

die hochschullehre – Jahrgang 7-2021 (9)

Herausgebende des Journals: Ivo van den Berk, Jonas Leschke, Marianne Merkt, Peter Salden, Antonia Scholkmann, Angelika Thielsch

Dieser Beitrag ist Teil des Themenheftes *Evidenzbasierte Hochschullehre – Verbindungslinien zwischen Forschung und hochschuldidaktischer Praxis* (herausgegeben von Martina Mörth, Julia Prausa, Nadine Bernhard und Rainer Watermann).

Beitrag in der Rubrik Forschung

DOI: 10.3278/HSL2109W

ISSN: 2199–8825 wbv.de/die-hochschullehre



Tiefenverarbeitung mit Multiple-Choice-Fragen: Entwicklung eines Lernstrategietrainings

NATALIE ENDERS

Zusammenfassung

In Veranstaltungen mit großen Lerngruppen werden zur Leistungsüberprüfung aus ökonomischen Gründen häufig Multiple-Choice-Fragen verwendet. Dieses Prüfungsformat wird jedoch dafür kritisiert, dass es reines Auswendiglernen begünstigt und nicht dazu geeignet ist, um ein tieferes Verständnis der Lerninhalte abzufragen. Am Beispiel einer Psychologievorlesung wird aufgezeigt, wie das Tiefenlernen mit MC-Fragen über ein speziell hierzu entwickeltes Fragetraining befördert werden kann. Die theoretischen Annahmen des Trainings basieren auf empirischen Befunden zum Testeffekt, retrieval practice und elaborativem Feedback. Die didaktische Umsetzung erfolgt über die Integration semesterbegleitender formativer Online-Quiz in die Selbstlernphase, in denen die Studierenden bei der Bearbeitung schrittweise zu lernförderlichen kognitiven Aktivitäten angehalten werden. Das Training wird über ein Blended Learning-Konzept an die Präsenzveranstaltung angebunden, um die studentischen Lernprozesse für die Lehre nutzbar zu machen und lernförderliche kommunikative Prozesse anzuregen.

Schlüsselwörter: MC-Fragen; Tiefenlernen; formative Leistungsdiagnostik; Testeffekt; Feedback

A learning strategy training for deep level processing of multiple-choice questions

Abstract

Multiple-choice-questions are frequently used for economic assessment of students' learning outcomes in lectures. However, they are criticized for promoting rote learning instead of evoking a deeper understanding of the learning material. Therefore, a training that promotes deep level processing of the learning material through practicing with multiple-choice-questions is to be integrated in a cognitive psychology lecture. The theoretical foundations are based on the testing effect, retrieval practice and elaborative feedback. Training is carried out through formative online quizzes in the self-study phase. Students answer multiple-choice-quizzes on the lecture content weekly during the semester. While quizzing, students are instructed step by step to engage in cognitive activities that promote their learning. Through a blended learning instructional design, quizzing is directly linked to the classroom activities in order to promote a fruitful discussion of the content.

Keywords: Multiple-choice test; deep level processing; formative assessment; testing effect; feedback

1 Merkmale und Messeigenschaften von MC-Fragen

Multiple-Choice-(MC-)Fragen sind dadurch gekennzeichnet, dass im Fragestamm Problemstellungen präsentiert werden, zu denen in mehreren Antwortalternativen Lösungen angeboten werden. Dieses geschlossene Frageformat ist in der Hochschullehre allgegenwärtig und kommt nicht nur in Abschlussprüfungen zum Einsatz, sondern wird auch zur Unterstützung des studentischen Lernens im Erarbeitungsprozess genutzt. Beispielsweise werden MC-Fragen während des Semesters eingesetzt, um Studierenden Übungsaufgaben bereitzustellen und ihnen Rückmeldungen zu ihrem Lernfortschritt zu geben. Bedingt durch die Covid-19-Pandemie erfreuen sich aktuell Onlinevarianten wie Live-Umfragen in synchronen Lehrveranstaltungen, Quiz für die Selbstlernphase oder E-Prüfungen zunehmender Beliebtheit.

Um MC-Fragen nach dem mit ihnen verbundenen Anspruchsniveau an die Lernenden zu klassifizieren, entwickelten Enders und Rothenbusch (im Druck) ein Klassifikationssystem zur Einordnung ebensolcher Fragen auf den Stufen *Erinnern*, *Verstehen* und *Anwenden* der kognitiven Lernzieltaxonomie (Anderson & Krathwohl, 2001). Denn obwohl MC-Fragen häufig lediglich zur Wissensabfrage (z. B. Definitionen, Bestandteile von Modellen) eingesetzt werden, kann mit diesem Fragetypus auch komplexes Wissen überprüft werden, das das Verständnis von Sachverhalten und/oder die Wissensanwendung ermöglicht (Higgins & Tatham, 2003; Lindner et al., 2015; Simkin & Kuechler, 2005). Dies kann über die Integration von Beispielen in den Fragestamm umgesetzt werden, welche den Transfer des Gelernten auf reale Fälle im Anwendungsbereich erfordern. Entsprechende Beispiele für geschlossene Fragen mit unterschiedlichen Anspruchsniveaus sind in Tabelle 1 dargestellt. In diesem haben *Wissensfragen* Fakten und Sachverhalte (z. B. Fachbegriffe oder Definitionen) zum Gegenstand. *Verständnisfragen* enthalten im Fragestamm grundlegende Konzepte, zu denen in den Antwortalternativen Erläuterungen oder Konkretisierungen (z. B. Beispiele oder Anwendungsfälle) gegeben werden. *Anwendungsfragen* sind in Fallbeispiele eingebettet, für die in den Antworten alltagsnahe Lösungsvarianten präsentiert werden. Dabei wird für die korrekte Beantwortung von Anwendungsfragen das Vorhandensein von Wissen und Verständnis vorausgesetzt.

