

Feedback

Bibliografie:

Marlit Annalena Lindner:

Lernbegleitende Tests in der Hochschullehre
als Feedback für Studierende und Lehrende:

Die Rolle des Aufgabenformats.

journal für lehrerInnenbildung, 19 (1), 62-68.

https://doi.org/10.35468/jlb-01-2019_05

Gesamtausgabe online unter:

<http://www.jlb-journallehrerinnenbildung.net>

<https://doi.org/10.35468/jlb-01-2019>

ISSN 1681-7028 (Printausgabe)

journal für lehrerInnenbildung
j l b
no. 1
2019

05

Marlit Annalena Lindner

Lernbegleitende Tests
in der Hochschullehre
als Feedback
für Studierende und Lehrende:
Die Rolle des Aufgabenformats

Hochschulen stellen hohe Anforderungen an das selbstregulierte Lernverhalten ihrer Studierenden. In der Regel muss eine stetige, eigenständige Initiierung sowie zielgerichtete Organisation von Lernprozessen stattfinden, um Studienziele zu erreichen. Als zentral gelten in diesem Zusammenhang summative Prüfungen, die neben der diagnostischen Einordnung der Studienleistung (Noten, Kreditierung etc.) ein wichtiges Instrument zur indirekten Beeinflussung studentischen Lernverhaltens im Hochschulalltag sind (*assessment drives learning*; vgl. z. B. Shepard, 2000). Allerdings ist das Bemühen, Studierende durch anspruchsvolle summative Prüfungen zu tiefem Lernverhalten zu motivieren sicherlich eine notwendige, womöglich jedoch keine hinreichende Bedingung für den Lernerfolg. Stellen Studierende nämlich erst in ihrer Abschlussprüfung fest, dass sie substantielle Lücken oder Verständnisprobleme haben, ist dieses Defizit oft nur durch zeitlichen (und damit letztlich auch finanziellen) Mehraufwand auszugleichen. Dies gilt selbstredend nicht nur für die Ausbildung angehender Lehrkräfte, sondern auch für Studierende im Allgemeinen.

Lernbegleitende, *formative Tests* könnten solchen negativen Studierenerfahrungen rechtzeitig entgegenwirken, sofern sie im Laufe des Semesters angeboten und von Studierenden für die Regulation ihrer Lerntätigkeiten im Sinne von Feedback genutzt werden (vgl. z. B. Roediger, Putnam & Smith, 2011). Formative Tests spielen in der Hochschulpraxis gegenüber summativen Prüfungen jedoch in vielen Fällen eine untergeordnete Rolle. Ein wichtiger Grund hierfür mag in dem hohen Aufwand für Lehrpersonen liegen, zusätzliche Aufgaben zu entwickeln, zu administrieren, auszuwerten und die Ergebnisse zurückzumelden. Fraglich ist möglicherweise auch, inwieweit Studierende von lernbegleitenden Tests Gebrauch machen, da diese in ihrem Grundverständnis ein freiwilliges Zusatzangebot ohne (verbindliche) Bewertung sind.

Mit Blick auf diese praktischen Herausforderungen skizziere ich in dem vorliegenden Artikel basierend auf aktuellen Forschungsbefunden, wie formative Tests an Hochschulen gewinnbringend eingesetzt werden können. Ein Schwerpunkt liegt auf der Beantwortung der Frage, warum insbesondere das *Aufgabenformat* bewusst gewählt werden sollte, um die Lernwirksamkeit formativer Tests zu erhöhen und gleichzeitig den Aufwand für Lehrende (und Studierende) in Grenzen

zu halten. Zur besseren Einordnung gehe ich vorab kurz auf wichtige Befunde und Annahmen zur Wirksamkeit formativer Tests ein.

