abschließende Leistungserfassung/Lernerfolgskontrolle (Monitoring)

Summatives eAssessment

- Bilanzierende Leistungserfassung
(zu summativem eAssessment siehe Ergebnisbericht Konzeptphase TP 3.1)

Anmerkung: Inhaltlich und konzeptionell besteht im optes-Projekt eine Trennung zwischen formativem eAssessment im Vorkurs und summativem eAssessment im Studium. So haben Erfolg oder Misserfolg im Vorkurs keine institutionelle Relevanz und die Inhalte im Vorkurs und in der Studieneingangsphase unterscheiden sich. Über die Interaktion mit der Lernplattform werden allerdings im Vorkurs erste Erfahrungen mit elektronischen Lern- und Prüfungsmethoden ermöglicht, die die Studieneingangs-phase sowie die spätere Teilnahme an eKlausuren erleichtern können (z.B. Hutchings et al., 2013).

2 Eigenschaften von Aufgaben

Je nach Sichtweise (Lerner-, Lehrer-, Autorensicht) existieren unterschiedliche Ansätze zur Charakterisierung und Kategorisierung von Aufgaben. Analog zur Unterscheidung zwischen formativem und summativem Assessment kann beispielsweise zwischen Lernaufgaben und Testaufgaben unterschieden werden. Lernaufgaben dienen dazu, Denkprozesse anzustoßen und zur Interaktion mit dem Lernmaterial anzuregen, während Testaufgaben Wissenszuwachs möglichst genau dokumentieren sollen (z.B. Leutner et al., 2008). Bei der Entwicklung von Testaufgaben sind Validität und Fairness darum von größerer Bedeutung als bei einer Lernaufgabe, die ohne Zeitdruck und im Austausch mit anderen (der Lehrperson oder Peers) gelöst werden kann. Da das Bearbeiten von Lernaufgaben aber (unter anderem) auf das Beantworten von Testaufgaben vorbereiten soll, kann die Entwicklung von Aufgaben für diese beiden Einsatzgebiete nicht losgelöst voneinander betrachtet werden. Nicht nur das Curriculum muss übereinstimmen, auch die Art der Fragestellung, die Art der Eingabe von Antworten, die Gestaltung des Feedback und der Schwierigkeitsgrad von Test- und Lernaufgaben sollten aufeinander abgestimmt sein (Dubs, 2008).

Die Komplexität oder Schwierigkeit einer Aufgabe bemisst sich dabei nach dem Grad der kognitiven Anforderung, die sie an die Lernenden stellt. In den Lernzieltaxonomien von Bloom werden sechs aufeinander aufbauende Stufen der kognitiven Verarbeitungstiefe (Erinnern, Verstehen, Anwenden, Analysieren, Evaluieren, Erschaffen) den Wissensdimensionen Faktenwissen, Konzeptwissen, prozedurales Wissen und meta-kognitives Wissen zugeordnet (vgl. Anderson & Krathwohl, 2001). Aufbauend auf diesem sehr komplexen Modell wurden Taxonomien zur Klassifizierung von Aufgaben vorgeschlagen, die mit drei oder vier Stufen der kognitiven Verarbeitungstiefe arbeiten, die auch für Online-Aufgaben nutzbar sind (z.B. Metzger et al., 1993; Hofmeister, 2005; Ivanov & Lehmann, 2005 oder Mayer et al., 2009).

Auf einem eher niedrigen kognitiven Niveau liegen jeweils Aufgaben, die sich auf das „Reproduzieren“, also z.B. das (auswendig) Kennen von Formeln oder das Wiederholen eines Rechenschritts beziehen. Die mittlere Verarbeitungstiefe ist bezeichnet mit „Anwenden und Verstehen“, d.h. das Gelernte kann in leicht veränderten Kontexten angewandt oder in eigenen Worten wiedergegeben werden. Die letzte(n) Stufe(n) werden mit „Reflektion“ des Gelernten sowie „Erschaffen“ beschrieben, also der „Fähigkeit, Elemente zu einem neuen, kohärenten, funktionierenden Ganzen zusammenzufügen“ (Mayer et al., 2009, S.101).

Aus kognitionspsychologischer Sicht wird die Fähigkeit, Gelerntes in ähnlichen oder komplett anderen Zusammenhängen anzuwenden, oft mit „Transfer“ bezeichnet (Resnick et al., 2010; Klauer, 2011). Je nachdem, wie ähnlich eine Aufgabenstellung den schon bekannten Lerninhalten ist, kann von nahem, mittlerem und weitem Transfer gesprochen werden (Stark et al., 2000). Andere Modelle unterscheiden zwischen „analogem“ und „metakognitivem“ Lerntransfer (Seel, 2000), der im Hinblick auf den Erwerb von Schlüsselqualifikationen noch auf außerfachliche Fähigkeiten und Kompetenzen erweitert werden kann (Weinert, 1998; Klieme, 2008).

Eine große Herausforderung ist die Operationalisierung komplexer Problemlösefähigkeiten oder gar überfachlicher Kompetenzen in Aufgaben, die die gewünschten Fähigkeiten „messen“. Ansätze hierzu finden sich in „komplexen Lernaufgaben“, deren Bearbeitung mehrere Stunden oder Tage dauern kann (Flehsig, 2008) oder, im eLearning Bereich, in Simulationen, die Lerninhalte in spielerische und/oder realitätsbezogene Rahmenhandlungen einbetten (van Eck & Dempsey, 2002). Die Beantwortung von Aufgaben auf dem Niveau Problemlösen/Transfer/Erschaffen verlangt allerdings die Eingabe von Texten, Bildern oder Formeln bis zur Erstellung ganzer Artefakte, daher kann weder die Rückmeldung noch die Bewertung automatisiert erfolgen.

Für die Entwicklung von Fragen für das formative eAssessment im Selbststudium ist eine Beschränkung auf die Verarbeitungsstufen „Üben und Reproduktion“ und „Anwenden und Verstehen“

sinnvoll (Mayer et al., 2009). Im Folgenden wird unterschieden zwischen einfachen Aufgaben, die zur Einübung und Verfestigung von Rechenverfahren dienen, und komplexeren Aufgaben, die das Nachdenken über ein mathematisches Problem anregen sollen – auch wenn eine deutliche Trennung nur bedingt möglich ist.

2.1 Aufgaben zur Übung/Reproduktion

Aufgaben zur Übung und Reproduktion zielen auf den mehr oder weniger „mechanischen“ Erwerb basaler Fertigkeiten (Renkl, 1991). Sie kommen zum Einsatz, wenn ein neu eingeführtes Verfahren wiederholt und konsolidiert werden soll. Die Wiederholung von Lerninhalten durch das Beantworten von Fragen oder das Rechnen von Aufgaben dient der Ausbildung eines gesicherten Wissens, das jederzeit abrufbar ist und bei Abruf keine große kognitive Last produziert (z.B. Sweller et al., 1998). Beim Üben geht es also weniger um die Zunahme des Wissens als um die Verminderung des Vergessens (Jacobs, 2008, S. 100). In der Mathematik kann sich diese „Automatisierung“ auf das Anwenden von Rechenschritten beziehen, aber auch den Umgang mit Maßeinheiten, Größen, mathematischer Syntax sowie die korrekte Anwendung der Fachsprache.

Aufgaben zur Reproduktion von Wissen können auf dem Niveau von „Drill & Practice“ liegen; ein Ansatz, der auf schnelle Erfolgserlebnisse bei den Lernenden zielt, indem relativ leichte Aufgaben so oft beantwortet werden, bis keine Fehler mehr gemacht werden. Das Bearbeiten solcher Aufgabensequenzen kann durchaus Lerneffekte erzielen, beispielsweise wenn nur oberflächlich gelernt wurde und der Lernstoff dann über das Bearbeiten einer größeren Anzahl von Aufgaben nochmals erarbeitet wird (Cook, 2001; Jacobs, 2008).

Problematisch bei sehr leichten Übungsaufgaben ist, dass sie schnell eintönig werden können und die Lernenden wenig anregen, über mögliche Anwendungsgebiete oder die Sinnhaftigkeit eines Verfahrens nachzudenken. In der Literatur wird in diesem Zusammenhang gern von „trägem Wissen“ gesprochen (z.B. Mayer, 2005, S. 13). Dies ist beispielsweise der Fall, wenn die Umstellung einer Gleichung dazu führt, dass Lernende einen zuvor eingeübten Lösungsweg nicht mehr durchführen können.

Das exzessive Bearbeiten von Online-Aufgaben kann außerdem kontraproduktiv wirken, wenn die Lernenden auf die Anzahl der Richtig-Rückmeldungen (scores) fokussiert sind und darüber hinaus keine weiteren Lernaktivitäten stattfinden (Hannah et al., 2013). Um ein tieferes Verständnis aufzubauen sind darum auch Aufgaben erforderlich, die das Nachdenken über Zusammenhänge fördern und die Anwendung mathematischer Routinen in leicht veränderten Kontexten verlangen.

2.2 Aufgaben auf dem Niveau Anwenden/Verstehen

Aufgaben, die Lernende zur tieferen Beschäftigung mit einem Sachverhalt anregen, bewegen sich auf der Ebene des „bedeutungshaltigen“ Lernens (Renkl, 1991). Während beim Üben und Trainieren von mathematischen Verfahren meist das Umformen eines Ausgangsterms oder das Lösen einer Gleichung verlangt wird, beschreiben Aufgaben, die das Verständnis von mathematischen Problemstellungen fördern sollen, komplexere Ausgangssituationen. Die Lösung kann die Anwendung eines oder mehrerer Verfahren oder aber die Auswahl des richtigen Verfahrens für eine gegebene Situation erfordern.

Beispiele für Anwendungsaufgaben können Textaufgaben sein, die ein mathematisches Problem in einen „realen“ Sachverhalt einkleiden, es kann aber auch eine Annahme oder Schätzung gefordert sein oder die Herleitung eines gegebenen Satzes. Je komplexer die Fragestellung, desto mehr

Vorkenntnisse müssen abgerufen und miteinander in Beziehung gesetzt werden. Der Übergang von der Anwendung zum Problemlösen ist dabei fließend, da der Komplexitätsgrad einer Aufgabe nicht unabhängig vom Vorwissen der Lernenden und der Lernumgebung betrachtet werden kann. Auch das Lösen von Problemen ist letztendlich eine Fertigkeit, die geübt werden kann (und muss): „Probably the only way to learn how to solve problems is to solve lots of problems“ (Gibbs & Simpson, 2004, S. 15)

Für den Einsatz im Vorkurs sind allerdings komplexe Fragestellungen auf dem Niveau Problemlösen schon aus Zeitgründen eher ungeeignet und werden im optes-Fragenpool die Ausnahme sein.

2.3 Praxisbezug als Motivation in den Ingenieurwissenschaften?

Für angehende Studierende in den Ingenieurwissenschaften, deren eigentliches Studieninteresse ja nicht die Mathematik ist, bietet sich die Verknüpfung zu Beispielen aus der Technik oder dem Ingenieursalltag als zusätzliche Motivation zur Beschäftigung mit mathematischen Fragestellungen an. Gerade zu Beginn des Studiums ist oft unklar, wofür bestimmte mathematische Verfahren benötigt werden und warum die Beherrschung der Grundlagen wichtig für den weiteren Studienerfolg ist.

Im Bereich der Studieneingangsphase finden sich Beispiele für den didaktischen Einsatz von Anwendungsaufgaben in großen Grundlagen-Vorlesungen in Physik oder Mathematik. Zur Einleitung eines Themas wird ein Problem aus der Praxis kurz vorgestellt und die Studierenden entscheiden per Abstimmung über den richtigen Lösungsansatz (Multiple Choice). Im Anschluss werden die Lösungsansätze in Kleingruppen diskutiert und am Ende der Vorlesung wird die Abstimmung wiederholt und das Ergebnis vom Dozenten kommentiert bzw. aufgelöst (Mazur, 1997; Crouch & Mazur, 2001; Preißler et al., 2010; Bender & Thiele, 2014).

Die Herausforderung bei diesem Ansatz ist es, Aufgaben zu finden, die gleichzeitig authentisch und für die Lernenden (in einem angemessenen Zeitraum) nachvollziehbar sind. Während einige Autoren dafür plädieren, die Komplexität zuzulassen, und für die Lösung authentischer Probleme mehrere Wochen Kleingruppenarbeit veranschlagen (Härterich et al., 2012; Rooch et al., 2014), weisen andere auf die knapp bemessene Lernzeit der Studienanfänger/-innen hin und schlagen eine (mäßige) Didaktisierung von Aufgaben vor (Wolf & Biehler, 2014).

Im Vorkursbereich wird die Problematik dadurch verschärft, dass die Auswahl an mathematischen Verfahren, die zur Lösung eines realistischen Problems zur Verfügung stehen, noch etwas geringer ist als im ersten Studienjahr. Die Lernenden sind zu diesem Zeitpunkt auch noch nicht mit ihren Kommiliton/-innen vernetzt, was den informellen Austausch erschwert. Um eine Verunsicherung der Lernenden zu vermeiden ist das Niveau der Aufgaben eher niedrig anzusetzen; die Praxisbeispiele sind eher als „Anker“ für das spätere Berufsziel zu sehen.

Im Bereich professioneller eLearning Software für Schüler/-innen existieren verschiedene Versuche, problemorientiert und anwendungsbezogen Grundlagenmathematik zu vermitteln. Studien zeigen allerdings, dass die Schüler/-innen den dort hergestellten Praxisbezug oft als unrealistisch und die Probleme als zu komplex und zeitaufwändig empfinden und dass solche Angebot in erster Linie wegen der Ablenkung vom „normalen“ Unterricht und der Möglichkeit des kooperativen Lernens geschätzt werden (Offer & Bos, 2009; Kebritchi et al., 2010).

Es wird darum Teil der Evaluation des Teilprojekts „formatives eAssessment“ sein, inwiefern Aufgaben mit Bezug zu technischen und (moderat) ingenieurwissenschaftlichen Fragestellungen zum Interesse und zur Lernmotivation beitragen können.

2.4 Aufgabenschwierigkeit

Auch wenn die kognitive Komplexität einer Aufgabe einen recht guten Anhaltspunkt für ihre Lösbarkeit bietet, ist dieser Zusammenhang nicht immer linear. So entstehen beispielsweise häufig Fehler bei der Umformung von Brüchen und Termen, weil die Regeln hierzu falsch angewendet werden. Die Schwierigkeit einer Aufgabe ist ein Maß, das die Wahrscheinlichkeit berücksichtigt, mit der eine Aufgabe korrekt beantwortet wird. Während die subjektive Schwierigkeit sich auf die Fähigkeit eines Lernenden bezieht, lässt sich objektive Schwierigkeit durch Zusammenstellen statistischer Angaben zu einer Aufgabe (annähernd) herstellen.

Dabei können neben der Lösungshäufigkeit Faktoren wie Anzahl der notwendigen Lösungsschritte, die Vertrautheit des Kontexts und eine eingeschränkte Bearbeitungszeit zur Schwierigkeit beitragen. Auch die Art der Aufgabenstellung, die Verständlichkeit der Formulierung, das verwendete Vokabular oder der Einsatz von Visualisierungen oder Hilfestellungen beeinflussen die Aufgabenschwierigkeit. Aus Lernericht sind das Vorwissen, aber auch emotional-motivationale Faktoren wie Selbstwirksamkeitserwartung oder Involviertheit bedeutsam (vgl. Nunan & Koepke, 1995).