Obwohl häufig angenommen wird, dass offene Fragen in der Regel schwieriger zu lösen sind als MC-Fragen, da bei ihrer Beantwortung die Lösung selbst abgerufen bzw. konstruiert werden muss, zeigen wissenschaftliche Untersuchungen, dass die Prüfungsergebnisse in MC-Fragen und offenen Fragen oftmals hoch miteinander übereinstimmen (Lindner et al., 2015). Dies bedeutet, dass beide Aufgabenarten – bei sorgfältiger Konstruktion – ähnlich gut zur Kompetenzmessung geeignet sind. Entscheidende Vorteile von MC-Fragen bestehen jedoch darin, dass ihre Auswertung objektiver ist (Kubinger, 2014), Lehrende durch automatisierte Auswertungsmechanismen Korrekturzeit einsparen (Higgins & Tatham, 2003) und Studierende ein zeitnahes Feedback erhalten können (Simkin & Kuechler, 2005).

Tabelle 1: Geschlossene Fragen mit unterschiedlichen kognitiven Anspruchsniveaus

Anspruchsniveau	Inhalte	Beispielfrage
Wissen	Fakten bzw. Sachverhalte	Welche Aussage über das sensorische Register ist richtig? a) Die Kapazität des sensorischen Gedächtnisses ist relativ klein. b) <i>Weiterverarbeitet werden nur die Informationen, auf die wir unsere Aufmerksamkeit lenken.</i> c) Die Informationen werden lange im sensorischen Register gespeichert. d) Alle Informationen aus der Umwelt werden weiterverarbeitet.
Verständnis	grundlegende Konzepte	Was kann das Arbeitsgedächtnis nicht? a) Mathematische Aufgaben lösen b) <i>Große Mengen an Informationen speichern</i>

(Fortsetzung Tabelle 1)

Anspruchsniveau	Inhalte	Beispielfrage
Verständnis		c) Bei einer Wegbeschreibung helfen d) Durch Verbindung zum Langzeitgedächtnis Sinneinheiten bilden
Anwendung	Fallbeispiele	Patrick hört sich Vokabeln über eine Audio-CD an. Welches Arbeitsspeichersubsystem ist für die Speicherung und Aktivhaltung dieser auditiven Informationen zuständig? a) visuell-räumlicher Notizblock b) zentrale Exekutive c) <i>phonologische Schleife</i> d) episodischer Puffer

Anmerkung: Tabelle adaptiert aus Enders & Rothenbusch (im Druck). Kategorisierung der Frageinhalte in Anlehnung an Levin und Arnold (2008). Korrekte Antworten wurden kursiv gestellt.

Je nachdem, wie viele Antworten in einer Frage vorgegeben werden und auszuwählen sind, werden verschiedene Fragetypen unterschieden (z. B. in Brauns & Schubert, 2008; Lindner et al. 2015). Während es sich bei den in Tabelle 1 abgebildeten Fragen um sogenannte Single-Choice-Fragen handelt, bei denen nur eine Antwortalternative auszuwählen ist, hat der vorliegende Artikel MC-Fragen mit vier Antwortalternativen zum Gegenstand, von denen mindestens eine bis maximal alle auszuwählen sind. Hierdurch ist die Ratewahrscheinlichkeit bei diesem Fragetypus im Vergleich zu den meisten anderen geschlossenen Fragetypen gering (Lindner et al., 2015). Wird eine Frage korrekt beantwortet, kann man daher hieraus mit größerer Zuverlässigkeit auf das Verständnis der Lernenden schließen.

2 Möglichkeiten zur Erhöhung des Lernerfolgs

Die Prüfungsform beeinflusst die Arbeits- und Prüfungsvorbereitungsstrategien der Lernenden (Biggs & Tang, 2011). Für MC-Klausuren zeigt sich, dass Studierende diese im Vergleich zu anderen Prüfungsformaten als leicht einschätzen und sich daher mit oberflächlichen Lernstrategien vorbereiten (Lindner et al., 2015; Scouller, 1998; Simkin & Kuechler, 2005). Dies birgt die Gefahr des Scheiterns in anspruchsvollen MC-Klausuren, in denen nicht nur Wissen abgefragt, sondern auch Verständnis und Wissensanwendung überprüft werden. Allerdings kann durch eine didaktisch angemessene Lehrplanung befördert werden, dass Studierende sich adäquat auf solche Prüfungen vorbereiten und auf ein tieferes Verständnis ausgerichtete Lernstrategien anwenden (Draper, 2009). Hierzu eignen sich folgende Maßnahmen:

