

# 9 Formatives E-Assessment und Diagnostik



Katja Derr

Der Einsatz von Übungsaufgaben und Tests zur Strukturierung des Lernprozesses, zur Aktivierung von Lernenden oder zur Anregung von Reflexion wird unter dem Begriff Formatives Assessment zusammengefasst, in Abgrenzung zum summativen Assessment, das Eingang in eine Gesamtbewertung oder Note findet (siehe auch Kapitel 11). Formative Tests sind auch geeignet, einen Eindruck über den aktuellen Lernstand einer Gruppe oder eines Individuums zu vermitteln und somit dem summativen Assessment vorgeschaltet. Formatives Assessment umfasst also „alle Formen von Leistungsbeurteilung, 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. 64 f). Der Abgleich der eigenen Kenntnisse mit dem erwarteten Wissensstand sollte möglichst häufig erfolgen, damit sich Missverständnisse und Fehlinterpretationen nicht verfestigen.

Lernmanagementsysteme (LMS) wie ILIAS sind sehr gut geeignet, um Tests und Übungsaufgaben bereitzustellen, automatisiertes Feedback zu geben und Lernfortschritte zu dokumentieren. 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 (Heritage 2007). Im Online-Selbststudium kommt dem formativen E-Assessment als strukturierendes Element und als Mittel zur Aktivierung eine zentrale Rolle zu. Dabei ist zu unterscheiden zwischen dem Zeitpunkt und der Funktion:

- 1) Diagnose: Tests zu Beginn des Lernprozesses machen „die relevanten Lernziele in operationalisierender Form transparent“ (Jacobs 2008, S.100). Die Lernenden werden über die Inhalte des Kurses informiert und gleichen diese mit dem eigenen Wissensstand ab. Diese „Kalibrierung“ (Winne 2004) hilft, Lernziele zu definieren und weitere Lernhandlungen zu planen.
- 2) Lernprozessbegleitung: Während des Lernprozesses kommt formatives E-Assessment in erster Linie in Form von kurzen Tests oder Übungsaufgaben zum Einsatz („On-the-Fly Formative Assessment“ Shavelson et al. 2008). Durch kurze, niederschwellige Übungen kann neu Gelerntes sofort probiert und angewendet werden. Neben der Vergewisserung, ob die Inhalte tatsächlich verstanden wurden, dienen solche Formen des E-Assessments zur Aktivierung der Lernenden.

- 3) Lernerfolgskontrolle: Tests zum Abschluss einer Lerneinheit oder des gesamten Kurses geben den Lernenden ein Werkzeug zur Selbstevaluation bzw. zum Monitoring an die Hand (Nicol und Macfarlane-Dick 2006).

Formatives E-Assessment bietet so immer wieder die Gelegenheit, Lernenden Rückmeldung und Handlungsempfehlungen zu geben. Zeitpunkt und Funktion des E-Assessment wirken sich auf die Gestaltung des (automatisierten) Feedbacks aus. Zu Beginn des Lernprozesses steht beispielsweise das Aussprechen von Lernempfehlungen und die Vor-Strukturierung des Lernprozesses im Vordergrund. Dies gilt vor allem für Kurse mit heterogenen Teilnehmergruppen, in denen nicht alle Lerninhalte für alle Teilnehmer\*innen relevant sind.

Während des Lernprozesses und bei der Lernerfolgskontrolle ist vor allem die Rückmeldung auf Aufgaben-Ebene von Bedeutung: Nur wenn das Feedback auch Hinweise auf mögliche Fehlkonzeptionen und Ansätze zum Lösen der Aufgabe beinhaltet, können die Lernenden von formativem E-Assessment profitieren.

## 9.1 Offene und geschlossene Fragen

Fragetypen lassen sich über den Grad der Offenheit des Antwort-Eingabeformats charakterisieren (Mayer, Hertnagel und Weber 2009). Geschlossene Fragen 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 Fragen erlauben eine „freie“ Antwort-Eingabe in einem sehr festgesteckten Rahmen, zum Beispiel 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.

