



Wie effektiv sind *Science Programs*?

21 Varianten im Praxistest

Infolge der Lehrplanreformen im naturwissenschaftlichen Unterricht haben Verlage und staatliche Institute in den USA einiges unternommen, um die aktuellen Bildungsstandards in der Praxis umzusetzen. Ein zentrales Angebot sind sogenannte *Science Programs*. Diese Komplett-Pakete enthalten sowohl Lern- bzw. Unterrichtsmaterial als auch begleitende Schulungen für die Lehrkräfte. Die Einführung derartiger Programme im Schulalltag ist mit einigem Aufwand verbunden. Die Metaanalyse¹ »Effective secondary science programs: A best-evidence synthesis« von Cheung, Slavin, Kim, & Lake (2017) untersucht, bei welchen *Science Programs* sich das für den Lernerfolg der SchülerInnen auszahlt.

EINLEITUNG. Aktuelle nationale wie internationale Reformen von Bildungsstandards für den naturwissenschaftlichen Unterricht haben eine gemeinsame Stoßrichtung: SchülerInnen sollen ein Verständnis von Konzepten und Zusammenhängen entwickeln sowie wissenschaftliche Kompetenzen erwerben, anstatt in erster Linie Fakten zu lernen. Verlage und staatliche Institutionen stellen eine Fülle von Angeboten zur Verfügung, mit denen die geforderte Umstrukturierung und Weiterentwicklung des naturwissenschaftlichen Unterrichts erleichtert werden soll. In den USA sind deshalb zahlreiche sogenannte *Science Programs* entstanden. Sie werden als Komplett-Pakete angeboten und umfassen an die aktuellen Bildungsstandards angepasstes, vollständig ausgearbeitetes Lern- bzw. Unterrichtsmaterial sowie mehr oder weniger umfangreiche begleitende Schulungen für die Lehrkräfte.

Die AutorInnen der vorliegenden Metaanalyse untersuchen, inwiefern und bei welchen *Sci-*

METAANALYSE IM ÜBERBLICK

Fokus der Studie	Lernwirksamkeit unterschiedlicher <i>Science Programs</i> , die Lern-/Unterrichtsmaterial mit Fortbildung für die Lehrkräfte kombinieren
Untersuchte Zielgruppe	Über 33.000 SchülerInnen der Sekundarstufe in den USA
Durchschnittliche Effektstärke	Im Mittel erreichen die <i>Science Programs</i> bei einem mind. 12-wöchigen Einsatz im Schulalltag einen kleinen positiven Effekt auf MINT-Leistungen im Vergleich zu regulärem Unterricht ($ES^2 = 0.17$)
Weitere Befunde	Deutliche Unterschiede in der Effektivität der <i>Science Programs</i> ; Programme, die sich auf <i>Digitale Lehrmaterialien</i> ($ES = 0.47$) und <i>Lehrstrategientraining</i> ($ES = 0.17$) konzentrieren, schneiden am besten ab

¹ Definitionen der unterstrichenen Fachbegriffe finden Sie per Mausklick in unserem [Online-Glossar](#).

² Die AutorInnen haben die Effektstärken in der Metaanalyse für die eingehenden Primärstudien neu berechnet. Dabei kommt jeweils diejenige Berechnungsformel zum Einsatz, die der Datenlage so exakt wie möglich entspricht. Aus diesem Grund verwenden die AutorInnen den Überbegriff Effektstärke (ES) anstelle einer spezifischen Formel (z.B. für Cohen's d).

ence Programs es sich im Hinblick auf den Lernerfolg von SchülerInnen auszahlt, diese Programme gegenüber regulärem Unterricht in den Schulalltag einzuführen. Damit erweitert die Metaanalyse den Wissensstand zur Effektivität innovativer Unterrichtsansätze (vgl. Kurzreview 12) in zwei Hinsichten:

1) Sie prüft die Effektivität ausgearbeiteter *Science Programs* im mehrwöchigen Einsatz unter Realbedingungen im Schulalltag und

2) sie berücksichtigt exklusiv SchülerInnen aus der Mittel- und Oberstufe. Zudem liefert sie Anhaltspunkte, welche Typen von Programmen Schülerleistungen besonders effektiv steigern. Entsprechend kann sie eine Orientierung geben, wie aktuelle Bildungsstandards effektiv umgesetzt werden können.