Wie bereits in der Einleitung beschrieben, können MC-Fragen nicht nur zur Überprüfung des Leistungsstands am Ende einer Lerneinheit (*summative* Leistungsdiagnostik, z. B. bei Abschlussklausuren) eingesetzt, sondern auch zur begleitenden Überwachung und Rückmeldung des Lernprozesses (*formative* Leistungsdiagnostik) genutzt werden (Deepröse & Armitage, 2004; Higgins & Tatham, 2003). Ein Beispiel für eine formative Leistungsdiagnostik wäre ein Live-Quiz in einer Vorlesung, dessen Ergebnisse unmittelbar in den sich anschließenden Vortrag eingebunden werden. Eine andere Variante wären wöchentlich in der Selbstlernphase zu bearbeitende Kurztests, die nicht bewertet werden, sondern den Lernenden helfen sollen zu erkennen, wo ihre Stärken und Verständnislücken liegen. Dementsprechend wird das Beantworten von MC-Fragen in einer formativen Testung als Lerngelegenheit (und nicht als Mittel der Leistungsüberprüfung) angesehen (Marsh et al., 2007). Die Effektstärke einer häufigen Lernstanderhebung im Sinne einer formativen Leistungsdiagnostik liegt für die Hochschullehre bei $d = .24$ (Schneider & Preckel, 2017). Um diesen statistischen Kennwert zu ermitteln, wurden in einer Metaanalyse die Lernleistungen von Studierenden in Lehrveranstaltungen, in denen die Lernenden durchschnittlich rund 14-mal im Semester (also ungefähr einmal pro Woche) getestet wurden, mit den Lernleistungen von Studie-

renden verglichen, die im Mittel nur rund zweimal pro Semester getestet wurden. Anhand des Effektstärkemaßes d (Cohen, 1988) kann nun eine Aussage über die praktische Bedeutsamkeit des statistisch signifikanten Leistungsunterschieds zwischen den beiden Lerngruppen getroffen werden (Schneider & Mustafić, 2015; Winteler & Forster, 2007). Zur Einordnung der praktischen Bedeutsamkeit unterscheidet man hier zwischen kleinen (ab $|d| = .20$), mittleren (ab $|d| = .50$) und großen (ab $|d| = .80$) Effekten. In unserem Beispiel handelt es sich um einen kleinen positiven Effekt zugunsten der Lerngruppen, in denen häufig getestet wurde. Dies bedeutet, dass rund 59 % der Studierenden, die im Semesterverlauf häufig getestet wurden, ein Lernergebnis erreichen, das mindestens genauso gut oder besser ist als das durchschnittliche Lernergebnis der Lernenden, die seltener getestet wurden.

Über diesen Effekt hinaus lässt sich ihre positive Wirkung über die Art und Weise, wie formative Tests gestaltet und eingesetzt werden, noch weiter erhöhen. Es wird empfohlen, die Lernenden vor einer formativen Leistungsdiagnostik darüber zu informieren, welches Ziel mit den regelmäßigen Überprüfungen verfolgt wird und wie sie die Testergebnisse zur Verbesserung ihres Lernprozesses nutzen können. Dies ist wichtig, um den Unterschied eines formativen Tests zu einer summativen Prüfung für die Lernenden herauszustellen und etwaige Prüfungsängste abzubauen (Schwieren et al., 2017). Je regelmäßiger formative Tests durchgeführt werden, desto mehr regen sie zeitlich verteiltes Lernen an (Batsell et al., 2017; Donovan & Radosevich, 1999). Zusätzlich kann es in Abhängigkeit des Tests zu einem sog. *Testeffekt* kommen. Hierunter versteht man das Phänomen, dass der Lernerfolg in einer Prüfung durch voriges Abprüfen des Wissensstandes erhöht werden kann (Marsh et al., 2007). Dieser Effekt erweist sich in Laborstudien mit unterschiedlichen Lernmaterialien und Prüfungsformen als sehr robust (Batsell et al., 2017) und hat auch in Metaanalysen von Untersuchungen in realen Lernsituationen, in denen das studentische Lernen mit Tests und ohne Tests verglichen wurde, eine mittelhohe Effektstärke ($d = .56$, Schwieren et al., 2017).

Damit der Testeffekt seine positive Wirkung entfalten kann, muss beim Informationsabruf im ersten Assessment kognitive Anstrengung aufgewendet werden (Rowland, 2014; Schwieren et al., 2017). Hierzu empfiehlt es sich, Informationen frei aus dem Gedächtnis zu rekonstruieren (engl. *retrieval practice*; Karpicke, 2012; Karpicke & Grimaldi, 2012). Für MC-Fragen bedeutet dies, dass keine Rate- oder Wiedererkennungsstrategien genutzt werden, um die Aufgaben zu bearbeiten. Stattdessen sollten mögliche Lösungen nach dem Lesen des Fragestamms ohne Beachtung der Antwortalternativen in Gedanken frei konstruiert und erst im Anschluss daran die Antwort unter den vorgegebenen Antworten herausgesucht werden (sog. *Produktionsstrategie*; Kubinger, 2014). Obwohl MC-Fragen diese Vorgehensweise grundsätzlich ermöglichen (Marsh et al., 2007), setzen Studierende entsprechende Bearbeitungsstrategien nur selten spontan und ohne vorige Anleitung ein (Grimaldi & Karpicke, 2014; Karpicke & Grimaldi, 2012). Daher sollten sie gezielt didaktisch angeregt werden (Grimaldi & Karpicke, 2014). Wie dies erfolgen kann, wird weiter unten exemplarisch für eine Vorlesungsveranstaltung beschrieben.