In digitalen Selbstlernumgebungen haben geschlossene Fragetypen wie Multiple Choice den Vorteil, dass sie niederschwellig zu bedienen sind und keine besonderen (E-Learning-) Vorkenntnisse verlangen. Bei diagnostischen Tests ist von Vorteil, dass Testteilnehmer\*innen bei der Eingabe keine Flüchtigkeits- oder Schreibfehler machen können, die die Interpretation des Ergebnisses verfälschen. Aus diesem Grund kommen bei größeren Studien und internationalen Leistungsmessungen fast ausschließlich geschlossene Fragen zum Einsatz – auch wenn ihre Aussagekraft immer wieder diskutiert wird. Die Einschränkung der Antwortmöglichkeiten auf eine überschaubare Auswahl lässt keinen Raum für eigene Ansätze und es besteht die Gefahr, dass sich die Lernenden weniger auf das Wissen der richtigen Antwort als das Ausschließen der Distraktoren konzentrieren (Snow 1993; Sangwin und Jones 2017). Die Beantwortung solcher Aufgaben kann auf Dauer eintönig sein und den Fokus auf die Leistungsmessung lenken, anstatt zum Nachdenken anzuregen (Stichwort „teaching to the test“).

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 Lernenden 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. Die Vorteile von Online-Fragetypen, wie schnelles und effizientes Feedback, können also bei offenen Aufgaben nicht genutzt werden.

Bei entsprechender Gestaltung können Multiple-Choice-Tests zu sehr ähnlichen Resultaten führen wie offene Formate (z. B. Hancock 1994; Kennedy und Walstad 1997; Lissitz, Hou und Slater 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 und Hassmén 1994; Sangwin 2013).

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.

## 9.2 Feedback

Ein großer Vorteil von E-Learning-Umgebungen ist das direkte Feedback, das Lernende auf ihre Eingaben erhalten. Grundsätzlich lässt sich zwischen „Verifikationsfeedback“ und „Elaborationsfeedback“ unterscheiden (Renkl 1991).

### 9.2.1 Verifikationsfeedback

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 weitere Informationen ergänzt.

### 9.2.2 *Elaborationsfeedback (Erweitertes Feedback)*

*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 ein möglicher Lösungsweg aufgezeigt, zum Beispiel 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 zum Beispiel möglich, Lernenden schon vor der Abgabe einen Tipp zur Lösung der Aufgabe zu geben, falls diese nicht mehr weiter wissen. Nach Abgabe der Antwort kann ein Kommentar abgestimmt auf die richtige oder falsche Antwort der Lernenden erfolgen.

*Bewertung:*

Bei der Bewertung des Ergebnisses wird ein Maß oder eine Bezugsnorm herangezogen, die den Lernenden spiegelt, wo sie im Verhältnis zur Erwartung der Autor\*innen, zur eigenen vorhergegangenen Leistung oder zur Leistung anderer stehen (Kluger und DeNisi 1996; Rheinberg 2002). Meist erfolgt die Bewertung über die Vergabe einer Punktzahl („Sie haben 3 von 10 Punkten erreicht“) oder über ein Notenschema, das den Lernenden aus der Schule vertraut ist.

## 9.3 **Erstellung von Online-Tests**

Bei der Entwicklung von webbasierten Tests ist zu berücksichtigen, dass die Lernenden nicht die Möglichkeit haben, spontan und informell Verständnisfragen zu stellen. Autor\*innen müssen also berücksichtigen, dass eine Aufgabe ohne zusätzliche Informationen „funktioniert“, um Demotivation oder gar Fehlkonzeptionen aufgrund missverständlich formulierter Aufgaben zu vermeiden. Die im Folgenden genannten Regeln sind weitgehend an die Vorschläge von Ebel und Frisbie (1991) bzw. Gronlund (1998) angelehnt (siehe auch Dubs 2008). Sie spiegeln Erfahrungswerte von Aufgabenautor\*innen wider, auch aus dem optes-Projekt, und sollen als Anregung zur Gestaltung von Aufgaben dienen – allerdings kann es auch immer wieder gute Gründe geben, gegen diese Regeln zu verstoßen!

