

**Nachhaltiges Lernen an der Hochschule: Untersuchungen zu Randbedingungen der  
positiven Effekte von digitalen Übungstests auf das Behalten von Lehrinhalten**

Julia Glaser & Tobias Richter

Universität Würzburg, Lehrstuhl für Psychologie IV

Erscheint in: Fütting-Lippert, A., Eisenmann, M., Grafe, S., Siller, H.-S. & Trefzger, T.  
(Hrsg.) (2024). *Digitale Medien in Lehr-Lern-Konzepten der Lehrpersonenbildung in  
interdisziplinärer Perspektive*. Berlin: Springer VS.

### Abstract

Lernen begleitet uns unser Leben lang, von der frühkindlichen Bildung über Schule, Ausbildung und Universität bis hin zum Lernen im Beruf. Vieles von dem, was wir lernen, wird aber rasch wieder vergessen. Die Nutzung von Übungstests ist eine Lernstrategie, die zum Erwerb *nachhaltigen Wissens* beiträgt, also Wissen, das dauerhaft zur Verfügung steht und abgerufen werden kann, wenn es benötigt wird. In einem Zyklus mit drei Experimenten im Rahmen regulärer Psychologievorlesungen im Lehramtsstudium wurde erforscht, ob der positive Effekt von digitalen (Online-)Übungstests auf das Behalten (Testungseffekt) von Personen- oder Situationsmerkmalen abhängt und auch das Behalten von Vorlesungsinhalten fördert, die nicht direkt getestet wurden (Transfer). In Experiment 1 lag der Schwerpunkt auf Lernermerkmalen wie Motivation, Prüfungsängstlichkeit oder Vorwissen als potenziellen Moderatoren. Experiment 2 befasste sich mit Transfereffekten auf Vorlesungsinhalte, die in den Übungstests nicht direkt enthalten waren. In Experiment 3 wurde untersucht, ob eine metakognitive Aktivierung den Testungseffekt noch verstärken kann. Alle drei Experimente wurden mit Lehramtsstudierenden durchgeführt und hatten einen ähnlichen Aufbau: Die Studierenden nahmen regulär an ihren universitären Veranstaltungen teil (Lernphase), beantworteten im Nachgang online Fragen zur Sitzung (Übungstests) und wurden zu einem späteren Zeitpunkt nochmals abgefragt (kriterialer Test). Der Testungseffekt konnte in allen drei Experimenten nachgewiesen werden, allerdings nur für direkt getestetes Wissen. Transfereffekte für verwandtes, nicht getestetes Wissen traten nicht auf. Weder Lernermerkmale noch die metakognitive Aktivierung schienen einen Einfluss auf die Effektivität des Testens zu haben. Der Testungseffekt scheint also eine sehr wirksame Lernstrategie zu sein, die sich sinnvoll und in Form von Online-Übungstests auch ökonomisch in der Hochschullehre einsetzen lässt und unabhängig von Lernermerkmalen zu besseren Lernergebnissen führt. Die Übungstests sollten aber die gesamte Bandbreite relevanter Inhalte abdecken, da Transfereffekte zu nicht getesteten Inhalten nicht zu erwarten sind.

Schlüsselwörter: Abrufübung, individuelle Lernermerkmale, metakognitive Aktivierung, Testungseffekt, Transfer, Übungstests

## Einleitung

Lernen begleitet uns ein Leben lang. Dabei geht es in der Regel darum, Wissen nachhaltig zu erwerben, sodass es dauerhaft zur Verfügung steht, die Basis für neues Lernen bilden kann und in den Situationen abgerufen werden kann, in denen es benötigt wird (Richter et al. 2022). Studierende nutzen spontan allerdings oft Lernstrategien, die lediglich das kurzfristige Behalten fördern, wie beispielsweise wiederholtes Lesen oder Auswendiglernen unmittelbar vor Prüfungen (Karpicke, Butler & Roediger 2009; Kornell & Bjork 2007; Taraban, Maki & Ryneason 1999). Eine vielversprechende Lernstrategie für das nachhaltige Lernen sind Übungstests, für die sich sowohl in Laborexperimenten als auch in Schule und Hochschule robuste Effekte auf das längerfristige Behalten gezeigt haben (für Metaanalysen, s. Yang, Luo, Vadillo, Yu & Shanks 2021; Rowland 2014). Übungstests können das nachhaltige Lernen zum einen direkt fördern, indem sie den Abruf aus dem Langzeitgedächtnis trainieren (Abrufübung, retrieval practice), was das gelernte Wissen festigt und leichter abrufbar macht (direkter Testungseffekt, z.B. Carpenter 2009). Zum anderen können Übungstests einen indirekten positiven Effekt auf den Lernprozess ausüben, indem sie den Lernenden Lücken im bereits Gelernten vor Augen führen, die dann gezielt nachgebessert werden können (indirekter Testungseffekt) (Arnold & McDermott 2013). Solche Übungstests lassen sich im Rahmen von Lehrveranstaltungen besonders gut in digitaler Form darbieten, etwa als ergänzende Online-Übungseinheit.

