

Tobias Böttger & Klaus Zierer

Effekte der pandemiebedingten Schulschließungen im Frühjahr 2020 auf fachlich-kognitive Leistungen von Schüler*innen im In- und Ausland

Ein narratives Review

Zusammenfassung

*Die weltweiten Einschränkungen aufgrund der Corona-Pandemie haben auch das Bildungssystem beeinflusst. Vielerorts fand Unterricht im Frühjahr 2020 aus der Distanz statt. Mittlerweile sind großangelegte Studien zu den Auswirkungen der Schulschließungen auf Schüler*innen erschienen. Deren Befunde sind in Bezug auf Schulstufe und Fachbereich heterogen. Im vorliegenden Review werden neun Studien näher beschrieben, um das Zustandekommen reflektierter Urteile über deren Ergebnisse zu ermöglichen.*

Schlüsselwörter: Schulschließungen, fachlich-kognitive Leistung, Pandemie-Effekte, Lernleistung

Effects of the Pandemic-related School Closures in Spring 2020 on the Cognitive Performance of Students at Home and Abroad

A Narrative Review

Abstract

The global restrictions due to the corona pandemic have also affected the education system. In many places, lessons were held from a distance in spring 2020. In the meantime, large-scale studies on the effects of school closings on schoolchildren have been published. Their findings vary in terms of school level and subject area. In the present review, nine studies are described in more detail in order to enable reflective judgements about their results.

Keywords: school closures, cognitive performance, effects of the pandemic, achievement

1 Einleitung

Zu den Auswirkungen der pandemiebedingten Schulschließungen im Jahr 2020 liegen bereits einige Studien vor. Die deutliche Mehrzahl fokussiert dabei auf fachlich-kognitive Leistungen der Schüler*innen. Da zu diesem Thema derzeit die meisten Befunde vorliegen, beschränkt sich auch das Erkenntnisinteresse dieses narrativen Reviews auf die fachlich-kognitiven Leistungen. Die Rahmenbedingungen, unter denen die Schulschließungen vorgenommen wurden, unterschieden sich zwischen den einzelnen Ländern stark. In Deutschland zum Beispiel schienen Lehrkräfte nicht optimal auf die verstärkte Nutzung digitaler Medien im Unterricht vorbereitet gewesen zu sein. Schon lange vor der Covid-19-Pandemie wurden Fortbildungen zu diesem Thema wenig besucht und im Rahmen der *International Computer and Information Literacy Study* (ICILS) 2013 wurde gezeigt, dass nur ein Drittel der Lehrer*innen regelmäßig auf digitale Medien zurückgriffen (Drossel & Eickelmann, 2018, S. 171 ff.). Im Rahmen der Studie ICILS 2018 konnte diesbezüglich zwar eine deutliche Steigerung festgestellt werden, im internationalen Vergleich erschien die Häufigkeit des Einsatzes digitaler Medien jedoch noch immer rückständig (Eickelmann et al., 2019). Es ist davon auszugehen, dass der Stand der Digitalisierung im Bildungsbereich in den Untersuchungsländern sehr unterschiedlich ist. Das wiederum kann einen Einfluss auf das Distanzlernen und somit die Studienergebnisse haben. Selbst in den Niederlanden, wo die Bedingungen günstig erschienen, deuteten Umfrageergebnisse jedoch auf ein hohes Maß an Unzufriedenheit mit dem Fernunterricht (De Haas, Faber & Hamersma, 2020) und auf erhebliche Unterschiede bei der Hilfeleistung bei Schulaufgaben und Lernressourcen (Bol, 2020) in der Zeit der Schulschließungen hin. Es ist folglich zu vermuten, dass die Effektivität des Distanzunterrichtes in Hinblick auf dessen Auswirkung auf die fachlich-kognitive Leistung von Schüler*innen in der Regel nicht an die Effektivität von Unterricht in Präsenz heranreichte.

2 Theoretischer Rahmen

Studien, in welchen die Effekte von Unterrichtszeit auf fachlich-kognitive Leistung untersucht werden, können aufschlussreich sein, um die Auswirkungen von Schulschließungen im Zusammenhang mit Covid-19 abzuschätzen. Carlsson, Dahl, Öckert und Rooth (2015) fanden, dass zehn Tage der Beschulung zu einer Zunahme der Punktzahl bei einem Intelligenztest von $d = 0.01$ führten. Lavy (2015) führte Leistungsunterschiede, die im Rahmen von PISA-Erhebungen gefunden wurden, auf Unterschiede bei der Unterrichtszeit in den untersuchten Ländern zurück. Bei nur einer Stunde Unterricht mehr pro Woche stieg die im Test erreichte Punktzahl um $d = 0.06$ (Tomasik, Helbling & Moser, 2020, S. 2).

Weitere Forschungsergebnisse weisen darauf hin, dass ein linearer Zusammenhang zwischen der Anzahl versäumter Schultage und den Prüfungsergebnissen am Ende eines Schuljahres besteht, wobei die Varianz der berichteten Effektstärken groß ist¹ (Tomasik et al., 2020, S. 3). Außerdem lassen Studien zum Thema Lernverluste aufgrund von Sommerferien, Lehrer*innen-Streiks, Reformen oder Naturkatastrophen (Bao, Qu, Zhang & Hogan, 2020; Eyles, Gibbons & Montebruno, 2020; Frenette, Frank & Deng, 2020; Kuhfeld, Ruzek, Johnson, Tarasawa & Lewis, 2020), darauf schließen, dass die pandemiebedingten Schulschließungen zu einem Verlust an fachlich-kognitiver Leistung geführt haben (Maldonado & De Witte, 2020, S. 1). Hansen (2011) fand beispielsweise heraus, dass jeder Tag schneebedingter Schulschließungen in Colorado zu einer Reduktion der Leistung um zwischen $0.013 \leq d \leq 0.039$ führte. Die Folgen des Hurrikans Katrina führten Sacerdote (2012) zufolge zu einem Lernverlust von insgesamt $d = 0.100$ (Tomasik et al., 2020, S. 3). Sämtliche der dargestellten Befunde zum Thema Schulschließungen bzw. Ausfall von Unterricht deuten auf negative Auswirkungen auf die fachlich-kognitive Leistung von Schüler*innen hin. Dennoch ist zu berücksichtigen, dass während der Covid-19-Pandemiebedingten Schulschließungen alternative Unterrichtsformen zum Einsatz kamen. Die zuvor präsentierten Forschungsergebnisse sind somit nur bedingt aussagekräftig, wenn es um die Vorhersage der Effekte der pandemiebedingten Schulschließungen auf fachlich-kognitive Leistung geht.

3 Methodisches Vorgehen

Zwar existieren zu den Effekten der pandemiebedingten Schulschließungen bereits Übersichtsarbeiten (z. B. Hammerstein, König, Dreisörner & Frey, 2021; Helm, Huber & Loisinger, 2021; Zierer, 2021), die sich vom methodischen Zugang her deutlich unterscheiden und von einer knappen Übersicht bis hin zu einem Rapid Review reichen. Allerdings liefern sie größtenteils nur stark verdichtete Informationen, so dass bis heute eine detaillierte Betrachtung der Forschungsergebnisse zu den Folgen der pandemiebedingten Schulschließungen fehlt. Aus diesem Grund wird im vorliegenden Beitrag ein narratives Review durchgeführt, mit dem ein umfassender Überblick über mögliche Effekte der Schulschließungen gegeben werden soll.² Es soll geklärt werden, ob sich anhand einer Auswahl an verfügbaren Studien die Hypothese bestätigen lässt, dass die pandemiebedingten Schulschließungen zu negativen Effekten auf die fachlich-kognitive Leistung von Schüler*innen im In- und Ausland führten. Ziel dieses narrativen Reviews ist es, umfassend über Studieninhalte zu infor-

1 Aucejo & Romano (2016), Gershenson, Jackowitz & Brannegan (2017) und Liu, Lee & Gershenson (2020) berichten Effektstärken von $.006 \leq d \leq 2.00$ (Tomasik et al., 2020, S. 3).