Dabei kann sowohl zu niedrige als auch zu hohe Aufgabenschwierigkeit zu De-Motivation oder Frustration führen, wobei die Befunde teilweise widersprüchlich sind, ob Über- oder Unterforderung lernhinderlicher sind. So wird für den Bereich des entdeckenden Lernens eine niedrige Aufgabenschwierigkeit empfohlen (Nußbaum & Leutner, 1986; Leutner et al., 2008), während bei der Wiederholung von bekannten Lerninhalten „mäßig schwierige“ Aufgaben deutlich geeigneter zu sein scheinen, allerdings auf individuellem Niveau (vgl. Schabram, 2007, „mäßig schwierig“ wurde hier modelliert als 30% Lösungswahrscheinlichkeit der Teilnehmer/-innen mit gleichem Fähigkeitsniveau).

Schwierigkeit ist also ein komplexes Konstrukt, sie entsteht „in der Interaktion zwischen Lernenden, der Aufgabe und dem Kontext“ (Astleitner, 2008, S. 66). Die perfekte Aufgabenschwierigkeit für jede/n Lernende/n zu jedem Zeitpunkt des Online-Vorkurses wird sicherlich kaum zu finden sein, dennoch sollte bei der Entwicklung von Aufgaben darauf geachtet werden, dass jedes Themengebiet durch Aufgaben unterschiedlicher Schwierigkeitsgrade abgedeckt wird.

D.h. schon bei der Aufgabenentwicklung werden alle Aufgaben einem von drei Schwierigkeitsgraden (leicht, mittel, schwer) zugeordnet. Diese Zuordnung des/der Autor/-in wird dann von einer zweiten Person überprüft und ggf. diskutiert. In einem dritten Schritt wird die Codierung mit den Teilnehmerergebnissen abgeglichen, d.h. der Schwierigkeitsindex einer Aufgabe kann sich ändern, wenn die Ergebnisse anders ausfallen als erwartet. Dabei ist zu berücksichtigen, dass Übungsaufgaben während des Lernprozesses mehrfach bearbeitet werden können. Um die Schwierigkeit von Übungsaufgaben auf Basis der Teilnehmerergebnisse zu bestimmen, wird vorgeschlagen auf die Ergebnisse ihrer jeweils ersten Bearbeitung zurück zu greifen.

Als „schwer“ werden im optes-Projekt Aufgaben klassifiziert, die von 40% und weniger Teilnehmer/-innen korrekt beantwortet werden können, als „mittel“ werden Aufgaben klassifiziert, die von 40-60% der Teilnehmer/-innen korrekt beantwortet werden können, und als leicht werden Aufgaben klassifiziert, die von 60% und mehr korrekt beantwortet werden können. Ggf. sind diese Werte auf Basis der Fähigkeiten der Lernenden anzupassen (Baumert et al., 1999).

Aufgaben, die von sehr vielen (>80%) oder sehr wenigen Lernenden (<20%) beantwortet werden können, sind wenig aussagekräftig und sollten aus diagnostischen Tests entfernt werden (es sei denn, eine sehr leichte Aufgabe soll als „Icebreaker“ im Test verbleiben).

3 Aufgabentypen Online

Fragetypen lassen sich über den Grad der Offenheit des Antwort-Eingabeformats charakterisieren (Mayer et al., 2009):

- Geschlossene Aufgabe: Auswahl der richtigen Antwort(en) aus einer Liste vorgegebener Antworten (z.B. Single/Multiple Choice)
- Halb offene Aufgabe: kurze Texteingabe, z.B. Lückentext
- Offene Aufgabe: Freitexteingabe, Abgabe eines Dokuments, Artefakts, etc.

Geschlossene Aufgaben umfassen alle Aufgaben, bei denen die Antwort in irgendeiner Form vorgegeben ist. Das bekannteste Aufgabenformat ist die Multiple (bzw. Single) Choice-Frage, bei der eine oder mehrere Antworten aus einer Liste ausgewählt werden.

Halb offene Aufgaben erlauben eine „freie“ Eingabe in einem sehr festgesteckten Rahmen, z.B. die Eingabe einer Zeichen- oder Buchstabenfolge in ein Lückentextfeld.

Bei offenen Fragen gibt es keine Einschränkung bei der Eingabe, es können Texte, Bilder, Dokumente etc. als Antwort verlangt und abgegeben werden. Eine automatisierte Rückmeldung oder Bewertung ist darum nicht möglich.

3.1 Pro und contra geschlossene/offene Aufgabenformate

Geschlossene Online-Aufgabentypen wie Multiple Choice sind niederschwellig zu bedienen und verlangen von den Lernenden keine besonderen eLearning-Vorkenntnisse, Testteilnehmer/-innen können bei der Eingabe keine Flüchtigkeits- oder Schreibfehler machen. Die Eindeutigkeit der Eingabe führt zu größerer Fairness bei der Bewertung; Benachteiligungen durch unterschiedliche Bewerterkriterien sind bei geschlossenen Aufgaben quasi unmöglich (z.B. Newstead, 2002).

Auf der anderen Seite lässt die Einschränkung der Antwortmöglichkeiten auf eine überschaubare Auswahl keinen Raum für eigene Ansätze. Es besteht die Gefahr, dass sich die Lernenden weniger auf das Wissen der richtigen Antwort als das Ausschließen der Distraktoren konzentrieren. Hinzu kommt, je nach Anzahl der Distraktoren, eine gewisse Ratewahrscheinlichkeit. Da die Antwort nur ein Begriff oder ein kurzer Satz sein kann, erzwingt das Format eine übersichtliche und kurze Aufgabenstellung. Die Beantwortung vieler kurzer MC-Aufgaben kann auf Dauer eintönig sein und den Fokus auf die Leistungsmessung lenken, anstatt zum Nachdenken anzuregen (Stichwort „teaching to the test“).

Auch wenn Multiple Choice-Fragen seit vielen Jahren in großen landesweiten und internationalen Leistungsmessungen zum Einsatz kommen, wurde immer wieder in Frage gestellt, ob über geschlossene Fragetypen überhaupt komplexere Gedankengänge, geschweige denn Problemlösestrategien getestet werden können (z.B. Hoffmann, 1962; Snow, 1993; Craven, 2009). Diese Kritik trifft allerdings auch auf halb offene Formate und damit letztlich auf alle Online-Fragetypen zu, lediglich die Ratewahrscheinlichkeit fällt bei halb offenen Aufgaben weg.

Nur über offene Aufgabenformate können Leistungen abgefragt und gewürdigt werden, die außerhalb eines vorgefertigten Antwortspektrums liegen. Durch die individuelle Bewertung ist die Vergabe von Teilpunkten möglich, beispielsweise wenn trotz falschem Endergebnis der Rechenweg korrekt war. Für den Bereich der Mathematik ergeben sich bei offenen Online-Aufgaben allerdings erhebliche technische Probleme. Die Eingabe mathematischer Ausdrücke und Formeln in eine Eingabemaske erfordert die Kenntnis bestimmter Syntax (z.B. LaTeX), die bei Schüler/-innen nicht vorausgesetzt

werden kann. Das Abgeben und Bewerten von ausführlichen Lösungswegen ist darum bislang nur über den Upload gescannter handschriftlicher Unterlagen (oder PDF-Dokumente) möglich. Dieses Verfahren ist auch in der Korrektur sehr aufwändig (siehe Einreichaufgaben, Abschnitt 5.3.2). Die Vorteile von Online-Fragetypen, wie schnelles und effizientes Feedback, können also bei offenen Aufgaben nicht genutzt werden.

Ansätze komplexe Denkvorgänge dennoch automatisiert zu prüfen finden sich beim „Two-tier-testing“ oder „Assertion-reason“. Die Testteilnehmer müssen bei diesen Fragetypen begründen, warum sie eine (MC-)Antwort gewählt haben (was wiederum die Korrektur über eine Lehrperson erfordert, oder über eine weitere MC-Frage erfolgen kann). Problematisch an diesen Fragetypen ist, dass die Gefahr von Verständnisproblemen beim Lesen und Interpretieren der Aufgabenstellung relativ groß ist, was dazu führen kann, dass die Ergebnisse noch weniger aufschlussreich sind als die einfacher MC-Tests (Jacobs, 2008).

Hinzu kommt, dass Multiple Choice-Tests in vielen Fällen zu sehr ähnlichen Resultaten führen wie offene Formate (z.B. Hancock, 1994; Kennedy & Walstad, 1997; Lissitz et al., 2012); zumindest für das Fach Mathematik und für ein Anforderungsniveau unterhalb des Problemlösens scheint der Einsatz geschlossener Aufgaben vertretbar (Traub, 1993; Martinez, 1999).

Zu beachten ist, dass die Formulierung geschlossener Aufgaben aus Autorensicht einen erheblichen Mehraufwand verlangt, da neben der korrekten Lösung mehrere zumindest auf den ersten Blick auch wahrscheinliche Lösungen gefunden werden müssen (Hunt & Hassmén, 1994; Sangwin & Badger, 2010). Diese Mehrarbeit lohnt sich darum vor allem für Fragen, die in Tests mit vielen Teilnehmer/-innen zum Einsatz kommen oder mehrfach genutzt werden können.

Angesichts der beschriebenen Vor- und Nachteile lässt sich zusammenfassend sagen, dass geschlossene bzw. halb offene Aufgaben eher für den Bereich der Lernsteuerung auf dem Anspruchsniveau Erinnern, Verstehen und Anwenden geeignet sind, dass sie aber durch entsprechend höheren Aufwand auf Seiten der Autoren auch für das Messen komplexerer kognitiver Prozesse, wie Problemlösen, geeignet sein können (Dubs, 2008, S.270).

Für die Entwicklung von Aufgaben für das Selbststudium sollte je nach Lernziel und Anspruchsniveau über den Aufgabentyp entschieden werden. Darüber hinaus erscheint eine Mischung der Aufgabentypen zur Auflockerung des Lernprozesses als sinnvoll.

3.2 Geschlossene Online-Fragetypen

ILIAS bietet im Testobjekt unterschiedliche geschlossene Fragetypen an (siehe auch Ergebnisbericht Konzeptphase TP 3.1):

- *Multiple Choice*: Liste von Antwortoptionen, eine oder mehrere korrekte Antworten/ein oder mehrere Distraktor(en)
- *Single Choice*: Liste von Antwortoptionen, eine korrekte Antwort/ein oder mehrere Distraktor(en)
- *Zuordnung*: „Definitionen“ werden per Drag & Drop „Termen“ zugeordnet, eine oder mehrere korrekte Zuordnungen sind möglich.
- *Anordnung*: Bilder oder Textfragmente werden per Drag & Drop horizontal oder vertikal in eine bestimmte Reihenfolge gebracht, nur eine korrekte Reihenfolge ist möglich.
- *Hotspot/Imagemap*: Auf einer Abbildung/Grafik befinden sich klickbare Bereiche, einer davon ist die korrekte Antwort, dazu ein oder mehrere Distraktor(en)
- *Fehlertext*: Innerhalb eines Textes werden fehlerhafte Wörter gefunden und angeklickt, ein oder mehrere Fehler müssen gefunden werden.

Der Vorteil der Niederschwelligkeit von Multiple Choice (bzw. Single Choice)-Aufgaben wurde schon genannt; bei den genannten interaktiven Aufgabentypen wie Zuordnung, Anordnung oder Imagemap wird dieser zu einem gewissen Grad wieder aufgegeben und die „Bedienung“ der Aufgabe rückt stärker in den Vordergrund. Diese Fragetypen sollten darum nur zum Einsatz kommen, wenn das Format tatsächlich zur Fragestellung passt, bzw. wenn die Aufgabenstellung so leicht ist, dass eine Ablenkung durch den Fragetyp eher unwahrscheinlich ist.

Regeln für die Formulierung von geschlossenen Aufgabentypen

Die im Folgenden genannten Regeln sind weitgehend an die Vorschläge von Ebel und Frisbie (1991), Jacobs (2000) bzw. Gronlund (1998) angelehnt (siehe auch Dubs, 2008). Sie spiegeln Erfahrungswerte von Aufgabenautoren wider und sollen als Anregung zur Gestaltung von Aufgaben dienen - allerdings kann es auch immer wieder gute Gründe geben, gegen diese Regeln zu verstoßen!

Aufgabenstellung:

- Alle Aufgaben die im Rahmen des formativen eAssessment zum Einsatz kommen, sollen für Lernende im Selbststudium ohne zusätzliche Erklärungen verständlich und beantwortbar sein. Es ist darum bei der Formulierung darauf zu achten, dass die Aufgabenstellung klar, eindeutig und komplett ist. Die verwendete Sprache sollte so einfach wie möglich sein und es sollten keine verschachtelten Sätze, doppelte Verneinungen etc. vorkommen. Die eigentliche Fragestellung sollte hervorgehoben und von den einleitenden Erläuterungen abgesetzt werden.
- Pro Aufgabe sollte möglichst nur ein Lernziel adressiert werden. Für Lernende wird so der Bezug zu einem Themengebiet leichter nachvollziehbar und bei fehlerhaften Antworten können im automatisierten Feedback eher passende Hinweise auf möglicherweise bestehende Wissenslücken und weiterführende Lernmaterialien gegeben werden.

Lösung:

- Es existiert eine eindeutige Antwort auf die Aufgabenstellung und es existiert eine Musterlösung, die begründet, warum diese Antwort korrekt ist/wie sie erzielt wurde.
- Weitere „auch richtige“ Antworten sollten aus Gründen der didaktischen Klarheit vermieden werden, d.h. Single Choice-Aufgaben sind Multiple Choice-Aufgaben vorzuziehen.

Inhaltliche Aspekte bei der Formulierung der Auswahlmöglichkeiten:

- Die Unterschiede zwischen den Auswahlmöglichkeiten sollten deutlich sein, nicht in Nuancen versteckt.
- Es sollten keine Tricks oder Doppeldeutigkeiten in den Distraktoren versteckt sein (keine Fallen stellen).
- Es sollten keine allgemeinen Ausdrücke, wie „immer, nie, ausschließlich“, verwendet werden, diese können ein Hinweis darauf sein, dass eine Antwort falsch ist. Eine Ausnahme von dieser Regel ist das Abfragen von mathematischen Definitionen.
- Jede der Auswahlmöglichkeiten sollte, zumindest auf den ersten Blick, sinnvoll erscheinen, unsinnige oder zusammenhangslose Distraktoren („Kängurus“) sollten vermieden werden
- Die Formulierung „keine der genannten Möglichkeiten ist korrekt“ sollte vermieden werden, sie trägt entweder zur Verwirrung der Lernenden bei oder dokumentiert „die mangelnde Fantasie des Aufgabenkonstruktors im Hinblick auf das Er- bzw. Auffinden echter Distraktoren.“ (Jacobs, 2000).

Formale Aspekte bei der Formulierung der Auswahlmöglichkeiten:

- Die Liste der Auswahlmöglichkeiten sollte übersichtlich und gut lesbar sein, damit die Lernenden nicht eine der Optionen übersehen (weil sie z.B. scrollen müssen).
- Alle Auswahlmöglichkeiten sollten ungefähr gleich lang sein, damit die Lernenden nicht versuchen, anhand äußerlicher Faktoren auf die richtige Antwort zu schließen.
- Die Distraktoren sollten übersichtlich angeordnet sein, also z.B. bei einer Liste von Jahreszahlen aufsteigend sortiert. Wenn sich keine Reihenfolge anbietet, sollten die Distraktoren nach dem Zufallsprinzip geordnet sein.
- Die Anzahl der Distraktoren liegt meist zwischen zwei und vier. Das Hauptkriterium sollte die Güte der Distraktoren sein, d.h. vor einer „Verlegenheits-Auswahl“ sollte lieber auf den Distraktor verzichtet werden. Mehr als fünf Distraktoren sind zu viele, das Erfassen der Liste mit einem Blick ist dann nicht mehr möglich.

Bewertung:

- Es sollten keine negativen Punkte vergeben werden, aus Gründen der Fairness sollte die falsche Beantwortung nicht zu einer niedrigeren Punktzahl führen als das Überspringen einer Aufgabe.

