

Berichte aus der Psychologie

**Michael Krämer, Siegfried Preiser,  
Kerstin Brusdeylins (Hrsg.)**

**Psychologiedidaktik und Evaluation XI**

Dem Wunsch mehrerer Autorinnen und Autoren folgend sei an dieser Stelle darauf hingewiesen, wenn in den folgenden Beiträgen entweder nur die männliche oder nur die weibliche Sprachform genutzt wird, so sind jeweils beide Geschlechter gemeint. Es geschieht ausschließlich der besseren Lesbarkeit halber.

Um den Datenschutz zu wahren, wurde auf die Veröffentlichung der Autorenadressen verzichtet. Wenn Sie Kontakt zu einer Autorin oder einem Autor aufnehmen wollen, schreiben Sie bitte eine e-mail an folgende Adresse. Der Herausgeber leitet Ihren Wunsch gerne weiter: [kraemer@fh-muenster.de](mailto:kraemer@fh-muenster.de)

# **Inhalt**

## **Aus-, Fort- und Weiterbildung in Psychologie**

MICHAEL KRÄMER

Novellierung des Psychotherapeutengesetzes – Stand der Dinge und Auswirkungen auf das Studium der Psychologie und die Profession 3

SIEGFRIED PREISER UND MICHAEL GIEBEL

Weiterbildungsmotivation von Studierenden der Psychologie 2012 - 2014 - 2016 11

LARS BEHRMANN

„Je höher der Frauenanteil eines Fachs, desto weniger Frauen promovieren ...“ 17

MIRIAM THYE, KATHARINA MOSEN, ULRICH WEGER UND DIETHARD TAUSCHEL  
Meditation und akademische Prokrastination – eine qualitative Studie 25

HANS-PETER LANGFELDT

Kulturelle Grenzen der (Psychologie)Didaktik  
– Erfahrungen an einer äthiopischen Universität 35

CARL P. A. KESSELER, STEFAN TROCHE UND MICHAELA ZUPANIC

Zur Konsistenz der Erwartungen an die ideale Persönlichkeit von Psychologie-Studierenden und die Auswirkungen auf Studienzufriedenheit und Studienleistung 43

CHARLOTTE VEHOFF, MICHAELA ZUPANIC, ROBIN JÖRN SIEGEL  
UND STEFAN TROCHE

Die Motivationsquellen studentischer GutachterInnen im Auswahlverfahren Psychologie: Neugier, Engagement oder doch das Bedürfnis nach Macht? 53

STEPHAN DUTKE, LENA KOEPCKE UND ELMAR SOUVIGNIER

Beiträge der Psychologie zum Praxissemester in der Lehramtsausbildung 61

## **Lehren und Lernen**

HANS-PETER NOLTING

Einführung in die Psychologie: themenbezogen oder systembezogen?  
Das Konzept der Integrativen Didaktik 73

MARLENE WAGNER, STEPHANIE MOSER, INES DEIBL UND JÖRG ZUMBACH

Psychologiedidaktik trifft Philosophiedidaktik:  
Der Einsatz von Wikis im interdisziplinären Unterricht 81