Die vorliegenden Metaanalysen belegen einen robusten positiven Effekt von Übungstests auf das nachhaltige Lernen (z.B. Hedges  $g = 0.51$ , Adesope et al. 2017;  $g = 0.50$ , Rowland 2014), werfen aber auch bestimmte Fragen auf, die für die Anwendung von Übungstests von Relevanz sind:

(1) Erstens zeigen sich in Studien zum Testungseffekt fast immer individuelle Unterschiede. Nicht alle Lernenden profitieren gleichermaßen von Übungstests. Es ist denkbar und

theoretisch plausibel anzunehmen, dass Lernende bestimmte Voraussetzungen – kognitiver, emotionaler oder motivationaler Art mitbringen müssen, um Übungstests bestmöglich für ihr Lernen zu nutzen (z.B. Carpenter 2009; Kubik, Gaschler & Hausman 2021; Pyc & Rawson 2010).

(2) Zweitens sind positive Effekte von Übungstests auf das Behalten bislang eindeutig nur für Inhalte nachgewiesen, die in den Übungstests selbst vorkommen. Inwieweit der Testungseffekt auch auf thematisch zusammenhängende Inhalte transferiert, die in derselben Lehreinheit vorgekommen sind, aber nicht direkt getestet wurden, ist bislang noch offen (s. die Metaanalyse von Pan & Rickard 2018).

(3) Drittens legt die Forschung zur Rolle von Metakognition beim Lernen (Schneider, Tibken & Richter 2022) nahe, dass Lernende Übungstests für das nachhaltige Lernen womöglich noch effektiver nutzen könnten, wenn sie sich bewusst sind, dass und wie der aktive Abruf aus dem Langzeitgedächtnis das langfristige Behalten fördert. Eine Intervention, die entsprechende metakognitive Überzeugungen kurz vor der Bearbeitung der Übungstests salient macht, könnte also den Testungseffekt noch verstärken.

Diese Fragen wurden in drei Feldexperimenten im Rahmen von Psychologievorlesungen für Lehramtsstudierende untersucht, die durch digitale (Online-)Übungseinheiten ergänzt wurden.

### **Experiment 1: Hängt der Testungseffekt in der Hochschullehre von individuellen Voraussetzungen der Studierenden ab?**

In Experiment 1 (Glaser & Richter 2023) wurden Übungstests in zwei regulären Universitätsvorlesungen für Lehramtsstudierende (psychologische Pflichtveranstaltungen im Rahmen des erziehungswissenschaftlichen Studiums) in Form einer minimalen experimentellen Intervention implementiert: Die Übungstests wurden den Studierenden als zusätzliche Online-Lehrangebote zur Verfügung gestellt und waren in wenigen Minuten zu bearbeiten. Zusätzlich wurde eine Woche vor der Vorlesung eine Reihe kognitiver,