Auch die Antwortalternativen einer MC-Frage enthalten Hinweisreize, die beim Beantworten als Abrufhilfen (engl. *retrieval cues*; Marsh et al., 2007) genutzt werden können. Diese Hinweisreize bereichern das Lernmaterial an und werden mitgelernt (Karpicke, 2012), was zur Elaboration des Lernstoffs führt (Weinstein & Mayer, 1986). Um diese günstigen Lernprozesse zu unterstützen, können Studierende dazu aufgefordert werden, zu jeder einzelnen Antwortalternative eine Begründungen zu formulieren, warum es (nicht) korrekt ist (Draper, 2009). Die Angabe von Unsicherheiten kann über die Bereitstellung einer „Ich-weiß-es-(noch)-nicht“-Antwortoption ermöglicht werden (Marsh et al., 2007), welche die Testung als Lerngelegenheit kennzeichnet und die Aufmerksamkeit der Lernenden auf noch vorhandene Wissenslücken lenkt.

Beim Einsatz von MC-Fragen können Studierende die in den Distraktoren präsentierten Fehlkonzepte fälschlicherweise erlernen, wenn sie kein Feedback zur Korrektheit ihrer Aufgabenlösung erhalten (Marsh et al., 2007; Roediger & Marsh, 2005). Diese Gefahr ist vor allem dann erhöht, wenn sie sich vor der Testung nicht mit dem Lernmaterial beschäftigt haben, sondern die

MC-Fragen selbst als Lernmaterial nutzen. Das Phänomen tritt auch vermehrt auf, wenn die Falschantworten auf gängigen Fehlschlüssen oder persönlichen Erfahrungen aufbauen, die mit einer geringeren Wahrscheinlichkeit als falsch identifiziert werden. Daher sollte nach jeder Lernstandüberprüfung Feedback zur Korrektheit der Fragenbeantwortung gegeben werden (Marsh et al., 2007). Auch Studierende geben an, von Feedback zu profitieren, da es sie motiviert und verteiltes Lernen fördert, bei dem kontinuierlich im Semesterverlauf gelernt und das Lernen nicht bis kurz vor der Prüfung aufgeschoben wird (Batsell et al., 2017; Deeptose & Armitage, 2004). Ferner können sie es nutzen, um ihren Leistungsstand abzuschätzen (Grimstad & Grabe, 2004).

Müssen große Lerngruppen unterrichtet werden, wird das Bereitstellen von Feedback zeit- und kostenintensiv (Deeptose & Armitage, 2004). Unter diesen Bedingungen stellen Online-Quiz mit automatischem Feedback eine ökonomische und effiziente Umsetzungsvariante dar. Dabei wird empfohlen, Lernende bei Falschbeantwortung einer Frage zum Beispiel durch eine Verlinkung direkt auf Lernhilfen hinzuweisen und zu einer erneuten Selbstüberprüfung aufzufordern (Grimstad & Grabe, 2004; Higgins & Tatham, 2003).

Empirisch lässt sich zeigen, dass das Bereitstellen von Feedback den wünschenswerten Testeffekt zusätzlich steigert (Schwieren et al., 2017). Als besonders effektiv hat sich dabei *elaboriertes Feedback* erwiesen, das über eine Rückmeldung zur Korrektheit einer Aufgabenlösung (sog. *Bearbeitungsfeedback*) hinaus weitere Informationen und Lösungshinweise enthält. Beispielsweise konnten Enders, Gaschler und Kubik (2020) zeigen, dass die unmittelbare Mitteilung eines Grundes bei falscher Beantwortung von Online-Quizaufgaben den Lernerfolg Studierender in einer zweiten Quizrunde signifikant erhöht. Metaanalytisch ermittelt hat hochwertiges häufiges Feedback (im Gegensatz zu seltenem Feedback) eine Effektstärke von $d = .47$ auf den Lernerfolg (mittlerer Effekt, Schneider & Preckel, 2017).

Die Lehrenden sollten die zusätzlichen Informationen, die sie anhand der formativen Tests über den Lernstand ihrer Gruppe erhalten, für die Gestaltung des nachfolgenden Unterrichts nutzen (Bücker et al., 2015). Daher muss die formative Leistungsdiagnostik sinnvoll in alle anderen didaktischen Aktivitäten eingebettet werden – im Falle des Einsatzes von Online-Quiz über ein gut durchdachtes Blended-Learning-Konzept (Erpenbeck et al., 2015).

3 Veranstaltungsbeschreibung

Die Vorlesung *Allgemeine Psychologie* wird jährlich angeboten und von etwa 120 bis 160 Nebenfachstudierenden besucht, die sie im Wahlpflicht-/Anwendungsfach oder im Studium Generale verschiedener Bachelor- und Masterstudiengänge belegen. Bezüglich der fachlichen Hintergründe und persönlichen Studienziele handelt es sich um eine heterogen zusammengesetzte Lerngruppe mit wenig bis gar keinem einschlägigen Vorwissen.