SIEGFRIED PREISER UND TAMARA TURASHVILI Wissenschaftskommunikation und Experten-Laien-Kommunikation: Kompetenzerwerb durch Trainingsseminare an deutschen und georgischen Universitäten	93
JULIA MENDZHERITSKAYA UND CAROLINE SCHERER Herausforderungen, Maßnahmen und Verbesserungsindikatoren guter Methodenlehre	101
LARS BEHRMANN UND STEFANIE VAN OPHUYSEN „Forschendes Lernen“ lernen – Die Methodenausbildung für Lehramtstudierende an der WWU Münster	109
DAGMAR TREUTNER Optimierung eines Kommunikationsseminars mit Videofeedback anhand von Videoannotation	119
BASTIAN HODAPP Medienbasiertes Forschendes Lernen – ein Modellprojekt	127
INGO JUNGCLAUSSEN UND SILVIANA STUBIG „Fack ju Pädda!“ – Neue Wege in der Didaktik der Pädagogischen Psychologie. Ergebnisse einer online-Umfrage zum Einsatz von Spielfilmen in der universitären Lehramtsausbildung am Beispiel der Schulkomödie „Fack ju Göhte“	135
INGO JUNGCLAUSSEN Die ‚Psychodynamik-Animation‘ – Ein mediengestützter Beitrag zur Didaktik der Psychoanalyse	145
MARKUS KNÖPFEL, FRANK MUSOLESI UND WILLI NEUTHINGER Konzeption eines PBL-Moduls im Rahmen des Psychologiestudiums	157
NICOLA BUCHHOLZ UND SUSANNE HILDEBRAND Selbstorganisiertes Lernen im Psychologieunterricht	167
NINA ZEUCH UND ELMAR SOUVIGNIER Wissenschaftliches Denken bei Lehramts- und Psychologiestudierenden	175
TOM ROSMAN, ANNE-KATHRIN MAYER UND GÜNTER KRAMPEN Die Förderung differenzierter epistemologischer Überzeugungen bei Studienanfängern der Psychologie: Empirische Befunde und fachdidaktische Implikationen	185

ANNE-KATHRIN MAYER, NIKOLAS LEICHNER UND GÜNTER KRAMPEN Förderung fachlicher Informationskompetenz von Psychologie- Studierenden durch ein curricular integriertes Blended Learning-Training	193
KATRIN B. KLINGSIECK, DANIEL AL-KABBANI, CARLA BOHDICK, JOHANNA HILKENMEIER, SEBASTIAN KÖNIG, HANNA S. MÜSCHE, SASKIA PRAETORIUS UND SABRINA SOMMER Gamebasiertes Lernen in der Lehrerbildung – spielend zur diagnostisch kompetenten Lehrkraft werden	203
NATHASHA BODONYI, VIKTORIA FALKENHORST UND ULRIKE STARKER, Planspiel – Papiersternmanufaktur	213
MIRIAM THYE, DÉsirÉE RITZKA, ROSE LINK UND DIETHARD TAUSCHEL Lernst du schon oder liest du noch? – Zu der Frage, wie man als Student das akademische Lernen lernen kann	219

## **Psychologie an Schulen**

PAUL GEORG GEIß Kompetenzorientierter Psychologieunterricht in Österreich	229
DOMINIK MOMBELLI Kompetenzorientierter Pädagogik- und Psychologieunterricht aus der gymnasialen Oberstufe in der Schweiz	239
JÜRGEN MALACH UND MARGRET PETERS Von der Input- zur Outputorientierung – Intention, Struktur und Implementation des kompetenzorientierten Kernlehrplans Psychologie für die gymnasiale Oberstufe NRW	247

## **Praxisbezogene Anwendung psychologischer Erkenntnisse**

KERSTIN BRUSDEYLINS UND JORINTHE HAGNER PENTApus Stressprävention für Oberstufenschüler und Studierende	257
TIMO BERSE Krank vor Sorgen – Ein Workshop zur klinischen Psychologie für Psychologielehrer/innen	263

SABINE FABRIZ, LUKAS SCHULZE-VORBERG UND HOLGER HORZ „Beratung und Betreuung von Studierenden im Studium“. Konzeption & Evaluation einer Schulungsreihe für schulische Betreuer/innen und Praktikumsbeauftragte im Praxissemester	271
--	-----

## **Evaluation**

MICHAEL KRÄMER Zufrieden und glücklich?! Zum Zusammenhang zwischen Studienzufriedenheit und Glücksempfinden	281
---	-----

ARNOLD HINZ Wie nützlich ist Lehrevaluation durch Studierende mittels Ratingskalen?	291
--	-----

DANIELA FEISTAUER UND TOBIAS RICHTER Wie zuverlässig sind studentische Einschätzungen der Lehrqualität? Eine Analyse mit kreuzklassifizierten Mehrebenenmodellen	299
--	-----