Effekte und Nutzen formativer Tests

Aus vielzähligen experimentellen psychologischen Studien wissen wir, dass die zwischenzeitige Bearbeitung von Testaufgaben nach einer Lerneinheit (im Sinne einer Lernstrategie) typischerweise zur Verbesserung späterer Prüfungsleistungen führt (vgl. Testing-Effekt; z. B. Adesope, Trevisan & Sundararajan, 2017; Roediger et al., 2011). Die Methode des Testens ist dabei sogar meist anderen gängigen Lernmethoden (z. B. wiederholtes Lesen oder Unterstreichen zentraler Textpassagen) hinsichtlich der Leistung in späteren Prüfungen überlegen. Vorteile ergeben sich für die im formativen Test abgefragten, aber auch für die nicht abgefragten Inhalte. Zudem mag es überraschen, dass sich ein positiver Effekt auch dann zeigt, wenn zu keinem Zeitpunkt (externes) Feedback zur Richtigkeit bzw. Falschheit gegebener Antworten erfolgt (Adesope et al., 2017). Basierend auf diesen Befunden lässt sich mutmaßen, dass vor allem die anspruchsvolle Übung des Abrufs zentraler Lerninhalte aus dem Gedächtnis bzw. die damit verbundene Aktivierung thematisch relevanter Gedächtnisnetzwerke ein bedeutsamer Erklärungsansatz für den Testing-Effekt ist (vgl. z. B. Endres & Renkl, 2015).

Abgesehen von dieser grundlegenden gedächtnispsychologischen Erklärung für die Wirksamkeit formativer Tests als Lerninstrument ist davon auszugehen, dass die Bereitstellung von Tests und Testergebnissen Studierenden hilft, ihr Lernverhalten besser zu adaptieren und einzelne Sachverhalte oder größere Themenfelder zu identifizieren, die vertieft werden müssen (vgl. z. B. Roediger et al., 2011). Formative Tests können und sollten somit im Fall der Hochschullehre spezifische Impulse für weitere Lernaktivitäten geben.

Die resultierenden Daten sollten sich auch Lehrende zu Nutze machen, indem sie die zusätzlichen Informationen über den Lernstand ihrer Studierenden in die Planung der Lehre einbeziehen. Stärken und Schwächen einer Kohorte können so früher erkannt und Ressourcen zur rechtzeitigen Aufarbeitung von Defiziten genutzt werden.

Die Rolle des Aufgabenformats in formativen Tests

Bei der praktischen Ausgestaltung formativer Tests spielen eine Vielzahl von Parametern eine Rolle, die den Erfolg des Testens hinsichtlich des Ausmaßes positiver Lernzuwächse beeinflussen können (für eine Übersicht siehe Adesope et al., 2017). In diesem Beitrag wird der Faktor des Aufgabenformats beleuchtet, da sowohl empirische Hinweise als auch pragmatische Gründe für eine bewusste Auswahl formativer Testaufgaben sprechen.

Im Wesentlichen lassen sich Aufgaben mit geschlossenem (z. B. Multiple-Choice-Aufgaben) und offenem (z. B. Aufsätze, Kurzantworten) Antwortformat unterscheiden. Beide Formate haben spezifische Vor- und Nachteile. Während die Erstellung von offenen Aufgaben vergleichsweise einfach ist, ist deren Auswertung je nach Länge der erwarteten Antwort aufwendig und fehleranfällig, da eine Bewertung individueller Texte von Studierenden erfolgen muss. Entgegengesetzt verhält es sich bei geschlossenen Aufgaben: Durch die Vorgabe von Antwortoptionen (im einfachsten Fall eine richtige und mehrere falsche Antworten) sind diese Aufgaben zeitintensiv und anspruchsvoll in der Erstellung. Die angemessene Konstruktion geschlossener Aufgaben erfordert zudem einiges Hintergrundwissen (für eine Übersicht siehe z. B. Lindner, Strobel & Köller, 2015). In der Administration und Auswertung sind geschlossene Aufgaben dagegen sehr ökonomisch und objektiv. Jedoch wird gerne diskutiert, ob geschlossene Formate das Erreichen höherer Lernziele überhaupt abbilden können. Während es mit offenen Aufgaben deutlich leichter fällt, anspruchsvolle Lernziele zu prüfen, kann dies aber in vielen Fällen ebenso mit geschlossenen Aufgaben gelingen (vgl. z. B. Rodriguez, 2003). In der Praxis werden geschlossene Aufgaben nichtsdestotrotz meist für klassische Wissensabfragen genutzt, was vermutlich die (falsche) Wahrnehmung begünstigt, dass diese nur für ebendiesen Zweck taugen. Anspruchsvolle geschlossene Aufgaben, die Transferleistungen erfassen und höhere Denkprozesse anregen, erfordern kreative Energie sowie eine Nutzung von Potenzialen verschiedener geschlossener Formate (Lindner et al., 2015; Waugh & Gronlund, 2013). Jedoch muss stets bedacht werden, dass insbesondere klassische Multiple-Choice-Aufgaben im Vergleich zu offenen Aufgaben typischerweise deutlich leichter zu lösen sind (vgl. Kastner & Stangl, 2011; Schult & Lindner, 2018). Dies hängt auch mit der Möglichkeit zusammen, bei der Be-