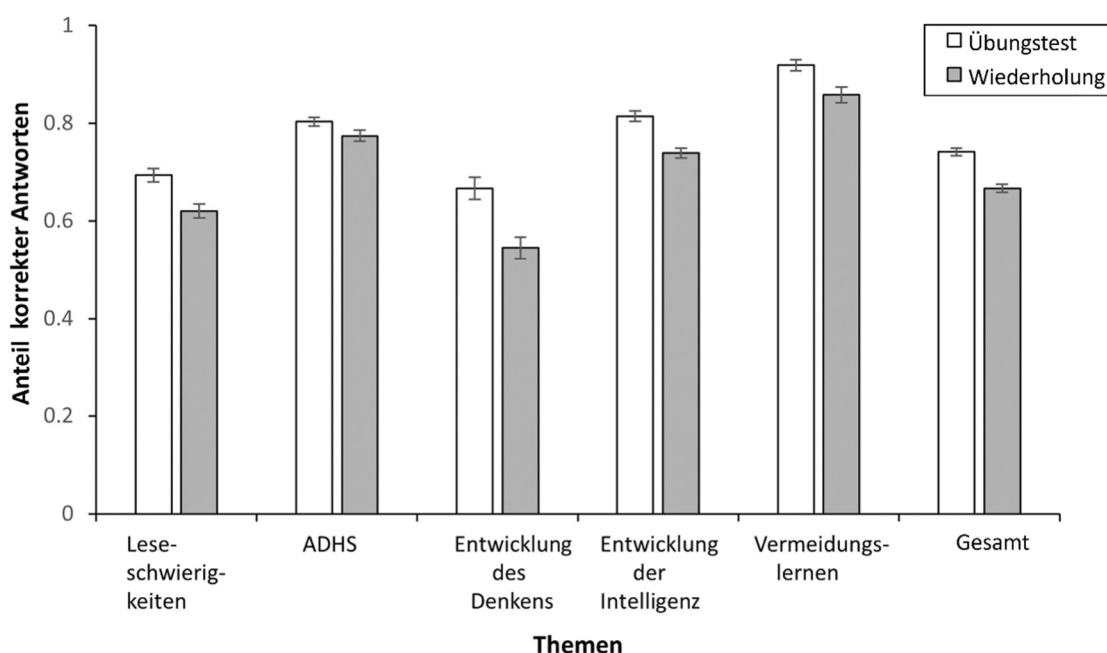
motivationaler und emotionaler Merkmale der Lernenden erfasst, die möglicherweise für eine sinnvolle Nutzung von Übungstests als Lernstrategie relevant sind. Dazu gehörten zum einen das einschlägige Vorwissen, das für das Verständnis der Vorlesungsinhalte und für die erfolgreiche Bearbeitung der Fragen im Übungstest erforderlich ist – beides potenzielle Moderatoren des Testungseffekts (Greving & Richter 2018; Rummer & Schweppe 2022). Zum anderen wurden motivationale und emotionale Merkmale wie Lernmotivation, Fehlerorientierung und Prüfungsängstlichkeit erfasst, die ebenfalls für eine lernförderliche Nutzung von Übungstests relevant sein könnten.

Unmittelbar nach einer Vorlesungssitzung, die die teilnehmenden Studierenden ( $N = 208$ ) selbst wählen konnten, absolvierten sie online eine Übungssitzung (auf der Plattform Unipark), in der Kurzantwortfragen (Testung mit korrektivem Feedback) zu beantworten und – für andere Inhalte – zusammenfassende Aussagen der Kerngedanken zu lesen waren (Wiederholung). Diese beiden Übungsformen – Testung mit Kurzantwortfragen vs. Wiederholung anhand zusammenfassender Aussagen – stellten die experimentelle Manipulation dar. Welche Inhalte der jeweiligen Vorlesungssitzung in Form von Kurzantwortfragen getestet oder in Form von Zusammenfassungen wiederholt wurden, wurde zwischen den Versuchspersonen per Zufall variiert. Eine Woche später wurde die Abrufbarkeit des Gelernten mit einem kriterialen Test geprüft, der Kurzantwort- oder Multiple-Choice-Fragen enthielt. Hier zeigte sich ein deutlicher Testungseffekt ( $\eta_p^2 = .07$ ): Vorlesungsinhalte, die in dem Übungstest angefragt worden waren, konnten besser erinnert werden als Inhalte, die in Form von Zusammenfassungen wiederholt worden waren. Der Lernvorteil der Übungstests zeigte sich für alle betrachteten Vorlesungssitzungen, also über verschiedenste Themen hinweg (s. Abb. 1). Belege für einen moderierenden Einfluss von kognitiven, motivationalen oder emotionalen Lernermerkmalen ergaben sich jedoch nicht (Glaser & Richter 2023a). Diese Ergebnisse legen nahe, dass der Testungseffekt ein sehr robustes Phänomen ist und sich Übungstests zur Unterstützung des nachhaltigen Lernens in

einer universitären Lehrveranstaltung auch ohne Rücksichtnahme auf individuelle Unterschiede gewinnbringend einsetzen lassen.

**Abb. 1:**

**Der Testungseffekt für verschiedene Vorlesungsthemen und über alle Vorlesungsthemen gemittelt (Experiment 1)**



*Anmerkung.* Fehlerbalken repräsentieren den Standardfehler des Mittelwerts. Die y-Achse zeigt den Anteil korrekter Antworten im kriterialen Test.