ELISABETH DALLÜGE, MICHAELA ZUPANIC, CORNELIA HETFELD UND MARZELLUS HOFMANN Wie bildet sich das Curriculum des Studiums im Progress Test Psychologie (PTP) ab?	307
---	-----

MICHAELA ZUPANIC, JAN P. EHLERS, THOMAS OSTERMANN UND MARZELLUS HOFMANN Progress Test Psychologie (PTP) und Wissensentwicklung im Studienverlauf	315
---	-----

JONATHAN BARENBERG, EVA SEIFRIED, BIRGIT SPINATH UND STEPHAN DUTKE Die Bearbeitung schriftlicher Problemaufgaben erhöht den Lernerfolg in einer Psychologie-Vorlesung	323
--	-----

JULIANE SCHWIEREN, JONATHAN BARENBERG UND STEPHAN DUTKE Testeffekt in Psychologie-Lehrveranstaltungen? Eine metaanalytische Perspektive	331
---	-----

DOROTHEA KRAMPEN, KARL SCHWEIZER, SIEGBERT REIß UND ANDREAS GOLD Erprobung einer Kurzsкала zur Erfassung von Impulsivität	339
---	-----

NIKOLAI ZINKE, STEFAN STÜRMER UND LAURA FROEHLICH Validierung einer deutschsprachigen Skala zur multidimensionalen Erfassung von interkulturellen Kompetenzen in der universitären Ausbildung	349
---	-----

# Testeffekt in Psychologie-Lehrveranstaltungen?

## Eine metaanalytische Perspektive

**Juliane Schwieren, Jonathan Barenberg und Stephan Dutke**

In einer Metaanalyse wurde untersucht, inwieweit der Testeffekt, ein über viele Kontexte hinweg stabiles Phänomen aus der empirischen Lehr-Lernforschung, auch im Rahmen von Psychologielehrveranstaltungen zu beobachten ist. Ein mittelgroßer Gesamteffekt von  $d = 0.56$ , der auf Basis von 72 Effektgrößen aus 19 Publikationen errechnet wurde, deutet darauf hin, dass der Einsatz von Tests im Rahmen der Psychologielehre wirksam ist. In weiteren Analysen wurden potentiell moderierende Einflüsse der Variablen Design, Art der Kontrollbedingung und Feedback untersucht. Die Abhängigkeiten zwischen Effektgrößen wurden kontrolliert. Die zentralen Ergebnisse werden in bestehende Forschungsbefunde eingeordnet und mit Blick auf Anwendungen in der Psychologielehre interpretiert.

Im Bildungskontext kommen Tests üblicherweise als Instrumente zur Leistungsabfrage und Leistungsbewertung zum Einsatz. Allerdings weisen zahlreiche Befunde zum sogenannten Testeffekt darauf hin, dass Testen auch eine lernförderliche Wirkung haben kann.

Der Testeffekt ist ein in der experimentellen Gedächtnispsychologie bereits früh beschriebenes, stabiles Gedächtnisphänomen (Abbott, 1909; Gates, 1917; Roediger & Karpicke, 2006): Mittels eines initialen Tests wiederholte Lerninhalte werden in einem späteren finalen Test besser erinnert als Inhalte, die zuvor nicht getestet oder aber in anderer Form (z.B. mittels erneuten Durchlesens) rekapituliert wurden. Trotz großer Heterogenität der Studien zum Testeffekt, lassen sich auf methodischer Ebene grundsätzlich drei Untersuchungsphasen ausmachen. Zunächst erfolgt in einer *initialen Lernphase* die Präsentation der Lerninhalte für alle Probanden. In der anschließenden *Interventionsphase* werden diese Inhalte entweder getestet, nicht getestet oder in anderer Form wiederholt. In der *finalen Testphase* werden sämtliche Lerninhalte aus der initialen Lernphase noch einmal abgefragt.