antwortung geschlossener Aufgaben im Zweifelsfall erfolgreich raten zu können. Während der Ratewahrscheinlichkeit von Multiple-Choice-Aufgaben in summativen Prüfungen angemessen begegnet werden muss (vgl. Lukas, Melzer & Much, 2017), ist dies in formativen Tests von geringerer Bedeutung. Hier stehen der Gedächtnisabruf und die Aktivierung lernrelevanter Netzwerke sowie eine potenzielle Wirkung des Feedbacks auf das selbstregulierte Lernen im Fokus.

Ein wichtiger metaanalytischer Befund von Adesope et al. (2017) zeigt, dass formative Multiple-Choice-Tests gegenüber Tests mit Kurzantworten mit signifikant größeren Lernzuwächsen im Sinne des Testing-Effekts einhergehen. Gleiches gilt für eine Mischung offener und geschlossener Aufgabenformate. Diese – über viele Studien hinweg aggregierten Befunde – suggerieren, dass sich die Konstruktion geschlossener Aufgaben für formative Tests besonders lohnt. Neben positiven kognitiven Effekten liegt ein weiterer großer Vorteil geschlossener Aufgaben in der einfachen Auswertung und Rückmeldung von Ergebnissen an Studierende. Insbesondere eine elektronische Darbietung geschlossener Aufgaben ermöglicht ein (nahezu) automatisches Feedback, das für die Lehrenden keine bedeutsame Belastung darstellt und Studierenden unmittelbar Anregungen zum weiteren Lernen geben kann. Die Bearbeitung von Aufgaben mit vorgegebenen Antwortoptionen ist auch für Studierende weniger zeitaufwendig, was bei freiwilligen Tests die Akzeptanz und Nutzungsfrequenz deutlich erhöhen könnte. Zusammenfassend gibt es also gute Gründe, geschlossene Aufgabenformate in formativen Tests bevorzugt in Erwägung zu ziehen.

Praxistipps für formative Tests an Hochschulen

Als Modell für die Praxis bietet sich basierend auf diesen Kenntnissen eine Mischung verschiedener Aufgabenformate an, wobei geschlossene Aufgaben bevorzugt werden sollten. Nachteil geschlossener Aufgabenformate bleibt jedoch der hohe Konstruktionsaufwand, der sich allerdings bei der Sammlung von Aufgaben über Jahre hinweg relativiert. Um die Qualität geschlossener Aufgaben zu sichern, ist es zudem wichtig, sich mit Richtlinien für die Erstellung geschlossener Aufgaben vertraut zu machen (vgl. z. B. Lindner et al., 2015). Ein innovativer Weg kann auch die Anleitung Studierender sein, formative

Testaufgaben (für andere Studierende) in Kleingruppen zu erstellen. Wie sich zeigt, ist nämlich nicht nur das formative Testen, sondern auch die Testkonstruktion selbst eine besonders effektive Lernmethode, da für die Erstellung und Abgrenzung richtiger und falscher Antworten substanzielles fachliches Wissen erforderlich ist (vgl. Fellenz, 2004; Haladyna, 2004).