2 „Aufgrund der zunehmenden Zahl an wissenschaftlichen Publikationen ist es wichtig, einzelne Studien zu einem Thema übersichtlich zusammenzufassen und gemeinsam zu bewerten“ (Ressing, Blettner & Klug, 2009, S. 456). Für einen Überblick über verschiedene Arten systematischer Überichtsarbeiten vgl. Ressing, Blettner und Klug, 2009.

mieren, ohne dass die zugrundeliegenden Forschungsarbeiten durch die Leser*innen gesichtet werden müssen. Somit ist es möglich, zu einem differenzierten Urteil über Qualität und Aussagekraft der vorgestellten Publikationen zu kommen. Dies wiederum kann dabei helfen zu klären, ob Lernrückstände entstanden sind und wie hoch diese gegebenenfalls ausfallen, sodass auch zukünftige Entwicklungen im Hinblick auf mögliche pandemiebedingte Lernrückstände besser abgeschätzt werden können. Besonders vor dem Hintergrund aktueller bildungspolitischer Diskussionen zum Thema „Aufholen nach Corona“, zum Beispiel durch Sommerschulskonzepte (BMFSFJ, 2021), scheint dies zentral. Es muss geklärt werden, ob ein Lernrückstand entstanden ist und wie hoch dieser gegebenenfalls ausfällt. Zur Steigerung der Übersichtlichkeit werden die Studienergebnisse nach Fachbezug und Schulstufe gegliedert präsentiert. Neben den Ergebnissen der einzelnen Studien werden auch Charakteristika wie Stichprobengrößen und Erhebungsinstrumente benannt. Solche und weitere Details können für die Bewertung der Befunde ausschlaggebend sein. Abschließend findet eine Diskussion der Studienergebnisse statt.

Um eine möglichst umfassende Darstellung möglicher Effekte pandemiebedingter Schulschließungen gewähren zu können, sollen im Rahmen dieses Beitrages Studien aus verschiedenen Ländern aufgegriffen werden. In den Entstehungsländern sollten einigermaßen vergleichbare Bedingungen im Hinblick auf gesellschaftliche und politische Ordnung und Schulsystem herrschen. Von Interesse waren Studien aus den Jahren 2020 und 2021, die sich mit Effekten der Schulschließungen im Jahr 2020 auf die fachlich-kognitive Leistung von Schüler*innen beschäftigen. Bei der Recherche nach Studien für dieses Review wurden die Datenbanken von *ResearchGate*, *SocArXiv*, *Google Scholar* und *PsyArXiv* mit den Schlagworten *Schulschließungen* beziehungsweise *school closure*, *Effekt* beziehungsweise *effect* und *kognitiv* beziehungsweise *cognitive*, in Verbindung mit *Covid-19* durchsucht. Im Rahmen des vorliegenden Reviews wurden neun der so gefundenen Studien aufgegriffen, welche aus den USA, Belgien, den Niederlanden, der Schweiz, England und Deutschland stammen (vgl. Tab. 1 auf S. 44). Trotz der Gemeinsamkeiten dieser Länder in Bezug auf gesellschaftliche und politische Ordnung sowie das Schulsystem sind die Kontexte, in denen die genannten Forschungsarbeiten entstanden sind, in einem Maße heterogen, dass im Rahmen dieses Reviews keine Synthese der Studienergebnisse vorgenommen wird. Es geht lediglich darum, durch eine Beschreibung der einzelnen Studien ein möglichst umfassendes Bild der Befundlage zu den möglichen Effekten pandemiebedingter Schulschließungen auf die fachlich-kognitive Leistung von Schüler*innen zu zeichnen. Die Tatsache, dass die Ergebnisse der präsentierten Studien nicht zuletzt von den Unterschieden zwischen den Schulsystemen in den Untersuchungsländern beeinflusst sein können, sollte bei der Lektüre stets berücksichtigt werden.

Die meisten der im Folgenden vorgestellten Studien bemessen Lernrückstände anhand von Standardabweichungen. Zur Einordnung dieser Maßzahl sollen die folgenden Ausführungen eine Hilfestellung bieten. Azevedo, Hasan, Goldemberg, Iqbal und Geven (2020) gehen in Bezug auf ein normales Schuljahr von einem Lernleistungszuwachs von 0.40 Standardabweichungen aus (Azevedo et al., 2020, zit. nach Engzell, Frey & Verhagen, 2021, S. 10). Engzell et al. (2021) beziffern die jährliche Zuwachsrate mit für gewöhnlich 0.31 bis 0.41 Standardabweichungen (ebd., S. 10). Maldonado und De Witte (2020) zufolge konnten Bloom, Hill, Black und Lipsey (2008) zeigen, dass Schüler*innen, Stand fünfte Klasse, in der Regel einen Leistungszuwachs um 0.4 Standardabweichungen pro Schuljahr zeigen (Maldonado & De Witte, 2020, S. 28). Schult, Mahler, Fauth und Lindner (2021) gehen in ihrer Studie von einem Kompetenzzuwachs von 0.65 Standardabweichungen in einem normalen Jahr aus (Schult et al., 2021, S. 16). Kogan und Lavertu (2021a, 2021b) setzten in ihrer Studie für ein normales Schuljahr einen Zuwachs der Leseleistung von 0.60 Standardabweichungen an (Kogan & Lavertu, 2021b, S. 2). Bei der Interpretation der Studienergebnisse kann die Kategorisierung der Effektgrößen nach Cohen (1988) irreführend sein. Da es bei der Bewertung von Effekten stets auf die Vergleichsgrundlage ankommt, sollte in Relation zu den gefundenen Effekten stets der oben dargestellte, eigentlich zu erwartende Lernzuwachs berücksichtigt werden.

4 Auswirkungen auf die Mathematikleistung in der Grundschule und Sekundarstufe

Kuhfeld et al. (2020) untersuchten im Rahmen ihrer Studie mehr als 300.000 Proband*innen aus verschiedenen Schulbezirken der USA (vgl. Kuhfeld et al., 2020, S. 9). Zur Leistungserhebung wurde das *NWEA MAP Growth Assessment* in den Kompetenzbereichen Mathematik und Lesen eingesetzt (ebd., S. 2). Da die Grundschule in den USA je nach Schuldistrikt unterschiedlich viele Jahrgangsstufen umfasst, werden an dieser Stelle nur die Ergebnisse bezüglich Dritt- und Viertklässler*innen berichtet. Im Rahmen der Studie wurden anhand verschiedener Stichproben sowohl intraindividuelle als auch interindividuelle Unterschiede untersucht (ebd., S. 1 ff.).

Über alle getesteten ethnischen Gruppen hinweg, ließen sich im Herbst 2020 niedrigere Werte feststellen als bei einer gleichaltrigen Gruppe von Schüler*innen im Herbst 2019. In der dritten Jahrgangsstufe ging die Mathematikleistung auf intraindividuelle Ebene um 9 Prozentpunkte zurück, in der vierten Jahrgangsstufe betrug der Rückgang 10 Prozentpunkte (vgl. Kuhfeld et al., 2020, S. 13)³. Beim Vergleich der erbrachten Leistungen auf interindividueller Ebene kamen die Autor*innen zu ähnlichen Ergebnissen (vgl. Schult et al., 2021, S. 5).