Eine große Anzahl von experimentellen Studien zeigte, dass der Testeffekt bei Stichproben unterschiedlichen Alters, mit unterschiedlich komplexen Lernmaterialien,

unter Einsatz variabler Testformate und Testfrequenzen und auch bei unterschiedlich anspruchsvollen finalen Tests eintritt (für eine Übersicht vgl. Dunlosky, Rawson, Marsh et al., 2013). Neben dem Bestreben, Aussagen über die Generalisierbarkeit des Testeffekts treffen zu können und Bedingungen zu identifizieren, die eine besonders effiziente und nachhaltige Wissenskonsolidierung erlauben, besteht ein weiterer Forschungsschwerpunkt im Auffinden der für den Testeffekt ursächlichen Mechanismen. Es wird angenommen, dass durch das wiederholte Abrufen von Informationen aus dem Gedächtnis Inhalte im Langzeitgedächtnis stärker elaboriert werden und verschiedene Zugangspfade zu den Inhalten entstehen (vgl. Dunlosky et al., 2013). Die Ergebnisse einer aktuellen Metaanalyse (Rowland, 2014) weisen außerdem darauf hin, dass auch die Anstrengung beim probeweisen Abrufen das Auftreten eines Testeffekts beeinflusst (schwierigere Tests in der initialen Lernphase führten zu einem größeren Testeffekt als leichtere Tests).

Wenn auch hinsichtlich der für den Testeffekt ursächlichen Mechanismen noch Forschungsbedarf besteht, besticht das Phänomen durch sein zuverlässiges Auftreten unter verschiedenen Untersuchungsbedingungen. Es überrascht daher nicht, dass das Testen von Lerninhalten als vielversprechende Technik im Rahmen des evidenzbasierten Lehrens aufgegriffen wurde (z. B. Dunn, Saville, Baker et al., 2013). Dieser Ansatz verfolgt das Ziel, den Einsatz empirisch validierter Lehr-Lernmethoden stärker in der schulischen und universitären Lehre zu verankern (Cranney, 2013).

Angesichts der stetig wachsenden Stofffülle im Psychologiestudium stellt sich die Frage, inwieweit Lehrende des Faches auf effektive Techniken wie beispielsweise das Testen zurückgreifen. Mit Hilfe einer Metaanalyse wurde untersucht, in welchem Ausmaß sich Testeffekte auch im Rahmen des Lernens psychologischer Inhalte beobachten und gewinnbringend nutzen lassen. Da eine steigende Anzahl von Studien auch in anwendungsnahen Settings lernförderliche Effekte von Tests aufgezeigt haben (z.B. Roediger, Agarwal, McDaniel et al., 2011), wurde ein positiver Gesamteffekt durch den Einsatz von Tests erwartet.

Die Metanalyse von Rowland (2014) hatte gezeigt, dass verschiedene Faktoren das Auftreten und die Größe des Testeffekts moderieren können. Daher wurde in einem zweiten Schritt der moderierende Einfluss folgender Merkmale überprüft:

*Studiendesign* (Zwischengruppen- vs. Innergruppen-Design), *Art der Kontrollbedingung* (Rekapitulation vs. kein Test) und *Feedback* (vorhanden vs. nicht vorhanden).

## **Methode**

Ausgangspunkt für die vorliegende Metanalyse bildete eine Literaturrecherche zum Testeffekt im Anwendungsfeld der Psychologielehre anhand der PsycINFO Datenbank und der Durchsicht von Literaturverzeichnissen diverser Übersichtsartikel. Am Ende dieser Recherche erfüllten 19 Publikationen im Erscheinungszeitraum von Juli 1984 bis Februar 2016 zwei Einschlusskriterien: (1) sämtliche Studien befassten sich mit der Vermittlung psychologischer oder psychologierelevanter Lerninhalte und ermöglichten (2) den direkten Vergleich einer Testbedingung mit einer adäquaten Kontrollbedingung (Rekapitulation oder kein Test).