3.3 Halb offene Online-Fragtypen

Der bekannteste halb offene Aufgabentyp ist der Lückentext, bei dem ein Begriff oder eine Zahl in ein Feld eingegeben werden muss. Die Bedienung ist ähnlich niederschwellig wie bei geschlossenen Formaten, ein Vorteil ist, dass die Rate-Wahrscheinlichkeit wegfällt. Ist die Lösung ein Wort, besteht allerdings die Gefahr, dass Testteilnehmer einen Begriff eingeben, der vom Autor nicht vorgesehen, aber dennoch eine korrekte Antwort auf die Frage ist. Diese Gefahr der Uneindeutigkeit ist bei numerischen Lücken nicht gegeben, Missverständnisse bei der Eingabe können aber dennoch entstehen, z.B. welcher Delimiter für Dezimalzahlen verwendet wird oder wie weit das Ergebnis gekürzt werden muss.

Beim Entwurf von halb offenen Aufgaben ist darum zu berücksichtigen, dass jede potentiell richtige Antwort auch in der Aufgabenstellung hinterlegt ist (z.B. alternative Schreibweisen, Reihenfolgen von Begriffen, Kürzungsregeln etc.). Oftmals muss auch hier eine „menschliche“ Nachkorrektur erfolgen.

Außer der numerischen Eingabe existieren für die Mathematik halb offene Fragetypen, bei denen auch Brüche und Formeln eingegeben werden können (durch entsprechende Syntax).

3.3.1 Numerische Eingabe

Im ILIAS-Testobjekt sind standardmäßig zwei Fragetypen vorgesehen, die die Eingabe von numerischen Werten erlauben:

- *Numerische Frage*: Eingabe eines numerischen Werts, mit der Möglichkeit einer vorgegebenen Wertespanne.
- *Lückentext*: Textlücken werden in einen bestehenden Text oder Kontext eingebaut. Sie können entweder aus einem Drop-Down-Menü bestehen oder müssen mit einem numerischen Wert bzw. einer Texteingabe gefüllt werden.

Die „numerische Frage“ und der „numerische Lückentext“ erlauben beide die Eingabe einer Zahl mit beliebig vielen Nachkommastellen (wobei als Konzession an die deutsche Schreibweise sowohl ein Punkt oder Komma als Delimiter gesetzt werden kann). Da andere Fragetypen die englische Schreibweise verlangen (siehe 3.3.3 ILIAS-Fragetyp STACK), wird in optes einheitlich ein Punkt als Delimiter gesetzt. Sowohl numerische Frage als auch Lückentext erlauben die Zuweisung eines Toleranzwertes (Ober- und Untergrenze), so dass potentielle Fehler beim Kürzen der Antwort vom Autor berücksichtigt werden können.

Die numerische Lückentextaufgabe hat den Vorteil, dass neben die Lücke eine Einheit (z.B. m²) geschrieben werden kann, so dass die Teilnehmer davon abgehalten werden, dies selbst zu tun. Auch erlaubt die numerische Lückentextaufgabe den Einsatz mehrerer Lücken in einer Aufgabe. Da die numerische Frage gegenüber der Lückentextaufgabe keinerlei Vorteile bietet, wird im optes-Projekt nur der numerische Lückentext für halb offene Online-Fragetypen eingesetzt.

Anmerkung: Lückentextaufgaben erlauben auch den Einsatz von „Drop-Down“-Lücken, in dem Fall sind sie wieder der Kategorie geschlossene Online-Fragetypen zuzuordnen.

3.3.2 ILIAS-Plug-In: Formelfrage

Als Erweiterung der Möglichkeiten, die die numerische Eingabe in ILIAS bietet, wurde das Plug-In „Formelfrage“ entwickelt. Aus Autorensicht erlaubt die Formelfrage die Verwendung von Variablen, es können also Zufallszahlen innerhalb bestimmter Wertebereiche definiert werden, so dass die Aufgabenstellung (und –lösung) bei jedem Abruf neu generiert wird (Dokumentation der Formelfrage auf www.ilias.de).

Die Formelfrage ist geeignet für Aufgaben zur Einübung von mathematischen Fertigkeiten, da durch relativ geringen Aufwand vielfach einsetzbare Aufgaben erstellt werden können. Die notwendige Syntax ist vergleichsweise schnell zu erlernen, bzw. durch Nutzung grafikfähiger Taschenrechner bekannt (ASCIIMATH). Dennoch sollte dieser Fragetyp nur bei erfahrenen Online-Lernenden zum Einsatz kommen.

Eine deutliche Einschränkung der Formelfrage für den Bereich des formativen eAssessment ist die fehlende Unterstützung des Feedback, d.h. es können keine Lösungswege generiert werden, die dann zu den Variablen der Aufgabenstellung passen. Ein weiterer Nachteil ist, dass es nicht möglich ist, die Variablen in der Aufgabenstellung vom Ergebnis aus zu berechnen. Dies führt dazu, dass für eine Aufgabe sowohl „glatte“ als auch „krumme“ Ergebnisse generiert werden und damit der Schwierigkeitsgrad deutlich schwanken kann. Da die Berechnung der Lösung nicht auf einem Computer Algebra System basiert, sondern die mathematischen Funktionen des Web-Servers ausnutzt, ist die Formelfrage außerdem relativ fehleranfällig.

3.3.3 ILIAS-Plug-In: STACK

Ein deutlich leistungsfähigeres Tool ist das Plug-In STACK¹, über das eine Anbindung des LMS an ein Computer Algebra System (CAS) erfolgt². Dies ermöglicht den Einsatz von Zufallszahlen in Mathematik-Aufgaben, so dass die Variablen bei jedem Aufruf der Aufgabe neu generiert werden. Die Rückmeldung (richtig/falsch/Lösungsweg etc.) basiert dann auf der Eingabe des Teilnehmers und der „zufällig“ erzeugten Variable (die Zufallszahlen können auf einen bestimmten Wertebereich eingeschränkt werden).

Aus Lernersicht ermöglicht STACK die Eingabe von Termen und Formeln via ASCIIMATH Syntax, verbunden mit einer direkten Vorschau (und der Möglichkeit zur Korrektur dieser Eingabe). Da die Werte einer Aufgabe immer wieder neu berechnet werden, sind STACK Aufgaben ideal für den Einsatz im Übungsbereich (Sangwin, 2010; Sangwin, 2012). Vielen Studienanfänger/-innen ist die ASCIIMATH Syntax von der Arbeit mit grafikfähigen Taschenrechnern her bekannt, allerdings kann dies nicht vorausgesetzt werden. CAS-Fragen sollten darum nicht kommentarlos bzw. ohne Übungsphase eingesetzt werden; für den Einstieg in das Lernprogramm, beispielsweise in diagnostischen Selbsttests, sind sie darum eher ungeeignet.

3.3.4 ILIAS-Objekt: FormATest

Das ILIAS Objekt FormATest erlaubt das schrittweise Prüfen einzelner Aufgabensequenzen. Eine detaillierte Dokumentation wird von Teilprojekt 1 Propädeutik erstellt (in Erstellung, Veröffentlichung im Materialienpool auf optes.de).

¹ STACK-Autor: Chris Sangwin (Loughborough University); STACK-Moodle Plug-In: Tim Hunt (Open University); STACK-ILIAS Plug-In: Fred Neumann, Jesus Copado (Universität Erlangen)

² STACK nutzt das Open-Source-CAS MAXIMA, <http://maxima.sourceforge.net>

3.3.5 Regeln für die Formulierung von halb offenen Aufgabentypen

Für die Aufgabenstellung und Lösung gelten die gleichen Regeln wie für geschlossene Aufgabentypen. Die Aufgaben sollten möglichst nur ein Lernziel adressieren, klar formuliert sein und eine eindeutige Lösung haben, für die ein Lösungsweg existiert. Da die Eingabe der Lösung von den Lernenden selbst vorgenommen wird, müssen potentielle Fehler nicht im gleichen Maße antizipiert werden wie bei geschlossenen Aufgaben. Dennoch sollten Aufgabenautoren sich darüber Gedanken machen, ob bei der Eingabe Missverständnisse entstehen könnten.

- Bei der numerischen Eingabe können Rundungsfehler entstehen. Wenn das Ergebnis eine nicht „runde“ Zahl ist, kann eine Toleranz eingeräumt werden.
- Es sollte deutlich auf die Maßeinheit hingewiesen werden, in der die Lösung anzugeben ist.
- Es sollte auf mögliche Syntax-Eingabefehler hingewiesen werden, wie z.B. Eingabe von Leerzeichen, Kommata, fehlende Multiplikations-Zeichen (z.B. $4x$ anstatt $4*x$) etc.
- Die Eingabe von Syntax sollte geübt werden können, bevor eine Aufgabe/ein Test abgegeben wird.

3.4 Offene Online-Fragtypen

Aktuell existieren in ILIAS zwei Möglichkeiten, einen mathematischen Text „abzugeben“:

- *Freitext*: Freie Texteingabe per LaTeX.
- *Datei-Upload*: Entweder eines handschriftlich erstellten und dann gescannten Dokuments oder eines Dokuments, das unter Verwendung eines Formel-Editors erstellt wurde (z.B. Word).

Beide Verfahren sind keineswegs niederschwellig und erfordern einen deutlich größeren Zeitaufwand als die „traditionelle“ Papiervariante. Auch aus Sicht der Lehrperson ist die Korrektur deutlich aufwändiger, da Kommentare entweder über ein PDF-Dokument erstellt oder von Hand in einen Ausdruck des Lösungswegs eingefügt, dann wieder eingescannt und an jede/n Teilnehmer/-in verschickt werden müssen.

Trotz des erheblichen Aufwands kann das didaktische Design die Bewertung eines längeren mathematischen Texts, z.B. eines Beweises oder eines detaillierten Lösungswegs, beinhalten. Dies wird im Rahmen des erweiterten eMentoring-Konzepts beschrieben.

3.4.1 Regeln für die Formulierung von offenen Online-Fragetypen

Offene Aufgaben verlangen ironischerweise sehr viel mehr Erläuterungen und Vorgaben als geschlossene Fragetypen. Damit die Lernenden der Aufgabenstellung gerecht werden können, müssen sie genau darüber informiert werden, was inhaltlich und formell von ihnen erwartet wird. Der Arbeitsauftrag muss dabei so eindeutig wie möglich formuliert sein und alle Anforderungen bezüglich Sprache, Gestaltung, Umfang, die in die Bewertung einfließen, sollten kommuniziert werden.

Zusatzinformationen können eine Aufgabe interessanter machen, erhöhen aber auch die Komplexität und Bearbeitungszeit. Je länger die Aufgabenstellung, desto anspruchsvoller ist auch die Aufgabe (vgl. Dubs, 2008, S. 270). Auch für offene Aufgabenstellungen sollte eine Musterlösung existieren, ggf. kann diese durch besonders gelungene Beispiele von Lernenden ergänzt werden.

3.5 Fazit zu Online-Aufgabentypen

„It is ironic that some of the more objective subject and sector areas bring with them particular technical challenges when e-assessment solutions are considered.“ (Craven, 2009), S.3)

Es wurde gezeigt, dass der Einsatz von Online-Tests im Bereich der mathematischen Grundlagen technisch keineswegs trivial ist. Bei der Entscheidung, welcher Aufgabentyp zum Einsatz kommt, sind neben dem Lernziel, sowie dem Komplexitäts- und Schwierigkeitsgrad der Aufgabenstellung darum auch die meta-kognitiven Anforderungen zu berücksichtigen, die durch das technische Handling entstehen können.

Die eher unflexiblen Multiple Choice-Aufgaben sind aufgrund ihrer Niederschwelligkeit bei der Eingabe für unerfahrene eLearning-Teilnehmer/-innen von Vorteil – und für viele Lerninhalte des Mathematikvorkurses vollkommen ausreichend (z.B. das Beherrschen mathematischer Routinen). Bei entsprechender Gestaltung können geschlossene Aufgaben zu Testergebnissen führen, die mit offenen Fragetypen vergleichbar sind (z.B. Martinez, 1999; Kastner & Stangl, 2010). Der Aufwand bei der Erstellung von geschlossenen Fragen kann allerdings deutlich höher sein, beispielsweise wenn jedem Distraktor ein Fehlerfeedback zugewiesen wird (siehe Abschnitt 4).

Halb offene Formate, bei denen ein mathematischer Ausdruck eingegeben werden muss (z.B. STACK), erfordern deutlich mehr Zeit (und/oder Übung) bei der Eingabe der Lösung. Sie sollten nur bei erfahrenen Lernern zum Einsatz kommen, für diagnostische Einstiegstests sind sie weniger gut geeignet.

Offene Aufgaben, die die Abgabe ganzer Rechenschritte oder mathematischer Lösungsansätze verlangen, erfordern die individuelle Korrektur und Rückmeldung einer Lehrperson, darum sind sie für das reine Online-Selbststudium nicht geeignet. Im Rahmen betreuter Online-Vorkurse ist aber auch der Einsatz offener Aufgaben denkbar und sinnvoll (siehe Abschnitt 5.3.3).

Konkrete Hinweise zur Gestaltung von Aufgaben im optes-Projekt finden sich in Kapitel 6.

4 Feedback

Das automatisch generierte Feedback auf eine Online-Aufgabe kann unterschiedliche Informationen umfassen, die verschiedene Dimensionen der Selbstreflexion adressieren. Grundsätzlich lässt sich zwischen „Verifikationsfeedback“ und „Elaborationsfeedback“ unterscheiden (Renkl, 1991):

4.1 Verifikationsfeedback (Richtig/Falsch)

Das Verifikationsfeedback signalisiert den Lernenden, ob die Aufgabe richtig, falsch, teilweise richtig oder gar nicht beantwortet wurde (Mayer et al., 2009). Während eine positive Rückmeldung („richtig“) meist keiner weiteren Ausführungen bedarf, zielt die negative Rückmeldung darauf, „die Persistenz des Fehlers“ zu vermeiden (Renkl, 1991, S. 52) und wird darum meist durch „anschließende, elaborative und remediale Maßnahmen“ (ebd., S. 51) ergänzt:

4.2 Elaborationsfeedback (Erweitertes Feedback)

Die folgenden Bestandteile eines ausführlichen Feedback auf eine Aufgabe (oder einen Test) sind mit den Standardmitteln des LMS ILIAS umsetzbar.

Lösung (Erwartete Antwort)

Wurde die Antwort auf eine Aufgabe als „richtig“ gewertet, ist die erwartete Antwort bekannt. Wurde sie als „falsch“ gewertet, interessiert die Lernenden, welche Antwort die richtige ist bzw. vom System erwartet wurde.

Musterlösung

Damit die Lernenden nachvollziehen können, warum die erwartete Antwort die richtige ist, wird der/ein möglicher Lösungsweg aufgezeigt, z.B. die einzelnen Rechenschritte, die zur erwarteten Antwort führen.

Kommentar

Der Kommentar zu einer Aufgabe kann unterschiedliche Informationen enthalten und zu unterschiedlichen Zeitpunkten der Bearbeitung angeboten werden. So ist es z.B. möglich, einem Lernenden schon vor der Abgabe einen Tipp zur Lösung der Aufgabe zu geben, falls dieser nicht mehr weiter weiß. Nach Abgabe der Antwort kann ein Kommentar abgestimmt auf die richtige oder falsche Antwort des Lernenden erfolgen.

Bei unterschiedlichen vorhersehbaren Falschantworten (Distraktoren bei Multiple Choice-Aufgaben) kann für jeden dieser Fehler ein passender Kommentar formuliert werden. So kann auf Rechenfehler oder Verständnisprobleme eingegangen werden, die sich häufig wiederholen. Die Erstellung von Fehlerfeedbacks ist allerdings sehr aufwändig und sollte von Autoren mit Lehrerfahrung geleistet werden (Lissitz & Hou, 2008).

