

Quizze im Unterricht und in der Lehre: Wie häufiges Testen zum Lernen beiträgt.

Quizze sind in der Unterrichtspraxis eine gerne genutzte Methode, um das Lernen und das Behalten von Wissen zu fördern. Doch wie tragen regelmäßige Quizze tatsächlich zur Förderung der Lernleistung bei? Welche Bedingungen und Formate sind dabei besonders erfolgversprechend? Diese und weitere Fragen wurden in der Metaanalyse¹ "Regarding class quizzes: A meta-analytic synthesis of studies on the relationship between frequent low-stakes testing and class performance" von Sotola und Crede (2021) untersucht, die sich mit der Wirksamkeit von Quizzen im Bildungskontext beschäftigt.

EINLEITUNG. Quizabfragen sind eine bewährte Methode im Klassenzimmer und erfreuen sich großer Beliebtheit bei Lehrkräften und SchülerInnen. Regelmäßiges Testen in Form von Quizzen bietet eine unkomplizierte Möglichkeit, SchülerInnen Rückmeldung darüber zu geben, wie gut sie den aktuellen Unterrichtsstoff verstehen.

Empirisch fundierte Lerntheorien, wie der Testeffekt (engl. testing effect), unterstützen den Ansatz, dass wiederholtes Testen – etwa durch unbenotete Quizze – das Verständnis und die Leistung der Lernenden verbessert. Der Testeffekt besagt, dass Lernende beim Beantworten von Testfragen Inhalte aktiv abrufen müssen, wodurch das Wissen langfristiger verankert wird. Darüber hinaus zeigt die Forschung, dass häufiges Testen auch das Erlernen neuen Wissens fördert (test-potentiated learning).

Diese Metaanalyse fasst erstmals die Ergebnisse von Feldstudien zum Einsatz von Quizzen zusammen. Die Erkenntnisse sind besonders relevant für die Lehrpraxis an Schulen und Hochschulen, da sie unter realen Bedingungen im Klassenzimmer – und nicht im Labor – gewonnen wurden.

METAANALYSE IM ÜBERBLICK

Fokus der Studie	Auswirkungen regelmäßiger Quizabfragen auf die Lernleistung
Untersuchte Zielgruppe	7.846 SchülerInnen und Studierende in Schule und Hochschule
Durchschnittliche Effektstärke	Der regelmäßige Einsatz von Quizzen zeigt einen positiven Effekt auf die Lernleistung (Cohen's $d = 0.42$)
Weitere Befunde	Quizze, die zur Abschlussnote zählten, erzielten höhere Effekte ($d = 0.51$) als Quizze, die nicht in die Abschlussnote eingingen ²

¹Die unterstrichenen Fachbegriffe sind auf unser Online-Glossar verlinkt.

²Weitere Trends zeigten höheren Effekt bei unmittelbarem Feedback (vs. verzögert); Offenen Fragen (vs. Multiple-Choice); Angekündigtes Quiz (vs. Unangekündigt)

WURUM GEHT ES IN DIESER STUDIE? In dieser Metaanalyse untersuchen die Autoren vier zentrale Fragestellungen: Erstens prüfen sie, ob regelmäßige Quizze im Unterricht im Vergleich zu Unterricht ohne Quizze einen Mehrwert bieten. Dafür analysieren sie 52 (quasi-)experimentelle Studien mit insgesamt 7864 Teilnehmenden, die zwischen 1932 und 2017 veröffentlicht wurden. Diese Studien vergleichen die Kursleistungen (Noten) in Unterrichtssettings mit und ohne Quizze.