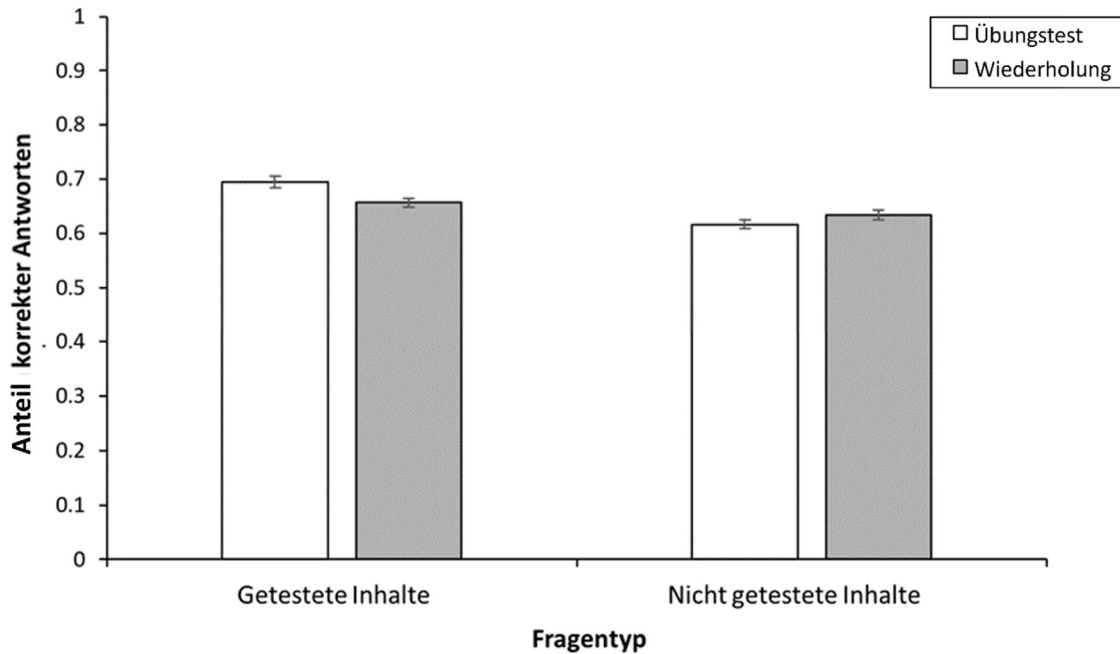
**Experiment 2: Fördern Übungstests auch das nachhaltige Lernen von Lehrinhalten, die nicht direkt getestet wurden?**

In Experiment 1 wurde der Testungseffekt für Lehrinhalte nachgewiesen, die in den Übungstests explizit vorkamen. Für die praktische Anwendung von Übungstests im Rahmen universitärer Lehrveranstaltungen wäre es jedoch wünschenswert, wenn Übungstests nicht nur das Behalten der direkt getesteten Inhalte fördern würden, sondern auch das Behalten nicht getesteter, aber thematisch mit den Übungsfragen verknüpfter Inhalte aus derselben Lehreinheit. Ein solcher Transfereffekt wäre theoretisch naheliegend (z. B. im Sinne der

Theorie des elaborativen Abrufs, Carpenter 2009), ist aber bislang noch nicht empirisch eindeutig belegt (Pan & Rickard 2018).