Basierend auf einem kompletten Testergebnis kann der Kommentar Interpretationsvorschläge liefern (fehlende Kenntnisse in einem Wissensgebiet) und Lernempfehlungen oder Links zu weiteren Aufgaben anbieten (Sangwin & Badger, 2010).

Bewertung

Bei der Bewertung des Ergebnisses wird ein Maß oder eine Bezugsnorm herangezogen, die den Lernenden spiegelt, wo sie im Verhältnis zur Erwartung der Autoren, zur eigenen vorhergegangen

Leistung oder zur Leistung anderer stehen (Kluger & DeNisi, 1996; Rheinberg, 2002). Meist erfolgt die Bewertung über die Vergabe einer Punktzahl („Sie haben 3 von 10 Punkten erreicht“) oder ein Notenschema, das den Lernenden aus der Schule vertraut ist. Diese Art der Bewertung wird oft kritisch gesehen, da die Aufmerksamkeit auf die individuelle Leistung gelenkt wird und damit andere Bestandteile des Feedback wie Lösungsweg, Hilfestellung oder Lernempfehlung in den Hintergrund rücken (Brennan, 2006; Black et al., 2003; Gibbs & Simpson, 2004). Die Fokussierung auf die Bewertung bestärkt Lernende mit starker Leistungsorientierung („performance goals“) in ihrer Haltung, dass ein gutes Ergebnis bzw. eine gute Note Ziel des Lernprozesses sind, die Lernziele rücken in den Hintergrund (Dweck, 1986). Insbesondere schwächere Schüler/-innen können durch eine negativ empfundene Bewertung demotiviert werden. Im Schulkontext wird darum oft vorgeschlagen, im Bereich des formativen Assessment auf Noten oder Punktzahlen zu verzichten und stattdessen individuelles Feedback in Form von Kommentaren oder Anmerkungen zu geben (Black et al., 2003).

Dem entgegen steht der Ansatz, erwachsene Lernende in die Gestaltung des Lernprozesses und damit auch in die Bewertung des Lernerfolgs einzubeziehen (Nicol & Macfarlane-Dick, 2006). Wichtig ist, dass das Feedback keine wertenden Aussagen über die Fähigkeiten der Lernenden beinhaltet: Sie sollten nicht verunsichert oder in ihrem Selbstwertgefühl angegriffen werden („threatening self-esteem“), aber auch übertriebenes Lob („praise“) im Feedback hat sich als wenig lernförderlich gezeigt (Kluger & DeNisi, 1996). Eine sachliche und übersichtliche Darstellung der Testergebnisse inklusive Punktzahlen und Interpretationshilfen scheint darum für Lernende im Selbststudium angemessen.

4.3 Fazit zu Feedback

Bei der Bearbeitung von Online-Aufgaben erwarten die Lernenden eine schnelle und eindeutige Rückmeldung (Kvale, 2007). Das Aufgabenfeedback soll sie aber neben der simplen Rückmeldung „richtig“ oder „falsch“ auch in die Lage versetzen, Fehler nachzuvollziehen und den Sinn der Aufgabenstellung zu erfassen. Gerade im Selbststudium können Missverständnisse und Fehlkonzeptionen sonst schnell zum Abbruch der Lernhandlungen führen.

Auf der anderen Seite fordert auch das Lesen und Verstehen von Feedback Konzentration und kann sehr zeitintensiv sein. Umfangreiche Lernhilfen sind sicherlich wenig hilfreich, wenn sie nicht gelesen werden (Nix & Wyllie, 2011). Bei der Gestaltung des Feedback sollte also abgewogen werden, zu welchem Zeitpunkt im Lernprozess welche Informationen besonders wichtig sind.

Im optes-Projekt wird unterschieden zwischen *diagnostischem* Feedback auf Tests, die zur Erfassung des Lernstandes zu Beginn oder zum Abschluss des Lernprozesses stehen und *formativem* Feedback, das sich auf die Bearbeitung von einzelnen Übungsaufgaben bezieht.

	Diagnostisches Feedback auf einen Test	Formatives Feedback auf eine Aufgabe
<i>Verifikationsfeedback</i> (Richtig/Falsch)	x	x
<i>Erweitertes Feedback</i>		
- Lösung (erwartete Antwort)		x
- Musterlösung		x
- Kommentar	x	x
- Bewertung	x	

Tab. 1 Feedback-Varianten im Lernprozess

Beim diagnostischen Feedback liegt der Fokus auf der Kalibrierung, d.h. der Einordnung der Kenntnisse eines Lernenden im Verhältnis zu den Lernzielen des Vorkurses. Die detaillierte Fehleranalyse ist hierbei weniger wichtig als das (An-)Erkennen von Wissensdefiziten (so denn vorhanden). Darum werden für diagnostische Selbsttests Gesamt-Ergebnisse interpretiert und als Basis für Lernempfehlungen genutzt.

Beim formativen Feedback steht dagegen die Fehleranalyse und damit das Verständnis eines einzelnen mathematischen Problems im Vordergrund. Neben dem Verifikationsfeedback ist darum eine detaillierte Musterlösung unerlässlich. Wenn sinnvoll und möglich, wird das Fehlerfeedback zu einzelnen Distraktoren durch Kommentare ergänzt.

Anmerkung: In Kapitel **Fehler! Verweisquelle konnte nicht gefunden werden.** werden konkrete Hinweise zur Gestaltung des Feedback in ILIAS gegeben.

5 Testkonzept im Projekt optes

5.1 Curriculum

Das Curriculum des optes-Vorkurses ist am Gymnasial-Lehrplan der Mittel- und Oberstufe orientiert. Dabei wurden Themengebiete berücksichtigt, bei denen Studienanfänger/-innen erfahrungsgemäß besonders häufig Schwierigkeiten haben (wie z.B. Termumformungen) und die Bestandteil der Lehre im ersten und zweiten Semester sind (wie Grenzwerte und Stetigkeit). Inhaltlich wurden die Aufgaben mit dem Vorschlag eines gemeinsamen Mathematik-Curriculums der „cosh“- Gruppe abgestimmt, das als Referenz für Lehrer und Hochschuldozenten in Baden-Württemberg dienen soll (siehe [cosh cooperation schule:hochschule, 2014](#)).

Auf Basis der relevanten Lerninhalte wurden zehn Kategorien gebildet, die als thematische Grundstruktur für die Sammlung und Klassifizierung der Aufgaben dienen:

Mathematische Kategorie	Präfix
1 Arithmetik	ari
2 Algebraische Gleichungen und Ungleichungen	gle
3 Potenzen, Wurzeln, Logarithmen	pwl
4 Funktionen	fun
5 Geometrie	geo
6 Trigonometrie	tri
7 Logik und Kombinatorik	log
8 Lineare Algebra	lal
9 Reelle Zahlen	ree
10 Grenzwerte und Stetigkeit	gre

Tab. 2 Strukturierung der mathematischen Themengebiete

Dabei erhält jede Aufgabe eine Bezeichnung, durch die sie eindeutig einem Themengebiet zugeordnet wird, z.B. „ari_ter_fak“ für eine Aufgabe im Bereich der Arithmetik, und dort in der Kategorie Termumformungen, wobei das Faktorisieren verlangt wird. Auch wenn eine eindeutige Abgrenzung zwischen den Kategorien oft nicht möglich ist, da viele Aufgaben mehr als ein mathematisches Verfahren adressieren, sind die Autor/-innen angehalten, sich auf eine Zuordnung festzulegen (und ggf. eine zweite in einer Übersichtsliste zu vermerken) (vgl. Aufgabenübersicht im optes-[Materialienpool](#) und Aufgabengestaltung in Kapitel 6.3.1). Die Zusammenstellung von Tests oder Übungsaufgaben innerhalb der Lernmodule erfolgt dann auf Basis dieser Kategorien.

5.2 Lernprozessinitiiierende Leistungserfassung: Diagnostischer Einstiegstest

Zu Beginn des Lernprozesses steht die Orientierung im System im Vordergrund, die Teilnehmer/-innen benötigen Informationen über Sinn und Ziel der Lernaktivitäten, über Art und Umfang der Lerninhalte, und natürlich darüber, ob sie in einem oder mehreren dieser Themen überhaupt Nachholbedarf haben (vgl. Lawson, 2003). Der diagnostische Einstiegstest dient zur Kalibrierung (Winne, 2004) und zur Vor-Strukturierung des Lernprozesses. Als *lernprozessinitiiierende Leistungserfassung* bezieht er sich auf das gesamte Themenspektrum des Selbstlernprogramms und wird nur einmal durchgeführt.

Jedes Themengebiet ist durch mehrere Aufgaben repräsentiert, damit Flüchtigkeitsfehler bei der Erstellung der Lernempfehlung nicht zu stark gewichtet werden. Um eine möglichst passgenaue Lernempfehlung zu erzielen, wird jede Aufgabe *einem* mathematischen Themengebiet zugeordnet d.h. bei der Lösung einer Aufgabe soll idealerweise nur ein mathematisches Prinzip zum Einsatz kommen (z.B. Geometrie: Strahlensatz). Die Aufgaben weisen darum wenig Komplexität auf und sollten relativ schnell zu lösen sein. Auch angesichts des großen Umfangs der mathematischen Inhalte sollte die durchschnittliche Bearbeitungszeit einer Aufgabe zwei Minuten nicht übersteigen.

Um die Überbewertung von aufgabenspezifischen oder technischen Einflüssen zu minimieren, kommen im diagnostischen Einstiegstest ausschließlich niederschwellige Online-Aufgabentypen zum Einsatz, die ohne Eingabe von Syntax beantwortet werden können, also Multiple Choice bzw. numerische Eingabe.

1	Zeitpunkt	Beginn des Selbstlernprogramms
2	Umfang	77 Aufgaben (basierend auf Curriculum)
3	Schwierigkeit	25% leicht, 50% mittel, 25% schwer
4	Ziel	Information über Curriculum/Kalibrierung/Lernempfehlung
5	Feedback	Erreichte Gesamtpunktzahl
		Interpretation des Ergebnisses: Kenntnisse pro Themengebiet vorhanden/nicht vorhanden
		Lernempfehlung: ein oder mehrere LoKs
6	Wiederholung	keine, wird nur einmal durchgeführt

Tab. 3 Merkmale des Diagnostischen Einstiegstests im optes-Vorkurs

Auf der Basis des Testergebnisses erhalten die Teilnehmer/-innen eine Bewertung pro Themengebiet, außerdem Kommentare in Form von Lernempfehlungen und Links zu den entsprechenden Lerninhalten. Beim diagnostischen Feedback steht die Übersichtlichkeit im Vordergrund, d.h. die Lernenden sollen sich einen allgemeinen Überblick verschaffen. Darum wird auf die Ausgabe von Verifikationsfeedback sowie Musterlösungen verzichtet. Stattdessen erhalten die Teilnehmer/-innen Links zu Lerninhalten und Hinweise zur Priorisierung, d.h. es werden die Lernzielorientierten Kurse besonders empfohlen, die zum Kerncurriculum gehören und zuerst bearbeitet werden sollten.

Trotz des Versuchs, das Feedback übersichtlich und einfach zu gestalten, kann es bei manchen Lernenden zu einem Gefühl der Überforderung kommen. Gerade Studienanfänger/-innen mit geringer Fähigkeit zu selbstreguliertem Lernen könnten zusätzliche Unterstützung bei der Planung des Lernprozesses benötigen, die von einem eLearning-System nicht geleistet werden kann (z.B. Azevedo, 2005; Artino & Stephens, 2009). D.h. an dieser Stelle sollte es möglich sein, Kontakt zu einem eMentor aufzunehmen und Beratung bezüglich des weiteren Vorgehens einzuholen.

5.3 Lernprozessbegleitende Leistungserfassung: Üben und Anwenden

Im Verlauf des Lernprogramms erhalten die Lernenden immer wieder die Möglichkeit, Übungsaufgaben oder Kurztests zu einem Themengebiet zu bearbeiten. Hier kommt dann auch der Einsatz komplexerer Online-Fragetypen, wie z.B. STACK, in Betracht, die gleichzeitig auf die Nutzung von Computer Algebra Systemen im Studium vorbereiten können.

5.3.1 Übungsaufgabe

Generell unterscheidet sich das formative Feedback innerhalb der Lernmodule vom diagnostischen Feedback zu Beginn/Abschluss des Vorkurses durch die Angabe von detaillierten Lösungswegen für jede Aufgabe. Hinzu kommen Interpretationshilfen in Bezug auf Rechenfehler, z.B. bei Distraktoren einer Multiple Choice-Frage.

1	Zeitpunkt	Innerhalb eines Lernmoduls in einem LoK (Lernzielorientierter Kurs)
4	Ziel	Aktivierung, Übung, Verständniskontrolle (On-the-Fly Formative Assessment)
	Feedback	Verifikationsfeedback (richtig/falsch) für jede Aufgabe
		<i>eigene Antwort/erwartete Lösung für jede Aufgabe*</i>
		<i>Musterlösung pro Aufgabe, ggf. Hinweis auf mögliche Rechenfehler *</i>
		Lernempfehlung: weitere Übungsaufgaben
6	Wiederholung	Unbegrenzt möglich

Tab. 4 Merkmale einer Übungsaufgabe im optes-Vorkurs

* Planung, in der aktuellen optes-Version noch nicht umgesetzt

5.3.2 Übung (Übungstest)

Eine Übung dient zur Festigung der Lerninhalte aus einem Lernmodul. Der Umfang sollte bei ca. 10 Aufgaben liegen, die mehrfach bearbeitet werden können. Zu jeder Aufgabe erhalten die Teilnehmer/-innen ein erweitertes Feedback sowie eine Musterlösung. Im Komplexitätsgrad adressieren diese Aufgaben unterschiedliche Niveaus, auch CAS-Items können hier zum Einsatz kommen. Nach Bearbeitung der Übung wird der qualifizierende Tests für dieses Lernziel freigeschaltet.

1	Zeitpunkt	Nach Bearbeitung eines Lernmoduls im LoK (Lernzielorientierter Kurs)
2	Umfang	Ca. 10 Aufgaben
3	Schwierigkeit	Leicht bis schwer
4	Ziel	Aktivierung, Übung, Verständniskontrolle, Freigabe qualifizierender Test
5	Feedback	Erreichte Gesamtpunktzahl
		Verifikationsfeedback (richtig/falsch) für jede Aufgabe
		<i>Eigene Antwort/erwartete Lösung für jede Aufgabe</i>
		<i>Musterlösung pro Aufgabe, ggf. Hinweis auf mögliche Rechenfehler</i>
6	Wiederholung	Unbegrenzt möglich

Tab. 5 Merkmale einer Übung (=Übungstest) im optes-Vorkurs

5.3.3 Einreichaufgabe (optional)

Es wurde schon erwähnt, dass geschlossene oder halb offene Online-Aufgabenformate nicht geeignet sind, komplexe Denkprozesse, wie z.B. Beweisführung oder Problemlösen, zu dokumentieren und zu bewerten. Wenn an einem Standort die Erweiterung des Kursangebots um Aufgaben von höherer Komplexität gewünscht ist und dementsprechend Lehrpersonal zur Korrektur zur Verfügung steht, sind offene Aufgaben eine sinnvolle Erweiterung.

1	Zeitpunkt	Parallel zur Bearbeitung eines Lernmoduls in einem LoK (Lernzielorientierter Kurs)
4	Ziel	Aktivierung, Übung, Verständniskontrolle
5	Feedback	Individuelle Rückmeldung durch eMentor
		Hinweise/Tipps für mögliche Lösungsansätze
		Verifikationsfeedback (richtig/falsch)
		Musterlösung pro Aufgabe, ggf. Hinweis auf mögliche Rechenfehler
	Lernempfehlung: weitere Übungsaufgaben	
6	Wiederholung	Keine, Abgabe und Feedback im Rahmen des Kurs-Zeitplans

Tab. 6 Merkmale einer Einreichaufgabe

5.4 Lernerfolgskontrolle

5.4.1 Qualifizierende Tests

Für jedes Lernziel eines LoK (Lernzielorientierter Kurs) wird ein qualifizierender Kurztest zur Lernzielerreichung angeboten. Dieser Test enthält tendenziell eher mittlere bis schwere Aufgaben und kann mehrfach durchgeführt werden.