Zweitens untersuchen die Autoren, unter welchen Bedingungen Quizze besonders wirksam sind. In sogenannten Moderatorenanalysen prüfen sie, ob die Effektstärke von bestimmten Faktoren wie Testhäufigkeit, sofortigem oder verzögertem Feedback, Fragetypen, der Durchführung als Hausaufgabe oder im Unterricht sowie einer vorherigen Ankündigung der Quizze beeinflusst wird. Ebenso betrachten sie Rahmenbedingungen wie die Bildungsstufe (Schule vs. Hochschule) und das Fach, wobei sie zwischen Psychologie (etwa die Hälfte der Studien) und anderen Fächern unterscheiden.

Drittens analysieren sie anhand von 5 Studien mit 1004 Teilnehmenden, ob mehr Lernende den Kurs bestehen, wenn Quizze häufiger eingesetzt werden, da insbesondere schwächere Lernende von regelmäßigen Tests profitieren könnten.

Abschließend wird untersucht, ob ein Zusammenhang zwischen den Quizleistungen und der späteren Kursnote besteht. Hierfür lieferten 20 Studien mit insgesamt 4268 Teilnehmenden die erforderlichen Daten³.

WAS FAND DIESE STUDIE HERAUS? Über alle Studien hinweg ergibt sich in dieser Metaanalyse ein signifikanter Gesamteffekt von $d = 0.42$ (Konfidenzintervall $d = 0.29$ bis $d = 0.55$) zugunsten der Leistung von Lernenden, die regelmäßig mit Quizzen getestet wurden.

In den Moderatorenanalysen zur zweiten Frage zeigte sich, dass die Effektstärken deskriptiv größer waren, wenn die Quizzergebnisse in die Endnote einfließen ($d = 0.51$) im Vergleich zu Quizzen ohne Einfluss auf die Note ($d = 0.33$). Höhere Effektstärken traten auf, wenn das Feedback unmittelbar erfolgte ($d = 0.41$) im Vergleich zu verzögertem Feedback ($d = 0.31$) und wenn die Quizze aus offenen Fragen bestanden ($d = 0.60$) im Gegensatz zu Multiple-Choice-Fragen ($d = 0.37$) oder Lückentext- und anderen Fragetypen ($d = .38$). Deskriptive Unterschiede in den Effektstärken waren auch für einige der Moderatoren erkennbar, die aus rein explorativen Gründen untersucht wurden. Das heißt, die Effektstärken waren größer, wenn die Quizze außerhalb des Klassenzimmers online stattfanden ($d = 0.62$) anstatt im Klassenzimmer ($d = 0.22$); angekündigt ($d = 0.43$) im Vergleich zu nicht angekündigt ($d = 0.15$) und wenn sie im Fach Psychologie stattfanden ($d = 0.47$) im Vergleich zu anderen Fächern ($d = 0.29$). In einer Metaregression wurden davon zwei Moderatoren signifikant: 1. Einfluss auf die Endnote: wenn die Quizzleistungen Einfluss auf die Endnote hatten, waren die Leistungen signifikant höher und 2. das Fach: Im Fach Psychologie zeigten sich signifikant höhere Effektstärken als in den anderen Fächern.

Die Befunde zur dritten Fragestellung zeigten, dass der Anteil der Lernenden, die den Kurs bestehen größer war, wenn Quizze eingesetzt wurden.

³Diese Studien sind Teil der oben genannten Gesamtstichprobe von 52 Studien und liefern die für diese Fragestellung notwendigen Daten.

Schließlich zeigte die Untersuchung des Zusammenhangs zwischen Quizeinsatz und späteren Prüfungsleistungen eine deutliche, statistisch signifikante Korrelation von $r = 0.55$.

Weitere Details und Einzelbefunde sind der Gesamtübersicht zu entnehmen.

