

Der Nutzen externer Intervention in der Grundschule: Die exemplarische Evaluation eines Stiftungsprogramms

Helbig, Marcel; Mehnert, Ulrike; Allmendinger, Jutta

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

Wissenschaftszentrum Berlin für Sozialforschung (WZB)

Empfohlene Zitierung / Suggested Citation:

Helbig, M., Mehnert, U., & Allmendinger, J. (2018). *Der Nutzen externer Intervention in der Grundschule: Die exemplarische Evaluation eines Stiftungsprogramms*. (Discussion Papers / Wissenschaftszentrum Berlin für Sozialforschung, P 2018-002). Berlin: Wissenschaftszentrum Berlin für Sozialforschung gGmbH. <https://hdl.handle.net/10419/179427>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Helbig, Marcel; Mehnert, Ulrike; Allmendinger, Jutta

Working Paper

Der Nutzen externer Intervention in der Grundschule: Die exemplarische Evaluation eines Stiftungsprogramms

WZB Discussion Paper, No. P 2018-002

Provided in Cooperation with:
WZB Berlin Social Science Center

Suggested Citation: Helbig, Marcel; Mehnert, Ulrike; Allmendinger, Jutta (2018) : Der Nutzen externer Intervention in der Grundschule: Die exemplarische Evaluation eines Stiftungsprogramms, WZB Discussion Paper, No. P 2018-002, Wissenschaftszentrum Berlin für Sozialforschung (WZB), Berlin

This Version is available at:
<http://hdl.handle.net/10419/179427>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WZB

Wissenschaftszentrum Berlin
für Sozialforschung



Marcel Helbig und Ulrike Mehnert

Der Nutzen externer Intervention in der Grundschule – Die exemplarische Evaluation eines Stiftungsprogramms

(Projektleitung Jutta Allmendinger und Marcel Helbig)

Discussion Paper

P 2018–002

Mai 2018

**Wissenschaftszentrum Berlin für Sozialforschung
Projektgruppe der Präsidentin**

Das Urheberrecht liegt bei den Autorinnen/dem Autor.

Discussion Papers des WZB dienen der Verbreitung von Forschungsergebnissen aus laufenden Arbeiten im Vorfeld einer späteren Publikation. Sie sollen den Ideenaustausch und die akademische Debatte befördern. Die Zugänglichmachung von Forschungsergebnissen in einem WZB Discussion Paper ist nicht gleichzusetzen mit deren endgültiger Veröffentlichung und steht der Publikation an anderem Ort und in anderer Form ausdrücklich nicht entgegen.

Discussion Papers, die vom WZB herausgegeben werden, geben die Ansichten des jeweiligen Autors wieder und nicht die der gesamten Institution WZB.

Marcel Helbig und Ulrike Mehnert

Der Nutzen externer Intervention in der Grundschule – Die exemplarische Evaluation eines Stiftungsprogramms

(Projektleitung Jutta Allmendinger und Marcel Helbig)

Discussion Paper P 2018–002

Wissenschaftszentrum Berlin für Sozialforschung (2018)

Inhaltsverzeichnis

Tabellenverzeichnis.....	4
Abbildungsverzeichnis	4
1 Einleitung	5
2 Ziele, Beschreibung und Ablauf des Programms	10
3 Evaluationsdesign.....	14
3.1 Schulauswahl.....	14
3.2 Evaluationskriterien und Befragungsablauf	15
3.3 Methodische Herausforderungen und Limitationen	18
4 Ergebnisse	22
4.1 Erwartungshaltung.....	22
4.2 Entwicklung der Schülerinnen und Schüler von Klasse 1 bis 4.....	24
4.2.1 Motivation und Zielorientierung stärken (Schülerinnen- und Schülersicht).....	24
4.2.2 Sozialkompetenzen.....	28
4.2.3 Das Selbst stärken	31
4.2.4 Sprachliche Kompetenz.....	35
4.2.5 Bildungsziele	37
4.2.6 Schulleistungen und Schulverlauf.....	39
4.3 Die Sicht der beteiligten Akteure	44
5 Zusammenfassung und Fazit	47
6 Literatur	49
7 Anhang	53
7.1 Operationalisierung der Konstrukte	53

Tabellenverzeichnis

Tabelle 1: Projektziele, deren Operationalisierung und Person, über die dies erfasst wurde.....	15
Tabelle 2: Überblick über den Zeitpunkt der Erhebung der eingesetzten Test- und Befragungsinstrumente.....	16
Tabelle 3: Entwicklung der Bildungsaspirationen (Schulform – Schülerinnen- und Schülersicht).....	38
Tabelle 4: Entwicklung der Bildungsaspirationen (Schulabschluss – Schülerinnen- und Schülersicht).....	39

Abbildungsverzeichnis

Abbildung 1: Erhebungsplan.....	18
Abbildung 2: Wachstenden pro Jahr in verschiedenen Kontexten.....	23
Abbildung 3: Entwicklung der Lernmotivation (Schülerinnen- und Schülersicht).....	26
Abbildung 4: Entwicklung von Anstrengungsbereitschaft und Arbeitsvermeidung (Schülerinnen- und Schülersicht).....	27
Abbildung 5: Entwicklung des Lernverhaltens (Lehrersicht).....	28
Abbildung 6: Entwicklung des Klassenklimas (Schülerinnen- und Schülersicht).....	29
Abbildung 7: Entwicklung der Skala Ausgrenzung (Schülerinnen- und Schülersicht).....	29
Abbildung 8: Entwicklung der Skalen selbst erlebte Ausgrenzung und Empathie (Schülerin- und Schülersicht).....	30
Abbildung 9: Entwicklung des Sozialverhaltens (Lehrersicht).....	31
Abbildung 10: Entwicklung des allgemeinen Selbstkonzepts (Schülerinnen- und Schülersicht).....	32
Abbildung 11: Entwicklung des schulischen Selbstkonzepts (Schülerinnen- und Schülersicht).....	33
Abbildung 12: Entwicklung Selbstwirksamkeitserwartung (Schülerinnen- und Schülersicht).....	34
Abbildung 13: Entwicklung Durchsetzungsvermögen (Lehrersicht).....	34
Abbildung 14: Schulische Kompetenzen am Ende der Klasse 3 (gemessen über die Vergleichsarbeiten der Schülerinnen und Schüler).....	35
Abbildung 15: Entwicklung der (eingeschätzten) Deutschnote (Lehrersicht).....	36
Abbildung 16: Entwicklung der (eingeschätzten) Mathematiknote (Lehrersicht).....	40
Abbildung 17: Entwicklung der (eingeschätzten) Sachkunde-Note (Lehrersicht).....	41
Abbildung 18: Klassenwiederholungen im Untersuchungszeitraum (Lehrerinformation).....	42
Abbildung 19: Übergang auf das Gymnasium (Lehrerinformation).....	43

1 Einleitung

Der Bildungserfolg hängt in kaum einem anderen Land so stark von der sozialen Herkunft ab wie in Deutschland. Bereits die breit rezipierten Studien von Peisert (1967) und Picht (1964) in den 1960er Jahren thematisierten dieses Problem. Auch die erste internationale Vergleichsstudie (Husén 1967) kam zu dem Ergebnis, dass in keinem anderen Land die soziale Herkunft und der Bildungserfolg stärker zusammenhängen als in Deutschland. In Folge dieser alarmierenden Befunde kam es spätestens Ende der 1960er Jahre zu regelrechten schulpolitischen Kämpfen. In einigen Bundesländern wurden Veränderungen des Schulsystems vorgenommen. Gesamtschulen wurden eingeführt und einige Bundesländer verlängerten das längere gemeinsame Lernen durch die Einführung schulartenunabhängiger Orientierungsstufen (Helbig und Nikolai 2015). Wie spätestens mit der ersten PISA-Studie deutlich wurde, haben diese Reformen nicht dazu geführt, dass der Zusammenhang zwischen sozialer Herkunft und Bildungserfolg in Deutschland abgenommen hätte (Deutsches PISA-Konsortium 2001). Zu diesem Befund kamen auch andere Analysen, die stärker erlangte Zertifikate in den Blick nahmen (Pollak 2009; Schimpl-Neimanns 2000).

Neben der Bundes-, Landes- und Kommunalpolitik hat sich in den letzten Jahren und Jahrzehnten ein weiterer großer Akteur in der Bildungspolitik etabliert: das Stiftungswesen (Höhne 2015). 2015 existierten bereits knapp über 20.000 Stiftungen in Deutschland und jedes Jahr kommen rund 700 neue Stiftungen hinzu. Viele dieser Stiftungen engagieren sich heute im Bildungssektor, um Lösungsansätze für Probleme zu finden, bei denen die deutsche Bildungspolitik (aus Sicht der Stifter) bisher versagt hat. Mittlerweile existiert in Deutschland eine Vielzahl von Programmen und Modellprojekten, die die Gruppe der sozial Benachteiligten in den Blick nehmen. In diesem Forschungsbericht wollen wir uns der Wirkungsweise eines dieser Modellprojekte zuwenden – dem Pilot-Programm „Vivo-Bildung von Grund auf!“ der Stiftung der Deutschen Wirtschaft (sdw) – und wie sich dieses auf die Entwicklung von Schülerinnen und Schülern in der Grundschule ausgewirkt hat.

Um die Frage zu beantworten, ob ein solches Programm es schaffen kann, sozial benachteiligten Schülerinnen und Schülern den erfolgreichen Weg durch das Bildungssystem zu ebnet, muss man zunächst die Frage stellen, an welchen Stellen des Bildungssystems es zu einer Benachteiligung von Kindern aus sozial niedrigen Schichten kommt und wie sich dies auf ihren Bildungserfolg auswirkt.

Aus verschiedenen Studien weiß man, dass die schlechteren Bedingungen von Kindern aus sozial benachteiligten Familien schon früh festzustellen sind. Im Vorschulalter sind Kinder aus sozial benachteiligten Haushalten stärker von verschiedenen belastenden Faktoren betroffen (z.B. Rauchen im Haushalt) und von förderlichen Aktivitäten stärker ausgeschlossen (z.B. sportliche Betätigung) (RKI 2008). So verwundert es nicht, dass Kinder aus sozial benachteiligten Familien bereits bei der Schuleingangsuntersuchung eher physische und psychische Defizite aufweisen als Kinder aus höheren Schichten (Oberwöhrmann et al. 2013).

Verschiedene Faktoren (Allmendinger 2012) führen außerdem zu schlechteren kognitiven Kompetenzen bei diesen Kindern. Aber auch bei gleichen Kompetenzen bekommen sie häufig schlechtere Noten und werden seltener für das Gymnasium empfohlen. Die schlechtere Benotung ergibt sich auch dadurch, dass Lehrer Kinder aus unteren Schichten als weniger motiviert und leistungsbereit wahrnehmen (siehe Helbig und Morar 2017).

Die geringere schulische Motivation ist freilich auch eine Folge unterschiedlicher Erziehungsstile und vorgelebter Arbeitsnormen in den Herkunftsfamilien. Zudem werden durch die ungünstige soziale Lage der Eltern auch die Bildungs- und Berufsaspirationen der Kinder negativ beeinflusst, für die „Hartz IV“ zu einer realistischen Lebensperspektive wird. Kinder aus sozial benachteiligten Familien haben schlechtere kognitive Kompetenzen, bekommen schlechtere Noten und werden dementsprechend am Ende der Grundschule seltener für das Gymnasium empfohlen. Die Gründe für diese sogenannten primären Herkunftsmerkmale (Kompetenzen und Noten) (Boudon 1974) sind vielfältig. Bei der Entscheidung, welche Schulform im Anschluss an die Grundschule besucht werden soll, vergrößert sich die soziale Schere allerdings durch sogenannte sekundäre Herkunftsmerkmale (Bildungsentscheidungen) weiter. Selbst, wenn Kinder aus

sozial benachteiligten Schichten eine Gymnasialempfehlung erhalten, entscheiden sich ihre Eltern häufiger gegen den Besuch des Gymnasiums. Im Gegensatz dazu korrigieren Eltern aus höheren Schichten eine Schulempfehlung häufig nach oben, also auf einen anspruchsvolleren Schultyp. Dies ergibt sich dadurch, dass Eltern aus niedrigeren Schichten die Kosten gymnasialer Bildung (inklusive Opportunitätskosten des längeren Schulbesuchs) höher einschätzen, die Erfolgswahrscheinlichkeit ihrer Kinder niedriger einschätzen und die Erträge gymnasialer Bildung geringer einschätzen als Eltern aus höheren Schichten (Becker 2000; Breen und Golthorpe 1997). Aus dem Zusammenwirken von primären und sekundären Effekten ergibt sich, dass Kinder aus hohen sozialen Schichten (Dienstklasse I) 4,6 mal häufiger das Gymnasium besuchen als Kinder aus Facharbeiterhaushalten (Deutsches PISA-Konsortium 2008).

Ziel unserer Untersuchung ist es, zu analysieren, ob das gezielte koordinierte Förderprogramm eines externen Akteurs es vermag, soziale Ungleichheiten auf dem Weg zum Gymnasium auszugleichen. Ein solches Förderprogramm muss – entsprechend der vorangegangenen Problembeschreibung – Lösungsansätze für drei Probleme bieten: Erstens muss ein solches Programm darauf ausgerichtet sein, Kompetenzunterschiede sozialer Gruppen, die bereits in der ersten Klasse vorliegen, im Laufe der Grundschulzeit auszugleichen. Konkret muss man vor allem die Sprach- und Lesekompetenzen von Schülern aus sozial benachteiligten Lagen fördern. Zweitens muss die Persönlichkeit von Schülern aus sozial benachteiligten Schichten gefördert werden. Wenn man das schulische Selbstkonzept, die Leistungsbereitschaft und die schulische Motivation dieser Schüler stärkt, dann sollte sich dies auch auf die Benotung der Schüler auswirken und zu häufigeren Gymnasialempfehlungen führen. Zum Aspekt der Persönlichkeitsentwicklung gehören auch soziale Kompetenzen. Drittens muss ein Programm zu Verringerung sozialer Differenzen beim Grundschulübergang im Bereich der Bildungsziele bzw. Bildungsaspirationen von Eltern und Kindern ansetzen. Das heißt die Eltern sozial benachteiligter Schüler müssen über die unterschiedlichen Bildungsgänge, deren jeweiligen Kosten und Nutzen informiert werden. Diese Informationen sollte man ebenso den Schülern zukommen lassen. Denn nicht nur die Bildungswünsche der Eltern, sondern auch ihrer Kinder

bestimmen maßgeblich den Übergang auf die weiterführende Schule (Wohlkinger und Ditton 2012).

Oft werden in Stiftungsprogrammen „begabte“¹ Kinder aus sozial benachteiligten Schichten herausgefiltert. Einen anderen Ansatz verfolgt die sdw in ihrem Pilot-Programm „Vivo – Bildung von Grund auf!“ mit der Förderung von Klassenverbänden in Schulen in sozial benachteiligten Nachbarschaften (überdurchschnittlicher Anteil von Transferleistungsbeziehern). Dass die sdw seit 2013 einen anderen Weg der Förderung geht, ermöglicht es überhaupt erst, dass man die Wirkung eines Stiftungsprogramms analysieren kann. Ein zentrales Problem bei der Evaluation von Stiftungsprogrammen (egal ob systemisch oder individuell angelegt) ist das Fehlen einer sinnvoll vergleichbaren Kontrollgruppe, aus deren Entwicklung (ohne Förderung) man die Wirkung des jeweiligen Programms an einer Experimentalgruppe beobachten kann. Betrachtet man z. B. eine Schülerin oder einen Schüler, der während seiner Schulzeit durch ein beliebiges Programm gefördert wird, und ermittelt, dass sich dessen Kompetenzen entwickeln oder seine Ziele eine andere Richtung genommen haben, so kann man dies nicht sicher diesem Programm zuschreiben. Die Veränderung dieser Aspekte sind zum einen Effekte des Alters als auch Effekte der Bildung eines öffentlichen Schulsystems. Um die Wirkung des Programms zu untersuchen, braucht man also eine Schülerin oder einen Schüler mit vergleichbaren Fähigkeiten, der keine Förderung erhält. Dies ist vor allem bei Programmen, die begabte Schülerinnen und Schüler fördern, ein großes Problem. Man braucht ebenso begabte Schülerinnen und Schüler, die diese Förderung nicht bekommen, obwohl sie bestenfalls sogar durch den gleichen Auswahlprozess gelaufen sind. Dass dies bei der wissenschaftlichen Begleitung eines Stiftungsprogramms geschehen wäre, ist uns aber nicht bekannt.

¹ Problematisch ist dabei, dass z. T. mit einem sehr hohen Mitteleinsatz begabte und engagierte Kinder gefördert werden, die zunächst einmal durch Lehrkräfte empfohlen werden müssen. Dass diese Kinder dann einen erfolgreichen Weg einschlagen, kann oft schwerlich auf die spezifische Förderung zurückgeführt werden. Es ist gerade ein Problem von Kindern aus sozial benachteiligten Familien, dass ihre „Begabung“ auf Grund mangelnder habitueller Passung nicht durch Lehrer erkannt wird. Da dies aber Voraussetzung der Förderung ist, stellt sich die Frage, ob es die geförderten Schülerinnen und Schüler (trotz ihrer sozial schwierigeren Lage) nicht auch ohne die Förderung zum Studium geschafft hätten.

Eine geeignete Kontrollgruppe braucht es aus unserer Sicht nicht nur, um Alterseffekte, Sozialisationseffekte (z. B. familiäres Umfeld) und Institutioneneffekte (z. B. Schulbesuch) konstant zu halten. Man braucht die Kontrollgruppe auch, um „Placebo-Effekte“ auszuschließen. Denn sowohl die Personen, die ein Programm konzipiert haben und durchführen, als auch diejenigen, die an einem Programm teilnehmen, neigen dazu, dieses positiv wahrzunehmen. Den „Erlebnisbericht“, der in der deutschen Bildungslandschaft oftmals als „Evaluation“ bezeichnet wird, halten wir deshalb für ungeeignet, um die Wirkung eines Bildungsprogramms einzuschätzen. Man kann uns an dieser Stelle (vielleicht zu Recht) einen zu technischen Blick auf Bildung vorwerfen, der die Wirkung eines Bildungsprogramms anhand messbarer Bildungs-Outcomes bewertet. Allerdings halten wir die Bewertung der Wirksamkeit von Bildungsprogrammen anhand von Erfahrungsberichten wie Evaluationen ohne Kontrollgruppe als nicht zielführend. Die wissenschaftliche Begleitung des Programms „Vivo - Bildung von Grund auf!“ der sdw ermöglicht es, den beschriebenen Problemlagen zu begegnen. Vivo wurde von Klasse 1 bis 4 an sechs Schulen in Düsseldorf, Hamburg und Neumünster durchgeführt. Ziel war es, Kinder aus sozial benachteiligten Lagen während ihrer Grundschulzeit zu fördern. Wir Forschenden konnten gleichzeitig die geförderten Schülerinnen und Schüler, die sich in einer Klasse befanden, als auch die Schülerinnen und Schüler der Parallelklassen, die die Förderung nicht erhalten hatten, wissenschaftlich begleiten. Dieses Grunddesign ermöglichte sehr günstige Voraussetzungen, um methodisch annähernd sauber die Wirkung des Stiftungsprogramms auf die Schülerinnen und Schüler der geförderten Schulen zu untersuchen. Allerdings ist auch die vorliegende Untersuchung nicht frei von Limitationen, die wir in Abschnitt 3.3 eingehend beschreiben werden.