Die metaanalytische Auswertung erfolgte mit dem R-Paket *metafor* (Viechtbauer, 2010). Für jeden Einzelvergleich wurde die Effektgröße Cohen's  $d$  als die standardisierte mittlere Differenz in der finalen Testleistung zwischen Testbedingung und Kontrollbedingung ermittelt. Für kleine Stichproben ( $N < 20$ ) wurde eine Korrektur der Effektgröße nach Hedges (1981) vorgenommen.

Die Berechnung einer mittleren Effektgröße über alle Einzelvergleiche hinweg erfolgte anhand eines Random-Effects-Modells, dem die Annahme zugrunde liegt, dass die Gesamtheit der beobachteten Effektgrößen auf einer Verteilung wahrer Werte beruht anstatt auf einem einzelnen wahren Wert. Folglich werden Unterschiede zwischen den beobachteten Effektgrößen nicht alleine als das Resultat von Stichprobenfehlern (Stichprobenvarianz) angesehen, sondern auch als das Ergebnis tatsächlicher Unterschiede zwischen den Effektgrößen (Parametervarianz). Im Lichte des vorliegenden heterogenen Stichprobenpools (unterschiedliche Implementierungen der Tests innerhalb der Lehre) erschien dieses statistische Modell angemessen. Zur Berechnung der Parametervarianz wurde im vorliegenden Fall ein Restricted-Maximum-Likelihood-Schätzer verwendet. Der berechnete Gesamteffekt ergab sich als Mittelwert aus der Verteilung der einzelnen, zuvor gewichteten Effektgrößen (diese Gewichtung berücksichtigte beide aufgeführten Varianzquellen, Stichprobenvarianz und Parametervarianz).

In einem weiteren Schritt wurde unter Verwendung eines Mixed-Effects-Modells eine Moderatorenanalyse mit den drei eingangs erwähnten jeweils zweistufigen Merkmalen *Studiendesign* (Zwischengruppen- vs. Innergruppen-Design), *Art der Kontrollbedingung* (Rekapitulation vs. kein Test) und *Feedback* (vorhanden vs. nicht vorhanden) durchgeführt.

Zur Korrektur von Abhängigkeiten zwischen einzelnen Effektgrößen, sowohl für das zuvor berechnete Random-Effects-Modell als auch für das Mixed-Effects-Modell, wurde eine im R-Paket *robumeta* (Fisher & Tipton, 2015) implementierte *Robuste Varianzschätzung* (RVE) in Form einer Metaregression vorgenommen. Dieses Korrekturverfahren wurde aus zwei Gründen angewandt: (1) einige Effektgrößen wurden zwar für verschiedene abhängige Variablen, aber auf Basis derselben Stichprobe ermittelt; (2) andere Effektgrößen wurden für dieselbe abhängige Variable anhand sich teilweise überlappender Substichproben ermittelt.

## Ergebnisse

Die Prüfung des Gesamteffekts über alle Studien hinweg ergab eine gewichtete mittlere Effektgröße von  $d = 0.56$ , 95% Konfidenzintervall (KI) [0.40, 0.71]. Dieses Ergebnis steht in Einklang mit dem typischen Befund bei Vorliegen eines Testeffekts, wonach die finale Testleistung in den Experimentalbedingungen signifikant höher ausfällt als in den jeweiligen Kontrollbedingungen. Es zeigte sich zudem, dass die Mehrheit der positiven Effektgrößen signifikant von Null verschieden war, während bei den negativen Effektgrößen nur ein einziger Wert signifikant von Null abwich und eine negative Auswirkung von Tests nahelegte. Tabelle 1 gibt einen Überblick über die Verteilung aller Effektgrößen in Bezug auf ihre Richtung (positiv vs. negativ) und Signifikanz.

**Tab. 1: Verteilung der einzelnen Effektgrößen nach Richtung und statistischer Signifikanz**

		95%-Konfidenzintervall		Gesamt
		Null ausschließend	Null einschließend	
Richtung	positiv	33	24	57
	negativ	1	13	14
	Gesamt	34	37	71

Anm.: Zusätzlich zu den 71 aufgeführten Effektgrößen wies eine den Wert  $d=0$  auf.

