

Förderung von Präsentationskompetenz: Schrittweise Implementation und Effektivitätsüberprüfung eines Präsentationstrainings für Grundschul Kinder

Evelin Herbein  · Jessika Golle · Benjamin Nagengast · Ulrich Trautwein

Online publiziert: 4. Februar 2020
© Der/die Autor(en) 2020

Zusammenfassung Erfolgreiche Interventionen sind sowohl wirksam als auch implementierbar. Um dies zu erreichen werden idealtypisch mehrere aufeinander aufbauende Phasen der Interventionsentwicklung und -überprüfung durchlaufen. Die damit einhergehenden Standards können jedoch oft nicht uneingeschränkt eingehalten werden, wenn Studien unter Praxisbedingungen durchgeführt werden. Die vorliegende Studie ist die vierte eines Forschungsprogramms, in dem eine Intervention zur Förderung der Präsentationskompetenz (PK) von Grundschulkindern systematisch entwickelt und evaluiert wurde. Sie untersucht, ob (1) das Training wie intendiert, d. h. mit hoher Durchführungstreue, in der Praxis implementierbar ist und (2) ob sich die zuvor beobachteten Effekte auf PK und Sprechangst replizieren lassen. Zehn Kursleitungen und 65 Kinder nahmen an der cluster-randomisierten Studie mit Wartekontrollgruppe und Prä-Post-Messung teil. Die Durchführungstreue wurde über Selbsteinschätzungen der Kursleitungen, PK über Videoratings und Sprechangst über Fragebögen erfasst. Es zeigte sich, dass das Training erneut mit hoher Durchführungstreue durchgeführt wurde. Für die PK fanden sich auf zwei

Zusatzmaterial online Zusätzliche Informationen sind in der Online-Version dieses Artikels (<https://doi.org/10.1007/s11618-019-00923-y>) enthalten.

Dr. E. Herbein (✉) · Prof. Dr. J. Golle · Prof. Dr. B. Nagengast · Prof. Dr. U. Trautwein
Hector-Institut für Empirische Bildungsforschung, Universität Tübingen,
Walter-Simon-Str. 12, 72072 Tübingen, Deutschland
E-Mail: evelin.herbein@uni-tuebingen.de

Prof. Dr. J. Golle
E-Mail: jessika.golle@uni-tuebingen.de

Prof. Dr. B. Nagengast
E-Mail: benjamin.nagengast@uni-tuebingen.de

Prof. Dr. U. Trautwein
E-Mail: ulrich.trautwein@uni-tuebingen.de

der 18 erfassten Präsentationsfähigkeiten signifikante Effekte: Körperspannung und persönliche Ansprache. Keine signifikanten Effekte fanden sich auf Sprechangst.

Schlüsselwörter Durchführungstreue · Präsentationskompetenz · Effektivitätsstudie · Grundschul Kinder

Fostering public speaking skills: Gradual implementation and effectiveness check of a public speaking training program for elementary school children

Abstract Putting an intervention successfully into practice requires an effective intervention and effective implementation. Based on the repeated demand for public speaking promotion programs for elementary school children, a corresponding program was developed. Next, we conducted a pilot, efficacy, and first effectiveness study. Across the studies, positive treatment effects were found. Herein, we present the second effectiveness study. Ten trained course instructors and 65 children participated in the study which was a cluster randomized trial with a wait-list control group and repeated measures. We assessed implementation fidelity, i.e., the extent to which the program was implemented as intended (via self-reports), the effects of the program on public speaking skills (via video ratings), and speech anxiety (via self-reports). High implementation fidelity was found. Positive treatment effects were found for two of the 18 assessed public speaking skills: posture (nonverbal public speaking skill) and addressing audience (language use). No treatment effect was found for speech anxiety. The results of this study are, inter alia, compared and discussed with respect to the first effectiveness study.