Der vorliegende Projektbericht ist so aufgebaut, dass wir in Abschnitt 2 auf den Ablauf und die Ziele des Programms eingehen und anschließend darlegen, wie wir die Evaluation angelegt haben (Abschnitt 3). In Abschnitt 4 gehen wir dann auf die Ergebnisse der wissenschaftlichen Begleitung ein. Dabei wollen wir neben der Darstellung der Ergebnisse aber auch allgemein die Frage stellen, was ein solches Bildungsprogramm überhaupt erreichen kann. Denn es wird häufig übersehen,

dass die Zeit, die die Kinder in diesen Programmen verbringen, nur ein Bruchteil der Zeit ist, die sie in der Schule oder der Familie verbringen. Eine realistische Einschätzung, was derartige Programme überhaupt zu leisten im Stande sind, hilft allen Seiten zu bewerten, ob ein solches Programm erfolgreich sein kann oder nicht.

2 Ziele, Beschreibung und Ablauf des Programms

Mit dem Projekt „Vivo – Bildung von Grund auf!“ will die Stiftung der Deutschen Wirtschaft (sdw) gGmbH einen Beitrag für mehr Bildungsgerechtigkeit leisten und die herkunftsbedingte Bildungsbenachteiligung von Kindern im Grundschulalter durch schulbegleitende Förderung verringern. Das Projekt soll die Bildungsorientierung von Grundschulkindern erhöhen und ihre Chancen auf gesellschaftliche Teilhabe und Verwirklichung ihrer individuellen Potenziale verbessern. Ziel ist, dass die Kinder selbstbestimmt einen Bildungsweg gehen, der ihren Anlagen und Fähigkeiten entspricht. und nicht von sozialen Faktoren gehemmt wird. Dafür bietet das Programm den Kindern zusätzliche Förderung während der gesamten Grundschulzeit und bereitet sie auf den Übergang an die weiterführende Schule vor. Zur Stabilisierung im neuen Lernumfeld der Sekundarstufe I begleitet das Programm die Kinder auch in der fünften Klasse. Die wissenschaftliche Begleitung endete jedoch am Ende von Klasse 4.

Um eine Stigmatisierung der Kinder zu vermeiden und die positive Wirkung des Klassenverbundes zu stärken und Synergieeffekte zu nutzen, findet die Förderung in den ersten vier Jahren im gesamten Klassenverband statt.

Die Zielgruppe des Vivo-Programms sind Kinder aus bildungsfernen und sozioökonomisch benachteiligten Elternhäusern. Um diese Zielgruppe zu erreichen, werden Grundschulen in sozial schwachen Stadtteilen ausgewählt. An diesen Schulen wiederum wurde im Pilotdurchlauf jeweils eine Klasse als „Vivo-Klasse“ gefördert. Einmal wöchentlich kamen zwei externe Trainerinnen bzw. Trainer für 90 Minuten in die Klasse und arbeiteten mit den Kindern. Besondere Formate wie Exkursionen oder Expertenbesuche ergänzten diese üblichen 90-

minütigen Fördereinheiten, die meist im Theater- oder Musikraum der Grundschule stattfanden, und boten den Kindern vielfältige (Kompetenz-) Erfahrungen über ihren bisherigen Alltagshorizont hinaus.

Diese von erfahrenen Trainerinnen und Trainern gestalteten Aktivitäten haben immer das Ziel, die Kinder in der Gruppe stark zu machen. Nach vier Jahren Förderung sollten die Kinder selbstbewusster sein und sich als selbstwirksam erleben, aber auch gelernt haben, mit anderen Kindern und Erwachsenen gut auszukommen. Ebenfalls sollten sie ihre Stärken und Talente kennen und eine Vorstellung von ihren Zielen und ihrer Zukunft bekommen. Wichtig war, dass die Kinder Bildung für sich selbst als positiv erachten und letztendlich als Chance begreifen, eine weiterführende Schule besuchen zu können, die ihren Potenzialen entspricht.

Ein nicht zu unterschätzendes Charakteristikum des Programmes ist es, dass Vivo kein Unterricht ist, der bewertet wird. So erleben gerade die Kinder, denen dem Unterricht zu folgen manchmal schwerfällt, ein offeneres Format, in dem sie auch andere Seiten von sich zeigen können. Die Kinder sind freier, sie können sich ausprobieren und haben ein Mitentscheidungsrecht. Wichtig ist außerdem, dass die Kinder sich positiv erleben; also nicht Defizite aufgezeigt bekommen, sondern vielmehr die jeweiligen Stärken und Interessen der Kinder gefördert werden.

Ablauf des Programmes

Um Kinder auf ihrem Bildungsweg zu stärken, setzte die sdw inhaltliche Themenschwerpunkte und formulierte Zielsetzungen auf der Kompetenzebene. Auf Basis dessen erhielten die Vivo-Trainerinnen und -Trainer halbjährliche Rahmenförderkonzepte, die sie in enger Zusammenarbeit mit der Lehrkraft, Erzieherin oder Erzieher sowie der Schulleitung auf die konkrete Situation in der jeweiligen Klasse anpassten und methodisch ausgestalteten.

Im ersten Förderjahr, das mit „Kultur erleben“ überschrieben ist, sollten die Kinder über theaterpädagogische und kreative Aktivitäten sich selbst besser kennenlernen, wichtige Gruppenerfahrungen machen und zahlreiche Sprachanlässe nutzen. Auch ein positives Selbstbild wurde hier bereits in den Blick genommen. „Forschen und Entdecken“ stand im folgenden Jahr auf dem

Programm: Durch einen partizipativen Ansatz wurden gemeinsam mit den Kindern Fragen entwickelt und Lösungswege gefunden, um damit Neugierde, Selbstvertrauen und Lernmotivation zu fördern. Dies sollte im dritten Jahr unter dem Motto „Lebenswelten erkunden“ auf einer anderen Ebene fortgesetzt werden, indem die Kinder andere Lebenswelten (fremde Kulturen o.ä.) und verschiedene Möglichkeiten der Freizeitgestaltung kennenlernten. Neben der Offenheit gegenüber Neuem ging es hier insbesondere darum, die Zielorientierung der Kinder zu fördern. Einen wichtigen Schwerpunkt bildete dabei die Auseinandersetzung mit verschiedenen Berufen und Berufsbildern. U.a. mit Exkursionen in nahegelegene Betriebe sollte die Bildungsaspiration der Kinder gefördert werden: Durch praktische Einblicke sollten sie frühzeitig erkennen, welchen Wert Bildung hat und dass die Dinge, die sie in der Schule lernen, in einen Beruf münden. So kamen sie auch in Berührung mit Arbeitsnormen und fragten sich, welche Kompetenzen für das Berufsleben notwendig sind. Darauf aufbauend wurde dann im vierten Jahr die dezidierte Vorbereitung des Überganges in den Fokus genommen. Selbstorganisation und Eigenständigkeit wurden zu ebenso wichtigen Bestandteilen wie Engagement und Anstrengungsbereitschaft. Unter dem Motto „Übergänge gestalten“ sollte das Bewusstsein für die eigene Stärke und Kraft geschärft werden. Auch hier benannte die sdw das Kennenlernen von Vorbildern als didaktischen Vorschlag – begleitet von erlebnis-(pädagogischen) Aktivitäten, die es den Kindern ermöglichten, sich selbst positiv und stark zu erleben.

Zu Beginn des Programms lag der Fokus darauf, die Bindung der Kinder zum Projekt zu stärken. Hierzu wurde unter anderem ein Projektmaskottchen (die Ameise „Vivo“) eingeführt. Ein weiterer zentraler Bestandteil des Programms waren regelmäßige Angebote zur Einbindung der Eltern. Sie hatten zum Ziel, deren Rolle als aktive Bildungspartner zu stärken. Dies geschah u.a. bei Elterncafés zum Halbjahr und Elternfesten am Ende des Schuljahres. Dort präsentierten die Kinder, was sie im vergangenen Jahr „mit Vivo“ erlebt und gelernt hatten. Dies bot einen niedrigschwelligen Anlass für die Eltern, sich am Schulleben ihrer Kinder zu beteiligen. Für die Vorbereitung des Übergangs an die weiterführende Schule

wurde ein besonderes Workshop-Format entwickelt, um die Eltern bei der passenden Schulwahl für ihr Kind zu unterstützen.

Neben den Eltern berücksichtigte das Programm auch die beteiligten Lehrkräfte, Erzieherinnen und Erzieher und Schulleitungen als wichtige Bildungspartner. Sie erhielten gemeinsam mit den Trainerinnen und Trainern speziell angepasste Fortbildungen, die in Kooperation mit dem TransferZentrum für Neurowissenschaften und Lernen (ZNL) angeboten wurden. So sollten Inhalte und Methoden nachhaltig in die Schulen getragen werden und zusätzlich positiv auf den Regelunterricht wirken. Damit zielte die sdw darauf ab, Schulen für die Chancenungleichheit am Übergang zu sensibilisieren und ihre Lehrkräfte dazu anzuregen, bildungsbenachteiligte Kinder unterrichtsbegleitend wirksam zu fördern. Auch die relativ freie Gestaltung der Vivo-Einheiten (im Gegensatz zum Regelunterricht) und die Zusammenarbeit mit Trainerinnen und Trainern, die pädagogische Expertise aus anderen Bereichen mitbrachten, sollte den Lehrkräften eine zusätzliche Perspektive auf ‚ihre‘ Kinder eröffnen. Denn das Vivo-Programm ermöglichte den Lehrkräften, ihre Schülerinnen und Schüler regelmäßig und über einen langen Zeitraum hinweg in anderen Lernkontexten kennenzulernen.

Mit dem Zusammenspiel all dieser Aktivitäten beabsichtigt die sdw auch sekundäre Herkunftseffekte am Übergang von der Grundschule auf die weiterführende Schule zu verringern. Wie eingangs erwähnt, war es Gesamtziel des Programmes „Vivo – Bildung von Grund auf!“, Bildungsungleichheit entgegenzuwirken und sozial benachteiligte Schülerinnen und Schüler auf ihrem Bildungsweg zu unterstützen. Dieses Ziel schlüsselte die sdw auf in die benannte Förderlogik und die ausgegebenen Rahmenförderkonzepte. Sie basierten auf der Grundannahme, dass folgende Dimensionen eine entscheidende Rolle beim Ausgleich von Startnachteilen spielen:

- Motivation und Zielorientierung steigern
- Soziales Zusammenleben verbessern
- Das Selbst stärken und Resilienz fördern
- Sprachliche Kompetenzen verbessern

- Realistische Einschätzung von Bildungszielen
- Schulleistungen verbessern

Gegenstand der vorliegenden Evaluation waren Veränderungen in diesen Dimensionen bis zum Ende der vierten Klasse. Im Fokus stand damit die Wirkung des Vivo-Programmes auf die Kinder und nicht die Veränderungen, die möglicherweise bei Lehrkräften und anderen Beteiligten eintraten. Zudem beruhte die wissenschaftliche Begleitung des Projekts auf den ersten vier Förderjahren.

3 Evaluationsdesign

3.1 Schulauswahl

Im Herbst 2013 wurden jeweils zwei Grundschulen an den drei Standorten Düsseldorf, Hamburg und Neumünster ausgewählt und in das Förderprogramm aufgenommen. Die Standorte wurden unter Berücksichtigung lokaler Gegebenheiten und in Kooperation mit den zuständigen Schulbehörden ausgewählt. In jeder Projektschule gab es eine Projektklasse (durch externe Trainerinnen und Trainer gefördert) und eine Kontrollklasse (vergleichbare Schülerinnen und Schüler, die nicht gefördert wurden). In einem Fall musste die Kontrollklasse an einer anderen, direkt angrenzenden Schule akquiriert werden, da die identifizierte Grundschule lediglich einzügig geführt wurde. Nach der ersten Erhebung konnte bestätigt werden, dass die Zielgruppe getroffen wurde.

3.2 Evaluationskriterien und Befragungsablauf

Tabelle 1: Projektdimensionen, deren Operationalisierung und Person, über die dies erfasst wurde

Projekt-dimensionen	Operationalisierung und Messung über...		
	Schülerinnen und Schüler	Lehrer	Eltern
Motivation und Zielorientierung steigern	Lernmotivation, Anstrengungsbereitschaft, Arbeitsvermeidung	Lernverhalten	
Soziales Zusammenleben	Klassenklima, Ausgrenzung, Empathie	Sozialverhalten	Kitabesuch, Gemeinsame Aktivitäten und Familientisch, Unterstützung durch andere, Wissen über Freunde
Das Selbst stärken, Resilienz, Persönlichkeit	Schulisches Selbstkonzept, Selbstwirksamkeitserwartung, allgemeines Selbstkonzept	Durchsetzungsvermögen, Persönlichkeitsmerkmale	Persönlichkeitsmerkmale
Sprachliche Kompetenzen	Kompetenz Lesen am Ende Klasse 3 ²	Note Deutsch	Familiensprache
Sonstige Schulleistungen	Kompetenzen in Mathematik am Ende Klasse 3 ³	Note Mathematik, Note Heimat- und Sachkunde	
Schulverlauf		Sitzenbleiben, Gymnasialübergang	
Bildungsziele	Schulaspirationen, Schulabschlussaspirationen		Schulaspiration
Unabhängige Variablen	Kognitive Fähigkeiten Anfang Klasse 2 ⁴ , Wortschatz ⁵	Fehltage, Stadt, Förderschüler, Geschlecht, Alter	Berufsprestige und Bildung der Eltern, Migrationsstatus

In Tabelle 1 haben wir festgehalten, über welche Konstrukte wir die einzelnen Projektdimensionen abbilden und ob wir diese Merkmale über Schülerinnen und

² Erfasst über die Ergebnisse der Vergleichsarbeiten (VERA) und „Kompetenz ermitteln“ KERMIT (in Hamburg) in der Domäne Lesen. Dabei stehen uns nur die Kompetenzstufen (1 bis 5) zur Verfügung.

³ Erfasst über die Ergebnisse der Vergleichsarbeiten (VERA) und KERMIT (in Hamburg) in den Domänen „Zahlen und Operationen“ und „Muster und Strukturen“. Dabei stehen uns nur die Kompetenzstufen (1 bis 5) zur Verfügung.

⁴ Erfasst über einzelne Subtests des Grundintelligenztest CFT 1-R (Weiß & Osterland, 2012).

⁵ Erfasst über ELFE 1-6 (Lenhard & Schneider, 2006).

Schüler, Lehrer oder Eltern erhoben haben. Wie die einzelnen Konstrukte operationalisiert wurden, haben wir im Anhang (Abschnitt 7.1) festgehalten. In Tabelle 2 haben wir überdies festgehalten, wann die einzelnen Konstrukte erfragt wurden. Wie ersichtlich wird, können wir im Längsschnitt überprüfen, inwieweit die Projektdimensionen von Vivo am Ende der 4. Klasse erreicht wurden. Die Haupterhebungen haben jeweils am Ende der einzelnen Schuljahre stattgefunden.⁶ Die Erhebungsinstrumente finden sich im Anhang in Abschnitt 7.2 wieder. Die Befragungen wurden von April 2014 bis September 2017 durchgeführt. Der detaillierte Erhebungsplan lässt sich aus Abbildung 1 entnehmen.

Tabelle 2: Überblick über den Zeitpunkt der Erhebung der eingesetzten Test- und Befragungsinstrumente

Konstrukt	W1 S	W1 L	TS	W2S	W2 L	W1 E	W3S	W3 L	W4 S	W4 L	W2 E
Lernmotivation	X			X			X		X		
Klassenklima	X			X			X		X		
Ausgrenzung	X			X			X		X		
Eigene Negativerfahrungen von Ausgrenzung							X		X		
Empathie							X		X		
Schulisches Selbstkonzept	X			X			X		X		
Selbstwirksamkeits- erwartung	X			X			X		X		
Allgemeines Selbstkonzept	X			X			X		X		
Anstrengungsbereit- schaft							X		X		
Arbeitsvermeidung							X		X		
Sozialverhalten											
Durchsetzungsver- mögen		X			X			X		X	
Lernverhalten		X			X			X		X	
Kompetenz Mathematik (Zahlen und Operationen)								X			
Kompetenz Mathematik								X			

⁶ Nicht festgehalten ist eine dritte Elternbefragung, die wir am Anfang von Klasse 4 durchgeführt haben. Hiermit wollten wir überprüfen, ob der Elternworkshop, der mit den Eltern der Projektklassen der Schülerinnen und Schüler durchgeführt wurde, das Übergangsverhalten verändert. Dadurch, dass der Rücklauf der Kontrollklasseneltern sehr niedrig war, haben wir diese Befragung nicht ausgewertet.

(Muster und Strukturen)											
Kompetenz Lesen							X				
(Eingeschätzte) Note Heimat- und Sachkunde		X			X		X		X		
(Eingeschätzte) Note Deutsch		X			X		X		X		
(Eingeschätzte) Note Mathematik		X			X		X		X		
Sitzenbleiben										X	
Schulaspirationen	X			X			X		X		
Schulabschlussaspirationen	X			X			X		X		
Gymnasialübergang										X	X
Unabhängige Variablen											
Alter		X									
Kognitive Fähigkeiten Anfang Klasse 2			X								
Wortschatz Ende Klasse 2			X								
Sprachstand		X									
Fehltage		X			X			X		X	
Stadt		X									
Förderschülerinnen und -schüler		X			X			X		X	
Geschlecht		X									
Alter		X									
Berufsprestige der Eltern						X					
Bildung der Eltern						X					
Migrationsstatus						X					

W= Welle, T=Testung, L=Lehrer, S=Schülerinnen und Schüler, E=Eltern,

In den Erhebungsinstrumenten für die Schülerinnen und Schüler haben wir als Antwort- und Bewertungsmöglichkeit eine 4-stufige Smiley-Skala vorgelegt.