Experiment 2 widmete sich dieser Forschungslücke. Auch dieses Experiment wurde als Feldexperiment in Psychologievorlesungen im Rahmen des Lehramtsstudiums durchgeführt. Im Rahmen einer sechswöchigen Experimentalphase mit sechs Vorlesungssitzungen besuchten die teilnehmenden Studierenden ( $N = 67$ ) regulär ihre Vorlesungen und erhielten jeweils zwei Tage nach jeder Vorlesungssitzung den Link zu einer Online-Übungssitzung (via SoSciSurvey). In diesen Übungssitzungen wurden den Studierenden entweder Multiple-Choice-Fragen (Testung mit korrektivem Feedback), die bei einer falschen Antwort einmalig wiederholt wurden, gestellt, oder sie lasen zusammenfassende Aussagen (Wiederholung). Wöchentlich wechselte die Aufgabe, sodass alle teilnehmenden Studierenden unter beiden Übungsbedingungen lernten. Nach Abschluss der Experimentalphase folgte in der siebten Woche ein kriterialer Test mit Multiple-Choice- oder Kurzantwortfragen zu allen sechs Vorlesungssitzungen. Ein Teil der Fragen waren dieselben wie in den Übungstests bzw. bezogen sich auf die Inhalte der zusammenfassenden Aussagen, die übrigen Fragen bezogen sich auf thematisch verwandte Inhalte, die in der Vorlesungssitzung, nicht aber in der dazugehörigen Übungssitzung vorgekommen waren. Das Ergebnismuster war sehr klar: Wie in Experiment 1 fand sich ein deutlicher Testungseffekt ( $\eta_p^2 = .08$ ; s. Abb. 2, helle Balken), allerdings nur für die Inhalte, die in den Übungssitzungen vorgekommen waren. Ein Transfereffekt auf gelernte, aber nicht direkt getestete Inhalte zeigte sich jedoch nicht (s. Abb. 2; dunkle Balken; Glaser & Richter 2023b).



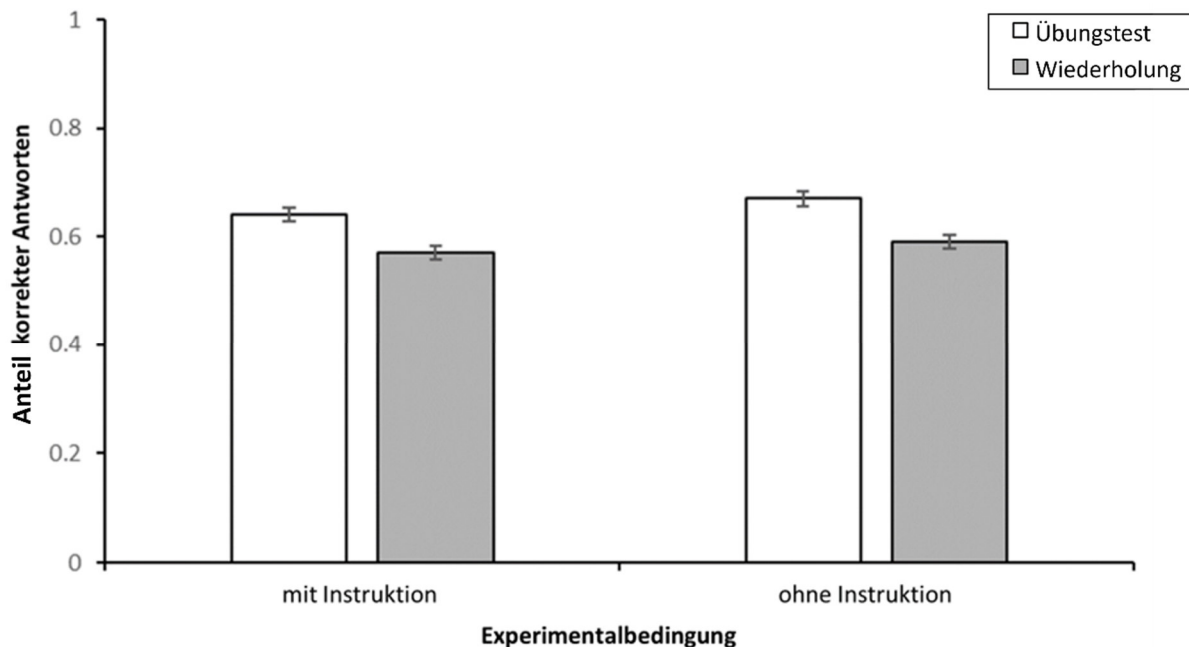
**Abb. 2:****Der Testungseffekte für getestete und nicht getestete, aber gelernte Inhalte (Experiment 2)**

*Anmerkung.* Fehlerbalken repräsentieren den Standardfehler des Mittelwerts. Die y-Achse zeigt den Anteil korrekter Antworten im kriterialen Test.