3 Die Autor*innen gaben keine Effektstärken an.

Tab. 1: Überblick über die im Review thematisierten Studien

Autoren*innen und Jahr	Titel ^a	Land	Messinstrument	Schulform	Fachbereich	N	Darstellungsform der Ergebnisse
Meeter, 2021	Primary School Mathematics During Covid-19	Niederlande	Snappet	Grundschule	Mathematik	25.336 SuS	Effektstärken grafisch dargestellt
Rose et al., 2021	Impact of School Closures and Subsequent Support Strategies	England	NFER-Assessment Tests	Grundschule	Mathematik	5.936 SuS	Effektstärken
Kogan & Lavertu, 2021	Covid and Student Achievement	USA (Ohio)	Third-Grade English Language Arts (ELA)	Grundschule	Lesen (Englisch)	5.931 SuS	Effektstärken
Kuhfeld et al., 2020	Technical Appendix for: Learning During COVID-19	USA	NWEA MAP Growth	Grundschule	Mathematik	> 325.000 SuS	Prozentual
Maldonado & De Witte, 2020	The Effect of School Closures	Belgien	Standardisierte Schulleistungstests	Grundschule	Mathematik	1.287 Schulen	Effektstärke
					Holländisch	1.480 Schulen	
					Französisch	1.325 Schulen	
Engzell et al., 2021	Learning Loss due to School Closures	Niederlande	Halbjährliche Tests in den Kernfächern	Grundschule	Mathematik	286.515 SuS	Mittelwerte und Signifikanz der Unterschiede
					Lesen	217.875 SuS	
					Rechtschreibung	284.499 SuS	

Autoren*innen und Jahr	Titel ^a	Land	Messinstrument	Schulform	Fachbereich	N	Darstellungsform der Ergebnisse
Depping et al., 2021	Kompetenzstände Hamburger Schüler*innen vor und während der Corona-Pandemie	Deutschland	LERNSTAND und KERMIT	Grundschule	Mathematik	13.238 SuS (2019) 13.165 SuS (2020)	Prozentual
					Deutsch (Leseverstehen)	13.238 SuS	
Tomasik et al., 2020	Educational Gains of In-person vs. Distance Learning in Primary and Secondary Schools	Schweiz	MINDSTEPS	Sekundarstufe	Mathematik	13.703 SuS (2019) 13.938 SuS (2020)	
					Deutsch (Leseverstehen)	13.716 SuS (2019) 13.936 SuS (2020)	
Schult et al., 2021	Did Students Learn Less During the COVID-19 Pandemic?	Deutschland	LERNSTAND	Primarstufe	Mathematik, Lesen und Grammatik	13.134 SuS	Veränderung von Chi-Quadrat, Abweichung der Slopes
				Sekundarstufe	Mathematik kombiniert	13.134 SuS	
				Sekundarstufe	Mathematik	250.109 SuS (2017-19) 81.550 SuS (2020)	Effektstärke, Mittelwerte
					Deutsch (Leseverstehen)	249.862 SuS (2017-19) 81.810 SuS (2020)	

Anm.: SuS = Schüler*innen, N = Stichprobengröße, ^aDie Studientitel werden hier nur verkürzt angegeben, vollständiger Titel siehe Literaturverzeichnis.

Quelle: eigene Darstellung

Maldonado und De Witte (2020) untersuchten Effekte siebenwöchiger Schulschließungen auf Grundschüler*innen flämischer Schulen in Belgien und stellten einen Rückgang der in Mathematik erreichten Testwerte um 0.19 Standardabweichungen fest. Die Heterogenität der Leistungen stieg sowohl innerhalb der untersuchten Schulen als auch zwischen den Institutionen (ebd., S. 1). Die Studiendaten wurden anhand von standardisierten Tests in der letzten Jahrgangsstufe der Grundschule gewonnen (vgl. Maldonado & De Witte, 2020, S. 6). Relevant ist hier der Vergleich zwischen den Jahren 2019 und 2020, zumal in beiden Jahren derselbe Test zum Einsatz kam (ebd., S. 10). Beim Vergleich der zu besagten Messzeitpunkten ermittelten Mathematiktestwerte zeigte sich, ohne die Berücksichtigung von Kontrollvariablen, ein signifikanter negativer Effekt von -0.19 ($p < 0.05$) Standardabweichungen bei einer Stichprobengröße von 1.287 Schulen (vgl. Maldonado & De Witte, 2020, S. 12). Der gefundene Effekt war robust gegenüber der Einführung von Kontrollvariablen in das Modell. Auch bei einer Beschränkung der Stichprobe auf Schulen, die in jedem Jahr der Testung, d.h. von 2015 bis 2020, teilgenommen hatten, blieb das Ergebnis ähnlich (vgl. Maldonado & De Witte, 2020, S. 24). Der Anteil der am Test teilnehmenden Schulen lag im Jahr 2020 deutlich unter der Beteiligungsrate in den Jahren zuvor (ebd., S. 32 ff.). Dieser Umstand könnte einen Einfluss auf die Validität der hier präsentierten Ergebnisse haben (vgl. Schult et al., 2020, S. 6).

Engzell et al. (2021) beschäftigten sich im Rahmen ihrer Studie mit den Effekten achtwöchiger Schulschließungen in den Niederlanden, wobei ihnen eine Stichprobe von etwa 350.000 Grundschüler*innen zur Verfügung stand. Datengrundlage waren Ergebnisse halbjährlich stattfindender Tests in den Kernfächern (vgl. Engzell et al., S. 1). In den Jahrgangsstufen 4 bis 7 wurden vom *Dutch National Institute for Educational Measurement* entwickelte standardisierte Mathematiktests durchgeführt (ebd., S. 5, S. 50). Bei einer Stichprobengröße von 286.515 erreichten die Schüler*innen, die nach den Schulschließungen getestet wurden, signifikant schlechtere Testwerte als die Vergleichsgruppe, die keine Schulschließungen erlebt hatte (vgl. Engzell et al., 2021, S. 85). Die Effektstärke betrug -0.14 Standardabweichungen, bei einem 95%-Konfidenzintervall von -0.15 bis -0.13 (eigene Berechnung nach Lenhard & Lenhard, 2016).⁴ Die Autor*innen berichteten, dass der gefundene Effekt in etwa einem Zeitraum von 7.9 bis 10.5 Wochen des Nicht-Lernens entsprach (vgl. Engzell et al., 2021, S. 10).

4 Die Berechnung der Effektstärke erfolgte unter Verwendung des Effektstärkerechners nach Lenhard & Lenhard, 2016, dem die folgende Formel zugrunde liegt: $d = \frac{\bar{x}_A - \bar{x}_B}{s_{AB}}$ (\bar{x}_A = Mittelwert der Gruppe A, \bar{x}_B = Mittelwert der Gruppe B, s_{AB} = gemeinsame Streuung) wobei $s_{AB} = \sqrt{\frac{s_A^2 n_A + s_B^2 n_B}{n_A + n_B}}$ (s_{AB} = Standardabweichung Gruppe A, s_B = Standardabweichung der Gruppe B). Bei der Berechnung der Effektstärke wird die unterschiedliche Gruppengröße bei der Berechnung der gepoolten Standardabweichung berücksichtigt.

Meeter (2021) berechnete Effekte von Schulschließungen auf Grundlage von Daten aus den Niederlanden. Diese wurden während der Zeit der Schulschließungen mithilfe der adaptiven Übungssoftware *Snappet* erhoben. Die Stichprobe bestand aus 53.656 Grundschüler*innen der Jahrgangsstufen 2 bis 6 und umfasste insgesamt 810 Schulen. In die Analysen einbezogen wurden nur solche Schulen, welche die Lernsoftware *Snappet* verwendeten. Aus dem Schuljahr 2018/19 standen die Daten von 48.815 Schüler*innen zur Verfügung, für das Schuljahr 2019/20 waren es 53.656 Fälle (vgl. Meeter, 2021, S. 4). Es wurde zwischen Effekten auf das *Lernen*, die *Leistung* und das *Üben* unterschieden. Außerdem wurden die Schüler*innen, ihrer Leistung in der ersten Hälfte des Jahres entsprechend, in drei gleichgroße Gruppen aufgeteilt (ebd., S. 5). Alle berichteten Effekte der Schulschließungen fielen überraschenderweise positiv aus und lagen im Bereich zwischen rund 0.05 und rund 0.85 Standardabweichungen (vgl. Meeter, 2021, S. 7). Genauere Angaben sind nicht möglich, da der Autor Effektstärken und Konfidenzintervalle lediglich grafisch präsentierte.