Keywords Effectiveness study · Elementary school children · Implementation fidelity · Public speaking

1 Einleitung

Eine Intervention wird dann als erfolgreich bezeichnet, wenn sie die intendierten Förderziele erreicht und gleichzeitig in der Praxis implementierbar ist. Um eine Intervention hinsichtlich dieser Eigenschaften einschätzen und optimieren zu können, werden idealtypisch mehrere Phasen der Interventionsentwicklung und -überprüfung systematisch durchlaufen: auf die Konzeption folgen die Pilotierung sowie verschiedene Effektivitätsstudien unter zunehmend realeren, d. h. weniger kontrollierten Durchführungsbedingungen, bevor die Intervention breiter in der Praxis eingesetzt werden kann (vgl. Gottfredson et al. 2015; Humphrey et al. 2016). Die Umsetzung der Studien, unter strenger Einhaltung der methodischen Standards, kann jedoch mit Herausforderungen verknüpft sein. Gründe hierfür sind Schwierigkeiten z. B. bei der Umsetzung von Datenerhebungen im schulischen Kontext (Richter et al. 2019) oder der Rekrutierung von Teilnehmer/inne/n für die Umsetzung von randomisierten Studiendesigns (vgl. Friedman et al. 2010) sowie Vorbehalte der teilnehmenden

Lehrpersonen gegenüber der Studie und beispielsweise der Beobachtung ihres Unterrichts (Krainer et al. 2012).

Die vorliegende Studie ist die vierte Studie eines Forschungsprogramms, in dem eine Intervention zur Förderung von Präsentationskompetenz (PK) von Grundschulkindern systematisch entwickelt und evaluiert wurde. Die Durchführung der einzelnen Phasen wurde bestmöglich an den bestehenden Standards der Interventionsforschung ausgerichtet. Kompetent präsentieren zu können ist bereits für diese Altersgruppe relevant, da Kinder meist ab der 2. Klasse Präsentationen halten müssen (z. B. Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, Ref. Öffentlichkeitsarbeit 2014; Ministerium für Kultus, Jugend und Sport Baden-Württemberg 2016). Das kann mit unterschiedlichen Herausforderungen einhergehen, da diese Aufgabe noch relativ neu ist, im Vergleich beispielsweise zum Erzählen (EDK 2010). Für besonders begabte und hochbegabte Grundschul Kinder, mit vertieftem fachlichen Wissen zum Präsentationsinhalt oder besonders hohen verbalen Fähigkeiten, kann es zudem eine Herausforderung sein, ihre Präsentation an das Wissen und die Interessen ihrer Mitschüler/innen anzupassen (vgl. Lee et al. 2012). Für die Förderung der PK finden sich jedoch kaum, hinsichtlich ihrer Wirksamkeit überprüfte, Unterrichtsmaterialien/Trainingsprogramme (vgl. Hunt et al. 2014; Herbein et al. 2018a).

Abgeleitet vom identifizierten Bedarf wurde ein extracurriculares Präsentations-training (11 Kurseinheiten à 90 min) entwickelt und anschließend in mehreren Studien überprüft. Nach einer Pilotierungsphase folgten zwei randomisierte kontrollierte Studien, um zu überprüfen, ob das Training die intendierten Effekte auf den Zielvariablen hat: In Studie 1 wurde das Training unter kontrollierten Bedingungen von der Entwicklerin selbst durchgeführt (Herbein et al. 2018a), in Studie 2 unter Praxisbedingungen von acht Kursleiterinnen (Herbein et al. 2018b). In beiden Studien fanden sich positive Interventionseffekte, jedoch waren die Effekte in Studie 2 teilweise geringer als in Studie 1. Die Abnahme der Interventionseffekte ist ein Muster das häufig beobachtet wird, wenn Interventionen unter weniger kontrollierten Bedingungen und von Personen aus der Praxis implementiert werden (vgl. Petermann 2014; Greene 2015). Folgestudien sind notwendig, um weitere Erkenntnisse über die Durchführbarkeit und Effektivität der Intervention zu erlangen. Ziel der vorliegenden Studie war deshalb nicht, den nächsten Schritt des Scaling-up des Trainings zu gehen, sondern stattdessen eine Replikation der Effektivitätsstudie. Untersucht wurde, ob die Intervention in der Praxis so durchgeführt werden kann, wie von den Entwickler/inne/n intendiert, d. h. mit hoher Durchführungstreue, und ob sich die erwarteten Effekte auf PK und Sprechangst zeigen. Die Studie wurde im Open Science Framework präregistriert.