### 9.3.1 Hinweise zur Formulierung von Fragen

- Alle Aufgaben, die im Rahmen des formativen E-Assessments zum Einsatz kommen, sollen für Lernende im Selbststudium ohne zusätzliche Erklärungen verständlich und beantwortbar sein. Bei der Formulierung ist 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 zum Themengebiet leichter nachvollziehbar und bei fehlerhaften Antworten ist es einfacher, passende Hinweise auf möglicherweise bestehende Wissenslücken und weiterführende Lernmaterialien zu geben.
- Auf Maßeinheiten und Vorgaben zur Rundung der Antwort wird in der Fragestellung hingewiesen.
- Jede Frage in einem Test oder einer Lernumgebung ist nach dem gleichen inhaltlichen und gestalterischen Schema aufgebaut, damit sich die Lernenden nicht bei jeder Aufgabe neu orientieren müssen. Gestalterische Elemente wie Abbildungen, Formeln, Tabellen sind einheitlichen Gestaltungsregeln unterworfen (siehe auch Kapitel 17).

#### *Offene Fragen:*

- Offene Fragen adressieren meist komplexere Probleme und benötigen darum meist umfangreichere Formulierungen. Zusatzinformationen können eine Aufgabe interessanter machen, erhöhen aber auch die Komplexität und Bearbeitungszeit. Je länger die Aufgabenstellung, desto anspruchsvoller ist auch die Aufgabe (Dubs 2008, S. 270). Die Lernenden müssen genau darüber informiert werden, was inhaltlich und formell von ihnen erwartet wird, und alle Anforderungen bezüglich Sprache, Gestaltung, Umfang, die in die Bewertung einfließen, sollten kommuniziert werden. Gegebenenfalls sollte die Möglichkeit eingeräumt werden, nachzufragen, ob und wie bei der Lösung vorzugehen ist.

### 9.3.2 Hinweise zur Formulierung von Antworten und Antwortmöglichkeiten

- Es existiert eine eindeutige Antwort auf die Frage und es existiert eine Musterlösung, die begründet, warum diese Antwort korrekt ist und wie sie erzielt wurde.

#### *Antworten auf geschlossene Fragen:*

- „Auch richtige“ Antworten sollten aus Gründen der didaktischen Klarheit vermieden werden. Auch aus Gründen der leichteren Punktevergabe sind Single-

Choice-Aufgaben (SC) grundsätzlich Multiple-Choice-Aufgaben (MC) vorzuziehen. Ausnahmen können zum Beispiel Logik- oder Knobelaufgaben sein, bei denen es um das Abwägen zwischen mehreren Lösungsansätzen geht.

- Bei geschlossenen Fragen wie SC oder MC sollten die Unterschiede zwischen den Auswahlmöglichkeiten 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 zusammenhanglose Distraktoren („Kängurus“) sollten vermieden werden.
- 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.
- Die Formulierung „keine der genannten Möglichkeiten ist korrekt“ sollte vermieden werden. In vielen Fällen wird sie nur aus Mangel an guten Distraktoren hinzugefügt und trägt dann nur zur Verwirrung der Lernenden bei.
- 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 zum Beispiel bei einer Liste von Jahreszahlen aufsteigend sortiert. Wenn sich keine Reihenfolge anbietet, sollten die Distraktoren nach dem Zufallsprinzip geordnet sein.

#### *Antworten auf halb offene Fragen:*

- Bei numerischen Eingaben können Rundungsfehler entstehen. Wenn das Ergebnis eine nicht „runde“ Zahl ist, sollte eine Toleranz eingeräumt werden und die Lernenden darauf hingewiesen werden, auf wie viele Stellen nach dem Komma zu runden ist.
- Es sollte deutlich auf die Maßeinheit hingewiesen werden, in der die Lösung anzugeben ist. Idealerweise wird die Maßeinheit direkt neben dem Eingabefenster der Antwort platziert.