Das Grobziel der Vorlesung besteht darin, die *Allgemeine Psychologie* als ein Teilgebiet des Fachs kennenzulernen. Hierzu werden die Studierenden anhand zentraler Inhaltsbereiche exemplarisch mit ausgewählten Fragestellungen, Theorien, Befunden und Forschungsansätzen vertraut gemacht. Die Feinziele der einzelnen Sitzungen können mehrheitlich den drei Stufen (*Erinnern*, *Verstehen* und *Anwenden*) der kognitiven Lernzieltaxonomie von Anderson und Kollegen (2001) zugeordnet werden.

Die Leistungsanforderungen bestehen bis dato in der regelmäßigen aktiven Veranstaltungsteilnahme und dem Bestehen einer Abschlussklausur (50 % Mindestkompetenz), welche zu gleichen Teilen aus *Wissens-* und *Anwendungsfragen* im MC-Format mit vier Antwortalternativen besteht, die in ihren Konstruktionsmerkmalen der Klassifikation nach Enders & Rothenbusch (im Druck) entsprechen (vgl. Tabelle 1). Die Inhalte der *Wissensfragen* sind Teil der Vorträge und daher überwiegend auf den Vorlesungsfolien aufgeführt, die den Studierenden vor jedem Termin online zur Verfügung gestellt werden. Bei den Inhalten der *Anwendungsfragen* handelt es sich oftmals um Analogien zu Beispielen, die in der Vorlesung vorgestellt werden, aber auch um neuartige Probleme.

Dabei wird davon ausgegangen, dass mit den Anwendungsfragen indirekt auch das Verständnis des Lernstoffs abgeprüft wird, und aus zeitökonomischen Gründen auf die Vorgabe von Verständnisfragen verzichtet. Für jede korrekt beantwortete MC-Frage wird insgesamt ein Punkt angerechnet.

Die Veranstaltung wurde in dieser Form erstmals im Wintersemester 2019/20 durchgeführt. Zu Semesterbeginn wurden die Studierenden über das Prüfungsformat informiert. Noch während des laufenden Semesters wurden zwei Probeklausuren geschrieben, um sie mit dem MC-Frageformat vertraut zu machen. Zuvor wurden die unterschiedlichen Fragetypen (Wissens- und Anwendungsfragen) erläutert und auf das mit den Anwendungsfragen verbundene Anspruchsniveau hingewiesen. In den Wissensfragen wurden statistisch signifikant mehr Punkte erzielt als in den Anwendungsfragen, $t(144)=9.62$, $p<.000$, $d_{\text{Cohen}}=0,76$, was annähernd einem großen Effekt entspricht. Dies korrespondiert aus theoretischer Sicht mit der der Fragenkonstruktion zugrunde liegenden Annahme, dass die lehrzieltaxonomisch höher angesiedelten Fragen schwieriger zu lösen sind und daher eine niedrigere Lösungswahrscheinlichkeit besitzen (vgl. Kap. 1). Allerdings verfehlten insgesamt 66 % der $N=145$ Prüfungsteilnehmer*innen die Mindestkompetenz, die allein durch das Lösen aller Wissensfragen zu erreichen gewesen wäre. Deswegen weist dieses Ergebnis auch auf ein unzureichendes grundlegendes Verständnis der Lerninhalte hin. Da die mit der Vorlesung verbundenen Lehrziele nicht nur das Erinnern, sondern auch das Verstehen und Anwenden der Inhalte beinhalten, wurde in der Konsequenz eine Intervention zur Förderung der Tiefenverarbeitung der Lerninhalte entwickelt.

4 Beschreibung des neuen Lehrkonzepts

Die Schwierigkeiten der Studierenden mit den Anwendungsfragen weisen unter anderem darauf hin, dass hinsichtlich des Transfers des Erlernten ein besonders hoher Unterstützungsbedarf durch die Lehrenden besteht. Da dieser jedoch ein zentrales Lernziel der Veranstaltung darstellt, soll das Anspruchsniveau der Prüfung nicht einfach abgesenkt werden. Ebenfalls kann die Prüfungsform aus organisatorischen Gründen nicht verändert werden. Daher besteht das Ziel der didaktischen Überarbeitung darin, die Lehr-Lernaktivitäten über eine Anpassung der Vermittlungsstrategie besser mit den Zielen und der Prüfungsform in Einklang zu bringen (Biggs & Tang, 2011). Hierzu wird die Art und Weise des MC-Frageeinsatzes verändert, um die Lernenden frühzeitig mit dem Prüfungsformat vertraut zu machen und Lernstrategien zum Fragenbeantworten zu vermitteln, welche die Studierenden anwenden können, um ein tieferes Verständnis der Inhalte zu erlangen. Im Fokus des Lehrkonzepts steht somit das Ziel, den Studierenden bessere Lernstrategien zur Prüfungsvorbereitung an die Hand zu geben, die sie zukünftig auch in anderen Lehrveranstaltungen einsetzen können.