2 Erfolgreiche Interventionen: wirksam und durchführbar

Für die Interventionsentwicklung und -überprüfung lassen sich sechs verschiedene Phasen unterscheiden (Abb. 1; vgl. Gottfredson et al. 2015; Humphrey et al. 2016; Herbein et al. 2018b). Nach der Ableitung des *Bedarfs* einer bestimmten Zielgruppe (Phase 1) erfolgt die *Konzeption* der Intervention (Phase 2). Es schließen sich

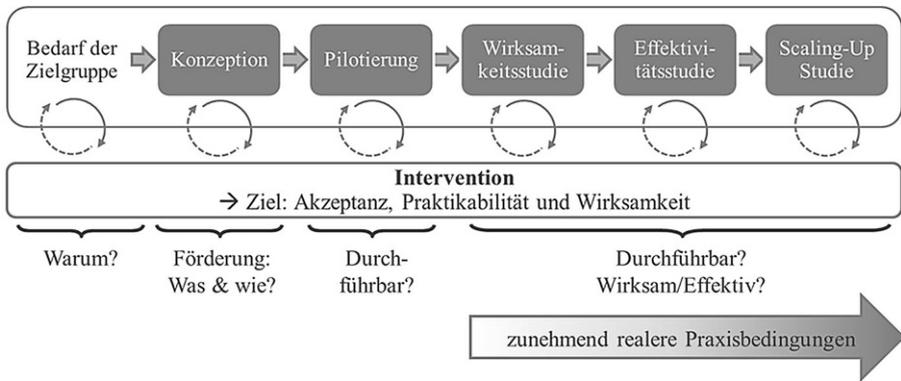


Abb. 1 Phasen der Interventionskonzeption, schrittweisen Implementation und Überprüfung in der Praxis (basierend auf Gottfredson et al. 2015; Humphrey et al. 2016; Herbein et al. 2018b)

eine *Pilotierungsphase* (3) und verschiedene Schritte der Effektivitätsüberprüfung an (Phase 4: *Wirksamkeitsstudie*, im Englischen *Efficacy-Study*; Phase 5: *Effektivitätsstudie*, im Englischen *Effectiveness-Study*). Dabei nähert sich die Interventionsdurchführung allmählich realen Praxisbedingungen an. Die Ergebnisse dieser Phasen werden dazu genutzt, die Intervention und deren Durchführbarkeit zu optimieren. Falls sinnvoll können die Phasen auch mehrmals durchlaufen werden (Gottfredson et al. 2015). Erweist sich eine Intervention als erfolgreich in den Phasen 1–5, kann sie weiter in der Praxis verbreitet werden (Phase 6: *Scaling-up*). *Scaling-up* meint einerseits den flächendeckenden Einsatz der Intervention, welcher durch Unterstützungsmaßnahmen begleitet sein kann (Baker 2006). Andererseits kann ein *Scaling-up* auch die Ausweitung des Implementationskontexts bedeuten (Gottfredson et al. 2015). In der Phase des *Scaling-up* sind erneut die Durchführbarkeit und Wirksamkeit zu überprüfen. Zudem können auch neue Fragestellungen, z. B. zur Nachhaltigkeit und Effizienz, adressiert werden (vgl. Gottfredson et al. 2015).

Während die verschiedenen Phasen durchlaufen werden, d. h. die Standardisierung der Durchführungsbedingung zunehmend abnimmt, ist häufig eine Abnahme der Stärke der Interventionseffekte beobachtbar (Petermann 2014; Greene 2015). Für die geringeren oder fehlenden Effekte gibt es verschiedene Gründe: die Unwirksamkeit des Trainings, Fehler bei der Studienplanung und -durchführung sowie Mängel in der Implementation (Greene 2015; Humphrey et al. 2016). Um einzuschätzen, ob das entwickelte Training effektiv ist, ist zunächst sicherzustellen, dass die Intervention so implementiert wurde, wie intendiert, d. h. mit hoher Durchführungstreue. Andernfalls ist eine entsprechende Interpretation und Diskussion der gefundenen Interventionseffekte nicht möglich (Siddiqui et al. 2018).