- Es sollte auf mögliche Syntax-Eingabefehler hingewiesen werden, wie zum Beispiel Eingabe von Leerzeichen, Kommata, fehlende Multiplikations-Zeichen (z. B.  $4x$  anstatt  $4*x$ ) etc.
- Die Eingabe von komplexerer Syntax (z. B. Fragetyp STACK, siehe Kapitel 10) sollte geübt werden können, bevor eine Aufgabe oder ein Test abgegeben werden.

### 9.3.3 Hinweise zur Gestaltung von Feedback auf Fragen

- Die Lernenden erhalten eine Rückmeldung, ob sie die Frage korrekt oder nicht beantwortet haben (Verifikationsfeedback).
- Die Lernenden können die eigene Antwort und die erwartete Lösung einsehen.
- Es existiert eine Musterlösung, die begründet, warum die erwartete Lösung korrekt ist. Analog zu den Regeln für die Aufgabenstellung ist die Musterlösung klar und einfach formuliert und basiert auf einheitlichen Gestaltungsregeln.

#### *Feedback auf geschlossene Fragen:*

- Bei unterschiedlichen Falschantworten (Distraktoren) kann für jeden dieser Fehler ein passender Kommentar formuliert werden. So kann auf Rechenfehler oder Verständnisprobleme eingegangen werden, die sich häufig wiederholen (Lissitz und Hou 2008). Die Erstellung von Fehlerfeedbacks ist allerdings sehr zeitaufwändig und darum ist zwischen Aufwand und Nutzen für die Lernenden jeweils abzuwägen.
- Bei der Bewertung 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.

#### *Feedback auf offene Fragen:*

- Auch für offene Aufgabenstellungen sollte eine Musterlösung existieren, gegebenenfalls kann diese durch besonders gelungene Beispiele von anderen Lernenden ergänzt werden.

### 9.3.4 Hinweise zur Gestaltung von Feedback auf Tests

- Die Lernenden erhalten eine Rückmeldung, wie sie im Test abgeschnitten haben. Zusätzlich zu einer Bewertung wird eine Interpretationshilfe zur Einordnung der Ergebnisse gegeben.

*Anmerkung:* Eine Bewertung nach Noten, Punkt- oder Prozentzahlen wird teilweise kritisch gesehen, da die Aufmerksamkeit auf die individuelle Leistung gelenkt wird und damit andere Bestandteile des Feedbacks wie Lösungsweg, Hilfestellung oder Lernempfehlung in den Hintergrund rücken (Brennan 2006;

Black et al. 2003; Gibbs und 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 Lernende können durch eine negativ empfundene Bewertung demotiviert werden. Ein Ansatz ist, im Bereich des formativen Assessments auf Noten oder Punktzahlen zu verzichten und stattdessen individuelles Feedback in Form von Kommentaren oder Anmerkungen zu geben (Black et al. 2003). Dementgegen steht der Ansatz, erwachsene Lernende in die Gestaltung des Lernprozesses und damit auch in die Bewertung des Lernerfolgs einzubeziehen (Nicol und Macfarlane-Dick 2006). Eine sachliche und übersichtliche Darstellung der Testergebnisse inklusive Punktzahlen und Interpretationshilfen scheint darum für Lernende im Selbststudium angemessen.

- Das Feedback beinhaltet keine wertenden Aussagen über die Fähigkeiten der Lernenden. Sie sollten nicht verunsichert oder in ihrem Selbstwertgefühl angegriffen werden, aber auch übertriebenes Lob hat sich als wenig lernförderlich gezeigt (Kluger und DeNisi 1996).
- Das Feedback macht Handlungsvorschläge, zum Beispiel durch Lernempfehlungen oder Links zur Bearbeitung ähnlicher Tests und Aufgaben.

## 9.4 ILIAS-Fragetypen

ILIAS bietet unterschiedliche Fragetypen zur Gestaltung von Tests und Übungsaufgaben an (siehe Tabelle 1).

Aus den genannten Gründen dominieren in webbasierten Lernumgebungen geschlossene Fragen. Die Art der Antworteingabe wird variiert, indem zum Beispiel die korrekte Antwort auf einer Abbildung anklickbar ist, Antwortpaare durch Drag & Drop gebildet werden oder Textelemente in die korrekte Reihenfolge gebracht werden.