In der Einführungsitzung werden die Leistungsanforderungen besprochen, die im neuen Konzept in der regelmäßigen Teilnahme an formativen Tests sowie der erfolgreichen Teilnahme an der Abschlussklausur bestehen. Die Funktion der formativen Assessments sowie die unterschiedlichen MC-Fragetypen und ihre Anspruchsniveaus werden den Studierenden erläutert. Es wird betont, dass die Herangehensweise bewusst ausgewählt wurde, um die Studierenden bestmöglich zu unterstützen und Schwierigkeiten bereits im Lernprozess sichtbar zu machen. Forschungsergebnisse zur Wirksamkeit der gewählten didaktischen Maßnahmen (vgl. Kap. 3) werden präsentiert, um das wissenschaftliche Fundament der Vermittlungsstrategie zu unterstreichen und Wissen über die Effektivität des Vorgehens zu vermitteln. Diese Art des Wissens kann als metakognitives Wissen beschrieben werden (Karpicke & Grimaldi, 2012). Metakognitives Wissen stellt Wissen über das eigene Wissen dar, in diesem Fall soll das Wissen über die Einsatzmöglichkeiten und Wirksamkeit der gewählten Lernstrategien die Nützlichkeit dieser Herangehensweise aufzeigen.

Die formative Testung erfolgt wöchentlich und wird über ein Online-Quiz umgesetzt, das von den Studierenden im Anschluss an jede Vorlesungssitzung in der Selbstlernphase bis spätestens 24 Stunden vor dem nächsten Präsenztermin bearbeitet wird. Es besteht aus jeweils fünf Wissens- und fünf Anwendungsfragen zu den Inhalten des vergangenen Termins. Als Lernstrategie wird retrieval practice (Karpicke, 2012) über die Vermittlung der Produktionsstrategie (Kubinger, 2014) eingeübt. Diese Lernstrategien werden dabei indirekt gefördert, indem sie durch eine entsprechende Aufgabenformulierung bei der Bearbeitung der MC-Fragen angestoßen werden. Dazu wird im ersten Bearbeitungsschritt des Quiz lediglich der Fragestamm in Kombination mit einem offenen Eingabefeld präsentiert, wie es in *Abbildung 1* exemplarisch für eine Anwendungsfrage dargestellt ist, und die Studierenden zur Aktivierung ihres Vorwissens aufgefordert. Um zum nächsten Bearbeitungsschritt zu gelangen, muss eine Eingabe erfolgen.

<p>Vier Jugendliche verhalten sich respektlos gegenüber ihrem Handballtrainer. Bei welchen der aufgeführten Konsequenzen handelt es sich nicht um eine negative Bestrafung?</p>
<p>Vorwissensaktivierung: Notieren Sie kurz das Wissen, das Sie benötigen, um die Aufgabe zu lösen.</p>
Empty input field

Abbildung 1: Anregung von retrieval practice im Online-Quiz

Der zweite Bearbeitungsschritt hat das sequenzielle Prüfen der Antwortalternativen sowie die Elaboration der Lerninhalte zum Ziel. Dazu geben die Studierenden für jede Antwortalternative eine Begründung in ein offenes Antwortfeld ein (*Abbildung 2*). Zur Kennzeichnung von Unsicherheiten wird die Antwortoption „Ich weiß es noch nicht“ angeboten, mit der die offene Eingabe übersprungen werden kann. Dies verhindert, dass unwissende Studierende zur Eingabe gezwungen werden.

<p>a) Der Trainer wird wütend und schimpft mit den Jugendlichen.</p>
<p>Begründung: Bitte schreiben Sie kurz auf, wieso Sie meinen, dass diese Antwort richtig bzw. falsch ist.</p>
Empty input field
<p><input type="checkbox"/> Ich weiß es noch nicht.</p>

Abbildung 2: Sequenzielles Prüfen und Begründen der Antworten im Online-Quiz

Der dritte Bearbeitungsschritt entspricht der klassischen Beantwortung geschlossener Fragen in einem Online-Quiz. Über das Anklicken eines Kästchens sind einige Antwortalternativen im Sinne der Aufgabenstellung zu markieren. Erneut wird zusätzlich die Antwortoption „Ich weiß es noch nicht“ angeboten (*Abbildung 3*).

Vier Jugendliche verhalten sich respektlos gegenüber ihrem Handballtrainer. Bei welchen der aufgeführten Konsequenzen handelt es sich nicht um eine negative Bestrafung? (Mehrfachantworten möglich.)

- Der Trainer wird wütend und schimpft mit den Jugendlichen.
- Die Jugendlichen werden für eine Woche vom Training ausgeschlossen.
- Der Trainer beschließt, sich nicht provozieren zu lassen und einfach keine Reaktion zu zeigen.
- Die anderen Mannschaftsmitglieder werfen den vier Jugendlichen böse Blicke zu.
- Ich weiß es noch nicht.

Abbildung 3: Präsentation der Quizfrage im klassischen MC-Format

Auf jede beantwortete Frage erhalten die Studierenden ein unmittelbares Feedback. Wurde das richtige Antwortmuster ausgewählt, wird dies rückgemeldet: „Prima, Ihre Auswahl ist korrekt!“ Für den Fall, dass eine oder mehrere Antworten falsch waren, erfolgt wie in *Abbildung 4* dargestellt eine differenzierte Rückmeldung. Sie enthält Angaben dazu, welche Auswahl falsch war (Bearbeitungsfeedback), eine Begründung (elaboriertes Feedback), Lösungshilfen und die Aufforderung zu einem neuen Versuch. Letztere dient dazu, die weitere Überwachung des Lernprozesses anzuregen.