(Kratzer, Cwielong 2013)

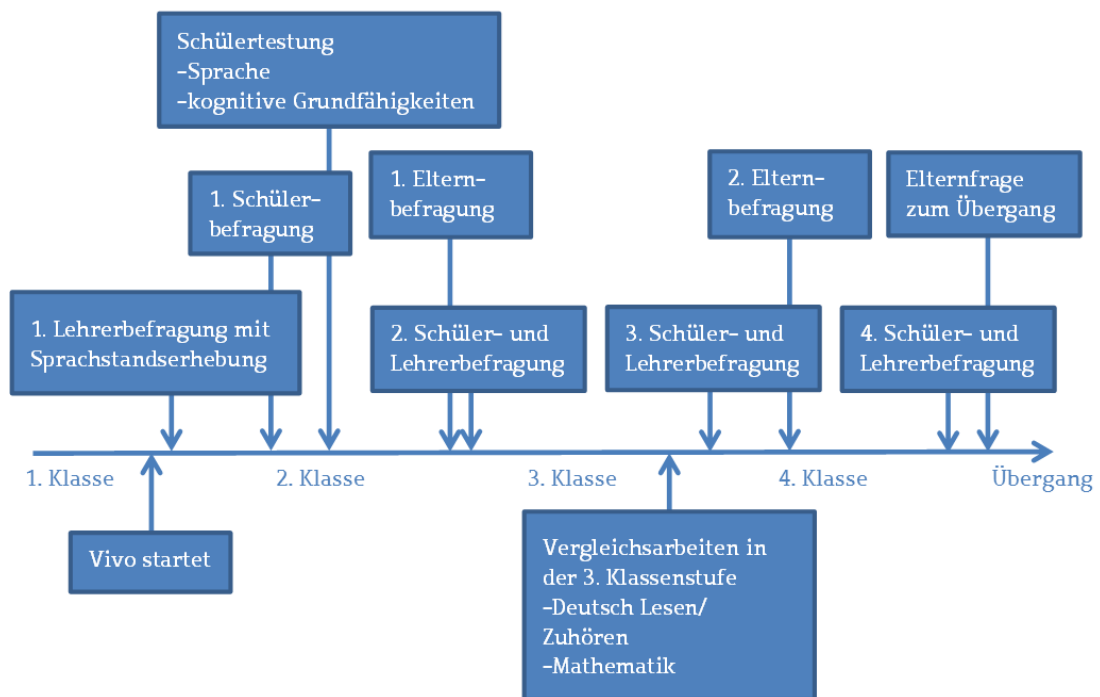
stimmt ganz genau = 😊, stimmt eher = 😊, stimmt eher nicht = ☹, stimmt gar nicht = ☹

Diese symbolische Ratingskala haben wir gewählt, um den Kindern die notwendige Abstraktion zu erleichtern (eigene Meinung in numerisches Relativ zu überführen) (Roebbers 2009).

Neben der quantitativen Erfassung der abgebildeten Konstrukte haben wir nach Ende des 4. Schuljahres eine Gruppendiskussion und einzelne Interviews mit den

beteiligten Akteuren (Schulleitern, Lehrern, Erziehern) durchgeführt. Auch wenn diese Interviews ergänzend dazu dienen, die Wirksamkeit von Vivo zu bewerten, so ergaben sich aus diesen Interviews sehr wichtige Aspekte, die uns einen anderen Blick auf das Programm ermöglichten.

Abbildung 1: Erhebungsplan



3.3 Methodische Herausforderungen und Limitationen

Wie eingangs angemerkt, bietet uns das Experimental- und Kontrollgruppendesign eine wichtige Grundvoraussetzung, um Vivo überhaupt in seiner Wirkung bewerten zu können. Allerdings ist die Evaluation auch nicht frei von methodischen Herausforderungen und Limitationen, auf die wir hier eingehen wollen.

1 Kleine Fallzahlen

Dadurch, dass das Programm nur an sechs Schulen durchgeführt wurde, ist die Anzahl der befragten Schülerinnen und Schüler eher gering. An der ersten Befragung am Ende der Klasse 1 nahmen 216 Schülerinnen und Schüler teil. Am Ende des Programms gab es nur noch 164 Fälle aus der Experimental- und Kontrollgruppe, die auch schon in Welle 1 dabei waren. Könnte man nur Fälle verwenden, zu denen alle notwendigen Informationen in allen Wellen vorliegen, würde unsere Fallzahl deutlich unter 100 liegen. Deshalb haben wir alle fehlenden Werte multipel mit dem Stata-ado „ice“ imputiert (30 Imputationen) (Royston 2004). So können wir zumindest alle 164 Fälle in unsere Analysen einbeziehen. Abweichende Fallzahlen verwenden wir nur bei der Analyse zur Frage, welche Kinder eine Klassenstufe wiederholen müssen⁷, und bei der Analyse der Bildungsaspirationen, da hier auch fehlende Werte eine Bedeutung haben. Mit der kleinen Fallzahl gehen drei weitere Probleme einher. Erstens erübrigen sich methodisch sinnvolle Verfahren wie beispielsweise Mehrebenenanalysen, in denen statistisch daraufhin korrigiert würde, dass die Schülerinnen und Schüler aus verschiedenen Klassen kommen. Hier ist die Anzahl der Klassen zu gering. Wir werden in all unseren Berechnungen einfache Regressionsanalysen durchführen, in denen die Schülerinnen und Schüler unsere Untersuchungsebene sind, und dabei auf eine Reihe von Variablen kontrollieren⁸. Zweitens ist es schwierig, bei derart kleinen Fallzahlen Ergebnisse zu erhalten, die sich eindeutig (statistisch signifikant) zwischen Kontroll- und Experimentalgruppe unterscheiden. Somit kann auch das Maß statistischer Signifikanz nicht das abschließende Kriterium sein, um festzustellen, wie das Programm gewirkt hat. Deshalb werden wir uns in unseren Analysen auf die Darstellung der Zusammenhangsstärke nach dem Cohens D beziehen. Hierbei werden die Mittelwertunterschiede (die sich bei uns aus der Regressionsanalyse ergeben) von Kontrollgruppe und Experimentalgruppe durch die Standardabweichung beider

⁷ Hier befinden sich auch alle Schülerinnen und Schüler in den Analysen, von denen wir wissen, dass sie sitzengeblieben sind.

⁸ Im Allgemeinen sind dies das berufliche Prestige und die Bildung der Eltern, Fehltag, Migrationsstatus, Stadt, Alter, Förderschülerinnen und -schüler, kognitive Fähigkeiten, Leseverständnis und Geschlecht.

Gruppen dividiert. Dabei wird ein Cohens D von 0,2 gemeinhin als schwacher Zusammenhang aufgefasst und ein Cohens D von 0,5 als mittlerer Zusammenhang. Drittens müssen sich die Klassen gerade bei so kleinen Fallzahlen in ihren sonstigen Randbedingungen sehr stark ähneln. Jede Ungleichheit zwischen Experimental- und Kontrollgruppe fällt bei kleinen Fallzahlen stärker ins Gewicht als bei punktuellen Unterschieden in größeren Untersuchungspopulationen.

2 Ähnlichkeit von Experimental- und Kontrollgruppe

Davon, dass diese Vergleichbarkeit von Experimental- und Kontrollgruppe gewährleistet ist, als auch von einer gewissen Zufälligkeit der Klassenauswahl, gingen wir zu Beginn unserer wissenschaftlichen Begleitung aus. Allerdings zeigte sich im Verlauf der Evaluation, dass diese Vergleichbarkeit nicht vollständig gewährleistet ist. Die ursprüngliche Grundvoraussetzung für Experimentalklassen, war im gebundenen Ganztags betrieben zu werden, d.h. dass alle Kinder die Klasse ganztags besuchen. Bei der Programm-Umsetzung zeigte sich jedoch, dass diese Bedingung nicht vollständig eingehalten werden kann. Zu Beginn der Evaluation ergab sich ein höhere Teilnahme der Vivo Schüler am Ganztagsprogramm als bei der Kontrollgruppe. Ansonsten zeigten sich kaum Gruppenunterschiede von Vivo Schülern und Kontrollgruppe. Obwohl man denken könnte, dass die Wahl einer Klasse im gebundenen und nicht gebundenen Ganztags z. B. von der Erwerbstätigkeit oder Bildung der Eltern abhängen könnte, konnten wir diesbezüglich keine Unterschiede zwischen Kontrollklasse und Experimentalklasse feststellen. Wenn man davon ausgeht, dass sich Schülerinnen und Schüler in Ganztagsbeschulung besser entwickeln (woran die ersten Ergebnisse der sogenannten StEG-Studie erhebliche Zweifel säen, StEG Konsortium 2016), könnten positive Ergebnisse des Programms auch auf die Ganztagsbeschulung zurückgeführt werden.

Im späteren Verlauf der wissenschaftlichen Begleitung wurde uns darüber hinaus bewusst, dass es noch andere extern angebotene Programme in einigen Schulen gab, die sich auf alle Schülerinnen und Schüler der Schule bezogen. Wenn diese ähnliche Förderziele wie Vivo haben, könnte der Einfluss von Vivo auf die Schülerinnen und Schüler in den Vivo-Klassen unterschätzt werden.

Ferner ist in diesem Kontext darauf hinzuweisen, dass es unbeobachtete Störvariablen gibt (Lehrerwechsel, unterschiedliche Ausstattung etwa durch Lehrkräfte und Erzieher, die sich um Kinder mit Förderbedarf kümmern), die unsere abhängigen Variablen für eine ganze Klasse beeinflussen können. Kontrollieren kann man aber diese wegen der kleinen Fallzahl nicht. Bei großen Fallzahlen würden sich solche Störvariablen darüber hinaus „ausgleichen“. Insgesamt ist es auch für zukünftige Evaluationen anderer Programme extrem wichtig, solche Störvariablen im Blick zu haben.

3 Konsistenz und Einheitlichkeit des Programms

Vor allem in wissenschaftlichen Kausalanalysen wird darauf geachtet, dass ein Programm oder ein Treatment alle Personen in der Experimentalklasse erreicht und alle in der Kontrollgruppe nicht. Dabei soll im Weiteren ein einheitliches Treatment verwendet werden, dass für alle Schülerinnen und Schüler der Experimentalklasse gleich ist. Die erste Voraussetzung kann in unserem Fall als gewährleistet aufgefasst werden, die zweite hingegen nicht in vollem Umfang. Wir haben es bei dem Vivo-Programm mit einem Pilotprojekt zu tun, das zwar feste Rahmendefinitionen bietet, aber in der Ausgestaltung der Einzelmaßnahmen Varianzen aufweist. Aufgrund des Pilotcharakters konnten keine standardisierten pädagogischen Einzelmaßnahmen in den Blick genommen werden, da die Erarbeitung solcher Standards das Ziel eines Pilotprojektes bildet und diese naturgemäß erst am Ende eines solchen vorliegen können. Es gibt eine Vielzahl von Elementen innerhalb des Programms, deren Wirkung innerhalb des Evaluationsdesigns nicht separat untersucht werden kann. Dies unterscheidet die vorliegende Evaluation auch von Evaluationen, die an wissenschaftlich ausgerichteten Bedürfnissen ausgerichtet sind und ein kleines begrenztes Treatment, wie z. B. eine Informationsveranstaltung, in seiner Wirkung untersuchen (z. B. Ehlert et al. 2017). Dass diese Treatments auch nur einen begrenzten Einfluss haben (können) und nicht dem Anspruch der „Praktiker“ entsprechen, liegt an den unterschiedlichen Prämissen beider Subsysteme. Dementsprechend kann eine Evaluation wie die vorliegende auch nicht die

Wirkung einzelner Programmteile nachweisen und auch keine Vorschläge machen, inwieweit man einzelne Programmelemente verändern müsste.

Des Weiteren war die Ausgestaltung des gesamten Programms von Klasse 1 bis Ende 4 (später bis Ende 5) zu Anfang noch nicht vollständig ausgearbeitet. Dies ist aus unserer Sicht nicht tragisch. Denn im Großen und Ganzen haben sich erstens die Ziele des Programms nicht verändert und zweitens konnten wir ohnehin nicht einzelne Programmelemente in ihrer Wirkung untersuchen.

4 Programm oder Aufmerksamkeit

Eine eher allgemeine Limitation derartiger Evaluationsuntersuchungen ist die Frage, ob man die potenzielle Wirkung des jeweiligen Programms in seiner Programmatik untersucht oder ob die etwaige Wirkung nicht einfach darauf zurückgeführt werden kann, dass den Schülerinnen und Schülern im Programm Aufmerksamkeit und Anregung zuteil wird. Auch wenn man herausstellt, dass sich eine Experimentalgruppe besser entwickelt hat als eine Kontrollgruppe, stellt sich die Frage, ob dies nicht auch der Fall gewesen wäre, wenn man sich einmal in der Woche z. B. zum Sport getroffen hätte. Um dieses Untersuchungsproblem zu umgehen, hätte es einer dritten Gruppe bedurft, der ein zeitlich und personell gleiches, aber unprogrammatisches Treatment zuteilwird. Dies war im Rahmen des vorliegenden Projekts nicht möglich.

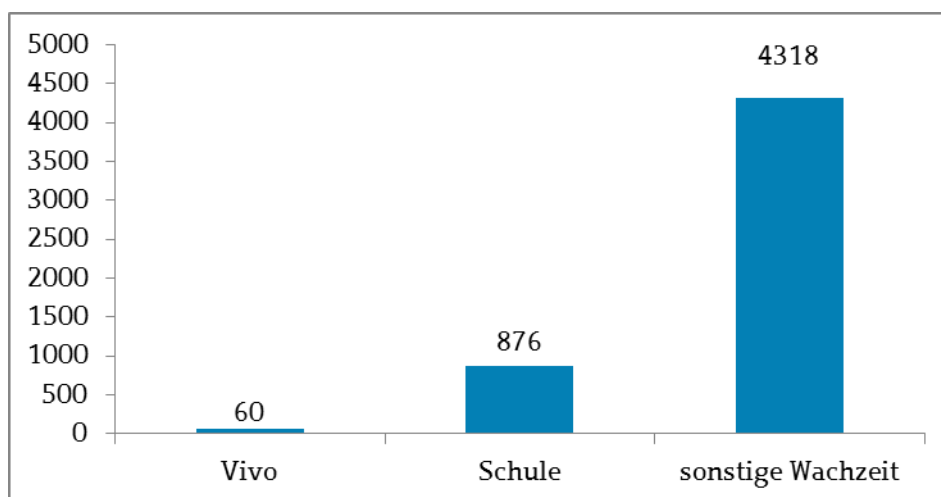
4 Ergebnisse

4.1 Erwartungshaltung

Bevor wir auf die Ergebnisse der Evaluation eingehen, wollen wir zunächst die Frage stellen, was externe Programme im schulischen Kontext überhaupt leisten können. Diesbezüglich ist die Überlegung der Bildungssoziologen Downey und Condon (2016) besonders interessant. In ihrer Studie weisen sie eindrucksvoll darauf hin, dass das, was wir heute von Schule erwarten, um soziale Ungleichheiten im Bildungssystem zu nivellieren, eine ideologische Überladung für

Schulen ist. Mit einem Verweis auf eine andere Studie stellen sie heraus, dass der durchschnittliche 18-jährige Amerikaner gerade einmal 13 % seiner Wachzeit in einer Schule verbracht hat. Dementsprechend könnte die Rolle, die Schule bei der Generierung oder Nivellierung sozialer Ungleichheiten einnimmt, bisher überzeichnet worden sein. Übertragen wir diesen Gedanken auf das Vivo-Programm. Bei einer Dauer von 1,5 Vivo-Kontaktstunden pro Woche und bei ungefähr 40 Schulwochen pro Jahr ergeben sich pro Jahr 60 volle Vivo-Kontaktstunden (siehe Abbildung 2). Dies entspricht im Verhältnis zur durchschnittlichen Schulzeit von knapp 900 Stunden nicht einmal 7 % und im Verhältnis zur gesamten Wachzeit (Schule und sonstige Zeit) eines Kindes in etwa einem Prozent.

Abbildung 2: Wachstunden pro Jahr in verschiedenen Kontexten



Eigene Berechnungen auf Basis von StBa (2015) anhand der 10- bis 17-jährigen.

Wir wollen damit nicht sagen, dass ein Programm wie Vivo nichts leisten kann. Wir wollen damit nur verdeutlichen, dass der Raum, den solche Programme im gesamten (Schul-)Leben eines Kindes einnehmen, relativ beschränkt ist.

Dementsprechend muss man aber auch mit den Erwartungen umgehen, die man an ein solches Programm richtet. Dabei ist Vivo aufgrund seiner Anlage schon als langfristiges Projekt zu bezeichnen.

Auch im Hinblick auf die angeführten methodischen Herausforderungen und Limitationen im vorhergehenden Abschnitt ist ein überhaupt gemessener Einfluss

des Programms auf einzelne Ziele schon als Erfolg zu bewerten. Dass das Programm alle seine Ziele erreicht, ist eher nicht zu erwarten. Aus unserer Sicht war es rückblickend ein womöglich unrealistisches Ziel, die sprachlichen Kompetenzen der Vivo-Kinder verbessern zu wollen. Pädagogisch geschulte Lehrkräfte unterrichten die Kinder vier Jahre lang in Deutsch; kann dann ein eher allgemein ausgerichtetes Programm zusätzliche Effekte bei der Schulung sprachlicher Kompetenzen erzielen?

4.2 Entwicklung der Schülerinnen und Schüler von Klasse 1 bis 4

Die quantitativen Ergebnisse der Evaluation stellen wir im Folgenden anhand der Projektdimensionen aus Tabelle 1 dar. Dabei werden wir in den Abbildungen die durchschnittlichen Werte für die Vivo-Kinder und die Kontrollgruppe in den jeweiligen Klassenstufen festhalten. Die Werte entstammen dabei linearen Regressionsanalysen, in denen wir folgende Variablen kontrolliert haben: berufliches Prestige und die Bildung der Eltern, Fehltage, Migrationsstatus, Stadt, Alter, Förderschülerinnen und -schüler, kognitive Fähigkeiten (bei Start des Programms), Wortschatz (bei Start des Programms) und Geschlecht. Zudem haben wir bei allen Analysen auf den jeweiligen Ausgangswert (meist aus Klasse 1) kontrolliert. In den jeweiligen Abbildungen haben wir überdies festgehalten, wie stark sich die Experimental- und Kontrollklassen als Cohens D in den jeweiligen Klassenstufen unterscheiden. Ist das Cohens D dabei fett gedruckt, liegt zudem ein statistisch signifikanter Unterschied auf einem 90-Prozent-Signifikanzniveau vor.

4.2.1 Motivation und Zielorientierung stärken (Schülerinnen- und Schülersicht)

Die erste Förderdimension des Programms war es, die Motivation und Zielorientierung der Schülerinnen und Schüler zu stärken. Dies haben wir über die Lernmotivation, die Anstrengungsbereitschaft und die Arbeitsvermeidung aus

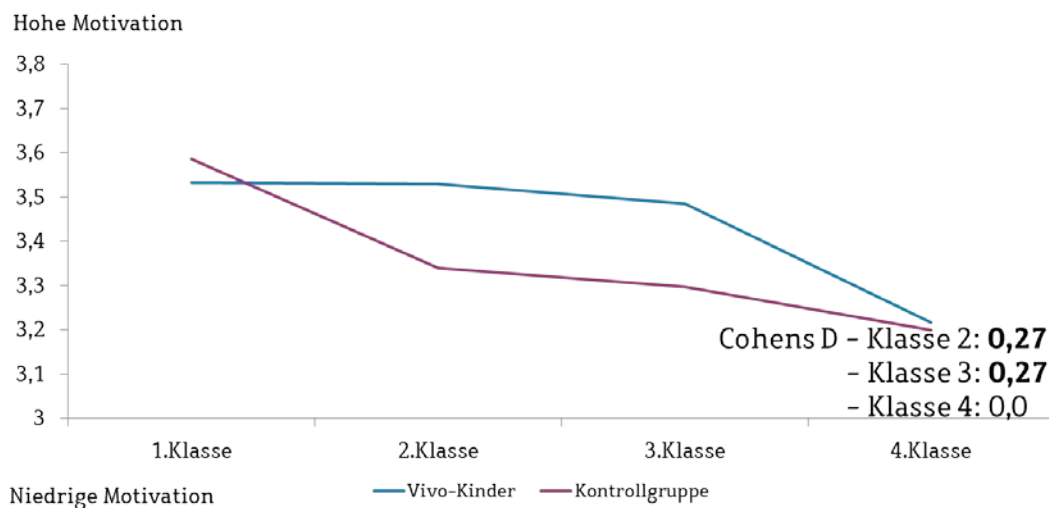
Schülerinnen- und Schülersicht und das Lernverhalten der Schülerinnen und Schüler aus Lehrersicht abgebildet. Hierbei zeigt sich für die Lernmotivation (Abbildung 3), dass diese bei der Kontrollgruppe von Klasse 1 zu 2 zunächst abfällt und bei den Vivo-Kindern stabil auf dem Niveau der 1. Klasse bleibt. Dadurch zeigt sich auch insgesamt ein statistisch eindeutiger Unterschied zu Gunsten der Vivo-Kinder. Dieser Unterschied bleibt auf dem gleichen Niveau in Klasse 3 bestehen. In Klasse 4 kommt es allerdings auch bei den Vivo-Kindern zu einem Einbruch der Lernmotivation, wodurch kein Unterschied zwischen Vivo-Kindern und der Kontrollgruppe mehr festgestellt werden kann.