Bei halb offenen Fragen beschränkt sich die Auswahl auf Lücken, in denen Buchstaben oder Zahlen eingegeben werden können. Der Fragetyp „Formelfrage“ erlaubt darüber hinaus die Verwendung von Variablen bei der Aufgabenerstellung, sodass die Aufgabe und das (numerische) Ergebnis bei jedem neuen Aufruf neu berechnet werden. Die Eingabe mathematischer Ausdrücke und Formeln ist mit den „Bordmitteln“ von ILIAS allerdings nicht möglich. Für eine sinnvolle Auswertung der Eingaben muss ein System darüber hinaus in der Lage sein, die mathematische Äquivalenz der Eingabe mit der Ausgangsformel zu prüfen. Gleichzeitig muss es den Unterschied zwischen beiden identifizieren, also beispielsweise überprüfen, ob ein Bruchterm vollständig gekürzt wurde. Aus diesem Grund wurde für optes nach einem Fragetyp gesucht, der die Anbindung eines Computeralgebra-systems (CAS) ermöglicht. Das Plug-in STACK erfüllt diese Anforderung



und wird seit 2016 im optes-Vorkurs eingesetzt (für eine detailliertere Beschreibung des Fragetyps siehe Kapitel 10).

**Tabelle 1:** Online-Fragetypen ILIAS (Version 5.4)

Geschlossen	Single, Multiple, Kprim Choice Fehler/Worte markieren Hotspot / Imagemap Auswahllücke (Drop down Menu) Begriffe benennen Anordnungsfrage: vertikal Anordnungsfrage: horizontal Zuordnungsfrage
Halb offen	Textlücke numerische Lücke Formelfrage STACK Frage
Offen	Freitext eingeben Datei hochladen

Die Auswertung von komplett freien Eingaben, wie zum Beispiel mehrzeiligen Formeln oder längeren Rechenwegen, ist aktuell nur über einen Datei-Upload möglich. Dokumente können handschriftlich erstellt und dann gescannt oder unter Verwendung eines Formel-Editors generiert werden (z. B. Word). Beide Verfahren sind einigermaßen aufwändig, auch aus Sicht der Lehrperson. Die Korrektur kann von Hand oder über die Kommentar-Funktion im PDF erfolgen, das Feedback-Dokument wird dann an die Teilnehmer\*innen verschickt oder über einen Link zum Download bereitgestellt. Bei großen Teilnehmerzahlen kann die Verwaltung dieser Dokumente sehr schnell unübersichtlich und mühsam werden. Trotz dieses Zusatzaufwands sind offene Fragen sehr sinnvoll, wenn es um die Bewertung individueller Lösungsansätze oder Beweise geht. Aus diesem Grund wurde im Rahmen des Betreuten E-Learnings an der DHBW Mannheim ein Konzept für das Arbeiten mit wöchentlichen Einreichaufgaben entwickelt (siehe auch Kapitel 3).

## 9.5 Konzept für die Erstellung von Aufgaben und Tests in optes

### *A. Diagnostischer Einstiegstest über alle Themengebiete*

Der Diagnostische Einstiegstest dient zur Information, Kalibrierung und Vorstrukturierung des Lernprozesses. Er umfasst das gesamte Themenspektrum des

Vorkurses und wird nur einmal durchgeführt. Jedes Themengebiet wird durch mindestens drei Aufgaben repräsentiert, damit Flüchtigkeitsfehler bei der Erstellung der Lernempfehlung nicht zu stark gewichtet werden. Die durchschnittliche Bearbeitungszeit pro Aufgabe sollte zwei Minuten nicht übersteigen. Die Aufgaben weisen wenig Komplexität auf, adressieren nur ein Thema und sind technisch einfach zu bedienen. Im Diagnostischen Einstiegstest (und damit auch im Abschluss-test) kommen ausschließlich niederschwellige Online-Fragetypen zum Einsatz, wie Single Choice oder numerische Eingabe.