Hinsichtlich der Darbietungsform formativer Tests gibt es viele Optionen. Die simpelste Variante ist die Vorgabe von (mündlichen) Fragen in Präsenzveranstaltungen. Um Studierende dabei tatsächlich zum Gedächtnisabruf anzuregen, sollte genügend Zeit für die Beantwortung der Fragen eingeräumt werden. Als digitale Variante sind – nicht nur in größeren Veranstaltungen – Online-Voting-Systeme mit anschließender Diskussion im Plenum empfehlenswert. Auch klassische schriftliche Tests oder E-Assessments können als formative Testinstrumente eingesetzt werden. Beides ist vergleichsweise aufwendiger, ermöglicht aber bei hinreichend vielen Aufgaben eine differenzierte, individuelle Rückmeldung an die Studierenden. Zur Rückmeldung formativer Testergebnisse sei angemerkt, dass die kognitionspsychologischen Effekte des Testens in bisherigen Studien zwar nicht signifikant positiver ausfallen, wenn Feedback zur Richtigkeit der Antworten gegeben wurde, dieser Befund aber vorsichtig interpretiert werden muss (vgl. Adesope et al., 2017). Bislang liegen zu wenig differenzierte Studien vor, zudem beziehen sich die Berichte vorrangig auf kognitive Effekte des Testens, nicht jedoch auf zu erwartende selbstregulatorische Effekte auf das Lernverhalten. Im Grundsatz kann damit wohl festgehalten werden, dass Testrückmeldungen grundsätzlich erstrebenswert sind, aber nicht zwangsläufig stattfinden müssen.

Wie auch immer die Gestaltung formativer Tests im Hochschulkontext konkret umgesetzt wird, so gibt es für jede Lehrveranstaltung passende Möglichkeiten die positiven Effekte des Testens zu nutzen. Es ist daher insbesondere für Hochschullehrende, aber auch für angehende Lehrkräfte empfehlenswert, sich weitergehend mit dieser vielversprechenden Lernmethode vertraut zu machen, da in diesem Beitrag nur erste Impulse gegeben werden können, die jedoch im besten Fall zu einer breiteren Nutzung lernbegleitender Tests beitragen.

Literatur

- Adesope, O. O., Trevisan, D. A. & Sundararajan, N. (2017). Rethinking the Use of Tests: A Meta-Analysis of Practice Testing. *Review of Educational Research*, 87 (3), 65-701.
- Endres, T. & Renkl, A. (2015). Mechanisms Behind the Testing Effect: An Empirical Investigation of Retrieval Practice in Meaningful Learning. *Frontiers in Psychology*, 6, 1054.
- Fellenz, M. R. (2004). Using Assessment to Support Higher Level Learning: The Multiple Choice Item Development Assignment. *Assessment & Evaluation in Higher Education*, 29 (6), 703-719.
- Haladyna, T. M. (2004). *Developing and Validating Multiple-Choice Test Items* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Kastner, M. & Stangl, B. (2011). Multiple Choice and Constructed Response Tests: Do Test Format and Scoring Matter? *Procedia – Social and Behavioral Sciences*, 12, 263-273.
- Lindner, M. A., Strobel, B. & Köller, O. (2015). Multiple-Choice-Prüfungen an Hochschulen? Ein Literaturüberblick und Plädoyer für mehr praxisorientierte Forschung. *Zeitschrift für Pädagogische Psychologie*, 29 (3-4), 133-149.
- Lukas, J., Melzer, A. & Much, S., unter Mitarbeit von S. Eisentraut (2017). *Auswertung von Klausuren im Antwort-Wahl-Format*. Verfügbar unter http://wiki.llz.uni-halle.de/images/3/38/Handbuch_-_PDF.pdf [22.01.2019].
- Rodriguez, M. C. (2003). Construct Equivalence of Multiple-Choice and Constructed-Response Items: A Random Effects Synthesis of Correlations. *Journal of Educational Measurement*, 40 (2), 163-184.
- Roediger, H. L. III, Putnam, A. L. & Smith, M. A. (2011). Ten Benefits of Testing and their Applications to Educational Practice. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation* (Vol. 55) (pp. 1-36). San Diego, CA, US: Elsevier Academic Press.
- Schult, J. & Lindner, M. A. (2018). Diagnosegenauigkeit von Deutschlehrkräften in der Grundschule: Eine Frage des Antwortformats? *Zeitschrift für Pädagogische Psychologie*, 32 (1-2), 75-87.
- Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29 (7), 4-14.
- Waugh, C. K. & Gronlund, N. E. (2013). *Assessment of Student Achievement* (10th ed.). Boston, MA: Pearson.

Marlit Annalena Lindner, Dr.,
Nachwuchsgruppenleitung am Leibniz-Institut
für die Pädagogik der Naturwissenschaften
und Mathematik (IPN) in Kiel.
Arbeitsschwerpunkt:
Pädagogisch-Psychologische Diagnostik



mlindner@ipn.uni-kiel.de