Auch wenn sich zum Messzeitpunkt am Ende der Grundschulzeit kein Effekt des Programms auf die Lernmotivation der Vivo-Kinder feststellen lässt, kann man dennoch zwei wichtige Aspekte ableiten. Erstens zeigt die Betrachtung der Lernmotivation auf, wie wichtig ein Experimental- und Kontrollgruppendesign ist. Bei einem Indikator wie der Lernmotivation, der im Allgemeinen im Schulverlauf sinkt (Fend 1997, Prenzel 2009), würde die alleinige Betrachtung einer Experimentalgruppe verhüllen, dass das Programm einen positiven Einfluss auf einen solchen Indikator hat. Erst der Vergleich mit einer Kontrollgruppe zeigt in diesem Fall auf, dass das Programm davor schützt, dass es zu einem Rückgang der Motivation kommt. Umgekehrt kann auch der Fall auftreten, dass man ohne Kontrollgruppe die Wirkung eines Programms überschätzt, wenn man einen Indikator hat, der im Lebens- oder Schulverlauf steigt (z. B. der Indikator Kompetenzen).

Ein zweiter wichtiger Aspekt bei dieser Betrachtung ist, dass das Programm zunächst einen positiven Einfluss auf die Lernmotivation nimmt, dieser Einfluss aber langfristig (Klasse 4) schwindet. Hierbei könnte man aufgrund der kleinen Fallzahl zwar auch von einem „statistischen Ausreißer“ sprechen. Nimmt man die Ergebnisse allerdings ernst, dann könnte die Interpretation auch lauten, dass das Programm kurzfristig erfolgreich ist und die Kinder eine gewisse Zeit davor schützt, ihre Lernmotivation zu verlieren. Dieser Effekt verläuft sich jedoch im Verlauf der Grundschulzeit. Positiv gesehen, bietet das Programm eine gewisse Unterstützung beim Eintritt in die Schule, zeigt am Ende der 4. Klasse aber keine messbare Wirkung im Hinblick auf die Lernmotivation. Wie dies nun abschließend

zu bewerten ist, liegt ein Stück weit im Auge des Betrachters. Aus einem Null-Effekt in Klasse 4 keinerlei Einfluss auf die Lernmotivation abzuleiten, erscheint aus unserer Sicht jedenfalls nicht richtig.

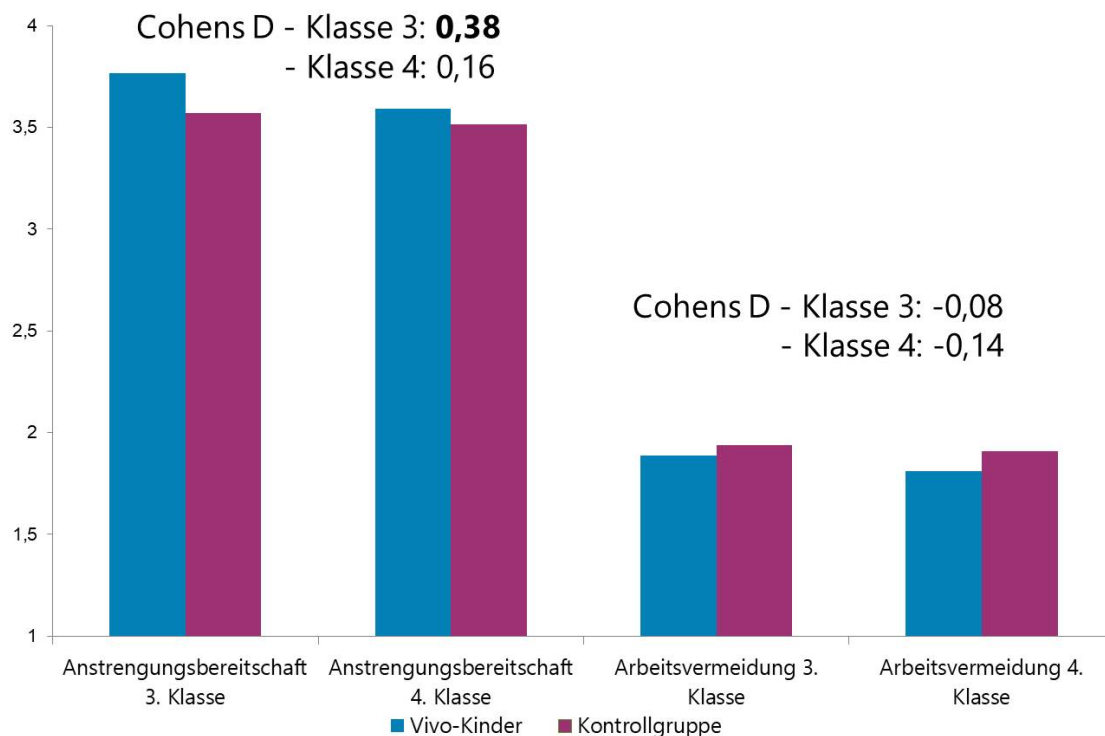
Abbildung 3: Entwicklung der Lernmotivation (Schülerinnen- und Schülersicht)



Auch bei der Anstrengungsbereitschaft (Abbildung 4), die wir erstmals in Klasse 3 gemessen haben, sehen wir analog zur Lernmotivation eine deutlich höhere Anstrengungsbereitschaft bei den Vivo-Kindern im Vergleich zu den Kontrollklassenkindern. Dieser Unterschied sinkt aber zur 4. Klasse und liegt unter der von uns formulierten Grenze eines Cohens D von >0.2 bei Cohens D 0,16. Die Arbeitsvermeidungsskala, reziprok zur Anstrengungsbereitschaft, unterscheidet sich weder in Klasse 3 noch 4 berichtenswert zwischen den beiden Gruppen. Allerdings gehen die Werte der Arbeitsvermeidung bei den Vivo-Kindern tendenziell zurück und der Unterschied bewegt sich auf die Grenze von Cohens D >0.2 zu.⁹

⁹ In diesem Fall bedeuten die tendenziell negativen Werte auch einen tendenziell positiven Effekt für die Vivo Kinder.

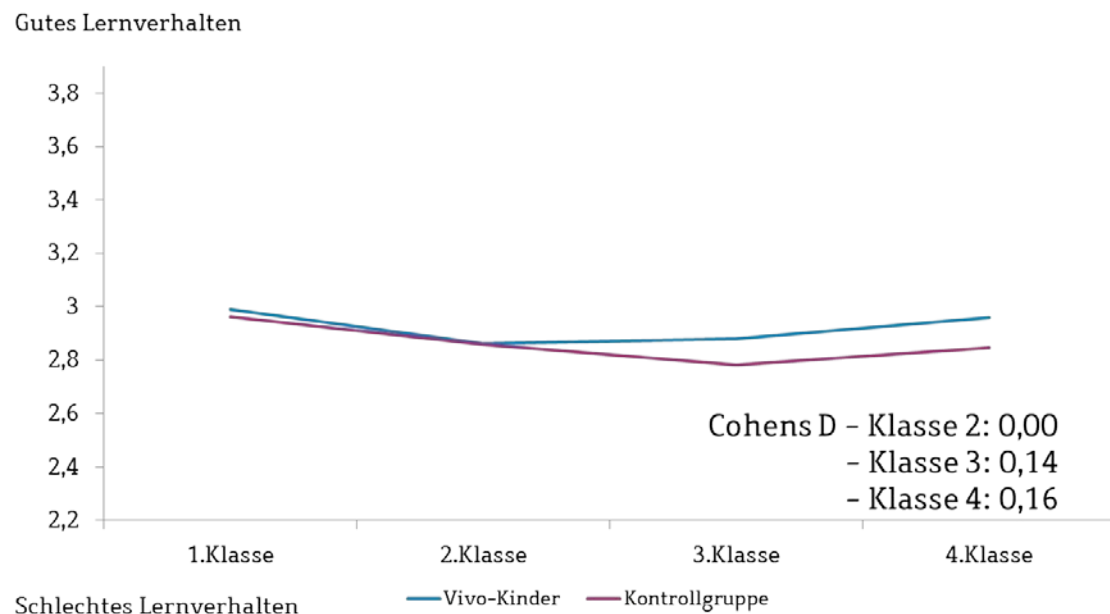
Abbildung 4: Entwicklung von Anstrengungsbereitschaft und Arbeitsvermeidung (Schülerinnen- und Schülersicht)



Auch beim Lernverhalten aus Lehrersicht zeigt sich am Ende von Klasse 4 keine statistisch bedeutende Abweichung des Lernverhaltens (Abbildung 5). Allerdings ist anders als bei der Lernmotivation auch kein kurzfristiger Anstieg zu beobachten, sondern eine tendenziell positivere Entwicklung der Vivo-Kinder bis Klasse 4.

Anhand dieser Indikatoren kann man vorsichtig formulieren, dass die Dimension, die Motivation und Zielorientierung zu stärken, kurzfristig geglückt ist: Die Vivo-Kinder weisen bis zum Ende der 3. Klasse höhere Werte bei Lernmotivation und Anstrengungsbereitschaft auf. Dies zeigt sich am Ende der 4. Klasse jedoch nicht mehr. Die beiden anderen Indikatoren, Arbeitsvermeidung und Lernverhalten, weisen in eine für das Programm positive Richtung, befinden sich aber in der 4. Klasse auf einem statistisch zu unsicheren Niveau (Cohens $D < 0,2$), als dass man sagen könnte, hier würde ein eindeutiger Zusammenhang zwischen dem Programm und der Motivation bzw. Zielorientierung vorliegen.

Abbildung 5: Entwicklung des Lernverhaltens (Lehrersicht)



4.2.2 Sozialkompetenzen

Die zweite Förderdimension des Programms war es, die Sozialkompetenzen zu verbessern. Dies haben wir über das Klassenklima aus Schülerinnen- und Schülersicht und über das Sozialverhalten aus Lehrersicht abbilden wollen. Bei den Items zum Klassenklima (siehe Anhang) zeigten sich allerdings bei der Analyse zwei Faktoren, die zum einen das Klassenklima beschreiben und zum anderen aufzeigen, inwieweit Kinder in einer Klasse ausgegrenzt werden.

Für das Klassenklima zeigt sich ein kurzfristiger Effekt des Programms in Klasse 2 (Abbildung 6), der allerdings ab Klasse 3 verschwindet. Überraschenderweise sind die Vivo-Kinder in der 2. Klasse der Meinung – obwohl in den Vivo-Klassen offensichtlich ein besseres Klassenklima herrscht –, dass in ihren Klassen Schülerinnen und Schüler häufiger ausgegrenzt werden (Abbildung 7). Dieser Unterschied bleibt dabei bis zum Ende der 4. Klasse bestehen, auf dann niedrigerem Niveau.

Abbildung 6: Entwicklung des Klassenklimas (Schülerinnen- und Schülersicht)

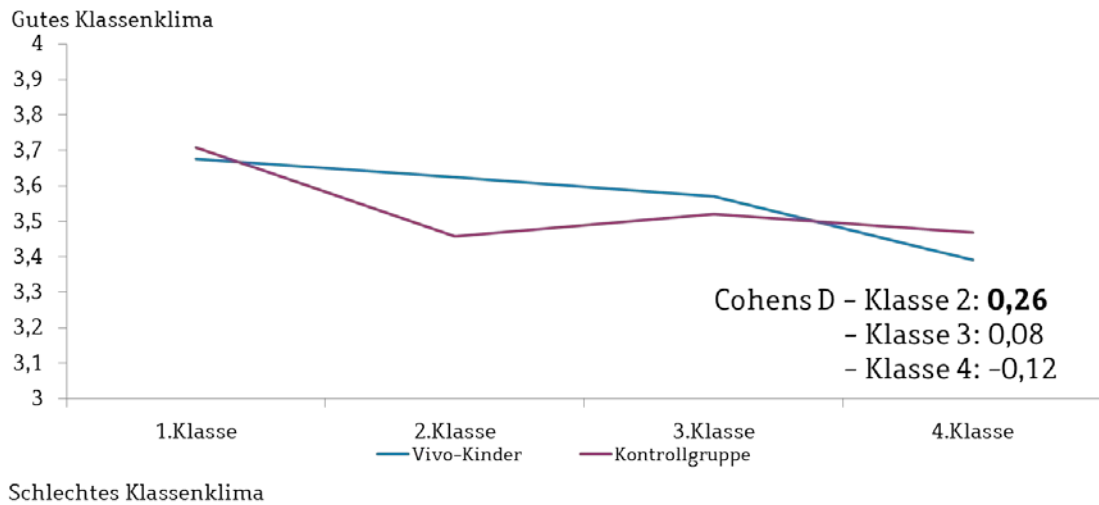
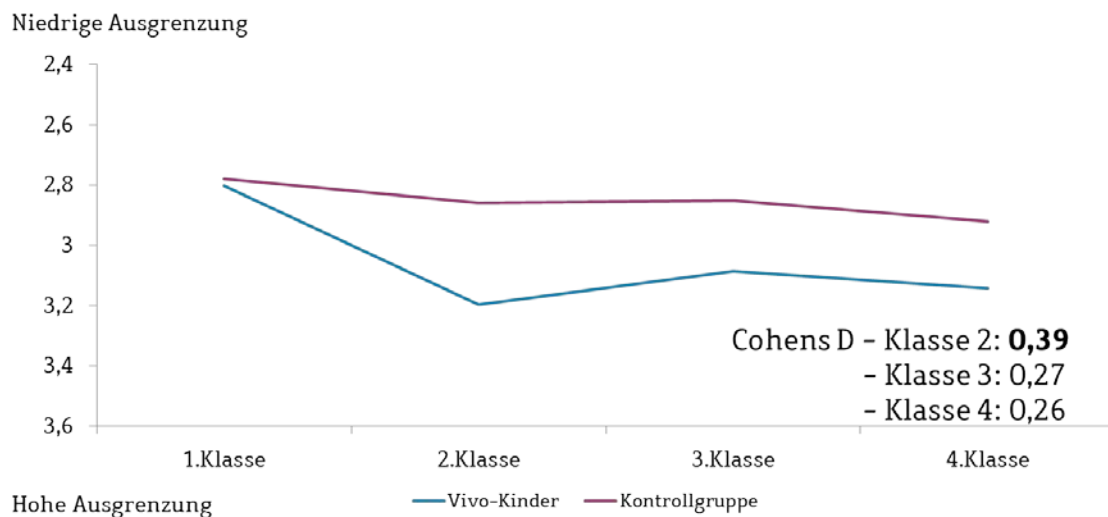


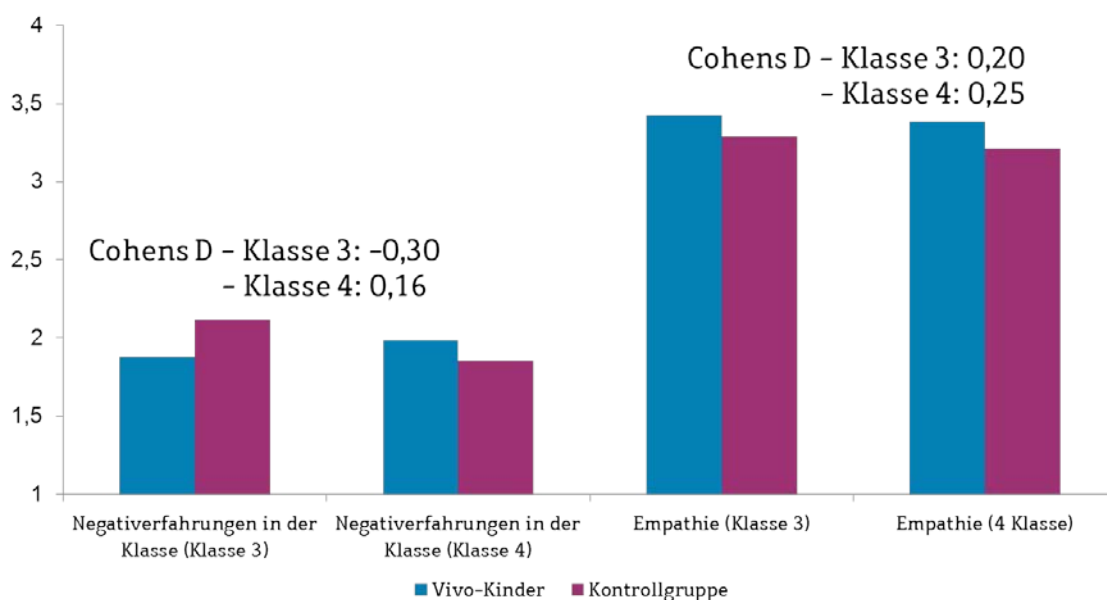
Abbildung 7: Entwicklung der Skala Ausgrenzung (Schülerinnen- und Schülersicht)



Dieses Ergebnis konnten wir uns zunächst überhaupt nicht erklären. Wie konnte das Ausmaß von Ausgrenzung höher sein, wenn gleichzeitig das Klassenklima besser war? Entspricht dies tatsächlich den Tatsachen oder nehmen die Schülerinnen und Schüler dies nur so wahr? Um dies zu überprüfen, haben wir ab Klasse 3 zum einen gefragt, ob die Schülerinnen und Schüler selber ausgegrenzt werden und dies nicht nur bei anderen wahrnehmen. Zum anderen haben wir eine Empathie-Skala eingesetzt. Vielleicht sind die Vivo-Schülerinnen und Schüler einfach nur sensibler, eine Ausgrenzung gegen andere wahrzunehmen, ohne dass

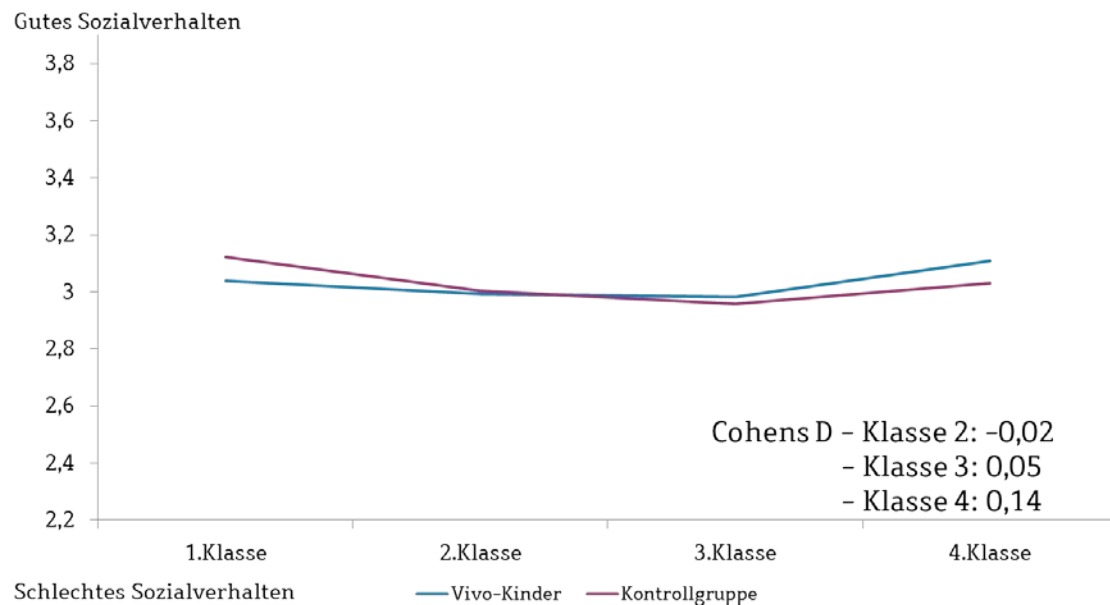
diese tatsächlich erfolgt? In Abbildung 8 zeigt sich dann tatsächlich, dass die Vivo-Kinder sogar seltener von der eigenen negativen Ausgrenzungserfahrung betroffen sind, dies aber für die eigene Klasse häufiger berichten. Allerdings sind diese Gruppenunterschiede am Ende der 4. Klasse nicht mehr statistisch bedeutsam. Ein Grund, warum es zu dieser zunächst widersprüchlich erscheinenden Wahrnehmung kommt, könnte tatsächlich an einer ausgeprägteren Sensibilität für die Nöte anderer liegen. Denn sowohl in Klasse 3 als auch in Klasse 4 haben die Vivo-Kinder höhere Empathie-Werte (Abbildung 8).