Die größten positiven Effekte wurden im Bereich *Lernen* in den niedrigeren Jahrgangsstufen verzeichnet. Genau in diesen Jahrgangsstufen zeigten sich im selben Bereich nach Wiedereröffnung der Schulen die größten negativen Effekte (Meeter, 2021, S. 6). In der Zeit geschlossener Schulen wurden auf der Dimension *Leistung* in allen Schüler*innengruppen positive Effekte zwischen rund 0.045 und rund 0.18 Standardabweichungen gefunden. Nach der Wiedereröffnung fielen diese kleiner, jedoch immer noch positiv aus (ebd.). Im Bereich *Üben* ließen sich während der Schulschließungen ebenfalls größere Effekte nachweisen als vor Schließung der Schulen. Nach Wiedereröffnung der Institute fielen diese Effekte sogar noch größer aus (vgl. Meeter, 2021, S. 7). Obwohl sich einige der gefundenen positiven Effekte nach der Wiedereröffnung der Schulen egalisierten, ließ sich über alle erfassten Domänen hinweg, bei einer Stichprobe mit 25.336 Fällen in der 4. Klasse, ein Zuwachs von 0.2 Standardabweichungen für das Schuljahr 2019/20 verzeichnen (vgl. Schult et al., 2021, S. 6). Diese Befunde lassen den Schluss zu, dass online-basierte formative Leistungserhebungen dabei helfen können, Schüler*innen im Distanzunterricht zu aktivieren. Allerdings bleibt die Kausalität unklar und die Intervention war gleichzeitig Instrument zur Datenerhebung, was Methodeneffekte begünstigt (ebd., S. 20).

Depping, Lücken, Musekamp und Thonke (2021) stellten in ihrer Studie Kompetenzstände von vergleichbaren Hamburger Schüler*innen-Kohorten in den Bereichen Deutsch-Leseverstehen und Mathematik zu Zeitpunkten vor und während der Schulschließungen gegenüber. Zur Datenerhebung kamen die Instrumente *LERNSTAND* und *KERMIT* in den Jahrgangsstufen 3 und 5 im Jahr 2019 und 4 und 5 im Jahr 2020 zum Einsatz (vgl. Depping et al., 2021, S. 51). Zunächst sind die Ergebnisse aus dem Grundschulbereich von Interesse. Die Schüler*innen wurden, je nach Sozialindex der Schule, in drei Belastungsgruppen eingeteilt (vgl. Depping et al., 2021, S. 59). Beim Vergleich der *KERMIT* 3-Gruppe mit 13.238 Fällen

mit der *LERNSTAND* 4-Gruppe mit 13.165 Fällen zeigte sich ein Effekt von -0.06 Standardabweichungen in Belastungsgruppe zwei (ebd., S. 62). Die Angabe von p-Wert und Konfidenzintervall für diesen Wert fehlte. Ein Vergleich der prozentualen Verteilung auf die Kompetenzstufen KSI bis KSIII+ zeigte, dass die Kohorten 2019 und 2020 in Mathematik in etwa gleichauf lagen (ebd., S. 63). Die Unterschiede ließen sich „als praktisch nicht bedeutsam [...] interpretieren und lassen sich [...] auf normale Leistungsschwankungen zwischen Kohorten verschiedener Jahrgänge zurückführen“ (ebd.). Die Autor*innen wiesen auf den Umstand hin, dass bei der Erhebung 2020 „deutlich mehr Schüler*innen [...] zu wenige Aufgaben für die Berechnung eines Kompetenzwertes bearbeitet haben“ (Depping et al., 2021, S. 74), als es 2019 der Fall war (ebd., S. 79). Dies könnte zu einer Verzerrung der Ergebnisse der Studie geführt haben (ebd., S. 74).

Depping et al. (2020) untersuchten auch die Kompetenzstände von Schüler*innen der Sekundarstufe. Im Jahr 2019 wurden 13.703 Schüler*innen an Hamburger Gymnasien und Stadtteilschulen mit *KERMIT* 5 und im Jahr 2020 13.938 Schüler*innen mit *LERNSTAND* 5 getestet (vgl. Depping et al., 2021, S. 65 ff.). Ein Vergleich der Verteilung der Schüler*innen auf die drei Kompetenzbereiche KSI, KSII und KSIII+ ließ keine größeren Abweichungen erkennen (ebd., S. 71 ff.). Dieser Befund wurde durch die Tatsache untermauert, dass auch ein Vergleich der im Mittel erreichten Testwerte keinen statistisch bedeutsamen Unterschied hervorbrachte (ebd., S. 70).⁵

Schult et al. (2021) werteten die Ergebnisse des Kompetenztests *LERNSTAND* 5 in den Bereichen Lesen und Mathematik aus. Dafür wurde eine Stichprobe bestehend aus über 80.000 baden-württembergischen Schüler*innen der 5. Jahrgangsstufe herangezogen. Grundschulen waren dort für acht Wochen geschlossen (vgl. Schult et al. 2021, S. 2 f.). Aus den Jahren 2017 bis 2019 lagen 250.109 Testwerte vor, aus dem Jahr 2020, nach den Schulschließungen, waren es 81.550 (ebd., S. 4, S. 29). Im Kompetenzbereich Mathematik wurde zwischen *Operationen* und *Zahlen* unterschieden. Beim Vergleich der gemittelten Testwerte aus den Jahren 2017 bis 2019 mit denen aus dem Jahr 2020 zeigte sich im Bereich *Operationen* ein negativer Effekt von -0.09 Standardabweichungen. Im Bereich *Zahlen* lag ebenfalls ein negativer Effekt von -0.03 Standardabweichungen vor (vgl. Schult et al., 2021, S. 29). Die Autor*innen kamen zu dem Schluss, dass – ausgehend von einem Kompetenzzuwachs von 0.65 Standardabweichungen in einem normalen Schuljahr – die Schüler*innen aufgrund der Schulschließungen bezüglich ihrer Kompetenzentwicklung um einen Monat in Rückstand gerieten (ebd., S. 16). Es wurde außerdem festgehalten, dass die Streuung der Kompetenzniveaus innerhalb der einzelnen Schulen die Lehrkräfte vor größere Herausforderungen stellen könnte als eine leichte pandemiebedingte Abschwächung des mittleren Kompetenzniveaus an der Schule (vgl. Schult et al., 2021, S. 18).

5 Die Autor*innen gaben keine konkrete Effektstärke an.

Rose, Twist, Lord, Rutt, Badr, Hope und Styles (2021) befassten sich mit der Frage, inwiefern sich die Schulschließungen auf die Leistung von britischen Schüler*innen in den Bereichen Lesen und Mathematik auswirkten. Im Fokus standen dabei Schüler*innen im Alter von sechs bis sieben Jahren, die im 2. Schuljahr getestet wurden. Insgesamt umfasste die Untersuchung 6.000 Schüler*innen aus 168 Schulen (vgl. Rose et al., 2021, S. 1). Als Vergleichsgrundlage wurde eine Stichprobe aus 2017 herangezogen, die zur Standardisierung des verwendeten Tests diente. In England waren die Schulen für Schüler*innen der zweiten Jahrgangsstufe vom 20. März 2020 bis September 2020 geschlossen (ebd., S. 2). An der Erhebung im November 2020 konnten nur solche Schulen teilnehmen, die als Kunden des *NFER-Assessment-Programms* gelistet waren, von dem auch die besagten Tests stammten (ebd., S. 3).

Der Mathematiktest bestand aus den Teilen *Arithmetik* und *logisches Denken* und wurde von 5.936 Schüler*innen aus 168 Schulen bearbeitet (vgl. Rose et al., 2021, S. 3). Die so gewonnen Testwerte wurden zunächst standardisiert und gewichtet, wobei die zur Gewichtung verwendete Leistungs-Variable eigentlich für eine andere Jahrgangsstufe gedacht war und sich nicht auf Schüler*innen-Ebene befand (ebd., S. 4). Die Mathematikleistung, welche die Schüler*innen im November 2020 zeigten, lag deutlich unter der der Stichprobe aus 2017. Der standardisierte Durchschnittswert betrug in der Herbststichprobe 2020 98,06, woraus sich eine Effektstärke von -0,14 beziehungsweise ein Fortschrittsverlust von etwa 2 Monaten errechnen ließ (vgl. Rose et al., 2021, S. 7). Ein Blick auf die Verteilung der standardisierten Testwerte offenbarte, dass 2020 im Vergleich zu 2017 mehr niedrigere und weniger höhere Testwerte erzielt wurden (ebd., S. 9).