WORUM GEHT ES IN DIESER STUDIE? Die AutorInnen der Metaanalyse tragen ausgearbeitete *Science Programs* für den Sekundarbereich zusammen. Sie analysieren deren Effektivität im Hinblick auf den Lernerfolg bei konzeptuellem Verständnis und wissenschaftlichen Kompetenzen der SchülerInnen. Inhaltlich berücksichtigen die AutorInnen nur Programme, die Lern- bzw. Unterrichtsmaterial, hier kategorisiert als *Digitale Lehrmaterialien*, *Experimentiersets* oder *Innovative Schulbücher*, mit einer begleitenden und teilweise recht umfangreichen Fortbildung für die Lehrkräfte, hier kategorisiert als *Lehrstrategie-entraining*, kombinieren.

Alle Programme haben den Anspruch, prinzipiell in jeder Schule umsetzbar zu sein. Zudem musste in den berücksichtigten Primärstudien die Wirksamkeit der Programme im Vergleich zu regulärem Unterricht im Schulalltag der Sekundarstufe über einen Zeitraum von mindestens zwölf Wochen untersucht werden. Nur Studien, die hohe wissenschaftliche Ansprüchen erfüllen – z.B. randomisierte, quasi-experimentelle Designs oder Kontrolle von Vorwissensunterschieden kamen für die Metaanalyse infrage. Insgesamt 21 solcher Programme konnten die AutorInnen identifizieren. An den Studien nahmen über 33.000 SchülerInnen teil. Zwei der Studien wurden vor dem Jahr 2000, acht zwischen 2000 und 2009 und elf zwischen 2010 und 2015 publiziert.

In einem ersten Systematisierungsschritt kategorisieren die AutorInnen die 21 Programme auf Grundlage der zentralen Lernaktivitäten. In *Tabelle 1* werden die vier verschiedenen Programmtypen mit ihren Charakteristika vorgestellt. Bei der statistischen Berechnung der Wirksamkeit der Programme (Effektstärke *ES*) beziehen die AutorInnen auftretende Unterschiede im Vorwissen der teilnehmenden SchülerInnen ein und berücksichtigen den Einfluss methodischer Faktoren auf die Effekte. Damit liegt der Metaanalyse ein bemerkenswert strenges Berechnungsverfahren zugrunde.

Table 1. Typen von MINT-Programmen mit Beschreibung und Anzahl der Studien.

PROGRAMMTYP	BESCHREIBUNG	ANZAHL STUDIEN
Lehrstrategietraining	<p>Diese Programme setzen vor allem auf das Training und Coaching von Lehrkräften beim Einsatz schülerzentrierter Unterrichtsformen wie dem Forschenden, Kooperativen oder projektbasierten Lernen und stellen zusätzlich entsprechendes Lern-/Unterrichtsmaterial bereit.</p> <p>Annahme zur Wirksamkeit: Umfassende Fortbildungen ermöglichen es Lehrkräften, neue Lehrstrategien effektiv im Unterricht einzusetzen, wodurch sich die Schülerleistung verbessert.</p> <p>Fortbildungsdauer: Mehrere Tage oder Wochen; häufig mit Live-Coaching im Klassenzimmer.</p>	6
Digitale Lehrmaterialien	<p>Diese Programme setzen vor allem auf den gesteigerten Einsatz von digitalen Anwendungen im Unterricht.</p> <p>Annahme zur Wirksamkeit: Digitale Anwendungen erhöhen die Lernwirksamkeit, indem sie wissenschaftliche Konzepte und Prozesse anschaulich visualisieren, eine stärkere Individualisierung des Lernangebots und schnelle Rückmeldung ermöglichen.</p> <p>Fortbildungsdauer: Je nach Programm; häufig Unterstützung im Klassenzimmer.</p>	5
Experimentiersets	<p>Mit Experimentiersets können SchülerInnen im Unterricht unter Anleitung selbst experimentieren. Lehrerfortbildungen zum Einsatz des jeweiligen Sets sind integraler Bestandteil dieses Programmtyps.</p> <p>Annahme zur Wirksamkeit: Durch das gut strukturierte Material und die Fortbildung können Lehrkräfte ihre SchülerInnen effektiv beim Experimentieren unterstützen. Diese lernen, durch eigenes Experimentieren wie Wissenschaftlerinnen zu denken und zu agieren.</p> <p>Fortbildungsdauer: In der Regel mehrere Tage.</p>	2
Innovative Schulbücher	<p>Diese Programme führen neu entwickelte Schulbücher ein, deren Texte, Inhalte und Aufbereitungsarten an die aktuellen Fächerstandards angepasst sind.</p> <p>Annahme zur Wirksamkeit: Die Lernwirksamkeit des Unterrichts wird durch den Einsatz besonders standardkonformer Schulbücher erhöht. Wie diese Inhalte vermittelt werden, bleibt den Lehrkräften überlassen.</p> <p>Fortbildungsdauer: Wenige Stunden; ohne Verbindung zum Unterricht</p>	8