### **Experiment 3: Kann eine metakognitive Aktivierung den Testungseffekt in der Hochschullehre noch verstärken?**

Die Wirksamkeit von Übungstests als eine Lernstrategie, die das nachhaltige Lernen fördert, hängt womöglich davon ab, wie Lernende sie nutzen (Soderstrom & Bjork 2014). In Analogie zu anderen Lernstrategien könnte insbesondere eine aktive, intentionale Nutzung besonders wirksam sein, bei der Lernende mehr kognitive Ressourcen investieren, über die Übungsfragen reflektieren und so mehr Inhalte aus dem Langzeitgedächtnis abrufen als bei einer eher passiven oder inzidentellen Nutzung (s. auch das Konzept des „constructive retrieval“, Hinze, Wiley & Pellegrino 2013). Wenn diese Überlegung zutrifft, sollte eine metakognitive Aktivierung, die den Lernenden den potenziellen Nutzen von Übungstests und die zugrundeliegenden psychologischen Mechanismen vor Augen führt, den Testungseffekt

noch verstärken. Diese Hypothese haben wir in Experiment 3 untersucht, das als ein Feldexperiment auf Basis eines ähnlichen Versuchsplans wie bei Experiment 1 umgesetzt wurde, wiederum in Psychologievorlesungen für Lehramtsstudierende ( $N = 157$ ). Im Unterschied zu Experiment 1 erhielt nun aber die Hälfte der teilnehmenden Studierenden zusätzlich vor dem Absolvieren der Online-Übungssitzung eine Instruktion, die der metakognitiven Aktivierung dienen sollte, und den Studierenden wurde im Laufe der Übungssitzung die Instruktion mehrfach in Erinnerung gebracht (metakognitive Prompts). Außerdem wurden Fragen im Übungstest, die zunächst falsch beantwortet wurden, im Laufe des Übungstests bis zu fünfmal präsentiert, bis die Teilnehmenden die richtige Antwort gaben (zur Wirksamkeit multipler Tests, s. z. B. Butler 2010). Eine Woche nach der Wiederholungssitzung erhielten alle teilnehmenden Studierenden wieder einen kriterialen Test, der aus Multiple-Choice- oder Kurzantwortfragen bestand. Die Ergebnisse zeigten einen deutlichen Testungseffekt ( $\eta_p^2 = .14$ ): Die getesteten Inhalte wurden besser behalten als die wiederholten Inhalte. Ein Effekt der metakognitiven Aktivierung zeigte sich jedoch nicht (s. Abb. 3). Eine mögliche Interpretation dieses Ergebnisses ist, dass sich Übungstests unabhängig davon, ob man sich ihrer positiven Lerneffekte bewusst ist oder nicht, positiv auf das Lernen auswirken. Es ist aber ebenso denkbar, dass die in Experiment 3 gewählte Umsetzung der Übungstests mit der mehrfachen Testung im Falle von Falschantworten die Übungstests bereits so wirksam gemacht hat, dass die metakognitive Aktivierung im Sinne eines Deckeneffekts keinen zusätzlichen Nutzen mehr haben konnte. Darauf deutet der sehr starke Testungseffekt hin, der in Experiment 3 auftrat.

**Abb. 3:****Der Testungseffekt mit und ohne Instruktion zur metakognitiven Aktivierung****(Experiment 3)**

*Anmerkung.* Fehlerbalken repräsentieren den Standardfehler des Mittelwerts. Die y-Achse zeigt den Anteil korrekter Antworten im kriterialen Test.

### **Fazit: Möglichkeiten und Grenzen der Nutzung von digitalen Übungstests in der Hochschullehre**

Unsere Experimente zeigen im Einklang mit einschlägigen Metaanalysen (z. B. Yang et al. 2021), dass Übungstests eine sinnvolle und effektive Möglichkeit sind, das nachhaltige Lernen im Hochschulkontext zu fördern. Sie zeigen weiterhin, dass selbst eine minimalistische Intervention, die in wenigen Minuten zusätzlich zur Vorlesung online umgesetzt wird, über verschiedene Vorlesungsthemen hinweg deutliche und robuste Effekte auf das Behalten von Vorlesungsinhalten haben kann. Dabei erwiesen sich sowohl offene Fragen (Kurzantwortfragen) als auch geschlossene Fragen (Kurzantwortfragen) als lernwirksam.

Studierende müssen offenbar keine besonderen kognitiven, motivationalen oder emotionalen Voraussetzungen mitbringen, um von Übungstests profitieren zu können. Besonders stark sind die positiven Effekte von Übungstests bei einer adaptiven Umsetzung: Bei zunächst falschen Antworten wird mehrfach getestet, bis die richtige Antwort erfolgt. Dies setzt neben einer digitalen Darbietung der Übungstests in der Regel die Verwendung von geschlossenen Fragen, in unserem Fall Multiple-Choice-Fragen, voraus.