Abbildung 8: Entwicklung der Skalen selbst erlebte Ausgrenzung und Empathie (Schülerin- und Schülersicht)



Entgegen dem sehr vielschichtigen Blick aus Schülerinnen- und Schülersicht zeigen sich aus Lehrersicht nur kleine Unterschiede im eingeschätzten Sozialverhalten. Tendenziell wird das Sozialverhalten der Vivo-Kinder am Ende der Klasse 4 zwar positiver eingeschätzt, signifikant ist dieser Unterschied jedoch nicht.

Abbildung 9: Entwicklung des Sozialverhaltens (Lehrersicht)



Insgesamt können wir auch für die Dimension „Verbesserung des Sozialverhaltens“ nicht eindeutig belegen, dass dieses erreicht wurde. Zwar hat sich das Klassenklima in den Vivo-Klassen in Klasse 3 und 4 besser entwickelt als in der Kontrollgruppe. Die Vivo-Kinder weisen auch noch in Klasse 4 höhere Empathie-Werte auf, weil sie wohl für die Problematik der Ausgrenzung sensibilisiert sind. Auf der anderen Seite ergibt sich am Ende der 4. Klasse statistisch aber kein eindeutig positiver Einfluss auf das Klassenklima und das Sozialverhalten.

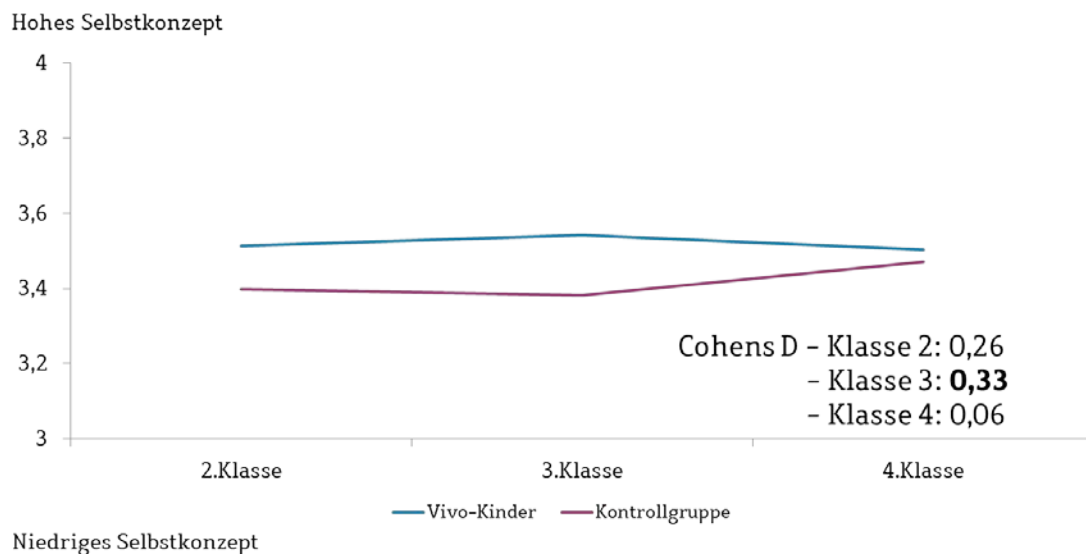
4.2.3 Das Selbst stärken

Die Stärkung des eigenen Selbst, wurde über das allgemeine oder soziale Selbstkonzept, das schulische Selbstkonzept und die Selbstwirksamkeitserwartung aus Schülerinnen- und Schülersicht und über das Durchsetzungsvermögen aus Lehrersicht gemessen.

Bezüglich des allgemeinen Selbstkonzepts zeigt sich in Abbildung 10, dass dieses in der 2. Klasse bei den Vivo-Kindern ausgereifter ist und der Unterschied zu den Kontrollgruppenkindern der Klasse 3 tendenziell sogar größer wird. Allerdings

verschwindet dieser Unterschied in Klasse 4. Der Grund für den Rückgang dieses Unterschieds ist hauptsächlich darauf zurückzuführen, dass sich die Entwicklung des allgemeinen Selbstkonzepts der Kontrollgruppenkinder dem der Vivo-Kinder zur 4. Klasse annähert.

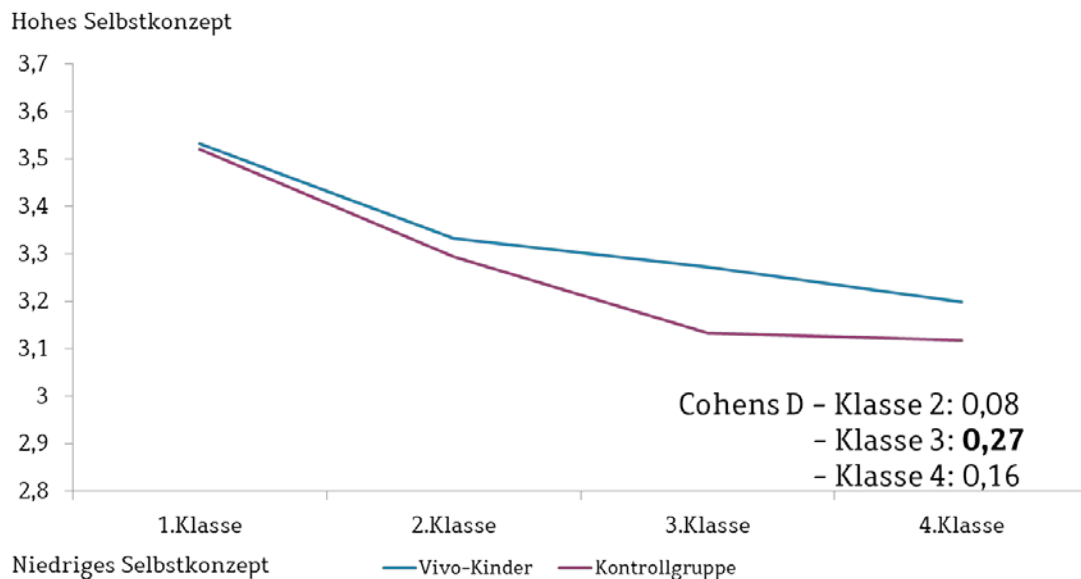
Abbildung 10: Entwicklung des allgemeinen Selbstkonzepts (Schülerinnen- und Schülersicht)¹⁰



Ein in der Quintessenz ähnliches Bild zeigt sich beim schulischen Selbstkonzept (Abbildung 11). Das schulische Selbstkonzept verringert sich zunächst in beiden Gruppen zur Klasse 2. Der weitere Rückgang in Klasse 3 ist jedoch nur noch bei den Kontrollklassen festzustellen, wodurch sich in der 3. Klasse ein statistisch bedeutender Unterschied zu Gunsten der Vivo-Kinder herauskristallisiert. Allerdings gehen die Unterschiede in Klasse 4 zwischen beiden Gruppen wieder zurück und sind statistisch nicht mehr bedeutsam.

¹⁰ Für das allgemeine Selbstkonzept haben wir auch in Klasse 1 ein Instrument eingesetzt. Die verwendete Skala entspricht jedoch nicht der hier präsentierten Skala. Um den Eindruck einer starken Niveauverschiebung von Klasse 1 zu 2 zu vermeiden, haben wir die Werte für Klasse 1 hier nicht dargestellt.

Abbildung 11: Entwicklung des schulischen Selbstkonzepts (Schülerinnen- und Schülersicht)



Auch bei der Selbstwirksamkeitserwartung ist (Abbildung 12) zunächst eine kurzfristig (2. Klasse) bessere Entwicklung der Vivo-Kinder zu beobachten, die auch in Klasse 3 stabil bleibt. Allerdings kommt es in der 4. Klasse zu einer Angleichung durch einen deutlichen Anstieg der Selbstwirksamkeitserwartung bei den Kontrollgruppenkindern. Wir haben überprüft, ob diese überraschenden Ergebnisse damit in Zusammenhang zu bringen sind, dass vielleicht nicht gegebene Schulempfehlungen oder schlechtere Notengebung dazu geführt haben, dass die Unterschiede im Selbstbild der Kinder in der 4. Klasse zurückgehen. Allerdings haben wir dafür keinerlei Belege gefunden. Warum das Programm auch hier nur einen kurzfristigen Effekt hatte und die Nivellierung der Unterschiede zu Klasse 4 vor allem auf ein Aufholen der Kontrollgruppe zurückzuführen ist, können wir dementsprechend nicht beantworten.

Abbildung 12: Entwicklung Selbstwirksamkeitserwartung (Schülerinnen- und Schülersicht)

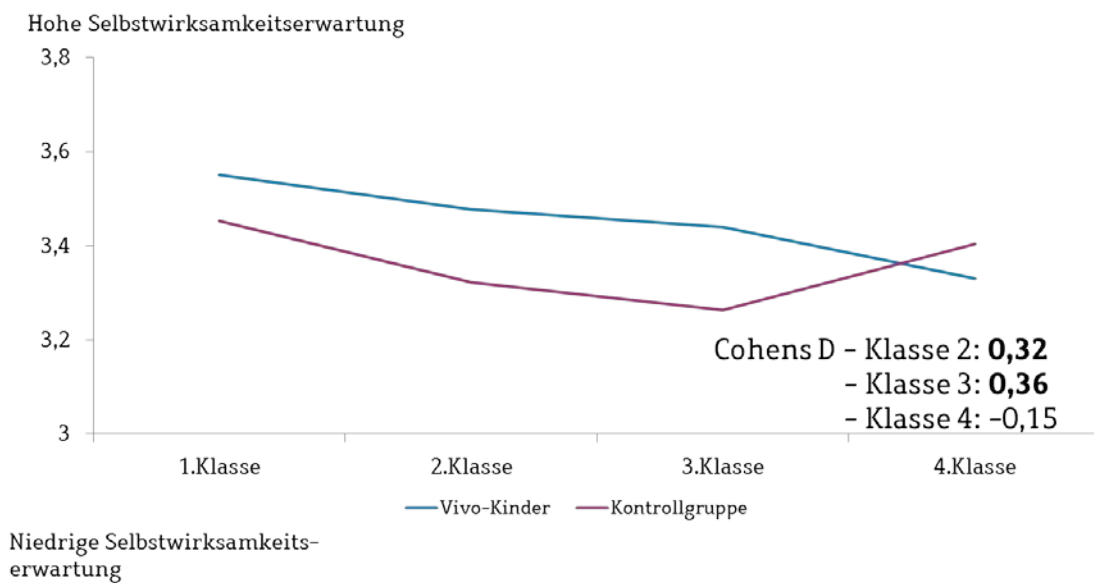
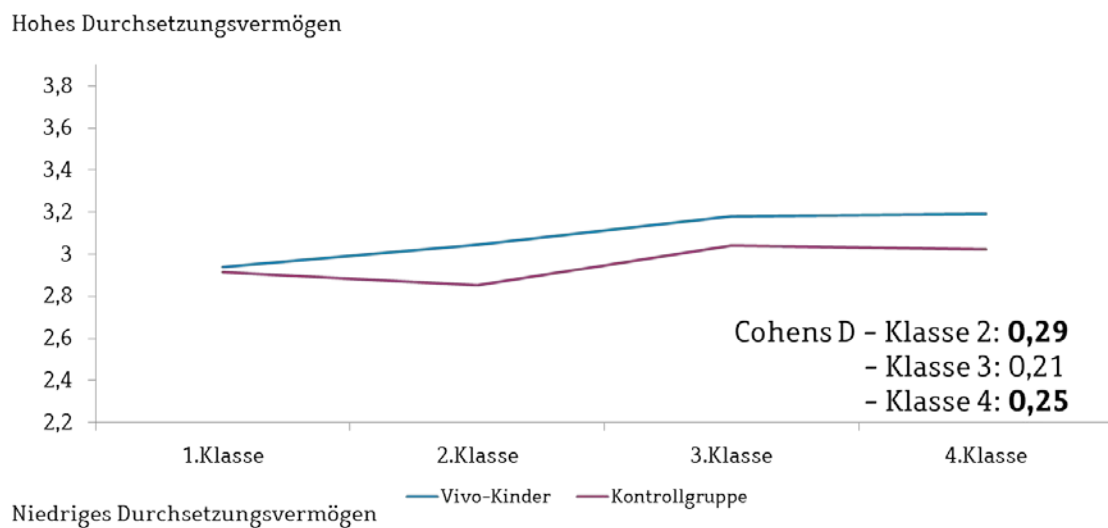


Abbildung 13: Entwicklung Durchsetzungsvermögen (Lehrersicht)

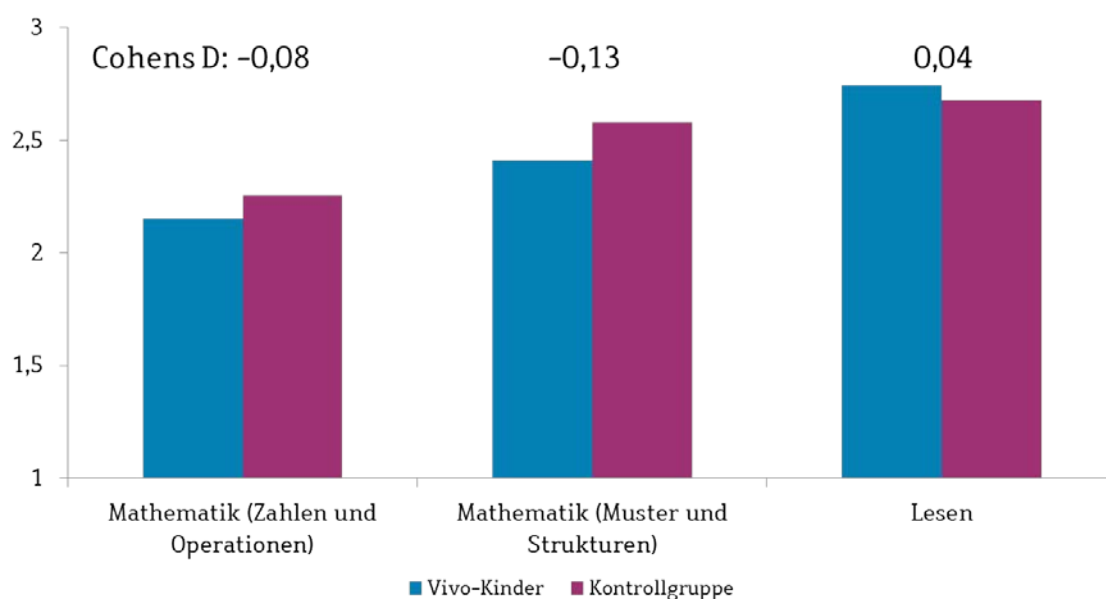


Betrachtet man jedoch die Einschätzung der Lehrkräfte, zeigt sich ein etwas anderes Bild, zumindest bei der Durchsetzungsfähigkeit (Abbildung 13). Hier ist die kurzfristig positivere Entwicklung der Vivo-Kinder auch konsistent bis zum Ende von Klasse 4 zu beobachten.

Im Hinblick auf die Dimension des „Selbst stärken“ haben sich die Vivo-Kinder bis in die 3. Klasse besser entwickelt als die Schülerinnen und Schüler der Kontrollgruppe. Dies ist in der 4. Klasse nicht mehr zu beobachten; aus Sicht der Lehrer wird den Vivo-Kindern allerdings auch in Klasse 4 eine höhere Durchsetzungsfähigkeit zugeschrieben. Das Selbst der Kinder zu stärken, wurde mit dem Programm also kurzfristig erreicht. Einen Effekt bis zum Ende des Programms können wir nur für das Durchsetzungsvermögen nachweisen.