Wurde im dritten Bearbeitungsschritt „Ich weiß es noch nicht“ angekreuzt, wird das Quiz durch die Rückmeldung als Lerngelegenheit gekennzeichnet: „Kein Problem! Sie befinden sich in der Lernphase und das Quiz hilft Ihnen, Ihre Lücken aufzudecken.“ Anschließend werden wie bei einer falschen Beantwortung Lösungshilfen gegeben und zur erneuten Bearbeitung aufgefordert.

<p>Sie haben Antwort b) „Die Jugendlichen werden für eine Woche vom Training ausgeschlossen“ ausgewählt. Dies ist leider nicht korrekt.</p>	}	Bearbeitungs-Feedback
<p>Denken Sie daran, dass die Jugendlichen durch den Ausschluss nicht mehr die Möglichkeit haben, an einer attraktiven Aktivität teilzunehmen, und von der Gruppe isoliert werden.</p>	}	elaboriertes Feedback
<p>Bitte lesen Sie den Inhalt der letzten Vorlesungssitzung auf den Folien Nr. X–Y nach, die Sie im Learnweb finden, oder sehen Sie sich Kapitel X im Lehrbuch von XY an.</p>	}	Lösungshilfen
<p>Versuchen Sie es im Anschluss daran noch einmal. Viel Erfolg!</p>	}	Aufforderung zur Wiederholung

Abbildung 4: Feedback bei falscher Fragenbeantwortung im Online-Quiz

Die Wissensfragen werden im Quiz zuerst präsentiert, um die Sicherung des Grundlagenwissens zu befördern. Erst im Anschluss daran wird der Transfer des Gelernten mit Anwendungsfragen überprüft. Dies regt Lernprozesse bei unzureichendem Vorwissen an und verhindert, dass die Studierenden durch eine direkte Bearbeitung der Anwendungsfragen ohne die notwendigen Kenntnisse demotiviert werden.

Um die Selbstlernphase sinnvoll an die Präsenzveranstaltung anzubinden, werden die Ergebnisse der formativen Tests in den ersten 15 Minuten der nachfolgenden Vorlesungssitzung im Plenum besprochen. Hierdurch werden der problembasierte Austausch zwischen Lehrenden und Lernenden sowie die Partizipation der Studierenden gefördert. Zur Vorbereitung des Unterrichtsgesprächs nutzen die Lehrenden die Ergebnisse des Online-Quiz und gehen verstärkt auf Fragen ein, die von einem Großteil der Studierenden nicht korrekt gelöst wurden. Vor allem die schwieri-

geren Anwendungsfragen bieten ein gutes Fundament für tiefergehende Diskussionen des Lernstoffs und gemeinsame Überlegungen zum Praxistransfer der Inhalte. Hierüber werden die anspruchsvollen Lernziele gemeinsam bearbeitet.

Bedingt durch die Covid-19-Pandemie wurde die Vorlesung, die ursprünglich immer im Winter stattfand, in das Sommersemester verlegt. Für das Sommersemester 2021 ist derzeit eine asynchrone videobasierte Onlinevorlesung geplant, in der auch eine erste Erprobung des neuen Lehrkonzepts erfolgen wird. Große Teile des beschriebenen Konzepts können problemlos auf die Onlinelehre übertragen werden, da das Training von vornherein im Onlineformat konzipiert wurde. Das Besprechen der Quizfragen kann gegebenenfalls in einer synchronen Videokonferenz, über die Bereitstellung eines kurzen Erklärvideos oder in einem Forum erfolgen. Zur Überprüfung der Wirksamkeit des Konzepts wird derzeit Begleitforschung geplant. Hierzu sind regelmäßige semesterbegleitende Studierendenbefragungen zur Akzeptanz und zum Nutzen des Trainings sowie die empirische Überprüfung des Lernerfolgs der Studierenden angedacht.