WIE BEWERTET DAS CLEARING HOUSE UNTERRICHT DIESE STUDIE? Die Clearing House Unterricht Research Group bewertet die Metaanalyse anhand der folgenden fünf Fragen und orientiert sich dabei an den Abelson-Kriterien (1995):

Wie substanziell sind die Effekte? Die Metaanalyse kommt bezüglich der Lernleistungen der SchülerInnen insgesamt zu einem signifikanten positiven Gesamteffekt $d = 0.42$. Diese Effektstärke bedeutet, dass etwa 66 % der SchülerInnen, die regelmäßig in realen Lehr-Lernsettings mit Quizzen getestet wurden, einen größeren Lernerfolg zeigten als der Durchschnitt der SchülerInnen in den Vergleichsgruppen. Zudem zeigen die Moderatorenanalyse, dass dieser Gesamteffekt über unterschiedliche Bedingungen weitgehend robust ist.

Vorherige Metaanalysen (z.B. Anderson et al., 1999; Vanhove & Harms, 2015) dokumentieren zudem, dass ähnliche Effektstärken auch unter Laborbedingungen beobachtet wurden. Zudem zeigte eine weitere aktuelle und umfassende Metaanalyse von Yang et al. (2020) ebenfalls einen positiven Effekt von $g = 0.50$, was die Ergebnisse dieser Metaanalyse bestätigt.

Erfahren Sie mehr über die Einschätzung von Effektstärken in unserem Handout.

Wie differenziert sind die Ergebnisse dargestellt? Die Differenziertheit der berichteten Effekte wird von der Clearing House Unterricht Research Group anhand der untersuchten Fachbereiche, Bildungsstufen und abhängigen Variablen eingeschätzt. In dieser Metaanalyse werden die Ergebnisse differenziert nach Bildungsstufe (Sekundarstufe vs. tertiäre Stufe) und Fachbereich untersucht. Bei den Fachbereichen wurde zwischen Befunden in Kursen der Psychologie und anderen Fächern (darin enthalten: u.a. naturwissenschaftliche Fächer oder Finanzen und Management) unterschieden, da viele der inkludierten Studien mit Studierenden des Fachbereichs „Psychologie“ durchgeführt wurden.

Sämtliche Befunde der Metaanalyse beziehen sich auf die Leistung der Lernenden; weitere mögliche Effektkategorien wie etwa Lernmotivation wurden nicht untersucht.

Wie verallgemeinerbar sind die Befunde? Die Autoren prüfen mithilfe zweier unterschiedlicher Analyseansätze bezogen auf die Moderatorvariablen, inwiefern der Gesamteffekt für unterschiedliche Bedingungen verallgemeinerbar ist. Als Erstes vergleichen sie deskriptiv die aggregierten Befunde der unterschiedlichen Stufen aller einbezogener Moderatorvariablen. Hier zeigten sich auf vielen Moderatorvariablen beachtenswerte Unterschiede (siehe Ergebnisse). Da diese Unterschiede jedoch nicht auf statistische Signifikanz hin getestet wurden und zum Teil auch erhebliche Unterschiede in der Zahl der eingeflossenen Primärstudien bestehen, sind diese Befunde als Trends zu verstehen und lassen nur bedingt Rückschlüsse auf Einschränkungen der Verallgemeinerbarkeit zu. Als Zweites führen die Autoren eine Metaregressionsanalyse durch. Hierbei zeigte sich, dass Quizze unterschiedlich wirksam sein können, je nachdem in welchen Fachbereichen sie durchgeführt werden und ob Quizzleis-

tungen für die Endnote gewertet werden. Für alle anderen Moderatoren ergab die Analyse keine Einschränkung der Verallgemeinerbarkeit. Bei der Auswahl der Moderatorvariablen zeigt sich, dass der Fokus vor allem auf Aspekten der Durchführung lag. Merkmale der Lernenden, wie beispielsweise deren Vorwissen, sowie Kontextvariablen wie die geografische Region, wurden hingegen nicht berücksichtigt.