1	Zeitpunkt	Nach Bearbeitung eines Lernmoduls und einer Online-Übung im LoK (Lernzielorientierter Kurs)
2	Umfang	5-8 Aufgaben
3	Schwierigkeit	Mittel bis schwer
4	Ziel	Information über Lernerfolg in Bezug auf das Lernziel
5	Feedback	Erreichte Gesamtpunktzahl
		<i>Verifikationsfeedback (richtig/falsch) für jede Aufgabe*</i>
		Rigene Antwort/erwartete Lösung für jede Aufgabe
		<i>Musterlösung pro Aufgabe, ggf. Hinweis auf mögliche Rechenfehler*</i>
		Interpretation des Ergebnisses: Anzahl erreichter/nicht erreichter Lernziele
6	Wiederholung	Unbegrenzt möglich

Tab. 7 Merkmale eines qualifizierenden Tests im optes-Vorkurs

* Planung, in der aktuellen optes-Version noch nicht umgesetzt

5.4.2 Zentraler Abschlusstest (optional)

Die Lernerfolgskontrolle im Abschlusstest ist optional, sie wird nicht von allen beteiligten Standorten gewünscht. Analog zum zentralen diagnostischen Einstiegstest umfasst der zentrale Abschlusstest alle zu behandelnden mathematischen Themengebiete (= Kontrolltest). Der zentrale Abschlusstest markiert das Ende des Vorkurses/den Beginn des Studiums. Um die Wichtigkeit der Mathematik für das Studium zu unterstreichen (und um evaluierbare Daten zu erhalten) kann es sinnvoll sein, diesen Test in der ersten Woche des ersten Semesters durchführen zu lassen, beispielsweise in den PC-Räumen der Hochschule.

Über diesen Test können dann auch die Vorkenntnisse der Studienanfänger/-innen erfasst werden, die am Vorkurs nicht teilgenommen haben.

1	Zeitpunkt	Nach Bearbeitung des Selbstlernprogramms
2	Umfang	45 Aufgaben (basierend auf Curriculum)
3	Schwierigkeit	25% leicht, 50% mittel, 25% schwer
4	Ziel	Information über Lernerfolg in Bezug auf den Vorkurs
5	Feedback	Erreichte Gesamtpunktzahl
		Interpretation des Ergebnisses: Kenntnisse pro Themengebiet vorhanden/nicht vorhanden
		Lernempfehlung: ein oder mehrere LoKs, gegebenenfalls weitere Maßnahmen, etwa studienbegleitende Tutorien
6	Wiederholung	Keine, wird nur einmal durchgeführt

Tab. 8 Merkmale eines zentralen Abschlusstests (Posttest)

6 Hinweise zur Gestaltung von Aufgaben

Die folgenden Hinweise zur Gestaltung von Aufgaben in ILIAS basieren auf den Empfehlungen der optes-internen Arbeitsgruppe „Aufgaben- und Testdesign“ (Abschlussdokumentation vom Oktober 2014, Überarbeitung für ILIAS 5 im Februar 2016).

6.1 Test-Startseite

Vor jedem Test sollten allgemeine Angaben zur Verwendung der Syntax gemacht werden.

Bitte beachten Sie bei der Eingabe von Zahlenwerten:

- 1) Dezimalzahlen: Dezimalzahlen können mit Komma *oder* Punkt eingegeben werden (2.75 oder 2,75).
- 2) Für mehr als dreistellige Zahlen geben Sie bitte keinen Punkt als Trennzeichen ein (also 1500 anstatt 1.500)
- 3) Negative Werte geben Sie mit einem Minus-Zeichen vor der Zahl ein (also -3)

Anmerkung: Anders als die Standard ILIAS Fragetypen lässt der Fragetyp STACK keine Eingabe von Komma als Delimiter zu. Kommen STACK-Aufgaben in einem Test zum Einsatz, wird empfohlen die Aussage in Punkt 1 umzuformulieren („Dezimalzahlen: Geben Sie bei Dezimalzahlen einen Punkt als Trennzeichen ein (2.75 und *nicht* 2,75“). Beim Einsatz von STACK-Fragen sollte außerdem ein Hinweis auf die von Teilprojekt 1 bereitgestellte Handreichung [Eingabesyntax in Aufgaben](#) sowie den [Übungstest zur Eingabesyntax](#) erfolgen.

6.2 Test-Einstellungen

Zur leichteren Navigation im Test werden in der Fragenliste alle Items angezeigt (Testeinstellungen: „Fragenübersicht anzeigen“).

The screenshot shows the ILIAS test interface for a course titled 'Optimierung der Selbststudiumsphase'. The breadcrumb trail is: Startseite » Angebote für Studienanfänger » Kurse » Quadratische Gleichungen » LZ 1: Unvollständige quadratische Gleichungen. The current test is 'LZ 1: Unvollständige quadratische Gleichungen'. The question list on the left shows 'Frage 2' selected. The question text is: 'Der Bremsweg eines Autos ist proportional zum Quadrat der Geschwindigkeit, mit der das Auto zu Beginn des Bremsvorgangs fährt. Ein Auto hat bei einer Geschwindigkeit von 100 km/h einen Bremsweg von 70 m. Wie schnell war das Auto unterwegs, wenn der Bremsweg 116 m beträgt? Hinweis: Geben Sie die Antwort als Dezimalzahl an und runden Sie, falls nötig, auf zwei Nachkommastellen. Das Auto war km/h schnell.' The interface includes navigation buttons like 'Zurück', 'Test unterbrechen', 'Fragenübersicht', and 'Weiter'.

Abb. 1 Screenshot eines qualifizierenden Tests im LoK „Quadratische Gleichungen“

6.3 Aufgabengestaltung

6.3.1 Bezeichnungen

Bei der Verwaltung der Aufgaben im Aufgabenpool wird der Aufgabentitel, z.B. ari_bru_100, immer sowohl in das Feld "Titel" als auch in das Feld "Beschreibung" eingetragen. Beim Verwenden einer Aufgabe in einem Test kann dann im Feld "Titel" ein alternativer Name verwendet werden, im Feld "Beschreibung" bleibt aber die ursprüngliche Bezeichnung stehen (die Vorlage für die Vergabe von Dateinamen findet sich im optes-[Materialienpool](#), siehe auch Übersicht über das Curriculum in Kapitel 5.1).

Abb. 2 Titelvergabe und Beschreibungstext nach mathematischer Kategorie (Autorensicht, Reiter „Frage bearbeiten“)

6.3.2 Typografie

- Als Standardtext wird einfacher Text ohne Auszeichnung verwendet. Die eigentliche Aufgabenstellung (z.B. Wie groß ist, Berechnen Sie, etc.) wird immer **fett** gesetzt, außerdem wird sie durch einen Absatz vom Text davor und danach (bei Lückentext) getrennt.
Ausnahme: Formeln/Tex-Inhalte werden *nicht* fett gesetzt.
- Eine Frage wird immer mit einem Fragezeichen abgeschlossen, eine Aufforderung mit einem Punkt oder einem Doppelpunkt (was passender ist).
- Abbildungen werden in die Aufgabenstellung eingebunden (blaues Feld), da sie sonst nicht mit-exportiert werden (bei Animationen ist das leider nicht möglich).

Zur Altersbestimmung organischer Materialien wird die ^{14}C -Methode verwendet. Dabei nutzt man aus, dass das Carbon-14-Isotop ^{14}C radioaktiv mit einer Halbwertszeit von 5730 Jahren ist (d.h. nach 5730 Jahren ist genau die Hälfte des ursprünglichen Materials zerfallen).

In einer lebenden Pflanze ist das Verhältnis zwischen radioaktivem ^{14}C und nicht radioaktivem ^{12}C immer konstant, in einer toten Pflanze nimmt der Anteil des ^{14}C -Isotops durch radioaktiven Zerfall konstant ab, während sich der ^{12}C -Anteil nicht ändert. Aus dem gemessenen Verhältnis lässt sich dann etwa zurückrechnen, wann ein Baum gefällt wurde.

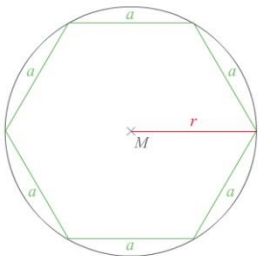


Bei Ausgrabungen wird ein Holzgegenstand gefunden, bei dem seit der Fällung des Baumes bereits 61 % der ursprünglich vorhandenen ^{14}C -Menge zerfallen ist.

Vor wie vielen Jahren t wurde der Baum, aus dem der Gegenstand gefertigt wurde, geschlagen?
(Runden Sie Ihr Ergebnis auf ganze Jahre)

$t \approx$ Jahre

Abb. 3 Beispiel für eine Aufgabe mit Abbildung und numerischer Eingabe



Ein *reguläres* n -Eck ist ein n -Eck, in dem alle Seiten die gleiche Länge a haben.

Es kann dadurch gewonnen werden, dass an einem Kreis mit Radius r die Kreislinie in n gleich lange Teile geteilt wird und benachbarte Endpunkte dieser Teilstücke durch ein Geradenstück miteinander verbunden werden.

Welche der folgenden Formeln ist richtig?

- $a = r \sin\left(\frac{360^\circ}{n}\right)$
- $a = r \sin\left(\frac{180^\circ}{n}\right)$
- $a = 2r \sin\left(\frac{180^\circ}{n}\right)$
- $a = \sqrt{\frac{r^2}{2}}$
- $a = \frac{2r\pi}{n}$

Abb. 4 Beispiel für eine Multiple Choice-Aufgabe mit Animation und LaTeX-Formeln in den Distraktoren

Wir setzen $r = 5^{2014} - \left(\frac{1}{5}\right)^{2014}$, $s = 5^{2014} + \left(\frac{1}{5}\right)^{2014}$

Berechnen Sie:

$r^2 - s^2 =$

Abb. 5 Beispiel für eine Aufgabe mit LaTeX-Formel und numerischer Eingabe

6.3.3 LaTeX-Eingabe

Ab ILIAS 4.4 werden html-Eingabeformate (z.B. [tex][[/tex]) vom System zwar noch akzeptiert, mittelfristig aber programmierseitig nicht mehr unterstützt. D.h. mathematische Formeln werden nur noch mit dem TinyMCE-Editor gesetzt, (bei Single Choice und Multiple Choice-Aufgaben muss dafür "mehrzeilige Antworten" ausgewählt werden):

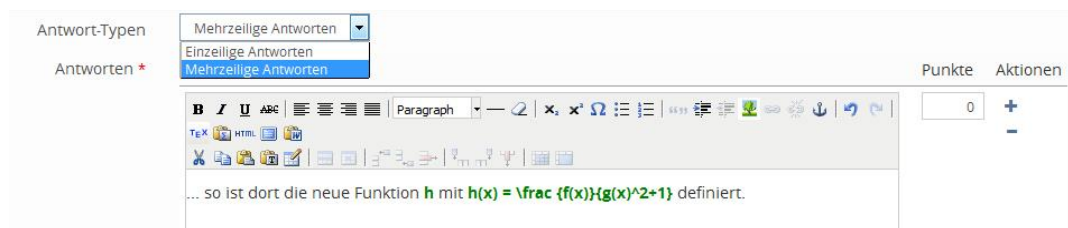


Abb. 6 Eingabefeld „Frage bearbeiten“

Anmerkung: komplexe Brüche sollten anstatt mit „\frac“ mit „\dfrac“ gesetzt werden (bessere Lesbarkeit):

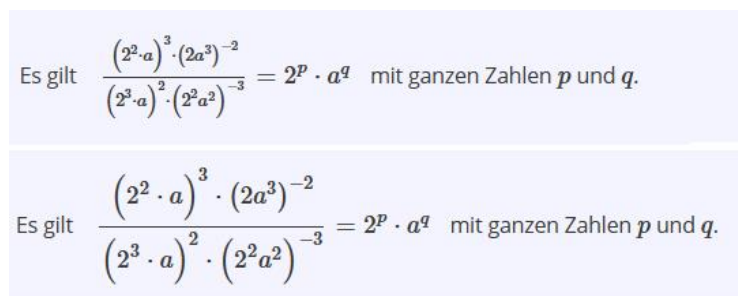


Abb. 7 LaTeX-Syntax „\frac“ und „\dfrac“

Im Fließtext der Aufgabenstellung werden alle Variablen und Formeln in LaTeX gesetzt, ebenso griechische Buchstaben oder physikalische Einheiten. Zahlen im Fließtext können ggf. auch ohne LaTeX gesetzt werden, wenn dadurch das Schrifbild sehr unruhig würde, z.B. bei einer einfachen Auflistung (20 Autos, 50 Fußgänger und 75 Fahrradfahrer).

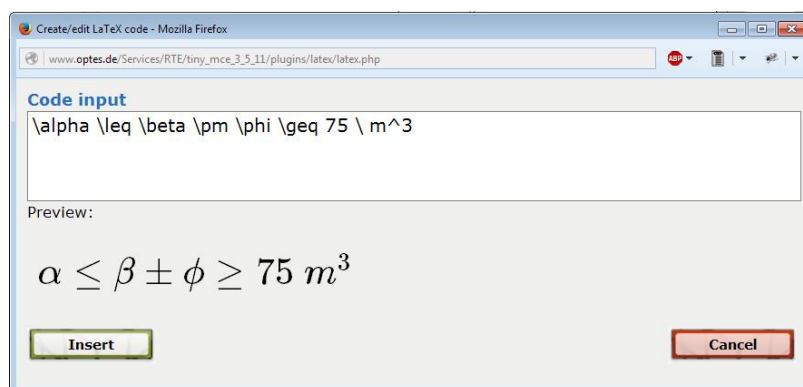


Abb. 8 ILIAS-LaTeX-Editor

6.3.4 Delimiter

- Dezimalzahlen werden in amerikanischer Schreibweise gesetzt (Punkt als Delimiter).
- Auf diese Schreibweise wird in jeder Testeinführung sowie Einführung zum LoK hingewiesen.
- Zahlen über 1000 werden ohne Tausendertrennzeichen gesetzt, damit die TN sich das für die Eingabe ins numerische Feld merken.
- Zahlen ab 10 000 werden mit Leerzeichen gesetzt (Achtung: es sollten keine Umbrüche innerhalb einer Zahl erscheinen!)
- Zahlen ab 1 000 000 werden abgekürzt (z.B. 3.2 Millionen).

6.3.5 Koordinaten

Angaben zu Koordinaten werden mit senkrechtem Strich „|“ getrennt, z.B. „P (3|4)“ (Tastenkombination: „`alt <`“ bzw. LaTeX-Syntax „`\vert`“)

6.3.6 Antwortliste SC/MC

Bei Multiple bzw. Single Choice Fragen mit Fließtext gilt die Rechtschreibung und der Punkt am Satzende wird gesetzt (Abb. 9). Bei Multiple bzw. Single Choice Fragen *ohne* Fließtext sollte der Punkt am „Satzende“ entfallen (Abb. 10 **Fehler! Verweisquelle konnte nicht gefunden werden.**).

Welche der folgenden Zuordnungen ist *keine* Funktion?

Die Zuordnung die jedem Tag...