WAS FAND DIESE STUDIE HERAUS? Insgesamt zeigt sich über alle Programme hinweg ein kleiner positiver Effekt der *Science Programs* gegenüber regulärem Unterricht ($ES = 0.17$). Die verschiedenen Programmtypen unterscheiden sich deutlich und signifikant voneinander. Gleichzeitig sind die Effekte der Studien innerhalb eines bestimmten Programmtyps sehr unterschiedlich. *Tabelle 2* zeigt sowohl die Mittelwerte als auch die Minimal und Maximal-Werte einzelner Programme innerhalb einer Kategorie. Auf dieser Grundlage ist davon auszugehen, dass sich die einzelnen Programme in mehr Merkmalen unterscheiden als durch die Kategorisierung repräsentiert werden kann.

Tabelle 2. Effekte der Typen von Programmen mit Mittelwert, Minimum und Maximum.

PROGRAMMTYP	Ø ES	MIN	MAX
Lehrstrategientraining	0.17	0.07	0.46
Digitale Lehrmaterialien	0.47	0.24	0.71
Experimentiersets	0.02	-0.06	0.05
Innovative Schulbücher	0.10	-0.02	0.39

Die größten durchschnittlichen Effekte zeigen Programme mit *Digitalen Lehrmaterialien* ($ES = 0.47$) gefolgt von *Lehrstrategientrainings* ($ES = 0.17$). *Digitale Lehrmaterialien* werden im Rahmen der Programme in erster Linie ergänzend als Tools zu Demonstrationen oder Simulationen von naturwissenschaftlichen Konzepten eingesetzt. Die *Lehrstrategientrainings* konzentrieren sich auf spezifische Lehrstrategien zur Unterstützung und Begleitung schülerzentrierter Unterrichtsformate. Beide Programmtypen fokussieren – aus Sicht der AutorInnen – auf die Unterstützung der SchülerInnen, während diese eigene Erfahrungen beim Entdecken und Forschen sammeln. Keinen nennenswerten Effekt zeigen dagegen Programme mit *Experimentiersets* ($ES = 0.05$) und *Innovativen Schulbüchern* ($ES = 0.10$). Die AutorInnen folgern aus diesem Befund, dass Programme, die in erster Linie auf die Bereitstellung von neuen Materialien setzen und diese in den Fortbildungen lediglich erklären und vorstellen, keine substanziellen Effekte erreichen. Nur diejenigen Programme, die zusätzlich das professionelle Training der Lehrkräfte zu Lehrstrategien und Unterrichtsgestaltung berücksichtigen, erreichen positive Ergebnisse.

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 gemittelte Effektstärke aller Programme liegt nach der üblichen Einteilung nach Cohen (1988) im kleinen Bereich (Gesamteffekt: $ES = 0.17$). Dieser kleine Gesamteffekt ist dennoch beachtlich, wenn man das strenge methodische Vorgehen sowohl in Bezug auf die Auswahl der Studien als auch bei der Berechnung der Effektstärken berücksichtigt. In einer methodisch vergleichbaren Metaanalyse zu den Effekten digitaler Anwendungen im Mathematikunterricht zeigt sich ein ähnlicher Effekt (siehe Kurzreview 7). Dies spricht für die Belastbarkeit der vorliegenden Befunde. Für einzelne Programmtypen ergeben sich auch größere Effekte: Der Effekt ($ES = 0.47$) von *Digitalen Lehrmaterialien* bedeutet beispielsweise, dass 68 % der SchülerInnen, die an einem entsprechenden Programm teilnahmen, bessere Leistungen erzielten als der Durchschnitt der SchülerInnen im regulären Unterricht.