Unabhängig von der Intervention lässt sich die Durchführungstreue auf acht Dimensionen erfassen (vgl. Durlak und DuPre 2008; Nelson et al. 2012; Humphrey et al. 2016). Diese können dem Verhalten der Lehrperson (Einhaltung, Dosis, Qualität, Adaption), dem Verhalten der Teilnehmenden (Aufnahmebereitschaft, Erreichbarkeit) und Eigenschaften der Intervention (Abgrenzbarkeit zu anderen Interventionen, Kontrolle der Vergleichsintervention) zugeordnet werden (Durlak und DuPre

2008; Humphrey et al. 2016). In verschiedenen Studien zur Effektivität von Interventionen zeigte sich, dass die Effekte auf den Zielvariablen umso höher sind, je geringer die Diskrepanz zwischen entwickelter und durchgeführter Intervention ist (vgl. Durlak und DuPre 2008; O'Donnell 2008). Wird das Training von Lehrpersonen aus der Praxis implementiert, ist es jedoch wahrscheinlich, dass diese Anpassungen vornehmen (Rimm-Kaufman et al. 2014; Greene 2015; Humphrey et al. 2016), was zu einer geringeren Durchführungstreue führt. Diese Anpassungen können unterschiedlich ausgeprägt sein. Sie reichen von der Auslassung ganzer Interventionselemente, bis zu feinen Anpassungen auf Umsetzungsebene, wie der Verwendung von spezifischen Beispielen aus dem Alltag der Teilnehmenden, um die Passung zur Zielgruppe zu erhöhen. Letztere müssen nicht zwingend zu einer Verringerung der Interventionseffekte führen (Blase und Fixsen 2013). Ausschlaggebend ist, dass die Kernkomponenten der Intervention weiterhin wie intendiert implementiert werden. Übertragen auf das vorliegende Präsentationstraining ist es deshalb bedeutsam, die Durchführungstreue jeweils empirisch zu überprüfen, während das Training schrittweise erprobt und evaluiert wird. Für das Training wurden bereits fünf der sechs Phasen durchlaufen. Sie werden im Weiteren genauer dargestellt.

2.1 Phase 1: Bedarf und Ziele eines Präsentationstrainings für Grundschul Kinder

Kompetent zu präsentieren, um Wissen weiterzugeben, ist bereits ab dem Grundschulalter bedeutsam und im Bildungsplan verankert (z.B. Common Core State Standards Initiative 2010; EDK 2010; Ministerium für Kultus, Jugend und Sport Baden-Württemberg 2016). Präsentationskompetenz wird als Kernkompetenz von gut ausgebildeten Schüler/innen erachtet (van Ginkel et al. 2015). Von der Grundschule an, über die Sekundarschule bis in den Hochschul- und beruflichen Kontext, gewinnt sie zunehmend an Bedeutung. Sie stellt eine fächerübergreifende Kompetenz dar, die nicht nur im Sprachunterricht, sondern auch in den MINT-Fächern relevant ist (Lee et al. 2013). Für Kinder im Grundschulalter ist das Präsentieren eine Aufgabe, mit der sie noch relativ unvertraut sind. Dazu kommt, dass die Themen, über die gesprochen wird, zunehmend komplexer und abstrakter werden (vgl. EDK 2010; Behrens und Eriksson 2011; Schleppegrell 2012). Haben Kinder bereits sehr hohe verbale Fähigkeiten und vertieftes Wissen über die Inhalte über die sie sprechen, wie das bei besonders begabten und hochbegabten Kindern der Fall sein kann, kann es für sie eine Herausforderung sein, ihre Präsentation an das Wissen und Interesse ihrer Mitschüler/innen anzupassen (vgl. Lee et al. 2012).