- ... die Anzahl der an einer gegebenen Maschine produzierten Teile zuweist.
- ... die Anzahl der an einer gegebenen Maschine produzierten defekten Teile zuweist.
- ... die an einer gegebenen Maschine produzierten Teile zuweist.
- ... die an einer gegebenen Maschine anfallenden Kosten (in €) zuweist.
- ... die Menge des für eine gegebene Maschine benötigten Schmiermittels (in Liter) zuweist.

Abb. 9 Beispiel für Satzzeichen am Ende von Distraktoren

In einem gleichschenkligen Dreieck mit Schenkellänge a sind die Basiswinkel 30° .

Wie lang ist die Basisseite dieses Dreiecks?

- $g = 3 \cdot a$
- $g = \sqrt{3} \cdot a$
- $g = \sqrt{2} \cdot a$

Abb. 10 Beispiel für kein Satzzeichen am Ende von Distraktoren

6.4 Fragetypen

6.4.1 Single Choice/Multiple Choice

Für Auswahl-Aufgaben, bei denen nur eine Antwort richtig ist, sollte nur das Format "Single Choice" verwendet werden. Das Format "Multiple Choice" sollte nur dann verwendet werden, wenn mehr als eine Antwort korrekt ist. Hier ist bei der Punktvergabe genau darauf zu achten, dass am Ende ein sinnvolles Ergebnis erzielt wird!

6.4.2 Single Choice

Als Standard für Single Choice-Aufgaben sollte "Antworten mischen" angeklickt werden. Es gibt allerdings Aufgaben, bei denen das Mischen der Antworten nicht sinnvoll ist, z.B. wenn die Antworten eine ansteigende Liste darstellen oder wenn die letzte Antwort "Es gibt keine Lösung" ist. In diesem Fall sollte „Aufgaben mischen“ nicht angeklickt werden!

6.4.3 Numerische Eingabe/Lückentext

Für Aufgaben mit numerischer Eingabe sollte nur das Format "Lückentext" verwendet werden (anstatt "Numerische Antwort").

Begründung: Der Fragetyp "Lückentext" bietet die gleiche Funktion (numerische Lücke), ist aber besser geeignet, da Einheiten hinter die Lücke oder "x = ..." vor die Lücke geschrieben werden können. Außerdem kann mehr als eine numerische Lücke pro Aufgabe angelegt werden.

Bei numerischen Lückentexten wird die Lücke so platziert, dass die Aufgabenstellung möglichst unmissverständlich ist, z.B. indem vor die Lücke immer die gesuchte Variable plus "=" gestellt wird. Falls der Antwort eine Einheit zugeordnet ist, wird diese immer rechts neben die Lücke geschrieben. Seit ILIAS 5 bestehen Lückentext-Aufgaben aus zwei Feldern, der Fragestellung (Feld „Frage“) und der Aufgabenstellung (Feld „Lückentextfrage“). Dementsprechend sollte die Aufgabenstellung in das obere und die Aufgabe in das untere Feld eingetragen werden:

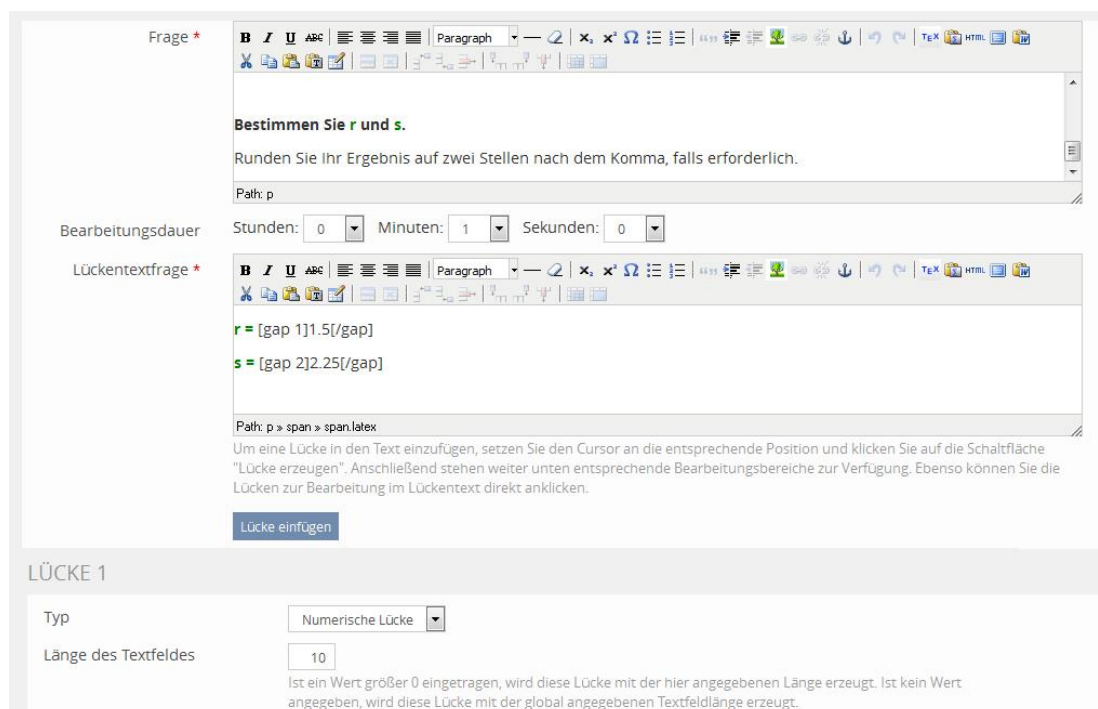
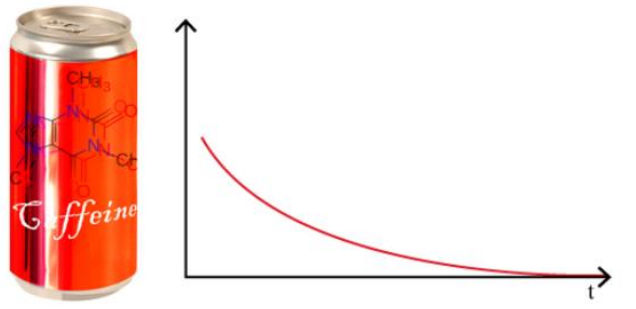


Abb. 11 Lückentextaufgabe numerisch (Autoersicht ILIAS 5)

Die Länge des Textfeldes wird stets auf „10“ gesetzt, da sich die Länge sonst an die Zahl/den erwarteten Wert anpasst und die TN von der Länge der Lücke auf die Länge des erwarteten Werts schließen können.

Ausnahmen: Wenn das Ergebnis eine sehr lange Zahl ist kann die Lücke dementsprechend verlängert werden; wenn mehrere Lücken im Fließtext stehen, sollten die Lücken nicht so lang sein, um den Lesefluss nicht zu stören.



Ein Dose eines Energy-Drinks enthält 64 mg Koffein. Diese gelangen jedoch erst eine Stunde nach Aufnahme des Getränks ins Blut. Nach Aufnahme in die Blutlaufbahn hat das Koffein eine Halbwertszeit von drei Stunden (d.h. der Koffeingehalt im Blut halbiert sich alle drei Stunden).

Nach wie vielen Stunden ist der Koffeingehalt unter $0,04 \text{ mg}$ gesunken?
Geben Sie die Antwort als Dezimalzahl an und runden Sie, falls nötig, auf zwei Nachkommastellen.

Nach ca. Stunden.

Abb. 12 Beispiel Fragetyp Lückentext numerisch, Lücke mit Länge = 10

Bestimmen Sie diese Konstante c , wenn der Fallschirmspringer 5 Sekunden nach dem Absprung eine Fallgeschwindigkeit von $32,54 \text{ m/s}$ erreicht hat.
Geben Sie die Antwort als Dezimalzahl an und runden Sie, falls nötig, auf zwei Nachkommastellen.

$c = \text{} \text{ kg/s}$

Abb. 13 Beispiel Fragetyp Lückentext numerisch, Lücke mit Länge = 10

Fassen Sie den folgenden Ausdruck zusammen:

$$t = 3(x - (2y - 4z)) - 2(2z - (2y - 3x))$$

$t = \text{} \cdot x + \text{} \cdot y + \text{} \cdot z$

Abb. 14 Fragetyp Lückentext numerisch, Lücke mit Länge = 4

6.4.4 Hinweis auf Rundung bei numerischer Eingabe

Wenn angemessen, erfolgt der Hinweis auf das Auf- bzw. Abrunden des Ergebnisses unter der Aufgabenstellung in folgender Form:

„Geben Sie die Antwort als Dezimalzahl an und runden Sie, falls nötig, auf zwei Nachkommastellen.“

Anmerkung für Autoren: Bei numerischen Lückentextaufgaben können Rundungsfehler über eine Erhöhung des Toleranzwerts (obere/untere Schranke) abgefangen werden.

Eine Hohlkugel (von gleichmäßiger Wandstärke) aus Beton hat einen Aussendurchmesser von 24 cm .
Der verwendete Beton hat eine spezifische Dichte von $\rho_b = 2,4\text{ g/cm}^3$.
Die Kugel versinkt zu einem Drittel im Wasser (die spezifische Dichte des Wassers ist $\rho_w = 1,0\text{ g/cm}^3$.)

Wie dick ist die Wandstärke d dieser Kugel?

Geben Sie die Antwort als Dezimalzahl an und runden Sie, falls nötig, auf zwei Nachkommastellen.
Das Gewicht der eingeschlossenen Luft in der Hohlkugel kann dabei vernachlässigt werden.

$d =$ cm

Abb. 15 Hinweis auf Rundung in der Aufgabenstellung

6.5 Feedback

Für Feedback (also adaptives Feedback und Lösungswege) wird in ILIAS nur der Reiter "Feedback" verwendet, da der Im-/Export von „Musterlösungen“ nicht möglich ist. Generell werden für das Verifikationsfeedback (richtig/falsch) die Formulierungen „Ihre Antwort ist korrekt.“, „Ihre Antwort ist nicht korrekt.“, bzw. „Ihre Antwort ist teilweise korrekt.“ verwendet.

6.5.1 Single Choice-Aufgabe, Option 1: Alle Distraktoren erhalten das gleiche Feedback

Feld „Richtige Lösung“:

Ihre Antwort ist korrekt. (fett, kursiv)

Erläuterung: (fett)

Hier dann der Lösungsweg (damit auch TN mit richtiger Antwort nochmal vergleichen können)

Feld „Mindestens eine Antwort ist nicht richtig“:

Ihre Antwort ist nicht korrekt. (fett, kursiv)

Lösung: (fett)

Hier die richtige Antwort.

Erläuterung: (fett)

Hier dann der Lösungsweg

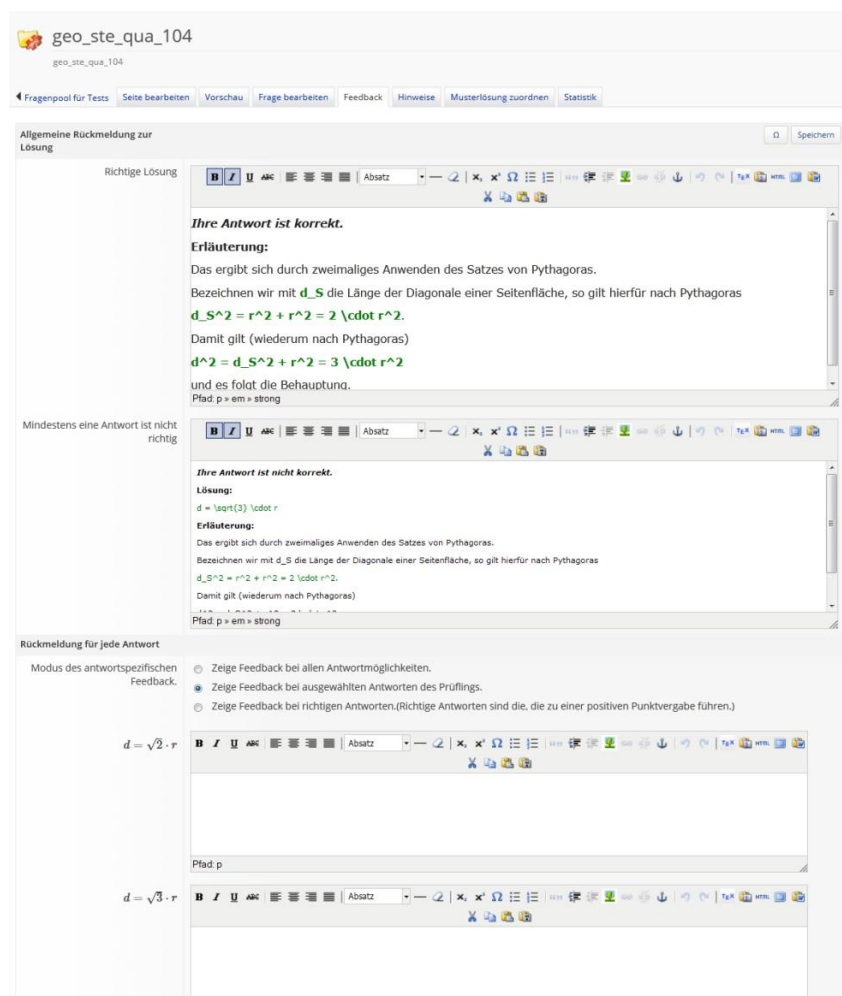
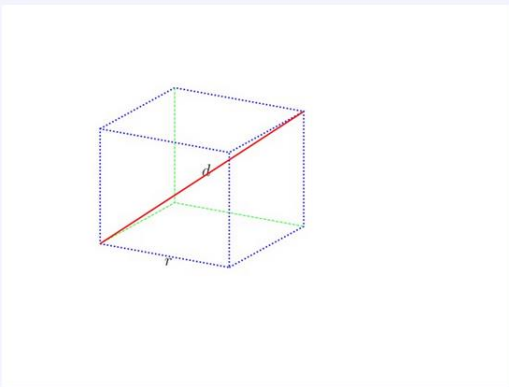


Abb. 16 Autorenansicht für Feedback bei einer Single Choice-Aufgabe (alle Distraktoren erhalten das gleiche Feedback)

Gegeben ist ein Würfel mit Seitenlänge r .



Nach welcher Formel berechnet sich die Länge d der Raumdiagonale?

- $d = \sqrt{2} \cdot r$
- $d = \sqrt{3} \cdot r$
- $d = 2 \cdot r$
- $d = r + \sqrt{2} \cdot r$
- d kann aus den Angaben nicht ermittelt werden.

Ihre Antwort ist nicht korrekt.
Lösung:
 $d = \sqrt{3} \cdot r$
Erläuterung:
Das ergibt sich durch zweimaliges Anwenden des Satzes von Pythagoras.
Bezeichnen wir mit d_S die Länge der Diagonale einer Seitenfläche, so gilt hierfür nach Pythagoras
 $d_S^2 = r^2 + r^2 = 2 \cdot r^2$.
Damit gilt (wiederum nach Pythagoras)
 $d^2 = d_S^2 + r^2 = 3 \cdot r^2$
und es folgt die Behauptung.

Abb. 17 Vorschau Feedback auf eine Single Choice-Aufgabe (alle Distraktoren erhalten das gleiche Feedback)

6.5.2 Single Choice-Aufgabe, Option 2: Jeder Distraktor erhält unterschiedliches Feedback

Adaptives Feedback mit Kommentaren zu (potentiellen) Fehlern:

Feld „Richtige Lösung“:
Ihre Antwort ist korrekt. (fett, kursiv)
Erläuterung: (fett)
 Hier dann der Lösungsweg (damit auch TN mit richtiger Antwort nochmal vergleichen können)

Feld „Mindestens eine Antwort ist nicht richtig“: bleibt leer!

Feld mit der richtigen Antwort: bleibt leer!