Das Feedback fokussiert auf das Aussprechen von Lernempfehlungen. Auf die Ausgabe von Verifikationsfeedback und Musterlösungen wird verzichtet. Stattdessen erhalten die Teilnehmer\*innen Links zu Lerninhalten und Hinweise zur Priorisierung.

*Hinweis:* Trotz des Versuchs, das Feedback übersichtlich und einfach zu gestalten, kann es bei Lernenden mit sehr vielen Lernempfehlungen 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 E-Learning-System nicht geleistet werden kann (z. B. Azevedo 2005; Artino und Stephens 2009). Das heißt, an dieser Stelle sollte es möglich sein, Kontakt zu einer Lehrperson oder E-Mentor\*innen aufzunehmen und Beratung bezüglich des weiteren Vorgehens einzuholen (siehe auch Kapitel 18).

### *B. Einstiegstest pro Kurs*

Analog zum Diagnostischen Einstiegstest über alle Themen ist jedem einzelnen Kurs ein kurzer Selbsttest vorgeschaltet, der den Lernenden dabei hilft, die passenden Lernmodule in diesem Kurs auszuwählen. So deckt der Kurs Arithmetik beispielsweise die Lernziele Rechnen mit ganzen Zahlen, Bruchrechnen und Termumformungen ab. Je nach Ergebnis des Einstiegstests für diesen Kurs werden den Lernenden Lernmodule für ein, zwei oder alle drei Themen vorgeschlagen. Analog zum Diagnostischen Einstiegstest finden sich im Einstiegstest pro Kurs nur niederschwellige Fragetypen wie SC oder numerische Eingabe. Im Unterschied zu Einstiegstestfragen wird zu jeder Aufgabe eine Musterlösung gegeben, die in der Auswertung des Tests eingesehen werden kann.

**Tabelle 2:** Formatives E-Assessment im Verlauf des optes-Vorkurses

Zeitpunkt	Test	Ziel der Interaktion	
Beginn / Initiierung des Lernprozesses	<b>A</b> Diagnostischer Einstiegstest über Kurscurriculum	Information + Kalibrierung	Wissen über Lerninhalte / Kurscurriculum vermitteln, Abgleich Vorwissen mit Kurscurriculum, Lernemp- fehlungen (Vorauswahl)
	<b>B</b> Einstiegstest pro Kurs		Detaillierte Lernempfeh- lungen
Konstant während des Lernprozesses	<b>C</b> Übungsaufgaben im Lernmodul	Monitoring + Aktivierung, Üben	Überprüfen des Lerner- folgs pro Lerneinheit
	<b>C</b> Training		Überprüfen des Lerner- folgs pro Lernziel
Abschluss des Lernprozesses	<b>D</b> Abschlusstest pro Kurs	Evaluation	Überprüfen des Lerner- folgs in einem Kurs / The- mengebiet
	<b>E</b> Abschlusstest über Kurscurriculum		Überprüfen des Lerner- folgs insgesamt, Vergleich vorher / nachher

### *C. Training (und Übungsaufgaben im Lernmodul)*

Wurde ein Lernmodul zur Bearbeitung empfohlen, gibt es während des Lernprozesses immer wieder die Möglichkeit, Übungsaufgaben zu bearbeiten und so das neu erlernte Wissen zu festigen. Viele Übungsaufgaben sind direkt im Lernmodul zu finden. Unabhängig davon kann zu jedem Lernziel ein Training absolviert werden. Aufgaben im Lernmodul oder im Training enthalten detaillierte Lösungswege für jede Aufgabe. Im Komplexitätsgrad adressieren Aufgaben im Training unterschiedliche Niveaus, auch CAS-Items können hier zum Einsatz kommen. Nach vollständiger Bearbeitung eines Trainings erhalten die Lernenden entweder weitere Lernempfehlungen oder werden auf den Abschlusstest des Kurses verwiesen.