Literatur

- Anderson, L. W. & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. Longman.
- Batsell, W. R., Perry, J. L., Hanley, E. & Hostetter, A. B. (2017). Ecological validity of the testing effect. *Teaching of Psychology*, 44(1), 18–23.
- Biggs, J. B. & Tang, C. (2011). *Teaching for quality learning at university. What the student does* (4th ed.). McGraw-Hill.
- Brauns, K. & Schubert, S. (2008). Qualitätssicherung von Multiple-Choice-Prüfungen. In S. Dany, B. Szczyrba & J. Wildt (Hg.), *Prüfungen auf die Agenda! Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (1. Aufl., S. 93–102). W. Bertelsmann Verlag.
- Bücker, S., Deimling, M., Durduman, J., Holzhäuser, J., Schnieders, S., Tietze, M., Sayeed, S. & Schneider, M. (2015). Prüfung. In M. Schneider & M. Mustafić (Hg.), *Gute Hochschullehre: Eine evidenzbasierte Orientierungshilfe: Wie man Vorlesungen, Seminare und Projekte effektiv gestaltet* (S. 119–152). Springer.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.
- Deepröse, C. & Armitage, C. (2004). Giving formative feedback in higher education. *Psychology Learning & Teaching*, 41(1), 43–46.
- Donovan, J. J. & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, 84(5), 795–805.
- Draper, S. W. (2009). Catalytic assessment: Understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology*, 40(2), 285–293.
- Enders, N., Gaschler, R. & Kubik, V. (2020). Online Quizzes with Closed Questions in Formal Assessment: How Elaborate Feedback can Promote Learning. *Psychology Learning & Teaching*. Advance online publication.
- Enders, N. & Rothenbusch, S. (im Druck). Frageaktivitäten als Lernstrategien im Studium: Ist das Beantworten oder das Generieren von Fragen effektiver? In S. Löffler & S. Alves Ferreira (Hg.), *Lehre - Beratung - Forschung: Lernprozesse im Hochschulkontext fördern*. Beltz Juventa.
- Erpenbeck, J., Sauter, S. & Sauter, W. (2015). *E-Learning und Blended Learning. Selbstgesteuerte Lernprozesse zum Wissensaufbau und zur Qualifizierung*. Springer.
- Grimaldi, P. J. & Karpicke, J. D. (2014). Guided retrieval practice of educational materials using automated scoring. *Journal of Educational Psychology*, 106(1), 58–68.
- Grimstad, K. & Grabe, M. (2004). Are online study questions beneficial? *Teaching of Psychology*, 31(2), 143–146.
- Higgins, E. & Tatham, L. (2003). Exploring the potential of multiple-choice questions in assessment. *Learning and teaching in action*, 2(1), 1–12.
- Karpicke, J. D. (2012). Retrieval-based learning. *Current Directions in Psychological Science*, 21(3), 157–163.
- Karpicke, J. D. & Grimaldi, P. J. (2012). Retrieval-based learning. A perspective for enhancing meaningful learning. *Educational Psychology Review*, 24(3), 401–418.
- Kubinger, K. D. (2014). Gutachten zur Erstellung „gerichtsfester“ Multiple-Choice-Prüfungsaufgaben. *Psychologische Rundschau*, 65(3), 169–178.

- Lindner, M. A., Strobel, B. & Köller, O. (2015). Multiple-Choice-Prüfungen an Hochschulen? *Zeitschrift für Pädagogische Psychologie*, 29(3–4), 133–149.
- Marsh, E. J., Roediger, H. L., Bjork, R. A. & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review*, 14(2), 194–199.
- Roediger, H. L. & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 31(5), 1155–1159.
- Rowland, C. A. (2014). The effect of testing versus restudy on retention. A meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432–1463.
- Schneider, M. & Mustafić, M. (2015). Hochschuldidaktik als quantitativ-empirische Wissenschaft. In M. Schneider & M. Mustafić (Hg.), *Gute Hochschullehre: Eine evidenzbasierte Orientierungshilfe. Wie man Vorlesungen, Seminare und Projekte effektiv gestaltet* (S. 1–12). Springer.
- Schneider, M. & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143(6), 565–600.
- Schwieren, J., Barenberg, J. & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching*, 16(2), 179–196.
- Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35, 453–472.
- Simkin, M. G. & Kuechler, W. L. (2005). Multiple-choice tests and student understanding. What is the connection? *Decision Sciences Journal of Innovative Education*, 3(1), 73–97.
- Weinstein, C. E. & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315–327). Macmillan.
- Winteler, A. & Forster, P. (2007). Wer sagt, was gute Lehre ist? Evidenzbasiertes Lehren und Lernen. *Das Hochschulwesen*, 55(4), 102–109.

Autorin

Dr. Natalie Enders. Universität Hildesheim, Institut für Psychologie, Hildesheim, Deutschland;
E-Mail: endersna@uni-hildesheim.de



Zitiervorschlag: Enders, N. (2021). Tiefenverarbeitung mit Multiple-Choice-Fragen: Entwicklung eines Lernstrategietrainings. *die hochschullehre*, Jahrgang 7/2021. DOI: 10.3278/HSL2109W. Online unter: wbv.de/die-hochschullehre



die hochschullehre

Interdisziplinäre Zeitschrift für Studium und Lehre



Die Online-Zeitschrift **die hochschullehre** wird Open Access veröffentlicht. Sie ist ein wissenschaftliches Forum für Lehren und Lernen an Hochschulen. Sie liefert eine ganzheitliche, interdisziplinäre Betrachtung der Hochschullehre.

Alles im Blick mit **die hochschullehre**:

- Lehr- und Lernumwelt für die Lernprozesse Studierender
- Lehren und Lernen
- Studienstrukturen
- Hochschulentwicklung und Hochschuldidaktik
- Verhältnis von Hochschullehre und ihrer gesellschaftlichen Funktion
- Fragen der Hochschule als Institution
- Fachkulturen
- Mediendidaktische Themen

Sie sind Forscherin oder Forscher, Praktikerin oder Praktiker in Hochschuldidaktik, Hochschulentwicklung oder in angrenzenden Feldern? Lehrende oder Lehrender mit Interesse an Forschung zu ihrer eigenen Lehre?

Dann besuchen Sie wbv.de/die-hochschullehre.

Alle Beiträge stehen kostenlos zum Download bereit.

➔ wbv.de/die-hochschullehre