Die Moderatorenanalysen erbrachten nur einen signifikanten Effekt für die Variable Feedback ( $z = 2.42$ ,  $p = .02$ ). Durch den Einsatz von Feedback ließ sich ein Anstieg der mittleren Effektgröße auf  $d = 0.60$ , 95% KI [0.29, 0.91] verzeichnen, wenn gleichzeitig alle anderen Moderatoren konstant gehalten wurden.

Nach Einsatz der RVE-Prozedur zur Korrektur von Abhängigkeiten zwischen Effektgrößen ergab sich eine gewichtete mittlere Effektgröße von  $d = 0.62$ , 95% KI [0.32, 0.93]. Die RVE-korrigierte Moderatorenanalyse lieferte keine signifikanten Effekte, wobei der Moderator Feedback der Signifikanzgrenze am nächsten lag ( $t = 1.84$ ,  $p = 0.10$ ). Insgesamt lässt sich festhalten, dass die zwischen mehreren Effektgrößen bestehenden Abhängigkeiten zu einer leichten Unterschätzung der mittleren Effektgröße führten und zu einer leichten Überschätzung des Einflusses von Feedback.

## **Diskussion**

Die vorliegende Metaanalyse verfolgte das Ziel zu untersuchen, in welchem Umfang sich ein Testeffekt auch im Rahmen von Psychologielehrveranstaltungen nachweisen lässt. Damit verbunden ist auch die Frage, inwieweit der Einsatz von Tests in der Psychologielehre zu empfehlen ist.

Als zentrales Ergebnis kann festgehalten werden, dass sich auch in Studien mit psychologischen bzw. psychologierelevanten Lerninhalten ein Testeffekt nachweisen lässt. Wurde zwischen initialer Lernphase und finaler Testphase ein Übungstest absolviert, führte dies im Mittel zu einer Verbesserung der Abrufleistung in der finalen Leistungsüberprüfung, wobei dieser Effekt sich unter Berücksichtigung einer Korrektur für abhängige Effektgrößen sogar noch erhöhte. Als Moderator kommt die Gabe von Feedback im Übungstest in Frage, die zu einem signifikanten Anstieg des Testeffekts führte, jedoch nur so lange, wie keine Korrekturprozedur für Abhängigkeiten zwischen einzelnen Effektgrößen angewendet wurde.

Diese Ergebnisse sind vergleichbar mit Befunden aus anderen Überblicksarbeiten zum Testeffekt. In einer früheren Metaanalyse von Bangert-Drowns, Kulik & Kulik (1991) wurden ebenfalls Einsatz und Wirkung von Tests in alltäglichen Lehrkontexten betrachtet, allerdings nicht ausschließlich in Psychologieveranstaltungen. Hier ermittelten die Autoren für eine Substichprobe von 11 Studien eine mittlere

Effektgröße von 0.54 Standardabweichungen und damit die nahezu gleiche Effektgröße wie in der aktuellen Analyse. Die Verstärkung des Testeffekts durch Feedback lässt sich gut mit den Befunden aus der Metaanalyse von Rowland (2014) vereinbaren. In dieser Analyse wurde eine mittlere Effektgröße von 0.73 beim Einsatz von Feedback festgestellt, im Vergleich zu einer mittleren Effektgröße von 0.39, wenn zuvor kein Feedback erfolgt war. Allerdings ist zu berücksichtigen, dass Rowland ausschließlich laborexperimentelle Studien eingeschlossen hatte.

Die Befundlage spricht also dafür, dass Tests auch im Kontext der Psychologielehre erfolgreich als lernförderliche Maßnahme eingesetzt werden können, zumal ihre Durchführung nur wenig Zeit beansprucht und sie leicht anzuleiten sind. Auch das Risiko einer Verschlechterung des Lernergebnisses durch den Einsatz von Tests scheint gering zu sein. In einer einzigen Studie ergab sich ein signifikant negativer Effekt nach vorherigem Übungstest.