4.2.4 Sprachliche Kompetenz

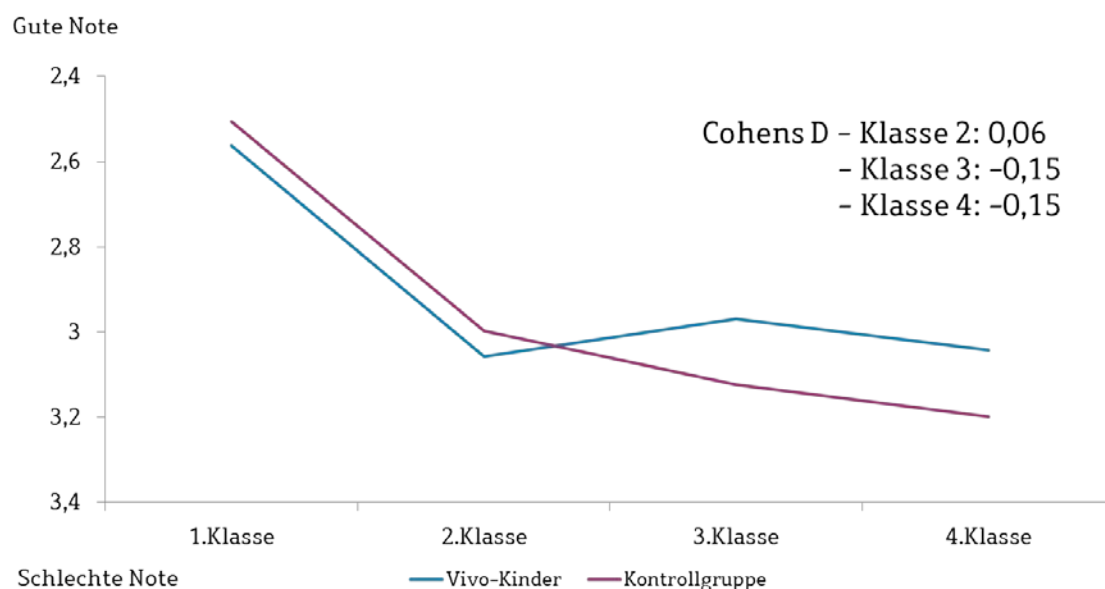
Abbildung 14: Schulische Kompetenzen am Ende der Klasse 3 (gemessen über die Vergleichsarbeiten der Schülerinnen und Schüler)



Ob die sprachliche Kompetenz der Kinder verbessert wurde, wollen wir anhand der Kompetenzmessung im Bereich Lesen am Ende der 3. Klasse sowie anhand der Deutschnoten beurteilen. Bei den Deutschnoten werden aber wie bei allen Noten neben fachspezifischen Kompetenzen auch Eigenschaften wie z. B. Anstrengungsbereitschaft und Arbeitsverhalten mitberücksichtigt (Helbig und Morar 2017). Dementsprechend kann auch die Veränderung dieser Eigenschaften die Deutschnote beeinflussen und nicht allein die Sprachkompetenz.

In Abbildung 14 ist neben den Kompetenzen in Mathematik (auf die wir später noch eingehen) auch die Lesekompetenz dargestellt. An dieser Stelle müssen wir uns aus Gründen der Datenverfügbarkeit und -vergleichbarkeit auf die Analyse der Kompetenzstufen beschränken, was sich natürlich auf die Varianz der Kompetenzwerte auswirkt. Für die Lesefähigkeiten können wir keine statistisch bedeutsamen Unterschiede zwischen den beiden Schülerinnen- und Schülergruppen herausarbeiten.

Abbildung 15: Entwicklung der (eingeschätzten) Deutschnote (Lehrersicht)



Für die (eingeschätzte)¹¹ Deutschnote zeigt sich zum einen, dass es besonders von Klasse 1 zu 2 zu einer massiv schlechteren Einschätzung der Noten beider Gruppen durch die Lehrer kommt. Dies spricht auch dafür, dass Grundschülerinnen und Grundschüler in den meisten Bundesländern erst ab Klasse 3 benotet werden (Helbig und Nikolai 2015). Denn so sehen zumindest Schülerinnen und Schüler diesen Bewertungsabfall am Anfang ihrer Schulkarriere nicht. Dabei zeigen sich in Klasse 2 noch keine Unterschiede zwischen den Vivo-Kindern und der Kontrollgruppe. Mit der 3. Klasse zeigt sich allerdings ein tendenzieller Vorteil der Vivo-Kinder bei der Deutschnote, der aber die Grenze von Cohens D >0,2 knapp

verfehlt. Insgesamt gibt es also keine deutlichen Hinweise darauf, dass durch das Programm die sprachliche Kompetenz gesteigert wurde.

4.2.5 Bildungsziele

Mit der Dimension „Bildungsziele“ war verknüpft, dass man Kinder und Eltern dahingehend fördert, den Wert von Bildung höher zu bewerten und die Erfolgswahrscheinlichkeiten eines bestimmten Bildungsweges besser einzuschätzen. Damit sollte einer Ursache der eingangs beschriebenen Bildungsungleichheit entgegengewirkt werden. Dementsprechend haben wir erfasst, welche Schulform die Kinder nach der Grundschule gerne besuchen würden und welchen Schulabschluss sie erreichen wollen. Dabei verwenden wir nicht die imputierten Daten, da auch die Kategorie „weiß nicht“ an dieser Stelle eine Bedeutung hat. So zeigt sich in Tabelle 3, dass naturgemäß die Schülerinnen und Schüler beider Gruppen in der ersten Klasse selten wussten, welche Schulform sie einmal besuchen wollen. Mit zunehmendem Alter sind die Kinder informierter, aber auch in Klasse 3 wissen es 30 % der Schülerinnen und Schüler nicht (ohne Gruppenunterschiede). Bei den Schülerinnen und Schülern, die schon bei Eintritt in die Schule wissen, welche Schulform sie einmal besuchen möchten, zeigen sich zunächst (Klasse 1) keine Unterschiede zwischen den Gruppen. In Klasse 3 wollen die Vivo-Kinder allerdings deutlich häufiger auf das Gymnasium wechseln als die Schülerinnen und Schüler der Kontrollgruppen. Wir fanden also deutliche Hinweise darauf, dass die Vivo-Kinder am Ende der 3. Klasse ambitioniertere Bildungsziele bezüglich der weiterführenden Schulform haben.

¹¹ Die Leistungen wurden nicht über die gesamte Zeit benotet. Wenn es keine regulären Noten gab (vor allem in Klasse 1 und 2), baten wir die Lehrer eine entsprechende Note zu vergeben.

Tabelle 3: Entwicklung der Bildungsaspirationen (Schulform – Schülerinnen- und Schülersicht)

	Vivo-Kinder	Kontrollgruppe
1. Welle: weiß nicht	46 %	73 %
2. Welle: weiß nicht	39 %	36 %
3. Welle: weiß nicht	30 %	32 %
1. Welle: nur Kinder, die sich festlegen (N=67) – <i>Gymnasium</i>	72 %	76 %
2. Welle: nur Kinder, die sich festlegen (N=102) – <i>Gymnasium</i>	84 %	77 %
3. Welle: nur Kinder, die sich festlegen (N=112) – <i>Gymnasium</i>	83%	68 %

Bei den Aspirationen für einen bestimmten Schulabschluss sind die Anteile von Schülerinnen und Schülern, die dazu keine Antwort haben, (naturgemäß) größer (Tabelle 4). Während sich die Schülerinnen und Schüler der Kontrollgruppen zunächst in Klasse 2 auf einen Schulabschluss festlegen, wissen die Vivo-Kinder am Ende von Klasse 3 häufiger, welchen Schulabschluss sie erwerben wollen. Und dies ist am Ende der 3. Klasse bei beiden Gruppen zu über 80 % das Abitur. Diese ambitionierten Ziele haben sich bei den Vivo-Kindern bereits zur 2. Klasse herausgebildet, bei den Kontrollgruppen erst zur 3. Klasse. Auch dies kann wieder als ein Hinweis verstanden werden, dass das Programm kurzfristig Wirkung erzielt, die Kontrollgruppen-Schülerinnen und Schüler die Vivo-Kinder im weiteren Verlauf aber wieder „einholen“.

Tabelle 4: Entwicklung der Bildungsaspirationen (Schulabschluss – Schülerinnen- und Schülersicht)

	Vivo-Kinder	Kontrollgruppe
1. Welle: weiß nicht	65 %	73 %
2. Welle: weiß nicht	55 %	39 %
3. Welle: weiß nicht	35 %	44 %
1. Welle: nur Kinder, die sich festlegen (N=50) – Abitur	48 %	58 %
2. Welle: nur Kinder, die sich festlegen (N=85) – Abitur	80 %	54 %
3. Welle: nur Kinder, die sich festlegen (N=100) – Abitur	82 %	84 %

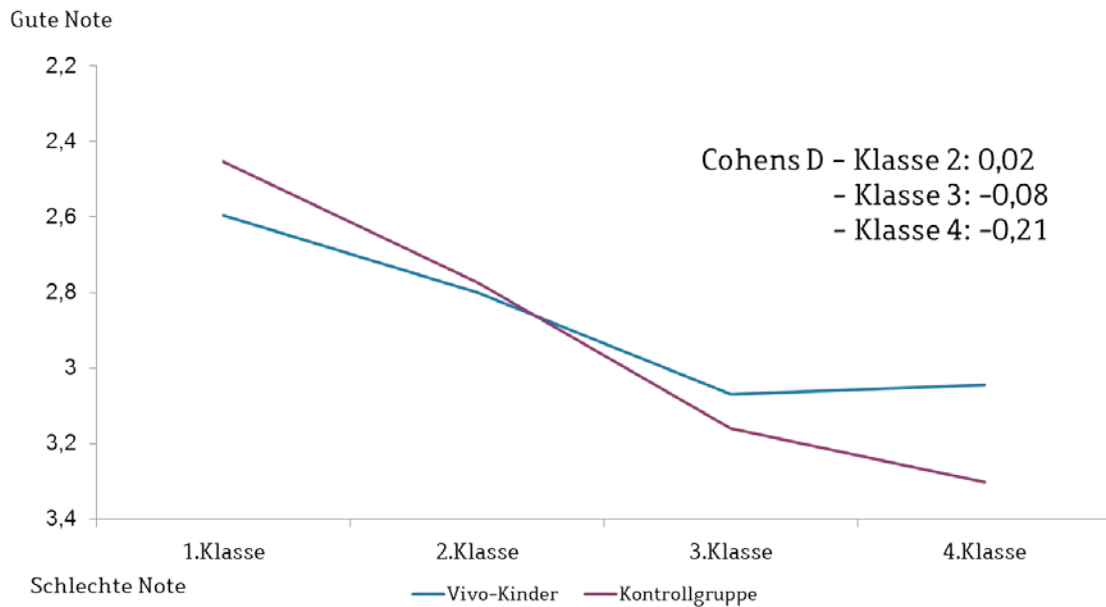
4.2.6 Schulleistungen und Schulverlauf

Neben den sprachlichen Kompetenzen und Noten in Deutsch haben wir auch noch die mathematischen Kompetenzen, die Mathematiknote und die Note in Sachkunde erhoben. Ferner haben wir als Indikatoren für den Schulverlauf den Gymnasialübergang und Klassenwiederholungen abgefragt.

Bei den mathematischen Kompetenzen zeigen sich am Ende von Klasse 3 keine Unterschiede zwischen den Vivo- und Kontrollgruppen-Kindern (Abbildung 14). Auch bei der Mathematiknote, die ebenfalls im Schulverlauf immer schlechter wird (Abbildung 16), zeigt sich bis zum Ende der 3. Klasse kein statistisch bedeutsamer Unterschied. Allerdings sind die Noten der Vivo-Kinder am Ende von Klasse 4 besser als die der Kontrollklasse. Da die mathematischen Kompetenzen der Vivo-Schülerinnen und -Schüler nicht besser sind als die der Kontrollgruppe, ist dies ein Hinweis darauf, dass andere Aspekte, die die Notengebung beeinflussen, für die

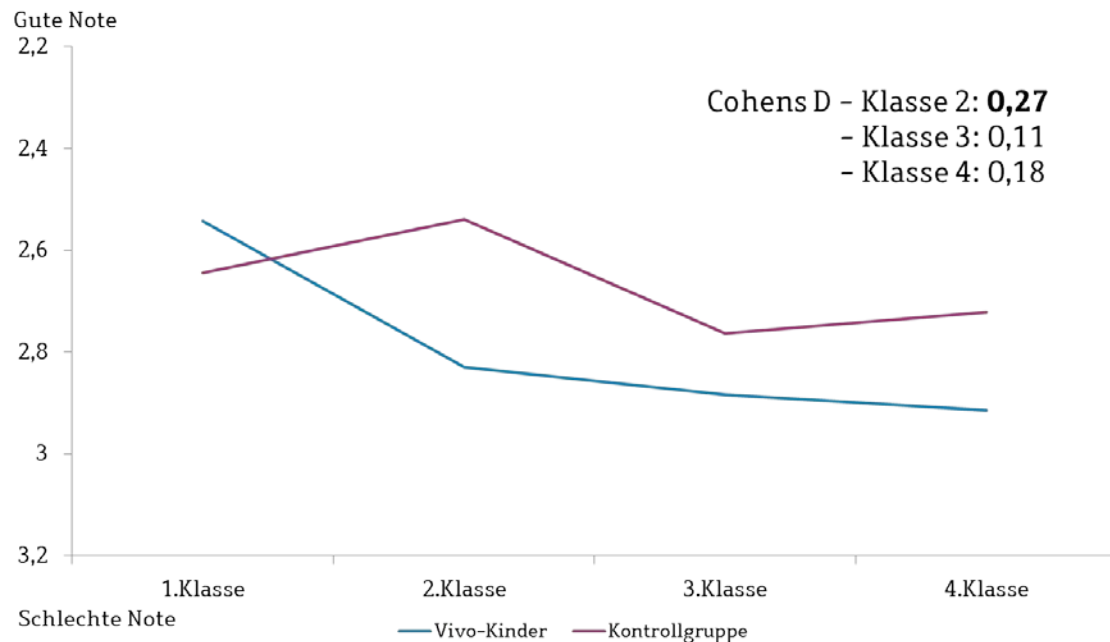
besseren Noten in Mathematik verantwortlich sind, wie z. B. das Arbeitsverhalten oder die Anstrengungsbereitschaft.

Abbildung 16: Entwicklung der (eingeschätzten) Mathematiknote (Lehrersicht)



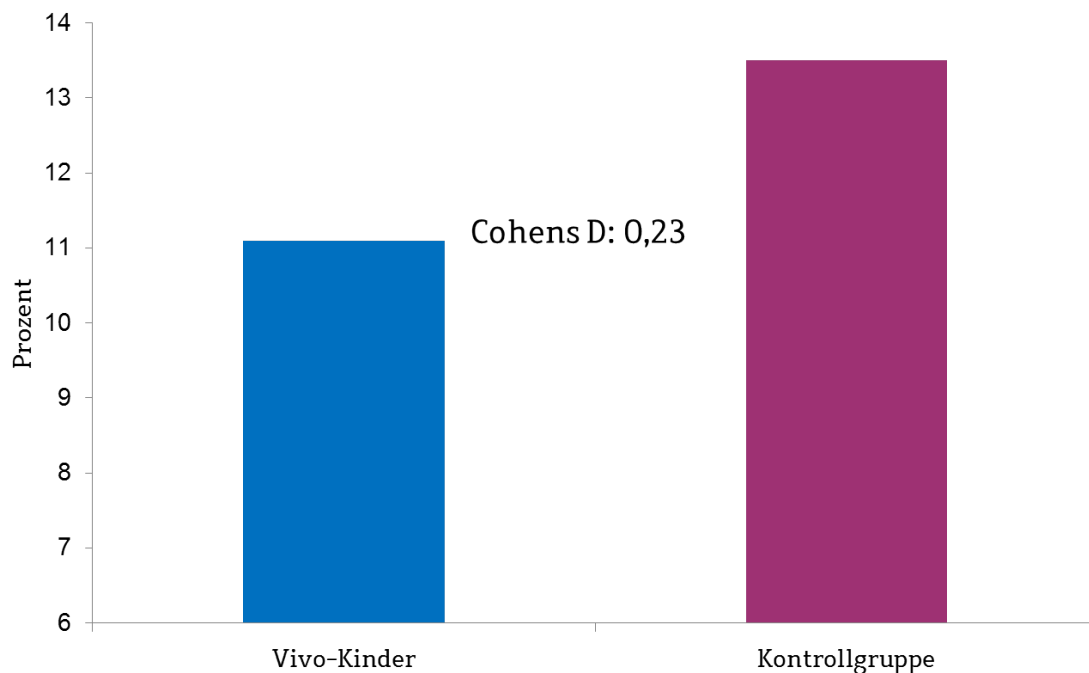
Bei den Noten in Sachkunde kommen wir überraschenderweise zu dem Befund, dass die Vivo-Kinder in Klasse 2 schlechter bewertet wurden als die Kontrollgruppe (Abbildung 17). Dieser Unterschied fällt in Klasse 3 und 4 zwar unter einen Wert von Cohens $D < 0.2$, ist tendenziell aber dennoch zu beobachten. Dieses Ergebnis, dessen Gründe sich uns nicht erschließen, stellt aber zugleich die einzige Dimension dar, in der die Vivo-Kinder statistisch bedeutsam schlechtere Werte aufweisen als die Kontrollgruppe.

Abbildung 17: Entwicklung der (eingeschätzten) Sachkunde-Note (Lehrersicht)



Eine eindeutig positive Entwicklung zeigt sich auch bei den Vivo-Kindern im Bereich der Klassenwiederholung; sie blieben seltener sitzen als die Kinder der Vergleichsgruppe (Abbildung 18). Bei dieser Betrachtung berücksichtigen wir nicht nur die Schülerinnen und Schüler, die am Ende von Klasse 4 noch in unseren Klassen sind, sondern alle Kinder, die in Klasse 1 anwesend waren und die jeweiligen Klassen verlassen haben, weil sie eine Klasse wiederholen mussten. Das Vivo-Programm hatte also den Effekt, dass weniger Kinder sitzen bleiben. Man könnte nun vermuten, dass sich dies auf alle anderen Indikatoren auswirken könnte. Immerhin haben ja in den Kontrollklassen mehr Kinder mit schlechten Schulleistungen die Klasse verlassen, als es bei den Vivo-Kinder der Fall war. Allerdings konnten wir in einer Selektionsanalyse, in der wir untersucht haben, ob die Ausfallquoten von Vivo-Kindern und Kontrollgruppen-Kindern sich nach bestimmten Merkmalen systematisch unterscheiden, keine bedeutsamen statistischen Unterschiede in beiden Gruppen finden.

Abbildung 18: Klassenwiederholungen im Untersuchungszeitraum (Lehrerinformation)



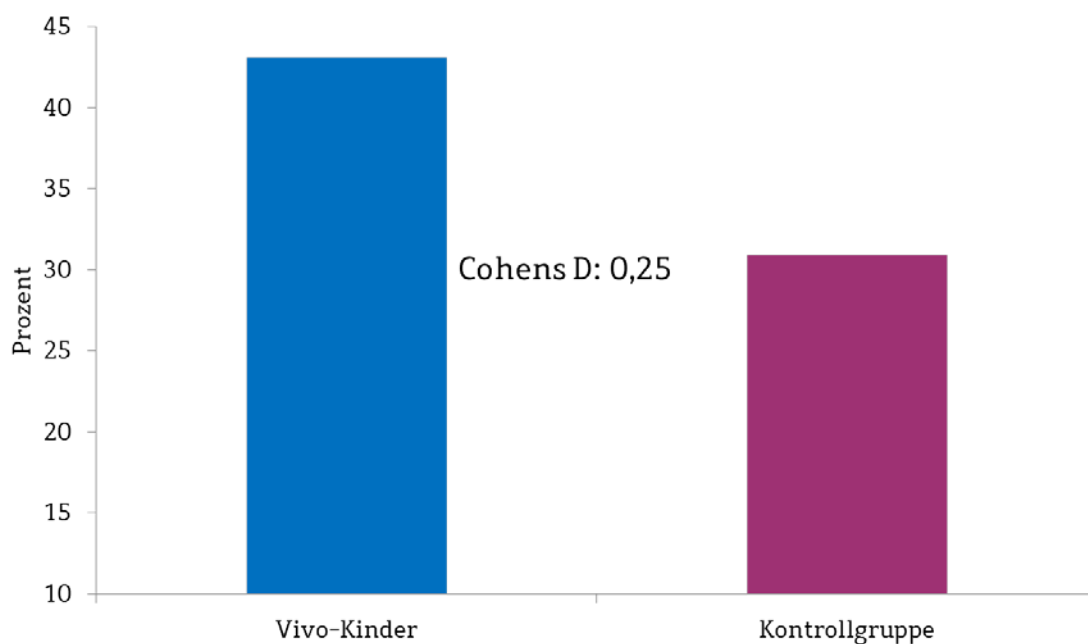
Auch wenn nicht vorrangiges Ziel des Vivo-Programms war, die Gymnasialübergangsquoten zu erhöhen und obwohl sich kaum statistisch bedeutende Effekte des Programms auf die Leistungen der Kinder nachweisen lassen, gehen dennoch deutlich mehr Vivo-Kinder nach der 4. Klasse auf das Gymnasium als Kinder der Kontrollgruppe (Abbildung 19). Während von den Kontrollgruppen-Kindern 30 % auf das Gymnasium übergehen, sind es bei den Vivo-Kindern rund 43 %. Der Unterschied beider Gruppen von 12,7 Prozentpunkten ist dabei gerade einmal zu einem Fünftel auf Leistungen (Verakompetenzen und Noten in Klasse 4) zurückzuführen (Analysen nicht gezeigt). Bessere Leistungen sind also nicht für die höheren Gymnasialquoten der Vivo-Kinder verantwortlich.