5 Auswirkungen auf Schulleistung im Bereich Sprache in Grundschule und Sekundarstufe

Kuhfeld et al. (2020) beschäftigten sich auch mit möglichen Veränderungen der Lesefähigkeit von amerikanischen Schüler*innen. Analog zur Mathematikleistung wurde die Lesefähigkeit mit dem *NWEA MAP Growth Assessment* erhoben (vgl. Kuhfeld et al. 2020, S. 2). Ein Perzentil-Werte-Vergleich zeigte, dass Schüler*innen, die im Herbst 2019 getestet wurden, im Bereich Lesen ähnlich hohe Testwerte erreichten wie Schüler*innen im Herbst 2020 (ebd., S. 5). Die prozentualen Anteile von Schüler*innen, die in Bezug auf Lesen eine Leistungssteigerung, keine Leistungsveränderung oder einen Abfall der Leistung zeigten, waren in den Jahren 2019 und 2020 in etwa gleich (ebd., S. 7). Die Steigerung der Leseleistung im Zeitraum vom Winter 2020 bis zum Herbst 2020 entsprach in etwa dem Zuwachs, der in einem typischen Jahr zu erwarten ist (Kuhfeld et al., 2020, S. 6)⁶.

6 Die Autor*innen gaben keine Effektstärken an.

Kogan und Lavertu (2021a, b) analysierten die Ergebnisse des *Third-Grade English Language Arts (ELA)* Tests, welcher im US-Bundesstaat Ohio mehrmals im Jahr durchgeführt wird (Kogan & Lavertu, 2021a, S. 2). Während für das Jahr 2019 die Testwerte von 124.816 Schüler*innen der dritten Jahrgangsstufe zur Verfügung standen, lieferte der Herbst-Test 2020 124.710 Beobachtungen (Kogan & Lavertu, 2021b, S. 5). Die von Drittklässler*innen erbrachte Sprachleistung ging im Zeitraum von Herbst 2019 bis Herbst 2020 signifikant um etwa 0.23 Standardabweichungen zurück (vgl. Kogan & Lavertu, 2021a, S. 2). Dieser Wert wurde adjustiert, indem fehlende Werte auf Grundlage von Daten aus der zweiten Jahrgangsstufe geschätzt wurden. Der nicht adjustierte Effekt lag bei -0.19 Standardabweichungen (Kogan & Lavertu, 2021b, S. 5). Der Rückgang der Leistung entsprach in etwa einem Drittel des Lernzuwachses, der normalerweise in einem Jahr zu verzeichnen ist (vgl. Kogan & Lavertu, 2021a, S. 2). Die Autor*innen wiesen darauf hin, dass die Schulschließungen nicht der einzige Grund für den Leistungsrückgang beim Sprachtest waren. Negative Effekte waren auch für Schüler*innen nachweisbar, die im Untersuchungszeitraum Hybrid- oder Präsenzunterricht erhielten (vgl. Kogan & Lavertu, 2021a, S. 3).

Maldonado und De Witte (2020) untersuchten neben der Mathematikleistung von Grundschüler*innen auch mögliche Veränderungen im sprachlichen Bereich. Fokussiert wurden die Fächer Niederländisch und Französisch an flämischen Schulen in Belgien (Maldonado & De Witte, 2020, S. 1). Zur Leistungserhebung kamen in beiden Fächern standardisierte Schulleistungstests zum Einsatz. Der Vergleich der Testwerte aus dem Jahr 2019 mit denen aus 2020 ließ einen signifikanten negativen Effekt der Schulschließungen erkennen, der mit -0.23 ($p < 0.01$) Standardabweichungen, bei 1.480 untersuchten Schulen, beziffert wurde (ebd., S. 13). Die Streuung der Leistungen im Fach Niederländisch stieg innerhalb der untersuchten Schulen um 20 und zwischen den Institutionen um 18 Prozent. Die Leistungsdefizite fielen bei Schüler*innen mit Migrationshintergrund größer aus (vgl. Maldonado & De Witte, 2020, S. 1).

Auch beim Vergleich der standardisierten Testwerte im Fach Französisch wurde ein negativer Effekt von -0.19 ($p < 0.01$) Standardabweichungen, bei 1.325 untersuchten Schulen, berichtet (ebd., S. 15). Bei der Berechnung dieser Effekte wurden keine Kontrollvariablen berücksichtigt.

Engzell et al. (2021) beschäftigten sich auch mit den Kompetenzbereichen Lesen und Rechtschreibung. Die verwendete Stichprobe enthielt Beobachtungen von 350.000 Grundschüler*innen aus den Niederlanden (vgl. Engzell et al., 2021, S. 1). Zur Erhebung der Leistung in den Bereichen Lesen und Rechtschreibung kamen standardisierte Tests zum Einsatz, die vom *Dutch National Institute for Educational Measurement* entwickelt wurden (ebd., S. 5, S. 50). In Bezug auf das Lesen ergab der Vergleich einer Kontrollgruppe mit 217.875 Fällen mit einer Versuchsgruppe bestehend aus 54.487 Beobachtungen einen signifikanten Mittelwertunterschied. Die

Schüler*innengruppe, die nach den Schulschließungen getestet wurde, zeigte im Mittel signifikant schlechtere Testwerte als die Vergleichsgruppe, die keine Schulschließungen erlebt hatte (ebd., S. 85). Die Stärke des gefundenen Effektes betrug -0.08 Standardabweichungen, bei einem 95%-Konfidenzintervall von -0.09 bis -0.07 (eigene Berechnung nach Lenhard & Lenhard, 2016).⁷

Auch im Bereich Rechtschreibung wurde für die Versuchsgruppe mit 58.627 Fällen im Vergleich zur Kontrollgruppe mit 284.499 Fällen ein signifikant niedrigerer Testmittelwert berichtet (vgl. Engzell et al., 2021, S. 85). Die Effektstärke betrug -0.04 Standardabweichungen, bei einem 95%-Konfidenzintervall von -0.05 bis -0.03 (eigene Berechnung nach Lenhard & Lenhard, 2016).⁸

Depping et al. (2021) erhoben in ihrer Studie auch die Kompetenzstände von Hamburger Grundschüler*innen im Bereich Deutsch-Leseverstehen. Anhand von Daten, die mit *LERNSTAND 4* und *KERMIT 3* erhoben wurden, konnten Kompetenzstände aus dem Schuljahr 2019/20 mit denen aus 2020/21 verglichen werden (vgl. Depping et al., 2021, S. 51). Analog zur Bestimmung der Kompetenzveränderung im Fach Mathematik erfolgte auch hier eine Einteilung der Ergebnisse in die drei Kompetenzstufen KSI, KSII und KSIII+ (ebd., S. 63). Die prozentualen Verteilungen auf die mit *KERMIT 3* ermittelten Kompetenzstufen aus dem Jahr 2019 wurden mit denen verglichen, die 2020 mit *LERNSTAND 4* ermittelt wurden. Dabei konnten keine bedeutsamen Unterschiede festgestellt werden (ebd., S. 63 f.). Genau wie bei der Untersuchung der Kompetenzen in Mathematik gilt es auch hier zu beachten, dass aufgrund zu weniger bearbeiteter Aufgaben in einigen Fällen kein Kompetenzwert berechnet werden konnte (ebd., S. 74).