Ein Problem, wie bei vielen praktischen Anwendungen des Testeffekts, besteht darin, Lernende für die lernförderliche Wirkung des Testens zu sensibilisieren. In einer Hochschullehre, in der Effizienz und Noten eine zentrale Rolle einnehmen, werden Tests eher als Mittel der Leistungsmessung verstanden und nur selten als eine Maßnahme zur Förderung von Lernen – und wenn, dann in erster Linie, weil Tests über den Erwartungshorizont von Prüfungen aufklären und nicht, weil sie als Übung des Gedächtnisabrufs verstanden werden. Andererseits sollte dieser Hintergrund gerade Psychologiestudierenden leicht zu vermitteln sein – schließlich sind das Ergebnisse der Forschung in ihrem eigenen Fach, die auf ihre eigenen Lernprozesse angewandt werden. Die wichtigste Begrenzung der vorliegenden Studie ist die geringe Anzahl von Studien, die in die Analyse einbezogen wurden. In der Folge ist die Anzahl abhängiger Effektgrößen beträchtlich, ein Problem, das post-hoc nur bedingt kontrolliert werden kann. Angesichts der wiederkehrenden Kritik, psychologische Forschung beruhe zu oft auf Daten, die an Psychologiestudierenden erhoben wurden, ist es erstaunlich, dass ausgerechnet in einem Feld, das die Psychologielehre unmittelbar betrifft, vergleichsweise wenig Daten von Psychologiestudierenden vorliegen. Es bleibt zu diskutieren, inwieweit die Psychologie das Konzept des evidenzbasierten Lehrens ernst nimmt. Hierzu gehört, nicht nur ihre lehr-lernbezogenen Forschungsergebnisse in der eigenen Lehre anzuwenden, sondern auch durch Erforschung der eigenen Lehre zum

Wissenskorpus evidenzbasierter Lehre beizutragen. Notwendig und nützlich wäre das allemal auch im Kontext des Testeffekts. Theoretisch wie praktisch bedeutsame Moderatoren wie Feedback, Art der begleitenden Lernaktivitäten oder Prüfungsangst sind nicht hinreichend erforscht, ebenso wenig wie notwendige oder förderliche Implementationsbedingungen des Testens zur Übung des Gedächtnisabrufs. Auch die Beantwortung der Frage, warum sich in der vorliegenden Analyse (trotz des signifikanten Gesamteffekts) gut die Hälfte der extrahierten Effektgrößen nicht signifikant von Null unterschied, gibt Anlass zu einer weiteren Erforschung des Testeffekts in der Psychologielehre.

## Literatur

- Abbott, E. E. (1909). On the analysis of the factors of recall in the learning process. *Psychological Monographs*, 11, 159-177.
- Bangert-Drowns, R. L., Kulik, J. A. & Kulik, C.-L. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, 85, 89-99.
- Cranney, J. (2013). Toward psychological literacy: A snapshot of evidence-based learning and teaching. *Australian Journal of Psychology*, 65, 1-4.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J. & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14, 4-58.
- Dunn, D. S., Saville, B. K., Baker, S. C. & Marek, P. (2013). Evidence-based teaching: Tools and techniques that promote learning in the psychology classroom. *Australian Journal of Psychology*, 65, 5-13.
- Fisher, Z. & Tipton, E. (2015). robumeta: Robust Variance Meta-Regression. R package version 1.6. Verfügbar unter: <http://CRAN.R-project.org/package=robumeta> (abgerufen am 25.07.2016).
- Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, 6(40).
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6, 107-128.

- Roediger, H. L., III, Agarwal, P. K., McDaniel, M. A. & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, 17, 382-395.
- Roediger, H. L., III. & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181-210.
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, 140, 1432-1463.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36, 1-48. Verfügbar unter: <https://www.jstatsoft.org/article/view/v036i03> (abgerufen am 25.07.2016).