Was macht die Metaanalyse wissenschaftlich relevant? Die vorliegende Metaanalyse untersucht im Gegensatz zu früheren Metaanalysen nur Befunde aus Primärstudien, in denen Quizze tatsächlich im Kontext des Schulunterrichts oder in der universitären Lehre eingesetzt wurden. Die Befunde lassen sich demnach mit höherer Wahrscheinlichkeit auf diese Kontexte übertragen verglichen mit Befunden reiner Laborstudien. Zudem wurden auf Basis verschiedener Theorien (z.B. Testeffekt, test potentiated learning) eine sehr umfassende Analyse von moderierenden Variablen vorgenommen.

Wie methodisch verlässlich sind die Befunde? Die Offenlegung und Begründung des methodischen Vorgehens entspricht überwiegend den Kriterien gängiger Anforderungskataloge (z.B. APA Meta-Analysis Reporting Standards). Die Schritte bei der Auswahl der Primärstudien und der Analyse der Befunde sind größtenteils transparent. Damit sind das Vorgehen bei der Suche nach relevanten Studien, Studiauswahl, Kodierung und statistische Analyse weitgehend offengelegt und nachvollziehbar beschrieben. Im Bereich der Kodierung der Primärstudien wären allerdings genauere Angaben zur Messung des Lernerfolgs, zur Kodierung der jeweiligen Studienqualität sowie zum Herkunftsland der Studie wünschenswert gewesen. Weitere Informationen zur methodischen Beurteilung finden Sie in unserem Rating Sheet.

FAZIT FÜR DIE UNTERRICHTSPRAXIS. Die Metaanalyse verdeutlicht, dass regelmäßige Quizze eine effektive Methode zur Förderung des Lernens sowohl im Schulunterricht als auch in der Hochschullehre darstellen. Die Ergebnisse untermauern die theoretischen Annahmen zum „Testing Effect“ und „Test-Potentiated Learning“ und zeigen, dass Quizze Lehrkräften eine einfache und kostengünstige Möglichkeit bieten, den Lernfortschritt der SchülerInnen gezielt zu verfolgen und Wissenslücken systematisch zu schließen. Dadurch wird das Verständnis der Lerninhalte verbessert und die gezielte Vorbereitung auf Prüfungen unterstützt.

Besonders wirksam erwies sich der Einsatz von Quizzen, wenn deren Ergebnisse in die Endnote einfließen. Neben einer allgemeinen Steigerung der Lernleistung trugen Quizze auch dazu bei, die Anzahl der SchülerInnen zu verringern, die einen Kurs nicht bestehen. Diese Methode erweist sich somit als praktisches und nachhaltiges Werkzeug zur Förderung des Lernerfolgs. Fortschritte in der künstlichen Intelligenz und der Entwicklung digitaler Quiztools erleichtern die Erstellung personalisierter Quizfragen, was den Einsatz dieser Methode weiter vereinfacht und deren Potenzial zur Lernförderung noch steigert.

STUDIENBEISPIEL

McDaniel et al. (2011) untersuchten in einer experimentellen Studie, wie regelmäßige Quizze die Lern- und Behaltensleistung im naturwissenschaftlichen Unterricht einer Middle School (vergleichbar mit der Sekundarstufe I) fördern können. Es wurden verschiedene Themenbereiche u.a. Genetik, Anatomie und Evolution in mehreren Unterrichtsstunden behandelt. In diesen sog. Unterrichtseinheiten wurde jeweils ein Thema behandelt, welches nach etwas 20 Tagen durch eine Prüfung abgeschlossen wurde. Die Studie wurde mit 139 AchtklässlerInnen in den USA in drei experimentellen Designs durchgeführt. Dabei erhielten die SchülerInnen zu drei Zeitpunkten Quizze im Multiple Choice Format:

- **Pre-Lecture-Quiz:** Quiz vor der Unterrichtsstunde zur Aktivierung des Vorwissens.
- **Post-Lecture-Quiz:** Quiz direkt nach der Unterrichtsstunde zur Festigung des im Unterricht behandelten Inhalts.
- **Review-Quiz:** Quiz vor Prüfungen zur gezielten Wiederholung der Inhalte.