Dementsprechend muss es vor allem die Bildungsentscheidung selbst sein, die dazu führt, dass Vivo-Kinder häufiger auf das Gymnasium wechseln. Dies könnte wiederum zum einen Folge der etwas höheren Schulformaspirationen der Vivo-Kinder sein (vgl. Tabelle 3). Zum anderen ist dies ein Hinweis darauf, dass die Bildungsaspirationen der Vivo-Eltern höher sind. Dass diese höher sind, kann wiederum eine Folge der kontinuierlichen Elternarbeit des Programms sowie der

intensiven Begleitung des Übergangs durch z.B. den Eltern-Kind-Workshops sein, der zu Beginn der 4. Klasse durchgeführt wurde. Dabei sollte nicht das Gymnasium idealisiert, sondern über die „passende“ Schulform informiert und von Eltern und Kindern gemeinsam Schulwahlkriterien erarbeitet werden. Es könnte zudem im Zuge des Vivo-Programms zu einer besseren Eltern-Kind-Beziehung gekommen sein, so dass nun die höheren Bildungswünsche der Kinder durch die Eltern häufiger berücksichtigt werden. Außerdem war es gängige Praxis, dass in der Vivo-Zeit verschiedene Projekte durch die Kinder erarbeitet und den Eltern präsentiert wurden. Die Berichte, die uns erreichten, sprachen hier von einer regen Teilnahme der Eltern bei diesen Veranstaltungen. Dadurch wurde möglicherweise auch die Bindung der Eltern an die Institution Schule insgesamt verbessert.

Wir können letztlich den Grund für die höheren Gymnasialübergangsquoten nicht aufklären. Es liegt aber nahe, dass ein Grund der höheren Übergangsquoten auf diese weiterführende Schulform unmittelbar (Workshop) oder mittelbar (Eltern-Kind-Arbeit) bei den Eltern der Vivo-Kinder zu suchen ist.

Abbildung 19: Übergang auf das Gymnasium (Lehrerinformation)



4.3 Die Sicht der beteiligten Akteure

Wie bereits erwähnt, haben wir am Ende des Programms eine Gruppendiskussion und zwei Einzelinterviews mit beteiligten Lehrern, Schulleitern und Erziehern geführt, um unserer rein quantitativen Sichtweise noch den Blick der beteiligten Akteure hinzuzufügen. Hierbei wurden zum einen einige für uns wichtige Aussagen zum Vivo-Programm gemacht und zum anderen einige bedeutungsvolle Punkte thematisiert, die wichtig sind, um Bildungsprogramme externer Anbieter im Allgemeinen besser bewerten zu können.

Kommen wir zunächst zu den Rückmeldungen zum Programm selbst. Insgesamt wurde das Programm von allen befragten Beteiligten positiv wahrgenommen. Als besonders positiv wurde bewertet, dass es durch das Programm zu einer Stärkung der Klassengemeinschaft gekommen sei. Ebenso wurde auf eine Verbesserung der Eltern-Kind-Beziehung durch die stärkere Integration der Eltern hingewiesen. Dies geschah zum einen dadurch, dass in regelmäßigen Abständen die Vivo-Kinder ihren Eltern verschiedene Projektergebnisse präsentierten. Zum anderen wurde ein Programmelement ausnahmslos am besten bewertet, welches wir leider nicht quantitativ evaluieren konnten: der Elternworkshop im ersten Halbjahr von Klasse 4. Hierbei wurde seitens zweier Schulen angegeben, dass man diesen Workshop in den normalen Schulbetrieb überführen wolle. Im Spiegel der gezeigten Ergebnisse aus dem vorhergehenden Abschnitt ist zu erwarten, dass der Eltern-Kind-Workshop jenes Puzzleteil sein kann, das die höheren Übergangsquoten auf das Gymnasium erklärt, obwohl sich die Leistungen der Vivo-Kinder nicht im selben Ausmaß verbessert haben. Bei dem Workshop, wie aber auch für das gesamte Programm, wurde es als positiv herausgestellt, dass Externe auf Schülerinnen, Schüler und Eltern einen anderen Blick hätten bzw. durch Schülerinnen, Schüler und Eltern auch als extern wahrgenommen würden. Dieser positive Aspekt trifft allerdings auch auf andere Programme von Externen zu.

Neben dem bereits erwähnten Elternworkshop wurden vier weitere Punkte des Vivo-Programms ebenfalls als positiv herausgestellt, die wir auch als allgemeine Gelingensbedingungen aus der Bildungsforschung kennen. Wie auch in der

inklusiven Pädagogik gefordert, ist es Vivo erstens besser gelungen, die Stärken von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf herauszustellen. Anscheinend kann intensiver auf die Stärken dieser Schülerinnen und Schüler eingegangen werden, als dies im normalen Unterricht möglich wäre. Vor allem im Vergleich zu anderen externen Programmen wurde es zweitens als positiv herausgestellt, dass das Programm langfristig angelegt ist und mit **gleichen Bezugspersonen** operiert. Drittens ist Vivo ein Programm, das das gleiche Angebot für alle Schülerinnen und Schüler am Nachmittag bereithält. Weder geschlechtstypische noch über Eltern vermittelte Interessen spielen bei der Art der nachmittäglichen Aktivitäten eine Rolle. Alle Kinder bekommen ein Programm, von dem die jeweiligen Trainerinnen und Trainer überzeugt sind und mit dem diese bestimmte Ziele erreichen wollen. Wenn man so will, könnte man diesen Aspekt als ein Merkmal des **rhythmisierten Ganztagsunterrichts** im Klassenverband bezeichnen. Viertens wurde herausgestellt, dass im Programm die Meinung der einzelnen Schülerinnen bzw. des einzelnen Schülers eher gefragt war als im sonstigen Schulalltag. Insoweit kann im Vivo-Programm auch ein Element **demokratischer Schule** gesehen werden, wie es heute im Schulbetrieb oft angemahnt wird. Man kann aus Sicht der Akteure herausarbeiten, dass Vivo in vielen Elementen so gesehen wird, wie Schule aus Sicht der Bildungsforschung heute sein sollte, es aber offensichtlich nicht ist.

Negative Aspekte gegenüber dem Vivo-Programm wurden in der Befragung nicht thematisiert. Allerdings äußerten sich die Akteure gegenüber extern angebotenen Programmen im Allgemeinen kritisch. Dies bezog sich vor allem auf drei Aspekte: Erstens wurde herausgestellt, dass es eine Vielzahl externer Programme im heutigen Schulsystem gibt, ob nun privat, öffentlich oder gemischt finanziert. Hierbei wurde angemahnt, dass sich diese Programme teilweise ähneln bzw. unter jeweils anderem Namen ähnliche Dinge angeboten werden. In diesem Zusammenhang wurde zweitens darauf hingewiesen, dass diese Programme häufig für maximal 2 Jahre aus dem jeweiligen Fördertopf finanziert würden und danach wegfielen. Hierdurch sei keine langfristige Wirkung oder personelle Kontinuität zu erreichen. Ein Schulleiter äußerte sich hierzu in der Weise, dass man noch Glück hätte, wenn der gleiche Träger im nächsten Jahr unter dem „Deckmantel“ eines

neuen Programms wiederkäme, damit man nicht wieder bei null anfangen müsse. Drittens wurde kritisch gefragt, inwieweit derartige Programme nachhaltig seien, ob eine Übertragung der Programminhalte über das Klassenleben hinaus gelänge und ob 90 Minuten pro Woche überhaupt ausreichen, um etwas zu bewirken. Diesbezüglich wurde es aber als positiv wahrgenommen, dass sich mit Vivo ein Programm einer externen Evaluation stellt.

Was uns im Gespräch mit den Akteuren sehr überrascht hat, ist die Aussage, dass man sich heute ganz genau aussuchen müsse, mit welchen externen Partnern man zusammenarbeiten wolle, da eine Vielzahl von externen Partnern anklopfe, um der Schulgemeinschaft „unter die Arme zu greifen“. Es ist zumindest zu vermuten, dass dies auch damit zusammenhängt, dass es etwa im Raum Hamburg und Düsseldorf in Folge wirtschaftlicher Stärke eine hohe Stiftungsdichte gibt und sich Stiftungen wiederum vor allem lokal engagieren. Ob diese Meinung auch in den strukturschwachen Regionen Brandenburgs, Sachsen-Anhalts oder Rheinland-Pfalz' geäußert worden wäre, kann bezweifelt werden.

5 Zusammenfassung und Fazit

Von 2013 bis 2017 haben wir das Pilot-Programm „Vivo – Bildung von Grund auf“ der Stiftung der Deutschen Wirtschaft begleitet. Dabei haben wir sowohl die Geldgeber als auch die operativ am Programm beteiligten Mitarbeiterinnen als sehr engagiert wahrgenommen. Sie hatten das Ziel, einen (kleinen) Beitrag dazu zu leisten, dass es in der Bildungsrepublik Deutschland etwas gerechter zugeht. Anders als der Großteil an extern finanzierten Programmen hat sich die sdw dabei einer quantitativen Evaluation (mit Experimental- und Kontrollgruppendesign) gestellt, um die Frage zu beantworten, ob Vivo wirkt.

Wie angesprochen (Abschnitt 3.3), hatten wir es mit einigen methodischen Herausforderungen zu tun, die es enorm erschwert haben, die Unterschiede zwischen der Experimentalgruppe (Vivo-Kinder) und der Kontrollgruppe eindeutig auf das Vivo-Programm zurückzuführen. Schlussendlich können wir „nur“ zeigen, inwieweit sich die Vivo-Kinder und die Kontrollgruppe von Klasse 1 bis 4 unterschiedlich entwickelt haben. Insgesamt haben wir Hinweise darauf gefunden, dass sich die Vivo-Kinder bezüglich der Dimensionen „Motivation und Zielorientierung“ sowie das „Selbst stärken“ zunächst besser entwickelt haben als die Kontrollgruppenkinder. Dieser Effekt lässt sich (bis auf das Durchsetzungsvermögen) allerdings am Ende von Klasse 4 nicht mehr eindeutig nachweisen. Insoweit scheint das Vivo-Programm dazu geeignet zu sein, Schülerinnen und Schüler beim Übergang in eine neue Institution zu unterstützen. Bezüglich der Förderdimensionen „Sozialkompetenzen“ und „sprachlichen Kompetenzen“ verbessern, zeigten sich insgesamt in der Tendenz positive Entwicklungen, ohne dabei statistisch bedeutsam zu sein.

Obgleich für einzelne Dimensionen lediglich kurzfristig höhere Werte erreicht wurden, zeigt sich bei den Schulleistungen und Schulverläufen ein deutlich positiveres Bild für die Vivo-Kinder: Sie haben bessere Mathematiknoten (teilweise auf Kosten der Sachkundenote), mussten seltener eine Klasse wiederholen und gehen deutlich häufiger nach der Grundschule auf das Gymnasium. Dass sich die höheren Gymnasialquoten der Vivo-Kinder minimal

über Leistungsindikatoren erklären lassen, kann als Hinweis verstanden werden, dass die Bildungsziele von Schülerinnen, Schülern (dies zeigt sich teilweise) und Eltern bei den Vivo-Kindern höher sind. Bei den Bildungszielen der Eltern könnte der Eltern-Kind-Workshop dazu beigetragen haben, sekundären Herkunftseffekten entgegenzuwirken.

Bei den einzelnen Förderdimensionen ist dies nur in einigen Fällen statistisch bedeutsam nachweisbar. Dieses Ergebnis mag, wie in Abschnitt 4.1 thematisiert wurde, auch darauf zurückgeführt werden, dass das Projekt nur einen zeitlich begrenzten Raum im (Schul-)Leben der Schülerinnen und Schüler darstellt. Allerdings kann es auch sein, dass die Fallzahl zu klein war, um statistisch bedeutsame Effekte herauszuarbeiten. In Hinblick auf die angeführten methodischen Herausforderungen und Limitationen im vorhergehenden Abschnitt (vor allem im Hinblick auf die Fallzahl) ist ein überhaupt gemessener Einfluss des Programms auf einzelne Ziele positiv zu bewerten. Interessant sind dabei die Rückmeldungen der beteiligten Schulakteure, die darauf hinweisen, dass Vivo in einigen Elementen so gesehen wird, wie Schule aus Sicht der Bildungsforschung heute sein sollte.

Als positive Aspekte kann man zudem festhalten, dass das Vivo-Programm den Schülerinnen und Schülern zu Beginn ihrer Bildungskarriere die Eingliederung in die Institution Schule erleichtert. Wenn man zudem weiß, dass die Vivo-Schülerinnen und -Schüler in Klasse 5 weiter begleitet werden, dann scheint die Stärke des Programms darin zu liegen, Übergänge ins und im Schulsystem zu erleichtern.

Ziel der Studie war es zu untersuchen, ob Stiftungsprogramme soziale Ungleichheiten auf dem Weg zum Gymnasium auszugleichen im Stande sind. Dass man es an Schulen in sozial benachteiligter Lage geschafft hat, die Gymnasialübergangsquoten signifikant zu erhöhen, kann als Erfolg gewertet werden. Damit wurde das übergeordnete Ziel von Vivo erreicht.

6 Literatur

Allmendinger, J. (2012): Schulaufgaben. Wie wir das Bildungssystem verändern müssen, um unseren Kindern gerecht zu werden München: Pantheon.

Becker, R. (2000): Klassenlage und Bildungsentscheidungen. Eine empirische Anwendung der Wert-Erwartungstheorie. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 52(3), pp. 450-474.

Boudon, R. (1974): Education, opportunity, and social inequality: Changing prospects in Western society New York, NY: Wiley.

Breen, R., & Golthorpe, J. H. (1997): Explaining educational differentials. Rationality and Society, 9, pp. 275-305.

Deutsches PISA-Konsortium. (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen: Leske+Budrich.

Deutsches PISA-Konsortium. (2008): PISA 2006. PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich Münster/ New York/ München/ Berlin: Waxmann.

Downey, D.B und Condrón, D. J. (2016): Fifty Years since the Coleman Report: Rethinking the Relationship between Schools and Inequality. In: Sociology of Education, Vol. 89, No. 3, S. 207-220.

Ehlert, M, Finger, C., Rusconi, A. und Solga, H. (2017): Applying to College. Do Information Deficits Lower the Likelihood of College-eligible Students from Less-Privileged Families to Pursue Their College Intentions? – Evidence from a Field Experiment. In: Social Science Research, Vol. 67, S. 193-212. (vorab online publiziert 29.04.2017).

Engel, E.-M. (2015): Selbstkonzeptfragebogen für Kinder (SKF). Freiburg: FEL.

Fend, H. (1997): Der Umgang mit der Schule in der Adoleszenz. Aufbau und Verlust von Lernmotivation, Selbstachtung und Empathie. Bern: Hans Huber.

Helbig, M. und Nikolai, R. (2015): Die Unvergleichbaren. Der Wandel der Schulsysteme in den deutschen Bundesländern seit 1949. Bad Heilbrunn: Julius Klinkhardt.

Helbig, M. und Morar, T. (2017): Warum Lehrkräfte sozial ungleich bewerten. Ein Plädoyer für die Etablierung tertiärer Herkunftseffekte im werterwartungstheoretischen Standardmodell der Bildungsforschung, WZB Discussion Paper, P 2017-005.

Helm, F., Pohlmann, B., Heckt, M., Gienke, F., May, P. und Möller, J. (2012): Entwicklung eines Fragebogens zur Einschätzung überfachlicher Schülerinnen- und Schülerkompetenzen. In: Zeitschrift für Lernforschung. Unterrichtswissenschaft 40. Jg., H. 3, S. 235-258.

Helmke, A. und Weinert, F.E. (1997): Bedingungsfaktoren schulischer Leistungen. In: Weinert, F.E. (Hrsg.): Enzyklopädie der Psychologie. Psychologie des Unterrichts und der Schule. Pädagogische Psychologie, Vol. 3. Göttingen: Hogrefe, S. 71–176.

Höhne, T. (2015): Stiftungen: Die neuen Player in der Bildungspolitik. Dossier Bildung. Bundeszentrale für politische Bildung. Online:
<http://www.bpb.de/gesellschaft/kultur/zukunft-bildung/199165/stiftungen?p=0>
(Stand 20.02.2018).

Hurrelmann, K. (2010): Kinder in Deutschland 2010: 2. World Vision Kinderstudie. Frankfurt am Main: Fischer Verlag.

Husén, T. (1967). International study of achievement in mathematics. A comparison of twelve countries Stockholm: Almquist & Wiksell.

Jerusalem, M. und Hopf, D. [Hrsg.] (2002): Das Konzept der Selbstwirksamkeit. In: Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen. Weinheim: Beltz, S. 28-53.

Kratzer, V. und Cwielong, I. (2013): Quantitative Befragung mit Kindern. In: Tillmann, A., Fleischer, F., Hugger, K.-U.: Handbuch Kinder und Medien. Wiesbaden: Springer VS, S. 183-198.