#### *Einreichaufgabe (optional)*

Da geschlossene oder halb offene Online-Aufgabenformate nur bedingt geeignet sind, komplexere Aufgabenstellungen und Rechenwege zu dokumentieren und zu bewerten, bietet sich die zusätzliche Arbeit mit Aufgaben an, die zwar online gestellt werden, aber von den Lernenden auf Papier oder mit Hilfe eines Text-Editors bearbeitet und dann hochgeladen werden. Über das ILIAS-Feature „Übung“ werden die Lernenden aufgefordert, zu einem bestimmten Termin ihren ausgearbeiteten Lösungsweg abzugeben.

*Hinweis:* Im Teilprojekt „Formatives E-Assessment und Propädeutika“ ist eine umfangreiche Sammlung an Einreichaufgaben erstellt und im Rahmen des „Betreuten E-Learnings“ erprobt worden.

#### D. Abschlusstest

Für jedes Lernziel eines Lernzielorientierten Kurses existiert ein Abschlusstest, der die Erreichung des Lernziels und damit das Bestehen des Kurses attestiert. Dieser Test kann mehrmals durchgeführt werden. Das Feedback auf den Test umfasst eine Übersicht über die Zahl der korrekt / nicht korrekt beantworteten Aufgaben. Zu jeder Aufgabe wird eine Musterlösung gegeben, die in der Auswertung des Tests eingesehen werden kann.

#### E. Abschlusstest über alle Themengebiete

Analog zum Diagnostischen Einstiegstest umfasst der Abschlusstest des Vorkurses das gesamte Themenspektrum. Der Test vergleicht die Kenntnisse zu Beginn des Vorkurses mit denen am Ende, misst also den Lernerfolg. Dementsprechend sind die Aufgaben vergleichbar mit den Aufgaben im Einstiegstest (aber nicht gleich).

Auch beim Abschlusstest ist das Feedback auf das Aussprechen von Lernempfehlungen beschränkt und enthält kein Verifikationsfeedback und keine Musterlösungen. Wie der Diagnostische Einstiegstest wird auch dieser Test nur einmal durchgeführt.

## Literatur

- 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.
- 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.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for Learning. Putting it into practice*. Open University Press.
- Brennan, R. L. (Hrsg.) (2006). *Educational Measurement*. Westport: Praeger.
- 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.

- 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.
- Heritage, M. (2007). Formative Assessment: What do teachers need to know and do? *Phi Delta Kappan*, 89(2), 140–145.
- Hunt, D. P. & Hassmén, P. (1994). Human self-assessment in multiple choice. *Journal of Educational Measurement*, 31(2), 149–160.
- 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.
- 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.
- 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). *Zeitschrift für Pädagogik*, 56. Beiheft. Weinheim: 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.
- Lissitz, R. W. & Hou, X. (2008). *Multiple Choice Items and Constructed Response Items: Does It Matter?* (Maryland Assessment Research Center for Education Success, Hrsg.). University of Maryland.
- 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.
- Martinez, M. E. (1999). Cognition and the Question of Test Item Format. *Educational Psychologist*, 34(4), 207–218.
- Mayer, H. O., Hertnagel, J. & Weber, H. (Hrsg.). (2009). *Lernzielüberprüfung im eLearning*. München: Oldenbourg.
- 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.
- Renkl, A. (1991). *Die Bedeutung der Aufgaben- und Rückmeldungsgestaltung für die Leistungsentwicklung im Fach Mathematik*. Dissertation: Universität Heidelberg, Heidelberg.
- Rheinberg, F. (2002). Bezugsnormen und schulische Leistungsbeurteilung. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (2. unv. Aufl., S. 59–72). Weinheim: Beltz.
- Sangwin, C. (2013). *Computer Aided Assessment of Mathematics Using Stack*. Oxford: Oxford University Press.
- Sangwin, C. & Jones, I. (2017). Asymmetry in student achievement on multiple-choice and constructed-response items in reversible mathematics processes. *Educational Studies in Mathematics*, 94(2), 205–222.

- Shavelson, R. J., Young, D.B., Ayala, C., Brandon, P.R., Furtak, E.M., Ruiz-Primo, M.A., Tomita, M.K. & 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.
- 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.
- 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.

**Open Access** Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