Verschiedene Disziplinen beschäftigen sich mit der Definition, Entwicklung und Förderung von PK, wobei über die Disziplinen hinweg bislang noch keine Einigkeit darüber herrscht, was unter PK verstanden wird. Im vorliegenden Projekt wurde deshalb aus verschiedenen Theorien und Modellen ein konzeptioneller Rahmen abgeleitet, der mit einer relativ großen Anzahl der bestehenden Konzeptionen vereinbar ist (vgl. De Grez und Valcke 2010; Backlund und Morreale 2015; Herbein et al. 2018a): Ein Sprecher/eine Sprecherin wird dann als kompetent wahrgenommen, wenn sein/ihr gezeigtes Präsentationsverhalten in der konkreten Präsentationssituation als effektiv und angemessen eingeschätzt wird. Effektiv bedeutet, dass

das Kommunikationsziel erreicht wird, zum Beispiel zu informieren. Angemessen meint, dass das gezeigte Verhalten angemessen für den Präsentationskontext ist.

Das gezeigte Präsentationsverhalten gliedert sich in verschiedene Präsentationsfähigkeiten, welche sich theoriegeleitet vier Dimensionen zuordnen lassen: nonverbal-visuelle Fähigkeiten (z. B. Blickkontakt, Gestik), nonverbal-auditive Fähigkeiten (z. B. Artikulation, Sprechflüssigkeit), Sprachgebrauch (z. B. Verwendung von Stilmitteln) und organisatorische Präsentationsfähigkeiten (z. B. Struktur, Informationsauswahl). Um diese Fähigkeiten in der Präsentationssituation kompetent einsetzen zu können, benötigt eine Person (i) *Wissen* über das Planen und Halten von Präsentationen, (ii) ein Repertoire an *Fähigkeiten* auf das er/sie zurückgreifen kann und (iii) *Motivation*. Letztere schließt die Bereitschaft tatsächlich kompetent präsentieren zu wollen, Selbstwirksamkeitserwartung im Präsentieren und Sprechangst mit ein (Croucher 2013). Alle drei Faktoren zusammen, Wissen, Fähigkeitsrepertoire und Motivation, interagieren und bilden zusammen die individuellen Voraussetzungen einer Person, um kompetent präsentieren zu können (De Grez und Valcke 2010; Backlund und Morreale 2015). Abgeleitet aus diesem Verständnis von PK wurden für das vorliegende, extracurriculare Präsentationstraining für besonders begabte und hochbegabte Grundschul Kinder, zwei Hauptziele festgelegt. Zum einen sollen die Kinder ihr Wissen zu und Repertoire an Präsentationsfähigkeiten erweitern und in der Lage sein, ihre Fähigkeiten angemessen in der Präsentationssituation einzusetzen. Zum anderen ist es das Ziel, Sprechangst zu reduzieren, da diese einen negativen Einfluss auf die gezeigte Leistung hat (Croucher 2013).

2.2 Phase 2: Konzeption des Trainings

Um eine erfolgreiche Intervention zu entwickeln ist ein theoretisch und empirisch fundiertes Konzept erforderlich. Dieses stützt sich auf ein *Interventionsmodell*, das eine Kombination aus zwei nacheinander abzuleitenden Modellen darstellt: dem *Change Model* und dem *Logic Model* (Nelson et al. 2012). Als erstes wird das Change Model (auch *Wirkmodell*) abgeleitet. Dieses besteht aus theoretisch und/oder empirisch fundierten Kernkomponenten, von denen man annimmt, dass sie die zugrundeliegenden Wirkmechanismen für die Intervention bilden und damit für die intendierten Effekte auf den Zielvariablen verantwortlich sind (Nelson et al. 2012; Blase und Fixsen 2013). Basierend auf dem Change Model wird das Logic Model abgeleitet. Es umfasst die konkreten Inhalte und Trainingsaktivitäten der Intervention (z. B. alle Aufgaben) sowie mögliche Unterstützungsmaßnahmen für die Interventionsdurchführung (vgl. Humphrey et al. 2016; Murrah et al. 2017; Coldwell und Maxwell 2018). Letzteres können unter anderem Schulungsmaßnahmen für die Lehrpersonen sein. Zusammen bilden das Change und das Logic Model das Interventionsmodell.