Die Pre- und Post-Lecture-Quizze wurden über den gesamten Zeitraum der jeweiligen Unterrichtseinheit an die SchülerInnen ausgegeben. Die Versuchsreihen endeten jeweils mit einem Review-Quiz, das etwa 24 Stunden vor der Prüfung am Ende der Unterrichtseinheit durchgeführt wurde. Die Prüfungen fanden je nach experimentellen Design am Ende der Unterrichtseinheit, am Ende des Semesters und am Ende des Jahres statt. Zur genaueren Untersuchung der Behaltensleistung wurde zwischen den Review-Quizzen vor der ersten Prüfung und den späteren Prüfungen keine weiteren Quizze durchgeführt. Die Hälfte der Testfragen waren inhaltlich übereinstimmend mit Fragen, die auch in den späteren Prüfungen abgefragt wurden. Die andere Hälfte an prüfungsrelevanten Inhalten wurde in Lektüren und im Unterricht besprochen. Dies ermöglichte es, den Effekt der Quizze auf die Lern- und Behaltensleistung im Vergleich zu regulärem Unterricht ohne Quizze zu untersuchen.

Die Ergebnisse zeigten, dass Review-Quizze die stärkste Verbesserung der Prüfungsleistungen erzielten: Gequizzte Inhalte wurden zu 86 % korrekt beantwortet, verglichen mit 64 % bei nicht gequizzten Inhalten. Post-Lecture-Quizze zeigten ebenfalls positive Effekte, während Pre-Lecture-Quizze keinen signifikanten Einfluss hatten. In einem weiteren Experiment wurde deutlich, dass sowohl einzelne Review-Quizze als auch Review- und Post-Lecture-Quizze kombiniert, langfristig den größten Nutzen für die Lernleistung der SchülerInnen bringen. Laut den Autoren, können strategisch platzierte Quizze eine effektive Methode sein, um sowohl kurzfristige Lernziele zu erreichen als auch das langfristige Behalten von Wissen zu fördern.

REFERENZEN.

- Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, C. A., Lindsay, J. & Bushman, B. (1999). Research in the laboratory: truth or triviality? *Current Directions in Psychological Science*, 8 (1), 3-9. doi: 10.1111/1467-8721.00002
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, NJ: Erlbaum.
- McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B. & Roediger III, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of educational psychology*, 103 (2), 399-414. doi: 10.1037/a0021782
- Sotola, M., L. K. und Crede. (2021). Regarding class quizzes: A meta-analytic synthesis of studies on the relationship between frequent low-stakes testing and class performance. *Educational Psychology Review*, 33, 407-426. doi: 10.1007/s10648-020-09563-9
- Vanhove, A. J. & Harms, P. D. (2015). Reconciling the two disciplines of organisational science: a comparison of findings from lab and field research. *Applied Psychology*, 64 (4), 637-673. doi: 10.1111/apps.12046
- Yang, C., Luo, L., Vadillo, M. A., Yu, R. & Shanks, D. R. (2021). Testing (quizzing) boots classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, 147 (4), 399-435. doi: 10.1037/bul0000309

LINKS.

Zur Metaanalyse von Sotola & Crede, 2021

Zum Studienbeispiel von McDaniel et al., 2011

ZITIEREN ALS.

Knogler, M., Arvaneh, B. & CHU Research Group (2024). Quizze im Unterricht und in der Lehre: Wie häufiges Testen zum Lernen beiträgt. www.clearinghouse-unterricht.de, *Kurzreview* 33.



Dieses Kurzreview ist lizenziert unter einer [Creative Common Namensnennung - Keine Bearbeitungen 4.0 International Lizenz](https://creativecommons.org/licenses/by-nc/4.0/). Verwendung und Verbreitung unter Namensnennung erlaubt, keine Veränderungen gestattet.