Depping et al. (2021) untersuchten außerdem die Deutsch-Leseverstehens-Kompetenz von Hamburger Fünftklässler*innen. Dabei kamen 2019 *KERMIT 5* und 2020 *LERNSTAND 5* als Erhebungsinstrumente zum Einsatz. Die Stichprobe aus dem Jahr 2019 umfasste 13.716 Schüler*innen, aus dem Jahr 2020 standen 13.936 Beobachtungen zur Verfügung (vgl. Depping et al., 2021, S. 65). Es konnten auch hier keine statistisch signifikanten Unterschiede zwischen den mittleren Punktwerten besagter Gruppen gefunden werden (ebd., S. 67).

Schult et al. (2021) erhoben Schüler*innen-Kompetenzen mit dem Instrument *LERNSTAND 5*. Im Leseverständnis-Teil dieses Tests wurden verschiedene Lesevorgänge, wie das Abrufen explizit angegebener Informationen, das Interpretieren von Ideen, die Evaluation von Inhalt und Textelementen und die Fähigkeit, Schlüsse zu ziehen, abgeprüft (vgl. Schult et al., 2021, S. 10). So ergab sich ein Gesamtestwert, dessen Entwicklung über die Jahre 2017 bis 2020 untersucht wurde. Beim Vergleich

7 Für Details zur Berechnung der Effektstärke siehe Fußnote 4.

8 Für Details zur Berechnung der Effektstärke siehe Fußnote 4.

der 249.862 standardisierten Testwerte aus den Jahren 2017 bis 2019, vor der Schulschließung, mit den 81.810 Werten aus 2020 ergab sich ein negativer Effekt von -0.07 Standardabweichungen (ebd., S. 29).

Rose et al. (2021) erhoben neben der Mathematikleistung von sechs- bis siebenjährigen britischen Schüler*innen der 2. Jahrgangsstufe auch Daten zur Leseleistung, wobei eine Stichprobe aus dem Jahr 2017 als Vergleichsgrundlage herangezogen wurde (vgl. Rose et al., 2021, S. 1). Am Sprachtest im Rahmen des *NFER Autumn Assessment* nahmen im November 2020 5.931 Schüler*innen teil (ebd., S. 4). Die Leseleistung, welche die Schüler*innen im Herbsttest zeigten, war signifikant niedriger als die in der Vergleichsgruppe gezeigte. Der gemessene Effekt betrug $d = -0.17$, was einem Verlust von zwei Monaten Lernfortschritt entsprach. Die Daten deuteten außerdem darauf hin, dass eine große Zahl von Schüler*innen im November 2020 nicht dazu in der Lage war, die Testaufgaben effektiv zu bearbeiten (vgl. Rose et al., 2021, S. 5 f.). Die Lücke, die sich in Bezug auf Schüler*innen-Leistung aufgetan hat, fiel sogar noch größer aus als prognostiziert (ebd., S. 12).

6 Auswirkungen auf die Schulleistung in Grundschule und Sekundarstufe – mehrere Fächer kombiniert

Mithilfe eines latenten Wachstumsmodells verglichen Tomasik et al. (2020) den Lernzuwachs von Primar- und Sekundarschüler*innen in den acht Wochen vor und während der Schulschließungen. Dies geschah auf Grundlage einer Stichprobe mit insgesamt 28.685 Schweizer Schüler*innen (Tomasik et al., 2020, S. 1). Zur Datenerhebung wurde das Online-Computersystem *MINDSTEPS* eingesetzt, das ursprünglich dafür entwickelt wurde, Lehrkräften und Schüler*innen formatives Feedback zu liefern. Es kann allerdings auch dazu genutzt werden, die Entwicklung von Fähigkeiten im Laufe der Zeit einzuschätzen. Es lieferte Werte für die Bereiche Mathematik, Lesen und Grammatik. Die Primarstufen-Stichprobe enthielt Beobachtungen von 13.134 Grundschüler*innen (vgl. Tomasik et al., 2020, S. 4 f.). Im Zuge der Analysen wurde nicht zwischen Fachbereichen unterschieden. Es konnte gezeigt werden, dass der Lernzuwachs der Primarschüler*innen, die in Präsenz unterrichtet wurden, mehr als doppelt so hoch war, wie der Lernzuwachs der Primarschüler*innen im Distanzunterricht (ebd., S. 6). Die Autor*innen wiesen darauf hin, dass besonders interindividuelle Unterschiede beim Lernen während der Zeit des Distanzunterrichts stark zugenommen hatten (ebd., S. 7 ff.). Der zur Untersuchung der Lernentwicklung bei Sekundarschüler*innen verwendete Datensatz enthielt 15.551 Fälle (ebd., S. 5). Im Gegensatz zum Primarbereich ließ sich hier kein signifikanter Unterschied zwischen dem Lernfortschritt der Schüler*innen im Präsenz- und derjenigen im Fernunterricht finden (vgl. Tomasik et al., S. 6). Die Autor*innen der Studie wiesen darauf hin, dass im Rahmen ihrer Analysen nicht der bestmögliche

Präsenzunterricht mit dem bestmöglichen Distanzunterricht verglichen wurde (ebd., S. 7).

Engzell und Kolleg*innen (2021) berichteten einen Gesamtwert, in dem die Testwerte aus den Bereichen Mathematik, Lesen und Schreiben zusammenliefen (Engzell et al., 2021, S. 5, S. 50). Während die Kontrollgruppe 289.189 Grundschüler*innen umfasste, enthielt die Versuchsgruppe 69.190 Fälle (ebd., S. 85). Die Testwerte der Schüler*innengruppe, die nach den Schulschließungen getestet wurde, fielen im Mittel signifikant niedriger aus als die der Kontrollgruppe, die keine Schulschließungen erfahren hatte (ebd., S. 85). Die Effektstärke lag bei -0.14 Standardabweichungen, mit einem 95%-Konfidenzintervall von -0.14 bis -0.13 (eigene Berechnung nach Lenhard & Lenhard, 2016).⁹ Die Autor*innen bezifferten den Lernverlust mit einem negativen Gesamteffekt von -0.08 Standardabweichungen, wobei die Verluste bei Schüler*innen aus bildungsfernen Elternhäusern um 60 Prozent größer ausfielen (vgl. Engzell et al., 2021, S. 1).

Maldonado und De Witte (2020) berechneten einen Effekt in Bezug auf den Gesamtnotendurchschnitt in den Fächern Mathematik, Niederländisch, Französisch, Naturwissenschaften und Sozialwissenschaften. Beim Vergleich der Werte von vor und nach den Schulschließungen ließ sich im vollständigen Modell ein negativer Effekt von -0.22 Standardabweichungen nachweisen. Die zugrundeliegende Stichprobe umfasste 513 belgische Grundschulen (Maldonado & De Witte, 2020, S. 16). Für das Modell, in dem keine Kontrollvariablen berücksichtigt wurden, berichteten die Autor*innen bei einer Stichprobe mit 719 Schulen einen ähnlichen Effekt von -0.22 Standardabweichungen (ebd.).

7 Diskussion

Die präsentierten Studienergebnisse sind aufgrund ihrer Vielfalt hinsichtlich Design, Stichprobengröße, Schulsystem im Land der Erhebung, Fachbezug, Schulstufe und Messinstrument differenziert zu betrachten und nur bedingt miteinander vergleichbar. Im Folgenden werden die Ergebnisse noch einmal überblicksartig zusammengefasst.

Im Fachbereich Mathematik wurden nur in einer der Forschungsarbeiten positive Effekte gefunden, und zwar in der Grundschule (Meeter, 2021). Keine bedeutsamen Unterschiede in diesem Bereich fanden die Autor*innen einer Studie sowohl in der Grund- als auch Sekundarstufe (Depping et al., 2021). Negative Effekte auf Mathematikleistung zeigten sich in fünf Studien in der Grundschule (Engzell et al., 2021; Kuhfeld et al., 2020; Maldonado & De Witte, 2020; Rose et al., 2021; Schult et

⁹ Für Details zur Berechnung der Effektstärke siehe Fußnote 4.

al., 2021). Bei Schult et al. (2021) wurden neben Grund- auch Sekundarschüler*innen untersucht.