Das vorliegende Präsentationstraining ist eine Multikomponenten-Intervention, die im Change Model sieben Komponenten berücksichtigt. Da es kaum Studien gibt, die die Effektivität von Präsentationstrainings im Primar- und Sekundarschulbereich untersucht haben, wurde bei der Konzeption auf Komponenten zurückgegriffen, die sich im Hochschulkontext als effektiv erwiesen hatten. Van Ginkel et al. (2015) leiten in ihrem Überblicksartikel für den Hochschulkontext sieben Prinzipien

Kurseinheit										
1	2	3	4	5	6	7	8	9	10	11
Prätest & Kennenlernen	Lampenfieber	Nonverbale Kommunikation		Verbale Kommunikation		Präsentationsvorbereitung & (Video)Feedback			Abchlusspräsentation	Posttest & Abschied IG/Willkommen KG
↓↑		↓↑	↓↑		↓↑			↓↑	↓↑	
Einzelpäsentationen zu einem naturwissenschaftlichen Thema → Themenwahl, Vorbereitung und Übung der Präsentation										
Schrittweise Gewöhnung an die Präsentationssituation → zunehmende Komplexität der Aufgaben										

Abb. 2 Inhalt und Aufbau des Präsentationstrainings (IG Interventionsgruppe, KG Kontrollgruppe; vgl. Herbein et al. 2018b)

ab: Theorieinput, Modelllernen, praktisches Üben, formatives Feedback, Selbsteinschätzungen, Fremdbewertung und Transfer. Diese wurden als Basis für das Change Model des Präsentationstrainings für Kinder im Grundschulalter verwendet.

Für das Logic Model wurden dann die Trainingsinhalte und -aktivitäten abgeleitet. Die Inhaltsauswahl orientierte sich an der oben beschriebenen Definition von PK und den Erwartungen, die beim Präsentieren an Kinder im Grundschulalter gestellt werden (vgl. Herbein et al. 2018a). Dementsprechend wurden für jede der vier Dimensionen – nonverbal-visuelle, nonverbal-auditive und organisatorische Präsentationsfähigkeiten sowie Sprachgebrauch – inhaltliche Schwerpunkte gesetzt, um das Vorwissen und das Interesse der Zielgruppe bestmöglich zu berücksichtigen. Insgesamt besteht der Kurs aus 11 wöchentlichen Kurseinheiten a 90 min. Abb. 2 bietet eine Übersicht über die Verteilung der Inhalte auf die Kurseinheiten. Basierend auf den sieben Komponenten des Change Modells und orientiert an bestehenden praktisch erprobten, jedoch nicht evaluierten Trainingsmaterialien für Schüler/innen der Grund- und Sekundarschule, wurden die Trainingsaktivitäten konzipiert. Diese umfassten Theoriebausteine, Lernen am Modell mittels Videobeispielen, Audiobeispielen und Beobachten der Lehrperson oder Peers sowie praktische Übungen, die schrittweise komplexer wurden. Zudem wurden verschiedene Phasen des Feedbacks, d. h. Selbst-, Peer-, Trainer- und Videofeedback, angewandt. Um den Transfer zu ermöglichen, arbeiteten die Kinder über den ganzen Kurs hinweg an ihrem eigenen MINT-Expertenthema, das sie selbst gewählt hatten.