Wie differenziert sind die Ergebnisse dargestellt? Die Differenziertheit der berichteten Effekte wird anhand der Bereiche Schulfächer, Jahrgangsstufen und des Erfolgskriteriums eingeschätzt. In der vorliegenden Metaanalyse sind die Ergebnisse nach Jahrgangsstufen (6.-8. Klasse und 9.-12. Klasse) differenziert dargestellt. Entsprechend lassen sich für diese beiden Altersgruppen spezifische Aussagen treffen. Für einzelne Schulfächer und nicht leistungsbezogene Lernerfolgsmaße – wie Motivation – werden keine gesonderten Effektstärken berichtet. (siehe dazu Kurzreview 12).

Wie verallgemeinerbar sind die Befunde? In der vorliegenden Metaanalyse wurden zahlreiche Moderationsanalysen durchgeführt. Dabei zeigt sich, dass die Befunde der Metaanalyse in sich nicht homogen sind. Das bedeutet, dass der Gesamteffekt nicht verallgemeinerbar ist, da er durch den Programmtyp beeinflusst wird. Um Aussagen über einen bestimmten Programmtyp zu treffen, ist es zuverlässiger, die spezifische Effektstärke des Programmtyps heranzuziehen. In Bezug auf die beiden unterschiedenen Altersgruppen (6.-8. Klasse und 9.-12. Klasse) zeigen sich keine signifikanten Unterschiede; die Ergebnisse sind demnach über beide Altersgruppen hinweg generalisierbar und bestätigen das sehr ähnliche Befundmuster für den Bereich der Grundschulen (Slavin et al. 2014). Einzelne Schulfächer, verschiedene Schulsysteme oder Länder – in der vorliegenden Metaanalyse stammen alle Primärstudien aus den USA – wurden nicht statistisch untersucht; entsprechend kann zur Verallgemeinerarbeit in dieser Hinsicht keine Aussage getroffen werden. In methodischer Hinsicht bestätigen sich drei Einflussfaktoren, die auch aus anderen Studien bekannt sind: Art der Testung: Studien mit standardisierten Tests weisen geringere Effekte auf als Studien mit eigens dafür erstellten Tests; Art der Publikation: Publierte Studien enthalten größere Effekte als nicht publizierte Studien; Art des Studiendesigns: Studien mit quasi-experimentellem Design weisen größere Effekte auf als Studien mit randomisiertem Untersuchungsdesign.

Was macht die Metaanalyse wissenschaftlich relevant? Die Metaanalyse ist für die Wissenschaft als bedeutsam zu werten. Sie untersucht erstmalig für den Bereich der Sekundarstufe, wie sich *Science Programs* auf Leistungen in MINT-Fächern auswirken. Damit erweitert sie die Ergebnisse früherer Befunde aus dem Primarschulbereich.

Wie methodisch verlässlich sind die Befunde? Die Transparenz und Begründung des methodischen Vorgehens entspricht weitgehend den Kriterien gängiger Anforderungskataloge (z.B. APA Metaanalysis Reporting Standards). Die teilweise eingeschränkte Transparenz ist der verkürzten Darstellung in der Metaanalyse geschuldet. Vertreter der Autorengruppe haben an diversen anderen Stellen (u.a. Slavin et al. 2014) jedoch bereits umfangreich ihr methodisches Vorgehen dokumentiert und damit international anerkannte Standards für metaanalytische Synthesen im Bildungsbereich geschaffen. Auf diese Arbeiten wird in der vorliegenden Studie lediglich verwiesen. Daraus erklärt sich das mittelmäßige Abschneiden im Qualitätsrating des *Clearing House Unterricht* (siehe Zusatzmaterial).