Im sprachlichen Bereich fanden zwei Studien keine bedeutsamen Unterschiede in der Grundschule (Depping et al., 2021; Kuhfeld et al., 2020). Depping et al. (2021) kamen auch für die Sekundarstufe zu diesem Befund. Negative Effekte im Bereich Sprache wurden in fünf Forschungsarbeiten berichtet (Engzell et al., 2021; Kogan & Lavertu, 2020; Maldonado & De Witte, 2020; Rose et al., 2021; Schult et al., 2021).

Über alle Schulstufen und Fachbereiche hinweg ließen sich in der Mehrzahl negative Auswirkungen der Schulschließungen finden. Dies belegen auch Befunde in Bezug auf die Kombination mehrerer Fachbereiche (Engzell et al., 2021; Maldonado & De Witte, 2020; Tomasik et al., 2021). Auch wenn die Höhe der in den Studien gefundenen Effekte durchaus schwankte und manche zunächst klein wirken (vgl. dazu Cohen, 1988), gilt es zu beachten, dass die Ausprägung der Effekte in Relation zu dem Lernfortschritt zu betrachten ist, der bei regulärer Beschulung in Präsenz zu erwarten wäre. Auch ein nach Cohen kleiner negativer Effekt fällt folglich ins Gewicht. In einigen Studien konnte gezeigt werden, dass die Heterogenität von Schüler*innen-Leistungen zugenommen hat. Dies deutet darauf hin, dass manche Schüler*innen in Hinblick auf die fachlich-kognitive Leistung stärker in Rückstand gerieten als andere. Dabei geben die Forschungsergebnisse aus allen hier berücksichtigten Ländern Hinweise darauf, dass die entstandenen Lernrückstände bei ohnehin schon benachteiligten Schüler*innen, sei es aufgrund des sozioökonomischen Status, der ethnischen Zugehörigkeit oder wegen eines Migrationshintergrundes, stärker ausgeprägt waren (vgl. zu dieser Problematik den Beitrag von Helm, Huber & Postlbauer in diesem Band).

Eine Limitation dieses narrativen Reviews besteht in der Tatsache, dass nur eine kleine Auswahl an Studien näher betrachtet wurde. Da es sich um ein junges Forschungsfeld handelt, werden vermutlich zahlreiche weitere Studien zum Thema erscheinen. Die Befundlage kann sich also durchaus noch ändern. Im Rahmen dieses Reviews wurden Studien aus Ländern betrachtet, in denen einigermaßen vergleichbare Bedingungen im Hinblick auf gesellschaftliche Ordnung und Schulsystem herrschen. Allerdings ist eine Verallgemeinerung der Forschungsergebnisse aufgrund der unterschiedlichen Rahmenbedingungen der Schulsysteme, die hier berücksichtigt wurden, nur unter Vorbehalt möglich. Der Lehrplan der Schulen in den USA, zum Beispiel, weist nicht die gleichen Inhalte in der gleichen Reihenfolge auf wie der in den Niederlanden oder in Deutschland. Auch ist beispielsweise eine Gliederung der Schulformen in Grund-, Mittel-, Realschule und Gymnasium nicht einmal in allen Bundesländern Deutschlands gegeben. Des Weiteren unterscheidet sich mit hoher Wahrscheinlichkeit die technische Ausstattung nicht nur zwischen den Ländern, sondern auch zwischen den einzelnen Schulen und Schüler*innen. Es liegt außerdem nahe, dass sich Lehrkräfte im Hinblick auf ihre Erfahrungen

mit und ihre Einstellungen zu digitalen Medien unterscheiden. Noch dazu variieren die beschriebenen Studien in Bezug auf die Darstellung der Ergebnisse und die eingesetzten Erhebungsinstrumente. Die Hypothese, dass die pandemiebedingten Schulschließungen zu negativen Effekten auf die fachlich-kognitive Leistung von Schüler*innen im In- und Ausland führten, kann daher nicht vollumfänglich bestätigt werden. Neben einigen negativen Effekten werden im Rahmen der präsentierten Studien auch positive Effekte berichtet. Wie bereits dargestellt, kann dies zum Teil auf Besonderheiten beim Studiendesign zurückgeführt werden (Depping et al., 2021; Meeter, 2021).¹⁰

Es gilt zu beachten, dass die gefundenen Effekte womöglich nicht allein auf geschlossene Schulen zurückzuführen sind. Ein Mangel an Orten zum Lernen zuhause, schlechte Internetverbindung und ein Mangel an geeigneter technischer Ausstattung könnten in der Folge zu Lernverlusten geführt haben (Schult et al., 2021, S. 5).

Wie schnell der entstandene Rückstand, auf den die präsentierten Studien hindeuten, wieder aufgeholt werden kann, ist fraglich. Es sollten langfristig Konzepte und Programme entwickelt werden, um dieser Herausforderung gerecht zu werden. Bestrebungen in diese Richtung, wie zum Beispiel das Sommerschulprogramm (vgl. BMFSFJ, 2021), sind bereits klar erkennbar. Dabei ist die Heterogenität der Forschungsbefunde zu berücksichtigen. Manche Schüler*innengruppen und Schulen weisen womöglich größere Rückstände auf als andere. Die Ermittlung potentieller Rückstände in Bezug auf die fachlich-kognitive Leistung von Schüler*innen gilt es unter Berücksichtigung von Fachbereich und Schulform beziehungsweise -stufe möglichst schulspezifisch vorzunehmen.

Literatur und Internetquellen

- Aucejo, E. M., & Romano, T. F. (2016). Assessing the Effect of School Days and Absences on Test Score Performance. *Economics of Education Review*, 55, 70–87. <https://doi.org/10.1016/j.econedurev.2016.08.007>
- Azevedo, J. P., Hasan, A., Goldemberg, D., Iqbal, S. A., & Geven, K. (2020). *Simulating the Potential Impacts of COVID-19 School Closures on Schooling and Learning Outcomes: A Set of Global Estimates*. World Bank Policy Research Working Paper. <https://doi.org/10.1596/1813-9450-9284>

10 Depping et al. (2021) schlossen Testhefte, die zu weniger als 20 Prozent bearbeitet waren, von ihren Analysen aus, wodurch vermutlich schwächere Schüler*innen von der Untersuchung ausgeschlossen wurden. Die positive Vorher-Nachher-Differenz beim Lesen könnte daher teilweise darauf zurückzuführen sein, dass im Jahr 2020 eine größere Anzahl Testhefte (5,5%) ausgeschlossen wurde als im Jahr 2019 (3,4%) (Schult et al., 2021, S. 7). Bei Meeter (2021) war die Stichprobe selektiv und die Intervention hatte gleichzeitig die Kompetenzbewertung zum Ergebnis, was zu Methodeneffekten führen kann (Schult et al., 2021, S. 20).