2.3 Phasen 3–5: schrittweise Überprüfung der Durchführbarkeit und Effektivität

Nach der Konzeption bildet die Pilotierung den ersten Schritt der Interventionsüberprüfung. Ziel ist es, die Durchführbarkeit der Intervention zu untersuchen (vgl. Humphrey et al. 2016). Dazu wurde das vorliegende Präsentationstraining zunächst von der Entwicklerin selbst, einer Psychologin und Sprechwissenschaftlerin, an zwei Gruppen in der Praxis pilotiert. Basierend auf den Ergebnissen wurden die Inhalte und Trainingsaktivitäten überarbeitet. Anschließend folgte eine erste Wirksamkeits-

studie, in der die Interventionsdurchführung unter stark kontrollierten Bedingungen erfolgte (vgl. Gottfredson et al. 2015; Humphrey et al. 2016). Das Training wurde erneut von der Entwicklerin, anhand eines Kursmanuals, durchgeführt. Die Durchführungstreue war dadurch maximal hoch. Ein randomisiertes Kontrollgruppendesign mit Messwiederholung wurde genutzt, um die Interventionseffekte zu überprüfen. Dabei erhielt die Kontrollgruppe ein alternatives Training, das die Förderung des Wissenschaftsverständnisses zum Ziel hatte. Es nahmen 65 Kinder teil. Es zeigten sich signifikant positive Effekte auf nonverbal-visuelle Präsentationsfähigkeiten (Cohen's d von 0,45–0,77), nonverbal-auditive Fähigkeiten ($d=0,54$ und 0,64) sowie organisatorische Präsentationsfähigkeiten ($d=0,54$ und 0,59) und das Gesamtaufreten der Kinder ($d=0,62$) während ihrer Präsentation zum Posttest, im Vergleich zur Kontrollgruppe. Zudem fand sich ein signifikant negativer Effekt auf Sprechatmung ($d=-0,70$). Keine signifikanten Effekte fanden sich für Sprechangst (Herbein et al. 2018a).

Aufgrund der positiven Ergebnisse folgte eine Effektivitätsstudie. Das Training wurde in dieser Studie nicht mehr von der Kursentwicklerin selbst, sondern von geschulten Kursleitungen aus der Praxis durchgeführt (vgl. Humphrey et al. 2016). Um die Effektivität des Trainings zu überprüfen wurde nun ein randomisiertes Wartekontrollgruppendesign mit Prä- und Posttest eingesetzt, d.h. die Kontrollgruppe erhielt kein alternatives Training sondern „business as usual“. An der Studie nahmen 61 Dritt- und Viertklässler/innen teil. Der Kurs wurde von acht Kursleitungen durchgeführt. Um die Durchführungstreue hoch zu halten, erhielten sie vor dem Kursstart eine eintägige, vorbereitende Schulung, ein detailliertes Kursmanual sowie alle Kursmaterialien. Die Ergebnisse zeigten, dass der Kurs mit hoher Durchführungstreue angeboten wurde. In Bezug auf die Effektivität des Trainings wurden signifikant positive Effekte auf organisatorische Präsentationsfähigkeiten (d von 0,61–0,89) und das Gesamtaufreten der Kinder ($d=0,56$), die das Training im Vergleich zur Wartekontrollgruppe besucht hatten, beobachtet. Außerdem war die Sprechangst der Trainingskinder geringer ausgeprägt als die Sprechangst der Kinder in der Kontrollgruppe ($d=-0,48$ und $-0,68$; Herbein et al. 2018b).

3 Fragestellungen

Ziel des vorliegenden Forschungsprogramms war es, ein Training zu entwickeln und zu überprüfen, das PK von Kindern im Grundschulalter fördert und Sprechangst reduziert. Das Training sollte sowohl in der Praxis implementierbar als auch effektiv sein. Die Trainingskonzeption basierte auf Erkenntnissen der Forschung zur Förderung von PK im Hochschulkontext (vgl. van Ginkel et al. 2015), welche für die Zielgruppe angepasst wurden. Nach einer Pilotierungsphase wurden Durchführbarkeit und Wirksamkeit des Trainings in zwei Studien überprüft. In beiden Studien konnte eine hohe Durchführungstreue gezeigt werden. Hinsichtlich der Interventionseffekte fand sich über beide Studien hinweg ein ähnliches, positives Ergebnismuster. Es zeigten sich jedoch teilweise geringere Effekte, wenn das Training von Kursleitungen aus der Praxis und unter weniger stark kontrollierten Bedingungen implementiert wurde.