FAZIT FÜR DIE UNTERRICHTSPRAXIS. Die Metaanalyse beruht ausschließlich auf Befunden zur Lernwirksamkeit von *Science Programs*, die über längere Zeiträume im Schulunterricht umgesetzt wurden und die prinzipiell an jeder Schule eingesetzt werden können. Daher haben die Ergebnisse grundsätzlich eine hohe Relevanz für die Weiterentwicklung von naturwissenschaftlichem Unterricht. Allerdings wurden alle berücksichtigten Programme in den USA entwickelt und getestet. Die Befunde haben für den deutschen Kontext in erster Linie Hinweischarakter, da mit der vorliegenden Metaanalyse nicht geklärt wurde, inwiefern die Lernwirksamkeit der Programme auch außerhalb der USA abgesichert ist.

Nimmt man – trotz dieser Einschränkung – die Hinweise aus dieser Metaanalyse ernst, so ist festzustellen, dass sich der Aufwand, der mit der Implementation dieser umfassenden Programme im Unterricht verbunden ist, positiv auf die Schülerleistungen auswirkt. Als besonders effektiv haben sich Programme erwiesen, in denen Lehrkräfte geschult werden, digitale Lehrmaterialien als zusätzliche Erklärungen, Veranschaulichungen und Illustrationen zum tiefergehenden Verständnis von naturwissenschaftlichen Konzepten in ihrem Unterricht einzusetzen. Außerdem scheinen sich solche Programme zu lohnen, bei denen die Lehrkräfte den Einsatz von bestimmten Lehrstrategien ausgiebig trainieren, um sie dann über unterschiedliche Themen hinweg zur Unterstützung ihrer SchülerInnen in schülerzentrierten bzw. offenen Lernsettings einsetzen zu können (siehe Studienbeispiel).

STUDIENBEISPIEL

Harris und Kollegen (2015) untersuchen in ihrer Studie einen umfassenden Modellversuch zum »project-based-inquiry science« (PBIS) Curriculum. Dieser wurde über ein komplettes Schuljahr durchgeführt. Alle 57 beteiligten MINT-Lehrkräfte der sechsten Klassen von 42 Schulen erhielten zu Beginn des Schuljahres eine zweitägige Schulung zu den Kernprinzipien des neuen Lehrplans. Daraufhin wurde in der Kontrollgruppe mit den üblichen Schulbüchern regulär unterrichtet und die Lehrkräfte erhielten keine weitere Fortbildung. Die Lehrkräfte in der Experimentalgruppe dagegen konnten auf die speziell nach dem neuen Lehrplan entwickelten Lern-/Unterrichtsmaterialien zurückgreifen. Zudem erhielten sie zu Beginn des Schuljahres eine grundlegende dreitägige Fortbildung zum Umgang mit den Lern-/Unterrichtsmaterialien, die auf das aktive Experimentieren von SchülerInnen und das Sammeln von Erfahrungen ausgerichtet waren. Auch in der Begleitung und Moderation von schülerzentriertem Unterricht wurden sie geschult. Im Verlauf des Schuljahres folgten zwei weitere eintägige Schulungen zu zwei Modulen mit spezifischen Unterrichtsinhalten. Am Ende des Schuljahres absolvierten alle SchülerInnen den selben standardisierten Test. Das PBIS erbrachte einen positiven Effekt von 0.24. Insbesondere benachteiligte SchülerInnengruppen profitierten stark von PBIS.

REFERENZEN.

Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Auflage). Hillsdale, NJ: Erlbaum.

Slavin, R. E., Lake, C., Hanley, P., and Thurston, L. (2014). Experimental evaluations of elementary science programs: A best-evidence synthesis. *Journal of Research in Science Teaching*, 51: 870–901. doi:10.1002/tea.

Cheung, A., Slavin, R. E., Kim, E., and Lake, C. (2017), Effective secondary science programs: A best-evidence synthesis. *Journal of Research in Science Teaching*, 54: 58–81. doi:10.1002/tea.21338.

Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., ... & Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching*, 52(10), 1362-1385. doi: 10.1002/tea.21263.

LINKS.

Zur Metaanalyse von [Cheung et al., 2017](#)

Zum Studienbeispiel von [Harris et al., 2015](#)

ZITIEREN ALS.

Hetmanek, A., Knogler, M. & CHU Research Group (2018). Wie effektiv sind *Science Programs*? 21 Varianten im Praxistest. www.clearinghouse-unterricht.de, Kurzreview 13.



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