- Langfeldt, H.-P. und Prücher, F. (2004): BSSK. Bildertest zum sozialen Selbstkonzept. Ein Verfahren für Kinder der Klassenstufen 1 und 2. Göttingen: Beltz.
- Lenhard, W. und Schneider, W. (2006): ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler. Göttingen: Hogrefe.
- Möller, J. und Trautwein, U. (2009). Selbstkonzept. In: E. Wild & J. Möller (Hrsg.) Pädagogische Psychologie. Wiesbaden: Springer, S. 180-202.
- Midgley et al. (2000). Manual for the Patterns of Adaptive Learning Scales (PALS). Ann Arbor, MI: University of Michigan.
- Oberwöhrmann, S., Bettge, S., Hermann, S., & Meinlschmidt, G. (2013): Migrationshintergrund als Einflussfaktor auf die kindliche Entwicklung im Einschulungsalter – ein multivariates Modell. Gesundheitswesen, 75, pp. 203-209.
- Peisert, H. (1967): Soziale Lage und Bildungschancen in Deutschland München: Piper.
- Picht, G. (1964): Die deutsche Bildungskatastrophe, Analyse und Dokumentation Freiburg: Walter-Verlag.
- Pollak, R. (2009): Chancengleichheit durch Bildung? Eine ländervergleichende Studie zum Einfluss der Bildung auf soziale Mobilität im Europa des 20. Jahrhunderts Universität Mannheim: Dissertation.
- Prenzel, M. (2009): Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms. Münster: Waxmann.
- Rauer, W. und Schuck, K.-D. (2003): FESS. Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen. Göttingen: Beltz.
- RKI. (2008): Kinder- und Jugendgesundheitsurvey (KiGGS) 2003-2006: Kinder und Jugendliche mit Migrationshintergrund in Deutschland. Beiträge zur Gesundheitsberichterstattung des Bundes. Berlin: Robert Koch-Institut.
- Roebbers, C. M., Schmid, C. und Roderer, T. (2009): Metacognitive Monitoring and Control Processes Involved in Primary School Children's Test Performance. In:

- British Journal of Educational Psychology, 79, S. 749-767.
- Royston, P. (2004): Multiple Imputation of Missing Values. *Stata Journal* 4: S. 227-241.
- Satow, L. (2000): Klassenklima und Selbstwirksamkeitsentwicklung.
- Schimpl-Neimanns, B. (2000). Soziale Herkunft und Bildungsbeteiligung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 52(4), pp. 636-669.
- Schwarzer, R. und Jerusalem, M. (1999): Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: Freie Universität Berlin.
- Stadler, C., Janke, W. und Schmeck, K. (2004): IVE. Inventar zur Erfassung von Impulsivität, Risikoverhalten und Empathie bei 9- bis 14-jährigen Kindern: Hogrefe. Göttingen.
- StEG Konsortium (2016): Ganztagschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote. Ergebnisse der Studie zur Entwicklung von Ganztagschulen 2012–2015. Online: https://www.projekt-steg.de/sites/default/files/StEG_Brosch_FINAL.pdf (Stand 20.02.2018).
- StBa (2015): Zeitverwendungserhebung Aktivitäten in Stunden und Minuten für ausgewählte Personengruppen 2012/13. Wiesbaden: Statistisches Bundesamt.
- Weiß, R. H. und Osterland, J. (2012): CFT 1-R. Grundintelligenztest Skala 1. Revision. Göttingen: Hogrefe.
- Wohlkinger, F., & Ditton, H. (2012): Entscheiden die Schüler mit? Der Einfluss von Eltern, Lehrern und Kindern auf den Übergang nach der Grundschule. In H. Solga & R. Becker (Eds.), *Soziologische Bildungsforschung. Sonderband der Kölner Zeitschrift für Soziologie und Sozialpsychologie* (pp. 44-63). Wiesbaden: Springer.

7 Anhang

7.1 Operationalisierung der Konstrukte

Lernmotivation

Die Skalen zur Lernmotivation beziehen sich auf die Ziele, die die Schülerinnen und Schüler verfolgen, wenn sie sich im Unterricht beteiligen. Wir haben lediglich abgefragt, ob die Motivation intrinsisch oder extrinsisch begründet ist. Liegt eine intrinsische Motivation vor, so handelt der Lernende um der Tätigkeit willen, Schülerinnen und Schüler lernen, weil das Lernen an sich Freude und Zufriedenheit bereitet. Bei der extrinsischen Motivation wird die Handlung ausgeführt, um bestimmte Ziele zu erreichen, d. h. die Handlung selbst dient lediglich als Zweck, das angestrebte Ergebnis zu erreichen. Schülerinnen und Schüler lernen also, um beispielsweise gute Noten zu erhalten. Wir gehen davon aus, dass für einen erfolgreichen Lernprozess die intrinsische Motivation erstrebenswerter und sinnvoller erscheint, da intrinsisch motivierte Personen bei ansonsten gleichen Rahmenbedingungen eine insgesamt höhere Leistung aufweisen als extrinsisch motivierte Personen. Denn bereitet eine Tätigkeit Freude, so setzt man sich vermehrt mit ihr auseinander, was wiederum zu einer tieferen Verarbeitung der erlernten Inhalte führt und somit einen Lernerfolg wahrscheinlicher werden lässt.

Die Skalen haben wir Midgley et al. (2000) entnommen.

- 1. Ich strenge mich in der Schule an, weil ich etwas dazulernen möchte.*
- 2. In der Schule strenge ich mich an, um besser als die anderen zu sein.*
- 3. Ich strenge mich in der Schule an, damit meine Mitschüler nicht glauben, ich sei dumm.*

Um eine additive Skala zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt Lernmotivation einen Faktor.

Eine 4 bedeutet eine hohe Lernmotivation, eine 1 deutet auf eine niedrige Lernmotivation hin.

Klassenklima

Das Klassenklima ist zugleich ein individuelles und kollektives Phänomen. Da angenommen wird, dass die Stimmung/das Klima in der Klasse einen wesentlichen Effekt auf die Selbstkonzeptentwicklung ausübt (Satow), war uns wichtig abzubilden, ob die Schülerinnen und Schüler gut zusammenarbeiten oder ob es Ausgrenzungen oder Gruppendruck gibt. Als Quelle diente uns der KlassenCheckUp! der Bundeszentrale für politische Bildung. Folgende 11 Items wurden über vier Wellen hinweg von den Schülerinnen und Schülern kommentiert.

- 1. In unserer Klasse helfen wir alle gerne, wenn jemand Hilfe braucht.*
- 2. Wir achten darauf, dass niemand mit seinen Problemen allein gelassen wird.*
- 3. Ich finde in der Klasse schnell jemanden, wenn wir mit jemandem zusammenarbeiten sollen.*
- 4. In unserer Klasse werden Streitereien, ohne jemandem wehzutun, gelöst.*
- 5. Einige in unserer Klasse suchen ständig Streit.*
- 6. Einige in unserer Klasse machen sich über ihre Mitschüler lustig.*
- 7. Bei uns darf jeder seine Meinung sagen.*
- 8. Bei uns werden einzelne Schüler ausgeschlossen.*
- 9. Die meisten sehen es nicht gern, wenn jemand allzu gut in der Schule ist.*
- 10. Kinder, die anders sind, haben es schwer in unserer Klasse.*
- 11. In meiner Klasse fühle ich mich wohl.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt Klassenklima drei Faktoren. Die Items 1, 2, 3, 7, 11 operationalisierten wir über positives Klassenklima, die Items 5 und 6 über Peer Pressure und die Items 8, 9 und 10 über negatives Klassenklima bzw. Ausgrenzung. Item 4 wurde ausgeschlossen.

Eine 4 bedeutet ein gutes Klassenklima, eine 1 deutet auf ein schlechtes Klassenklima hin.

Bei den Berechnungen nach Welle 2 wurden wir von den Zahlen überrascht: Das Klassenklima nahm in den Vivo-Projektklassen ab. Daher nahmen wir in Welle 3 und 4 den Bezug der Items, die eher eine Ausgrenzung aufzeigen, zusätzlich zum Zusammenhang zur eigenen Person auf. Operationalisiert wurden sie über die eigene Negativerfahrung von Ausgrenzung und beinhalteten folgende Items.

- 1. Einige in unserer Klasse suchen ständig Streit mit mir.*
- 2. Einige in unserer Klasse machen sich über mich lustig.*
- 3. Schülerinnen und Schüler unserer Klasse schließen mich aus.*

Ebenfalls wurde in Welle 3 das Konstrukt Empathie hinzugenommen und die Lehrer zusätzlich einmal zum Klassenklima befragt.

Empathie

Empathie ist die Fähigkeit zu kognitivem Verstehen und affektivem Nachempfinden der vermuteten Emotion eines anderen Lebewesens. Man kann davon ausgehen, dass bei einer hohen Empathiefähigkeit verstärkt die Tendenz zu prosozialem Verhalten zu beobachten ist und eine Ausgrenzung, die über das Klassenklima erhoben werden konnte, eventuell so erklärt werden kann. Geeignete Skalen für den Grundschulbereich haben wir dem Inventar zur Erfassung von Impulsivität, Risikoverhalten und Empathie bei 9- bis 14-jährigen Kindern (IVE) entnommen. Die 8 Items liefen in der 3. und 4. Erhebungswelle.

- 1. Es bedrückt mich, wenn ich sehe, dass jemand ausgelacht wird.*
- 2. Es beschäftigt mich sehr, wenn ich jemanden weinen sehe.*
- 3. Kinder, die niemanden zum Spielen finden, tun mir sehr leid.*
- 4. Wenn ich Entscheidungen treffe, mache ich mir viele Gedanken darüber, wie sich andere dabei fühlen.*
- 5. Ich spüre oft Mitgefühl für Leute, denen es schlechter geht als mir.*
- 6. Ich kann mir gut vorstellen, wie sich jemand fühlt, der sehr einsam ist.*
- 7. Manchmal ist mir zum Weinen, wenn ich im Fernsehen einen traurigen Film sehe.*
- 8. Kinder, die oft gehänselt werden, tun mir leid.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt Empathie einen Faktor.

Eine 4 bedeutet eine hohe Ausprägung von Empathie, eine 1 deutet auf eine geringe Empathie hin.

Schulisches Selbstkonzept

Neben der Einschätzung und Bewertung der eigenen Person existieren auch Einschätzungen der eigenen Eigenschaften, Fähigkeiten und Kompetenzen, das sogenannte schulische oder akademische Selbstkonzept. Der Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern (FEES) weist geeignete Skalen sowohl für die erste und zweite, als auch für die dritte und vierte Klasse auf. Nachfolgende 13 Items erhoben wir über die gesamten vier Wellen.

- 1. Ich bin ein guter Schüler.*
- 2. Ich schaffe immer meine Aufgaben im Unterricht.*
- 3. Ich kann gut rechnen.*
- 4. Ich merke mir, was ich gelernt habe.*
- 5. Ich kann gut lesen.*
- 6. Ich verstehe immer, was der Lehrer erklärt.*
- 7. Ich lerne sehr schnell.*
- 8. Ich verstehe das meiste in der Schule.*
- 9. Ich mache in der Schule das meiste richtig.*
- 10. Ich kann meine Aufgaben meistens alleine lösen.*
- 11. Ich kann ganz gut lernen.*
- 12. Ich kann gut vor der ganzen Klasse erzählen.*
- 13. Ich kann den anderen gut etwas erklären.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt schulisches Selbstkonzept einen Faktor.

Eine 4 bedeutet ein hohes schulisches Selbstkonzept, eine 1 deutet auf ein geringes schulisches Selbstkonzept hin.

Selbstwirksamkeit

Die Selbstwirksamkeit beschreibt die Erwartungshaltung, dass man Widerstände bewältigen und sich gegen sie durchsetzen kann. (Jerusalem/ Schwarzer 1981/1999) Nicht vergessen sollte man dabei aber auch die Selbstwirksamkeitserwartung im Umgang mit sozialen Anforderungen. Also sollte man auch das Augenmerk darauf haben, dass man gemeinsam mit anderen Hindernisse überwinden kann, dass man Rückhalt fühlt und trotzdem mit seiner eigenen Haltung in der Gruppe bestehen oder sich durchsetzen kann. (Satow/Mittag 1999) Eine Skala, die sowohl diese allgemeine Selbstwirksamkeit als auch die soziale Selbstwirksamkeit abdeckt, haben wir der World Vision Kinderstudie entnommen. Folgende 9 Items wurden in den Schülerinnen- und Schülerbefragungen über die 4 Wellen abgefragt:

1. *Ich glaube, mein Leben wird richtig schön.*
2. *Ich glaube, dass ich viele Dinge gut kann.*
3. *Ich habe viele Menschen, die mir helfen.*
4. *Wenn mich jemand ärgert oder ungerecht behandelt, kann ich mich dagegen wehren.*
5. *Schwierige Sachen kriege ich, wenn ich mich anstrenge, gut hin.*
6. *Ich traue mich zu sagen, was ich denke, auch wenn die anderen anderer Meinung sind.*
7. *Ich finde schnell neue Freunde.*
8. *Wenn andere Kinder was machen, kann ich immer dabei sein.*
9. *Egal was ist, ich weiß mir immer zu helfen.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt Selbstwirksamkeit einen Faktor. Eine 4 bedeutet eine hohe Selbstwirksamkeitserwartung, eine 1 deutet auf eine geringe Selbstwirksamkeitserwartung hin.

Allgemeines/Soziales Selbstkonzept

In der ersten Klasse haben wir einen Bildertest zum sozialen Selbstkonzept (BSSK) durchgeführt. Per Definition ist das soziale Selbstkonzept die kognitive Repräsentation der eigenen Person, beim BSSK wird der Schwerpunkt auf die Facette „Peers“ als Teil des sozialen Selbstkonzepts gelegt. Die Idee dahinter ist die Wahrnehmung und Beurteilung der eigenen sozialen Kompetenz und des Verhaltens in Bezug auf die soziale Interaktion mit anderen. Der Selbstkonzeptfragebogen für Kinder (SKF), der in den Klassenstufen 2 bis 4 eingesetzt wurde, basiert auf dem gleichen Modell. Mit dem Begriff Selbstkonzept werden Vorstellungen, Einschätzungen und Bewertungen bezüglich verschiedener Aspekte der eigenen Person bezeichnet. (Moschner, 2001-Aus Möller, J. & Trautwein, U. (2009). Selbstkonzept. In E. Wild & J. Möller (Hrsg.)).

1. *Ich helfe anderen gerne.*
2. *Ich verstehe mich gut mit anderen Kindern.*
3. *Andere Kinder spielen gerne mit mir.*
4. *Ich bin gerne mit anderen Kindern zusammen.*
5. *Andere Kinder hören mir zu, wenn ich etwas erzähle.*
6. *Ich habe oft Streit mit anderen.*
7. *Ich bin ein fröhliches Kind.*
8. *Andere Kinder sind nett zu mir.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das soziale Selbstkonzept einen Faktor. Item 6 wurde ausgeschlossen.

Eine 4 bedeutet ein hohes soziales Selbstkonzept, eine 1 deutet auf ein geringes soziales Selbstkonzept hin.

Anstrengungsbereitschaft mit Arbeitsvermeidung

Die Skala Anstrengungsbereitschaft erfasst das Ausmaß, in dem ein Kind bereit ist, sich in der Schule auf Neues einzulassen und Anforderungen in der Schule durch eigenes Bemühen zu bewältigen. Auch diese Skala mit 13 Items entnahmen wir

dem Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern in der dritten und vierten Klasse (FEES).

1. *Ich strenge mich an, damit ich besser schreiben kann.*

2. *Es macht mir keinen Spaß, neue Dinge zu lernen.*

3. *Ich will jeden Tag in der Schule etwas Neues lernen.*

4. *Ich gebe mein Bestes in der Schule.*

5. *Ich strenge mich ungern an beim Lernen.*

6. *Ich strenge mich an, damit ich besser lesen kann.*

7. *Ich mag nur leichte Aufgaben.*

8. *Wenn mir etwas schwerfällt, gebe ich schnell auf.*

9. *Ich versuche alles richtig zu machen.*

10. *Ich versuche, auch ganz schwierige Aufgaben zu lösen.*

11. *Ich gebe schnell auf, wenn ich Probleme habe.*

12. *Ich strenge mich an, damit ich besser rechnen kann.*

13. *Ich freue mich auf neue Aufgaben.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt Arbeitsbereitschaft zwei Faktoren.

Die Items 1, 3, 4, 6, 9, 10, 12, 13 operationalisierten wir über

Anstrengungsbereitschaft, die Items 2, 5, 7, 8, 11 über Arbeitsvermeidung.

Eine 4 bedeutet eine hohe Arbeitsbereitschaft, eine 1 deutet auf eine niedrige Arbeitsbereitschaft hin.

Sozialverhalten

Das Sozialverhalten soll die Fähigkeiten und Einstellungen abbilden, die die individuelle Haltung zur Welt auch in Bezug auf das Arbeiten ausdrücken, so u. a. das Vertreten eigener Interessen. Das Konstrukt des Sozialverhaltens bilden Kenntnisse, Fertigkeiten oder Fähigkeiten, die dazu beitragen, dass in den Beziehungen zu Mitmenschen situationsadäquat gehandelt werden kann. Darunter werden auch Kompetenzen wie Kooperationsfähigkeit oder Konfliktfähigkeit verstanden. Hier konnten wir aus der Kompetenzerfassung in Kindergarten und Schule (KEKS) 10 Items für die Lehrkräfte ableiten.

1. *Das Kind arbeitet in Gruppen kooperativ.*
2. *Das Kind kann eigene Fähigkeiten realistisch einschätzen.*
3. *Das Kind geht mit Kritik oder Misserfolgen angemessen um.*
4. *Das Kind kann sich in andere hineinversetzen.*
5. *Das Kind hilft anderen.*
6. *Das Kind setzt sich gegenüber anderen Kindern durch.*
7. *Das Kind entwickelt eine eigene Meinung, trifft eigene Entscheidungen und vertritt diese gegenüber anderen.*
8. *Das Kind ist mit vielen Kindern in der Klasse befreundet.*
9. *Das Kind hält vereinbarte Regeln ein.*
10. *Das Kind kann in Konfliktsituationen eigene Interessen vertreten, aber auch nachgeben und Kompromisse finden.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt Sozialverhalten zwei Faktoren. Die Items 1, 2, 3, 4, 5, 8, 9, 10 operationalisierten wir über Sozialverhalten, die Items 6 und 7 über Durchsetzungsfähigkeit.

Eine 4 bedeutet ein positives Sozialverhalten, eine 1 deutet auf ein negatives Sozialverhalten hin.

Lernverhalten

Das Lernverhalten umfasst die Fähigkeiten, die es ermöglichen, Aufgaben und Probleme durch die Auswahl, Planung und Umsetzung sinnvoller Lösungsstrategien zu bewältigen (Lernstrategien, Fähigkeit zur Informationspräsentation). Unsere Items wählten wir aus der Kompetenzerfassung in Kindergarten und Schule (KEKS). Folgende 10 Items wurden über alle 4 Wellen von den Lehrkräften abgefragt.

1. *Das Kind strengt sich in der Schule an.*
2. *Das Kind beschäftigt sich konzentriert mit einer Sache.*
3. *Das Kind arbeitet und lernt selbständig.*
4. *Das Kind arbeitet gründlich.*
5. *Das Kind ist beharrlich und ausdauernd.*

6. *Das Kind ist motiviert, etwas zu schaffen oder zu leisten, es ist zielstrebig.*
7. *Das Kind zeigt Interesse und Neugier am Lernen.*
8. *Das Kind arbeitet im Unterricht eifrig mit.*
9. *Das Kind geht sorgfältig mit Arbeitsmaterialien um.*
10. *Das Kind gibt schnell auf, wenn ihm etwas schwerfällt.*

Um additive Skalen zu bilden, führten wir eine Faktorenanalyse mit Varimax-Rotation durch und erhielten für das Konstrukt Lernverhalten einen Faktor. Eine 4 bedeutet ein gutes Lernverhalten, eine 1 deutet auf ein schlechtes Lernverhalten hin.

Discussion Papers der Projektgruppe der Präsidentin 2018

Marcel Helbig und Stefanie Jähnen

P-2018-001

Wie brüchig ist die soziale Architektur unserer Städte? Trends und
Analysen der Segregation in 74 deutschen Städten