Feld „erste falsche Antwort“:
Ihre Antwort ist nicht korrekt. (fett, kursiv)
Lösung: (fett)
 Hier die richtige Antwort.
Erläuterung: (fett)
 Hier dann der Lösungsweg

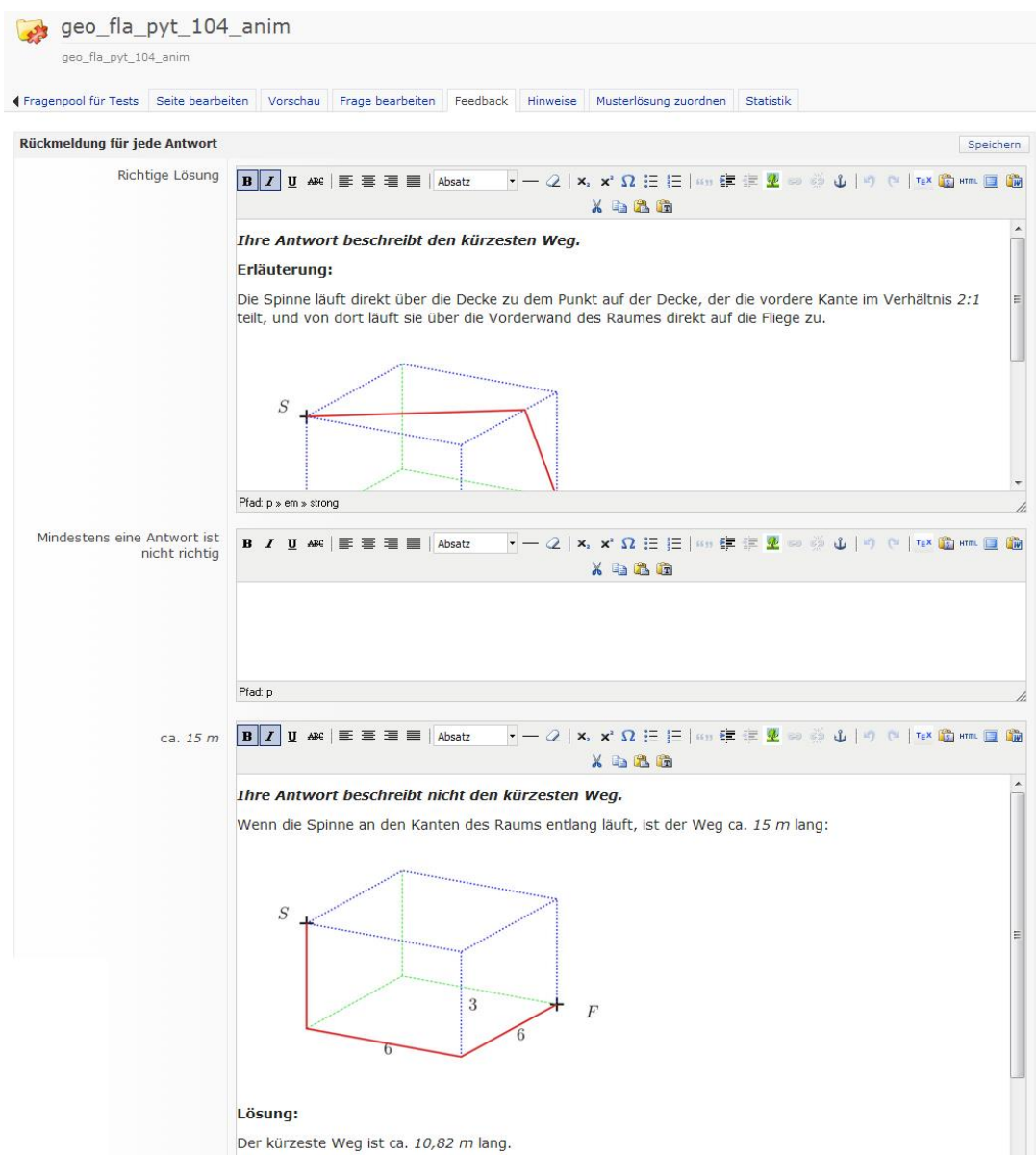
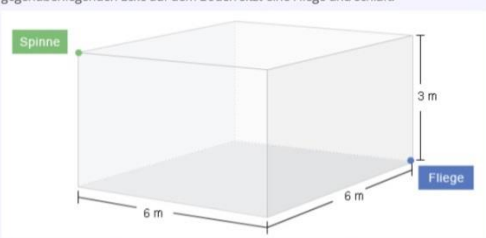


Abb. 18 Autorensicht für Feedback bei einer Single Choice-Aufgabe (jeder Distraktor erhält unterschiedliches Feedback)

geo_fla_pyt_104

Eine Spinne sitzt an der Decke eines Raums mit quadratischem Grundriss der Seitenlänge 6 m und der Höhe 3 m in einer Ecke. In der diagonal gegenüberliegenden Ecke auf dem Boden sitzt eine Fliege und schläft.



Die Spinne will möglichst schnell und auf kürzestem Weg zur Fliege, um sie zu fressen, bevor sie aufwacht. Dabei kann die Spinne an der Decke, an den Seitenwänden und über den Boden laufen.

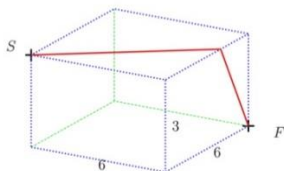
Welches ist der kürzeste Weg?

- ca. 12,71 m
- ca. 15 m
- ca. 10,82 m
- ca. 11,49 m

Ihre Antwort beschreibt den kürzesten Weg.

Erläuterung:

Die Spinne läuft direkt über die Decke zu dem Punkt auf der Decke, der die vordere Kante im Verhältnis 2:1 teilt, und von dort läuft sie über die Vorderwand des Raumes direkt auf die Fliege zu.



Der Weg, den sie dabei zurücklegt, ist nach dem Satz von Pythagoras genau

$$s = \sqrt{6^2 + 4^2} + \sqrt{2^2 + 3^2} \approx 10,82$$

Meter lang.

Abb. 19 Vorschau für Feedback auf eine Single Choice-Aufgabe (jeder Distraktor erhält unterschiedliches Feedback)

6.5.3 Lückentext-Aufgabe (numerisch)

Feld „Richtige Lösung“:

Ihre Antwort ist korrekt. (fett, kursiv)

Erläuterung: (fett)

Hier dann der Lösungsweg (damit auch TN mit richtiger Antwort nochmal vergleichen können)

Feld „Mindestens eine Antwort ist nicht richtig“:

Ihre Antwort ist nicht korrekt. (fett, kursiv)

Lösung: (fett)

Hier die richtige Antwort.

Erläuterung: (fett)

Hier dann der Lösungsweg

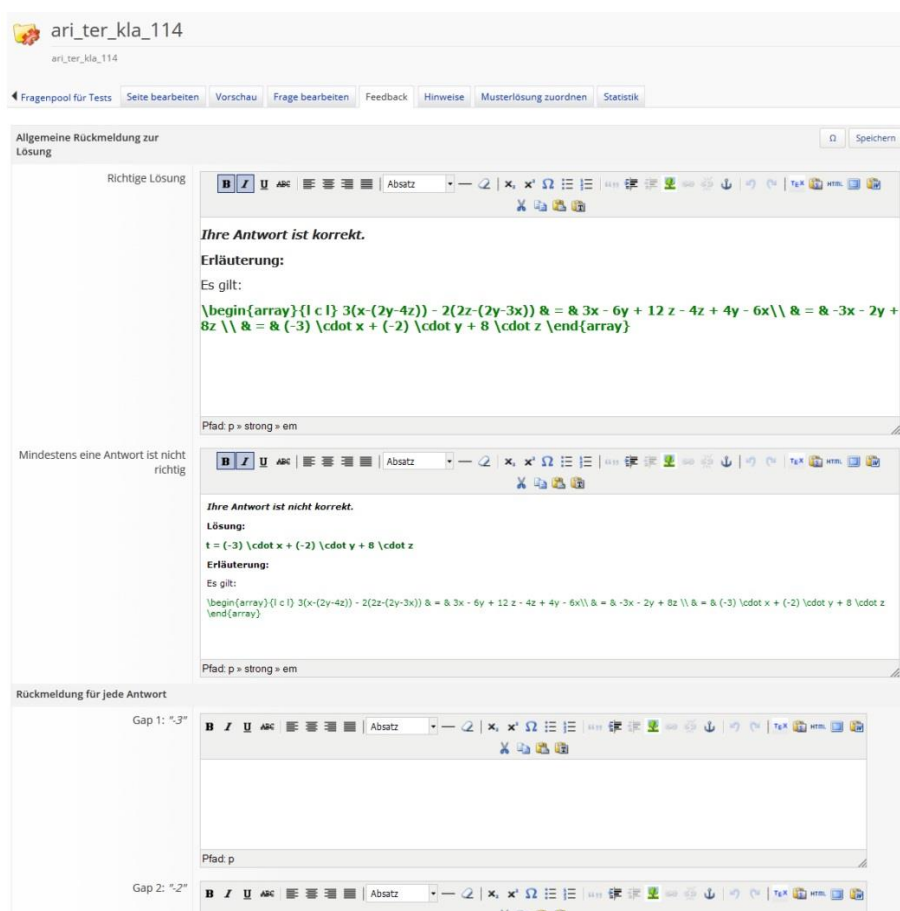


Abb. 20 Autorsicht für Feedback bei einer numerischen Lückentext-Aufgabe

ari_ter_kla_114

Fassen Sie den folgenden Ausdruck zusammen:

$$t = 3(x - (2y - 4z)) - 2(2z - (2y - 3x))$$

$$t = 2 \cdot x + 3 \cdot y + 4 \cdot z$$



Ihre Antwort ist nicht korrekt.

Lösung:

$$t = (-3) \cdot x + (-2) \cdot y + 8 \cdot z$$

Erläuterung:

Es gilt:

$$\begin{aligned} 3(x - (2y - 4z)) - 2(2z - (2y - 3x)) &= 3x - 6y + 12z - 4z + 4y - 6x \\ &= -3x - 2y + 8z \\ &= (-3) \cdot x + (-2) \cdot y + 8 \cdot z \end{aligned}$$

Abb. 21 Vorschau Feedback Feedback auf eine numerische Lückentext-Aufgabe

7 Glossar

grün = optes-spezifische Begriffe

Begriff	Aktuelle Beschreibung im Glossar
Assessment	übergeordnete Bezeichnung für verschiedene Formen des Erfassens und Bewertens von Lernprozessen und -ergebnissen in elektronischer oder analoger Form, wobei zwischen bilanzierender (summativer) und lernprozessbegleitender (formativer) Leistungserfassung unterschieden wird (vgl. Klieme et al., 2010, S. 64)
Formatives Assessment	Assessment zum Zwecke der Lernförderung bzw. lernprozessbegleitenden Leistungserfassung (vgl. Klieme et al., 2010, S. 64-65)
Summatives Assessment	Assessment zum Zwecke der bilanzierenden Leistungserfassung (vgl. Klieme et al., 2010, S. 64-65)
Test	eine Sammlung offener oder geschlossener Aufgaben, die der Übung oder Erfassung des aktuellen Lernstandes dienen
Formatives eAssessment	Durchführung von Online- (Selbst-) Tests zum Zwecke der Lernförderung. In optes wird unterschieden zwischen: Diagnostischen Tests/Übungsaufgaben/Lernerfolgskontrolle
Diagnostischer Einstiegstest	Test zur Selbst-Diagnose im formativen eAssessment, wird zu Beginn des Lernprozesses durchgeführt, um Feedback über den eigenen Kenntnisstand sowie Lernempfehlungen zu erhalten
Übungsaufgabe	Kurzes "On-the-Fly Formative Assessment" (Shavelson et al., 2008) während des Lernprozesses zur Anwendung und Einübung des gerade Gelernten. In optes beinhaltet das Feedback auf eine Übungsaufgabe die Informationen: Verifikationsfeedback (richtig/falsch), Lösung (erwartete Antwort), Musterlösung und ggf. aufgabenspezifisches Feedback (z.B. Hinweise auf mögliche Rechenfehler oder Links zu weiteren Aufgaben)
Übung	Zu einem Kurztest zusammengefasste Übungsaufgaben, einem Lernziel im Lernzielorientierten Kurs zugeordnet.
Qualifizierender Test	Zu einem Kurztest zusammengefasste Übungsaufgaben, einem Lernziel im Lernzielorientierten Kurs zugeordnet, attestiert ein Lernziel als „bestanden“. Wurden alle Lernziele in einem Kurs bestanden, gilt auch der Kurs als „bestanden“.
Feedback (auf Aufgaben bzw. Tests)	Rückmeldung auf eine Aufgabe oder ein Testergebnis. Ein Feedback kann sich zusammensetzen aus: Verifikationsfeedback (richtig/falsch), Lösung (erwartete Antwort), Musterlösung, Bewertung (Einordnung der erzielten Leistung), Kommentar (ergebnis- oder aufgabenspezifisches Feedback)
Verifikationsfeedback	Rückmeldung, ob eine Aufgabe richtig, falsch, teilweise richtig oder gar nicht beantwortet wurde.
Formatives Feedback	Formatives Feedback wird im Lernzielorientierten Kurs (LoK) in Form von Übungsergebnissen und Ergebnissen von Ausgangstests erstellt. Es umfasst neben dem Verifikationsfeedback (richtig/falsch) die erwartete Antwort, eine Musterlösung (= Lösungsweg) und ggf. Hinweise auf mögliche Rechenfehler
Diagnostisches Feedback	Diagnostisches Feedback ist das Ergebnis des Diagnostischen Tests im Zentralkurs und den Einstiegstests jedes LoK. Es dient zur Einschätzung der Vorkenntnisse. Basierend auf dem Ergebnis wird eine Bewertung des aktuellen Lernstands vorgenommen und ggf. Lernempfehlungen ausgesprochen.
CAS	Computer Algebra System
LMS	Learning Management System
Zentraler Abschlusstest	Test zur Lernerfolgskontrolle im formativen eAssessment, wird zum Abschluss des Lernprozesses durchgeführt, um Feedback über den Lernfortschritt (plus ggf. weitere Lernempfehlungen) zu erhalten. Im optes-Vorkurs optional

8 Quellen

optes-Anwendernetzwerk/Materialienpool

http://www.optes.de/goto.php?target=wiki_4929_Einstieg

Formelfrage Plug-In

Link zu Formelfrage Download und Dokumentation (Autor: Helmut Schottmüller):

http://www.ilias.de/docu/goto_docu_cat_2848.html

STACK Plug-In

Link zu STACK Demo und Dokumentation:

<http://stack.bham.ac.uk/>

STACK sample questions (Autor: Chris Sangwin):

https://github.com/math/moodle-qtype_stack/blob/master/samplequestions/STACK-demo.mbz

Authoring quick start:

https://github.com/math/moodle-qtype_stack/blob/master/doc/en/Authoring/Authoring_quick_start.md

Link zu ILIAS STACK Plug-In Spezifikation:

http://www.ilias.de/docu/goto_docu_wiki_1357_CAS_Question_Plugin.html

ILIAS STACK Plug-In Download:

http://www.ilias.de/docu/goto_docu_dcl_3342_358.html

9 Literatur

- Anderson, L. W. & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives (Abridged ed.). New York: Longman.
- Artino, A. R. & Stephens, J. M. (2009). Academic motivation and self-regulation: A comparative analysis of undergraduate and graduate students learning online. *Internet and Higher Education*, 12, 146–151.
- Astleitner, H. (2008). Die lernrelevante Ordnung von Aufgaben nach der Aufgabenschwierigkeit. In: J. Thonhauser (Hrsg.), *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, allgemeiner Didaktik und Fachdidaktik* (S. 65–80). Münster/München/Berlin: Waxmann.
- Azevedo, R. (2005). Using Hypermedia as a Metacognitive Tool for Enhancing Student Learning? The Role of Self-Regulated Learning. *Educational Psychologist*, 40 (4), 199–209.
- Baumert, J., Bos, W., Klieme, E., Lehmann, R., Lehrke, M., Hosenfeld, I., Neubrand, J. & Watermann, R. (1999). *Testaufgaben zu TIMSS/III. Mathematisch-naturwissenschaftliche Grundbildung und voruniversitäre Mathematik und Physik der Abschlußklassen der Sekundarstufe II (Population 3)*
- Bausch, I., Biehler, R., Bruder, R., Fischer, P. R., Hochmuth, R., Koepf, W., Schreiber, S. & Wassong, T. (Hrsg.) (2014). *Mathematische Vor- und Brückenkurse. Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik*. Wiesbaden: Springer.
- Bender, G. & Thiele, K. (2014). Feedback und formative Assessments in der Mathematikvorlesung. *Zeitschrift für Hochschulentwicklung*, 9 (4), 155–167.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5 (1), 7–71.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability* (21(1)). Zugriff am 9.5.2012
<http://dx.doi.org/10.1007/s11092-008-9068-5>
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for Learning. Putting it into practice*: Open University Press.
- Bloom, B. S. (1956). *Taxonomy of educational objectives. The classification of educational goals. Handbook 1, Cognitive domain*. London: Longman.
- Boekaerts, M., Pintrich, P. R. & Zeidner, M. (Hrsg.) (2000). *Handbook of self-regulation*. New York, NY: Academic Press.
- Brennan, R. L. (Hrsg.) (2006). *Educational Measurement*. Westport: Praeger.
- Cook, A. (2001). Assessing the Use of Flexible Assessment. *Assessment & Evaluation in Higher Education*, 26 (6), 539–549.
- cosh cooperation schule:hochschule (2014). *Mindestanforderungskatalog Mathematik (2.0) der Hochschulen Baden-Württembergs für ein Studium von WiMINT-Fächern*
http://www.mathematik-schule-hochschule.de/images/Aktuelles/pdf/MAKatalog_2_0.pdf
- Craven, P. (2009). History and Challenges of e-assessment. The 'Cambridge Approach' perspective - e-assessment research and development 1989 to 2009. Zugriff am 5.9.2013
<http://www.cambridgeassessment.org.uk>
- Crouch, C. H. & Mazur, E. (2001). Peer Instruction: Ten Years of Experience and Results. *American Journal of Physics*, 69 (9), 970–977.