Unsere Untersuchungen zeigen aber auch eine wesentliche Einschränkung der Wirksamkeit von Übungstests: Gefördert wird offenbar nur das Behalten der direkt getesteten Inhalte, Evidenz für einen Transfer des Testungseffekts auf gelernte, aber nicht getestete Inhalte haben wir nicht gefunden (s. auch Pan & Rickard 2018). Es ist allerdings möglich (und theoretisch plausibel, s. Carpenter 2009), dass solche Transfereffekte auftreten, wenn die Inhalte in einer Lerneinheit semantisch sehr eng miteinander zusammenhängen. Diese Frage sollte in der zukünftigen Forschung näher untersucht werden, nicht zuletzt wegen der hohen Relevanz und dem geringen Aufwand von digitalen Übungstests zur Unterstützung des nachhaltigen Lernens im Hochschulkontext und darüber hinaus.

## 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), 659–701. <https://doi.org/10.3102/0034654316689306>
- Arnold, K. A. & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 940-945. <https://doi.org/10.1037/a0029199>
- Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118–1133. <https://doi.org/10.1037/a0019902>
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1563–1569. <https://doi.org/10.1037/a0017021>
- Glaser, J. & Richter, T. (2023a). The testing effect in the lecture hall: Does it depend on learner prerequisites? *Psychology Learning and Teaching, 22*(2), 159–178. <https://doi.org/10.1177/14757257221136660>
- Glaser, J. & Richter, T. (2023b). The testing effect in the lecture hall: Does it transfer to content studied but not practiced? *Teaching of Psychology*. <https://doi.org/10.1177/00986283231218943>
- Greving, S. & Richter, T. (2018). Examining the testing effect in university teaching: Retrievability and question format matter. *Frontiers in Psychology, 9*, Article 2412. <https://doi.org/10.3389/fpsyg.2018.02412>
- Hinze, S. R., Wiley, J. & Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *Journal of Memory and Language, 69*(2), 151–164. <https://doi.org/10.1016/j.jml.2013.03.002>

Karpicke, J. D., Butler, A. C. & Roediger, H. L., III (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, 17(4), 471–479. <https://doi.org/10.1080/09658210802647009>.

Kornell, N. & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, 14(2), 219–224. <https://doi.org/10.3758/BF03194055>

Kubik, V., Gaschler, R. & Hausman, H. (2021). Enhancing student learning in research and educational practice: The power of retrieval practice and feedback. *Psychology Learning & Teaching*, 20(1), 1-20. <https://doi.org/10.1177/1475725720976462>

Pan, S. C. & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, 144(7), 710–756. <https://doi.org/10.1037/bul0000151>

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330(6002), 335. <https://doi.org/10.1126/science.1191465>

Richter, T., Berger, R., Ebersbach, M., Eitel, A., Endres, T., Borromeo Ferri, R. et al. (2022). How to promote lasting learning in schools: Theoretical approaches and an agenda for research. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology*, 54(4), 135–141. <https://doi.org/10.1026/0049-8637/a000258>

Roediger, H. L., III & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves longterm retention. *Psychological Science*, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432–1436. <https://doi.org/10.1037/a0037559>

Rummer, R. & Schweppe, J. (2022). Komplexität und der Testungseffekt: Die mögliche Bedeutung der Verständnissicherung für den Nutzen von Abrufübung bei

komplexem Lernmaterial. *Unterrichtswissenschaft*, 50, 37–52.

<https://doi.org/10.1007/s42010-021-00137-4>

Soderstrom, N. C. & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, 73(1), 99–115.

<https://doi.org/10.1016/j.jml.2014.03.003>

Schneider, W., Tibken, C. & Richter, T. (2022). The development of metacognitive knowledge from childhood to young adulthood: Major trends and educational implications. In J. Lockman (Hrsg.), *Advances in child development and behavior* (Vol. 63, S. 273–307).

Elsevier. <https://doi.org/10.1016/bs.acdb.2022.04.006>

Taraban, R., Maki, W. S. & Ryneerson, K. (1999). Measuring study time distributions: Implications for designing computer-based courses. *Behavior Research Methods, Instruments & Computers*, 31(2), 263–269. <https://doi.org/10.3758/BF03207718>

Yang, C., Luo, L., Vadillo, M. A., Yu, R. & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, 147(4), 399–435. <https://doi.org/10.1037/bul0000309>