- Bao, X., Qu, H., Zhang, R., & Hogan, T. (2020). *Literacy Loss in Kindergarten Children During COVID-19 School Closures*. <http://dx.doi.org/10.31235/osf.io/nbv79>
- Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions. *Journal of Research on Educational Effectiveness*, 1 (4), 289–328. <https://doi.org/10.1080/19345740802400072>
- BMFSFJ (Bundesministeriums für Familie, Senioren, Frauen und Jugend). (2021). *Aktionsprogramm „Aufholen nach Corona für Kinder und Jugendliche“*. Zugriff am 23.08.2021. Verfügbar unter <https://www.bmfsfj.de/bmfsfj/themen/corona-pandemie/aufholen-nach-corona>.
- Bol, T. (2020). *Inequality in Homeschooling During the Corona Crisis in the Netherlands. First Results from the LISS Panel*. SocArXiv. <https://doi.org/10.31235/osf.io/hf32q>
- Carlsson, M., Dahl, G. B., Öckert, B., & Rooth, D.-O. (2015). The Effect of Schooling on Cognitive Skills. *Review of Economics and Statistics*, 97, 533–547. https://doi.org/10.1162/REST_a_00501
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Hillsdale, NJ: Erlbaum. <https://doi.org/10.4324/9780203771587>
- De Haas, M., Faber, R., & Hamersma, M. (2020). How COVID-19 and the Dutch ‘Intelligent Lockdown’ Changed Activities, Work and Travel Behaviour. Evidence from Longitudinal Data in the Netherlands. *Transportation Research Interdisciplinary Perspectives*, 100150. <https://doi.org/10.1016/j.trip.2020.100150>
- Depping, D., Lücken, M., Musekamp, F., & Thonke, F. (2021). Kompetenzstände Hamburger Schüler*innen vor und während der Corona-Pandemie. In D. Fickermann & B. Edelstein (Hrsg.), *Schule während der Corona-Pandemie. Neue Ergebnisse und Überblick über ein dynamisches Forschungsfeld* (Die Deutsche Schule, 17. Beiheft) (S. 51–79). Münster: Waxmann. <https://doi.org/10.31244/9783830993315.03>
- Drossel, K., & Eickelmann, B. (2018). Die Rolle der Lehrerprofessionalisierung für die Implementierung neuer Technologien in den Unterricht. Eine Latent-Class-Analyse zur Identifikation von Lehrertypen. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 31, 166–191. <https://doi.org/10.21240/mpaed/31/2018.06.04.X>
- Eickelmann, B., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., et al. (2019). *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking*. Münster: Waxmann. <http://dx.doi.org/10.25656/01:18166>
- Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning Loss due to School Closures During the COVID-19 Pandemic. SocArXiv. <https://doi.org/10.31235/osf.io/ve4z7>
- Eyles, A., Gibbons, S., & Montebruno, P. (2020). Covid-19 School Shutdowns: What Will They Do to Our Children’s Education? *LSE CEP COVID-19 Analysis, 001*. Zugriff am 23.08.2021. Verfügbar unter <https://cep.lse.ac.uk/pubs/download/cepcovid-19-001.pdf>
- Frenette, M., Frank, K., & Deng, Z. (2020). School Closures and the Online Preparedness of Children during the COVID-19 Pandemic. (Vol. 103; Tech. Rep.). In Statistics Canada (Hrsg.), *Economic Insights* (Catalogue 11-626-X, Issue 2020001, No. 103). Zugriff am 23.08.2021. Verfügbar unter https://www.researchgate.net/publication/340720825_School_Closures_and_the_Online_Preparedness_of_Children_during_the_COVID-19_Pandemic.
- Gershenson, S., Jackowitz, A., & Brannegan, A. (2017). Are Student Absences Worth the Worry in US Primary Schools? *Education Finance and Policy*, 12, 137–165. http://dx.doi.org/10.1162/EDFP_a_00207

- Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). *Effects of COVID-19-related School Closures on Student Achievement. A Systematic Review*. PsyArXiv. <https://doi.org/10.31234/osf.io/mcnvk>
- Hansen, B. (2011). *School Year Length and Student Performance: Quasi-experimental Evidence (SSRN Working Paper No. 2269846)*. New York: Social Science Research Network. <https://dx.doi.org/10.2139/ssrn.2269846>
- Helm, C., Huber, S., & Loisinger, T. (2021). Was wissen wir über schulische Lehr-Lern-Prozesse im Distanzunterricht während der Corona-Pandemie? – Evidenz aus Deutschland, Österreich und der Schweiz. *Zeitschrift für Erziehungswissenschaft*, 24, 237-311. <https://doi.org/10.1007/s11618-021-01000-z>
- Kogan, V., & Lavertu, S. (2021a). *Covid and Student Achievement: Early Evidence from Ohio's Fall Tests*. Zugriff am 23.06.2021. Verfügbar unter: <https://www.educationnext.org/covid-student-achievement-early-evidence-from-ohios-fall-tests/>.
- Kogan, V., & Lavertu, S. (2021b). *The COVID-19 Pandemic and Student Achievement on Ohio's Third-Grade English Language Arts Assessment*. The Ohio State University. Zugriff am 23.06.2021. Verfügbar unter: http://glenn.osu.edu/educational-governance/reports/reports-attributes/ODE_ThirdGradeELA_KL_1-27-2021.pdf.
- Kuhfeld, M., Ruzek, E., Johnson, A., Tarasawa, B., & Lewis, K. (2020). *Technical Appendix for: Learning During COVID-19: Initial Findings on Students' Reading and Math Achievement and Growth*. NWEA. Zugriff am 23.06.2021. Verfügbar unter: <https://www.nwea.org/research/publication/learning-during-covid-19-initial-findings-on-students-reading-and-math-achievement-and-growth/>
- Lavy, V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries. *Economic Journal*, 125, F397–F424. <https://doi.org/10.1111/econj.12233>
- Lenhard, W., & Lenhard, A. (2016). *Berechnung von Effektstärken*. <http://dx.doi.org/10.13140/RG.2.2.17823.92329>
- Liu, J., Lee, M., & Gershenson, S. (2020). *The Short- and Long-run Impacts of Secondary School Absences (EdWorking-Paper no. 20-125)*. Annenberg Institute. <https://doi.org/10.1016/j.jpubeco.2021.104441>
- Maldonado, J. E., & De Witte, K. (2020). *The Effect of School Closures on Standardised Student Test Outcomes*. KU Leuven – Faculty of Economics and Business. <http://dx.doi.org/10.1002/berj.3754>
- Meeter, M. (2021). *Primary School Mathematics During Covid-19. No Evidence of Learning Gaps in Adaptive Practicing Results*. PsyArXiv. <https://doi.org/10.31234/osf.io/8un6x>
- Ressing, M., Blettner, M., & Klug, S. (2009). Systematic Literature Reviews and Meta-Analyses. *Deutsches Ärzteblatt*, 106 (27), 456-463. <https://doi.org/10.3238/arztebl.2009.0456>
- Rose, S., Twist, L., Lord, P., Rutt, S., Badr, K., Hope, C., & Styles, B. (2021). *Impact of School Closures and Subsequent Support Strategies on Attainment and Socio-emotional Wellbeing in Key Stage 1. Interim Paper 1*. NFER. Zugriff am 23.06.2021. Verfügbar unter: https://educationendowmentfoundation.org.uk/public/files/Publications/Covid-19_Resources/Impact_of_school_closures_KS1_interim_findings_paper_-_Jan_2021.pdf.
- Sacerdote, B. (2012). When the Saints go Marching Out: Long-term Outcomes for Student Evacuees from Hurricanes Katrina and Rita. *American Economic Journal: Applied Economics*, 4, 109–135. <http://dx.doi.org/10.1257/app.4.1.109>
- Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2021). *Did Students Learn Less During the COVID-19 Pandemic? Reading and Mathematics Competencies Before and After the First Pandemic Wave*. <https://doi.org/10.31234/osf.io/pqgtf>
- Tomasik, M. J., Helbling, L. A., & Moser, U. (2020). Educational Gains of In-person vs. Distance Learning in Primary and Secondary Schools. A Natural Experiment During

the COVID-19 Pandemic School Closures in Switzerland. *International Journal of Psychology*. <https://doi.org/10.1002/ijop.12728>
Zierer, K. (2021). Effects of Pandemic-Related School Closures on Pupils' Performance and Learning in Selected Countries. A Rapid Review. *Education Sciences*, 11 (6), 252. <https://doi.org/10.3390/educsci11060252>

Tobias Böttger, geb. 1989, Wissenschaftlicher Mitarbeiter am Lehrstuhl für Schulpädagogik der Universität Augsburg.
E-Mail: tobias.boettger@phil.uni-augsburg.de

Klaus Zierer, Prof. Dr., geb. 1976, Ordinarius für Schulpädagogik, Universität Augsburg.
E-Mail: klaus.zierer@phil.uni-augsburg.de

Korrespondenzadresse: Universität Augsburg, Philosophisch-Sozialwissenschaftliche Fakultät, Lehrstuhl für Schulpädagogik, Universitätsstraße 10, 86159 Augsburg