- Dubs, R. (2008). Qualitätsvolle Aufgaben als Voraussetzung für sinnvolles Benchmarking. In: J. Thonhauser (Hrsg.), *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, allgemeiner Didaktik und Fachdidaktik* (S. 259–279). Münster/München/Berlin: Waxmann.
- Dweck, C. S. (1986). Motivational Processes Affecting Learning. *American Psychologist*, 41 (10), 1040–1048.
- Ebel, R. L. & Frisbie, D. A. (1991). *Essentials of educational measurement* (5. Aufl.). Englewood Cliffs, NJ: Prentice Hall.
- Fischer, P. R. & Biehler, R. (2011). Über die Heterogenität unserer Studienanfänger. Ergebnisse einer empirischen Untersuchung von Teilnehmern mathematischer Vorkurse, *Beiträge zum Mathematikunterricht* (Jahrestagung der Gesellschaft für Didaktik der Mathematik, S. 255–258). Münster: WTM.
- Flehsig, K.-H. (2008). Komplexere Lernaufgaben in der beruflichen Aus- und Weiterbildung. In: J. Thonhauser (Hrsg.), *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, allgemeiner Didaktik und Fachdidaktik* (S. 241–256). Münster/München/Berlin: Waxmann.
- Gibbs, G. & Simpson, C. (2004). Conditions under which assessment supports student learning. *Learning and Teaching in Higher Education*, 1 (1), 3–31.
- Gronlund, N. E. (1998). *Assessment of Students Achievement*. Boston, MA: Allyn&Bacon.
- Hancock, G. R. (1994). Cognitive complexity and the comparability of multiple-choice and constructed-response test formats. *Journal of Experimental Education*, 62 (2), 143–157.
- Hannah, J., James, A. & Williams, P. (2013). Does computer-aided formative assessment improve learning outcomes? *International Journal of Mathematical Education in Science and Technology*, 45 (2), 269–281.
- Härterich, J., Kiss, C., Roach, A., Mönnigmann, M., Schulze Darup, M. & Span, R. (2012). MathePraxis – connecting firstyear mathematics with engineering applications. *European Journal of Engineering Education*, 37 (3), 255–266.
- Heritage, M. (2007). Formative Assessment: What do teachers need to know and do? *Phi Delta Kappan*, 89 (2), 140–145.
- Hoffmann, B. (1962). *The tyranny of testing*. New York: Crowell-Collier.
- Hofmeister, W. (2005). Erläuterung der Klassifikationsmatrix zum ULME-Kompetenzstufenmodell. *bwp@ Berufs- und Wirtschaftspädagogik online*, 8
http://www.bwpat.de/ausgabe8/hofmeister_bwpat8.pdf
- Hunt, D. P. & Hassmén, P. (1994). Human self-assessment in multiple choice. *Journal of Educational Measurement*, 31 (2), 149–160.
- Hutchings, M., Quinney, A., Galvin, K. & Clark, V. (2013). The Yin/Yang of Innovative Technology Enhanced Assessment for Promoting Student Learning. In: S. Greener (Hrsg.), *Case Studies in e-Learning Research for Researchers, Teachers and Students* (S. 62–79). Reading: Academic Publishing International.
- Ivanov, S. & Lehmann, R. H. (2005). Mathematische Grundqualifikationen zu Beginn der beruflichen Ausbildung. *bwp@ Berufs- und Wirtschaftspädagogik online*, 8
http://www.bwpat.de/ausgabe8/ivanov_lehmann_bwpat8.pdf

- Jacobs, B. (2000). Richtlinien zur Erstellung von einfachen Multiple-Choice-Aufgaben nach Gronlund. Zugriff am 5.1.2015
<http://psydok.sulb.uni-saarland.de/volltexte/2005/516/html/mcguideline.html>
- Jacobs, B. (2008). Was wissen wir über die Lernwirksamkeit von Aufgabenstellungen und Feedback. In: J. Thonhauser (Hrsg.), *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, allgemeiner Didaktik und Fachdidaktik* (S. 99–114). Münster/München/Berlin: Waxmann.
- Kastner, M. & Stangl, B. (2010). Multiple-Choice and Constructed Response Tests: Do Test Format and Scoring matter?, *International Conference on Education and Educational Psychology (ICEEPSY)* (S. 263–273).
- Kebritchi, M., Hirumi, A. & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, 55, 427–443.
- Kennedy, P. & Walstad, W. B. (1997). Combining Multiple-Choice and Constructed-Response Test Scores: An Economist's View. *Applied measurement in education*, 10 (4), 359–375.
- Klauer, K. J. (2011). *Transfer des Lernens: Warum wir oft mehr lernen als gelehrt wird*. Stuttgart: Kohlhammer.
- Klieme, E. (2008). Assessment of competencies in educational contexts. In: J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 3–22). Göttingen: Hogrefe.
- Klieme, E., Bürgermeister, A., Harks, B., Blum, W., Leiß, D. & Rakoczy, K. (2010). Leistungsbeurteilung und Kompetenzmodellierung im Mathematikunterricht. Projekt Co2CA. In: E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung* (S. 64–74). Weinheim, Basel: Beltz.
- Kluger, A. N. & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119 (2), 254–284.
- Krumke, S. O., Roegner, K., Schüler, L., Seiler, R. & Stens, R. (2012). Der Online-Mathematik Brückenkurs OMB. Eine Chance zur Lösung der Probleme an der Schnittstelle Schule/Hochschule. DMV Mitteilungen Juni 2012. Zugriff am 21.1.2014
<http://page.math.tu-berlin.de/~seiler/publications/OMB-eine-Chance.pdf>
- Kvale, S. (2007). Contradictions of assessment for learning in institutions of higher learning. In: D. Boud & N. Falchikov (Hrsg.), *Rethinking assessment in higher education: learning for the longer term* (S. 55–71). Abingdon, Oxon: Routledge.
- Lawson, D. (2003). *Diagnostic testing for mathematics. LTSN MathsTEAM Project*. Zugriff am 29.8.2012
<http://www.ltsn.ac.uk/mathsteam>
- Leutner, D., Fischer, H. E., Kauertz, A., Schabram, N. & Fleischer, J. (2008). Instruktionspsychologische und fachdidaktische Aspekte der Qualität von Lernaufgaben und Testaufgaben im Physikunterricht. In: J. Thonhauser (Hrsg.), *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, allgemeiner Didaktik und Fachdidaktik* (S. 168–181). Münster/München/Berlin: Waxmann.
- Lissitz, R. W. & Hou, X. (2008). *Multiple Choice Items and Constructed Response Items: Does It Matter?*: University of Maryland. Zugriff am 30.12.2014 marces.org/article.htm
- Lissitz, R. W., Hou, X. & Slater, S. (2012). The contribution of constructed response items to large scale assessment: measuring and understanding their impact. *Journal of Applied Testing Technology*, 13 (3), 1–50 <http://www.testpublishers.org/jatt-volume-13>.

- Martinez, M. E. (1999). Cognition and the Question of Test Item Format. *Educational Psychologist*, 34 (4), 207–218.
- Mayer, H. O., Hertenagel, J. & Weber, H. (Hrsg.) (2009). *Lernzielüberprüfung im eLearning*. München: Oldenbourg.
- Mayer, R. E. (2005). Introduction to Multimedia Learning. In: R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 1–16). New York: Cambridge University Press.
- Mazur, E. (1997). *Peer Instruction: A User's Manual*. Upper Saddle River, NJ: Prentice Hall.
- Metzger, C., Waibel, R., Henning, C., Hodel, M. & Luzi, R. (1993). *Anspruchsniveau von Lernzielen und Prüfungen im kognitiven Bereich*. St. Gallen: IWP.
- Newstead, S. (2002). Examining the examiners: Why are we so bad at assessing students? *Psychology Learning and Teaching*, 2, 70–75.
- Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31 (2), 199–218. Zugriff am 15.5.2012 http://tltt.strath.ac.uk/REAP/public/Resources/DN_SHE_Final.pdf
- Nix, I. & Wyllie, A. (2011). Exploring design features to enhance computer-based assessment: Learners' views on using a confidence indicator tool and computer-based feedback. *British Journal of Educational Technology*, 42, 101–112.
- Nunan, D. & Koepke, K. (1995). Task difficulty from the learner's perspective: Perceptions and reality. *Hong Kong Papers in Linguistics and Language Teaching*, 18 (1), 1–12.
- Nußbaum, A. & Leutner, D. (1986). Entdeckendes Lernen von Aufgabenlösungsregeln unter verschiedenen Anforderungsbedingungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 18 (2), 153–164.
- Offer, J. & Bos, B. (2009). The design and application of technology-based courses in the mathematics classroom. *Computers & Education*, 53 (4), 1113–1137.
- Pintrich, P. R. & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82 (1), 33–40.
- Preißler, I., Müller, R., Hammerschmidt, J. & Scholl, S. (2010). Treibstoff für die Ingenieurausbildung - fachübergreifende Didaktik. *Zeitschrift für Hochschulentwicklung*, 5 (3), 105–115.
- Pressley, M., Borkowski, J. G. & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *Journal of Educational Research* (2), 857–867.
- Renkl, A. (1991). *Die Bedeutung der Aufgaben- und Rückmeldungsgestaltung für die Leistungsentwicklung im Fach Mathematik*. Dissertation, Universität Heidelberg.
- Resnick, L. B., Michaels, S. & O'Connor, M. C. (2010). How (well structured) talk builds the mind. In: R. J. Sternberg & D. Preiss (Hrsg.), *From genes to context: New discoveries about learning from educational research and their applications* (S. 163–194). New York, NY: Springer.
- Rheinberg, F. (2002). Bezugsnormen und schulische Leistungsbeurteilung. In: F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen*. Beltz Pädagogik. Weinheim: Beltz, 59–72.
- Rooch, A., Kiss, C. & Härterich, J. (2014). Brauchen Ingenieure Mathematik? Wie Praxisbezug die Ansichten über das Pflichtfach Mathematik verändert. In: I. Bausch et al. (Hrsg.), *Mathematische Vor- und Brückenkurse. Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik* (S. 389–409). Wiesbaden: Springer.
- Sadler, R. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science*, 18, 119–144.

- Sangwin, C. (2010). *Who uses STACK? A report on the use of the STACK CAA system. Tech. Rep.:* University of Birmingham. Zugriff am 16.10.2012
<http://web.mat.bham.ac.uk/C.J.Sangwin/Publications/2010-3-1-STACK.pdf>
- Sangwin, C. (2012). Computer Aided Assessment of Mathematics using Stack, *12th International Congress on Mathematical Education COEX* .
- Sangwin, C. & Badger, M. (2010). Core Skills in Entry Level Mathematics. Zugriff am 1.8.2013
<http://web.mat.bham.ac.uk/C.J.Sangwin/projects/2010STEM/CoreSkills2.pdf>
- Schabram, K. (2007). *Lernaufgaben im Unterricht: Instruktionspsychologische Analysen am Beispiel der Physik*, Universität Duisburg-Essen.
- Schwenk, A. & Berger, M. (2006). Zwischen Wunsch und Wirklichkeit: Was können unsere Studienanfänger? *Die Neue Hochschule* (2), 36–40.
- Seel, N. M. (2000). *Psychologie des Lernens. Lehrbuch für Pädagogen und Psychologen*. München; Basel: Reinhardt.
- Shavelson, R. J., Young, D., Ayala, C., Brandon, P., Furtak, E., Ruiz-Primo, M., Tomita, M. & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied measurement in education* (21), 295–314.
- Snow, R. E. (1993). Construct Validity and Constructed-Response Tests. In: R. E. Bennett & W. C. Ward (Hrsg.), *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (S. 45–60). Hillsdale, NJ: Lawrence Erlbaum.
- Stark, R., Gruber, H., Renkl, A. & Mandl, H. (2000). Instruktionale Effekte einer kombinierten Lernmethode. Zahlt sich die Kombination von Lösungsbeispielen und Problemlöseaufgaben aus? *Zeitschrift für Pädagogische Psychologie*, 14 (4), 206–218.
- Sweller, J., van Merriënboer, J. G. & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10 (2), 251–296.
- Traub, R. E. (1993). On the equivalence of the traits assessed by multiple-choice and constructed-response tests. In: R. E. Bennett & W. C. Ward (Hrsg.), *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (S. 75–106). Hillsdale, NJ: Lawrence Erlbaum.
- van Eck, R. & Dempsey, J. (2002). The effect of competition and contextualized advisement on the transfer of mathematics skills a computer-based instructional simulation game. *Educational Technology Research and Development*, 50 (3), 23–41.
- Weinert, F. E. (1998). Vermittlung von Schlüsselqualifikationen. In: S. E. Matalik & D. Schade (Hrsg.), *Entwicklungen in Aus- und Weiterbildung: Anforderungen, Ziele, Konzepte* (S. 23–43). Baden-Baden: Nomos.
- Winne, P. H. (2004). Students' calibration of knowledge and learning processes: Implications for designing powerful software learning environments. *International Journal of Educational Research*, 41 (6), 466–488.
- Wolf, P. & Biehler, R. (2014). Entwicklung und Erprobung anwendungsorientierter Aufgaben für Ingenieurstudienanfänger/innen. *Zeitschrift für Hochschulentwicklung*, 9 (4), 169–190.
- Zimmerman, B. J. (1989). Models of self-regulated learning and academic achievement. In: B. J. Zimmerman & D. H. Schunk (Hrsg.), *Self-regulated learning and academic achievement: Theory, research and practice* (S. 1–26). New York, NY: Springer.