Basierend auf diesen Ergebnissen und aufgrund kleiner Stichproben in den vorausgegangenen Studien erschien es sinnvoll, zunächst eine zweite Effektivitätsstudie, als Replikationsstudie, anzuschließen, bevor Überlegungen hinsichtlich eines Scaling-up vorgenommen werden können (vgl. Gottfredson et al. 2015; Erdfelder und Ulrich 2018). Ziel der vorliegenden Studie ist es, die Ergebnisse zur Durchführbarkeit und Effektivität des Trainings zu bestätigen. Um die Ergebnisse möglichst gut vergleichen zu können, wurden der Trainingskontext, die Zielgruppe der Kinder, das Studiendesign sowie die Erfassung der Ziel- und Kontrollvariablen parallel gehalten (Erdfelder und Ulrich 2018). Zudem wurde das Präsentationstraining erneut von Kursleitungen aus der Praxis angeboten. Für das Trainingskonzept wurden teilweise Änderungen, basierend auf den Rückmeldungen der Kursleitungen, vorgenommen. Ein randomisiertes Wartekontrollgruppendesign (mit Randomisierung auf Individual- oder Kursebene) mit Prätest und Posttest wurde eingesetzt. Die Durchführungstreue wurde anhand von Kursleiterfragebögen erhoben. Die Präsentationsfähigkeiten wurden mittels Videoratings und die Sprechangst mittels Fragebogen erfasst.

Drei Fragestellungen wurden betrachtet:

1. Wird das Training, wie in den zuvor durchgeführten Studien, mit hoher Durchführungstreue implementiert?
2. Lassen sich die positiven Effekte auf die Präsentationsfähigkeiten (nonverbal, organisatorisch, Sprachgebrauch, Gesamtaufreten) der am Training teilnehmenden Kinder replizieren?
3. Lassen sich die negativen Effekte auf die Sprechangst replizieren?

4 Methode

4.1 Teilnehmerinnen und Teilnehmer

Das Präsentationstraining ist Teil eines Förderprogramms für besonders begabte und hochbegabte Grundschulkinder in Baden-Württemberg, den Hector Kinderakademien (<http://www.hector-kinderakademie.de>). Über ganz Baden-Württemberg hinweg werden an 65 Akademiestandorten (Stand Januar 2019) verschiedene, extracurriculare Enrichment-Kurse angeboten. Die Kurse werden dabei sowohl von Lehrkräften als auch Nicht-Lehrkräften mit unterschiedlichem beruflichem Hintergrund durchgeführt. An den Kursen können besonders begabte, interessierte, motivierte und kreative Kinder teilnehmen, die von ihren Lehrkräften für das Programm nominiert wurden.

Das Präsentationstraining wurde im Rahmen der Studie, welche im Schuljahr 2017/2018 stattfand, von 10 Kursleitungen angeboten (90% Frauen). Alle Kursleitungen willigten ein, an der Studie teilzunehmen. Im Durchschnitt waren sie 42,4 Jahre ($SD=9,72$) alt und hatten unterschiedliche berufliche Hintergründe, wobei acht Personen angaben, eine pädagogische Ausbildung zu haben. Das Präsentationstraining hatten sie im Mittel bereits 2,20 Mal angeboten ($SD=1,69$), wobei nur eine Person das Training zum ersten Mal gab. Weitere Informationen zum Hintergrund der Kursleitungen siehe Appendix A im Zusatzmaterial online.

Auch von den Eltern und Kindern wurde vor Studienstart die schriftliche Einwilligung zur Studienteilnahme eingeholt. Von den 65 Dritt- und Viertklässlern die teilnahmen, waren 48% Mädchen. Im Schnitt waren die Kinder 8,91 Jahre alt ($SD=0,65$). Ihr mittlerer verbaler IQ lag bei 112,27 ($SD=13,48$). Sie besuchten das Präsentationstraining an einer von 10 Akademien.

4.2 Ablauf der Studie

4.2.1 Vorbereitung der Implementation

Um hohe Durchführungstreue zu erzielen, wurde auf das bestehende Unterstützungsverfahren aufgebaut, das für die vorausgegangene Effektivitätsstudie entwickelt und seitdem kontinuierlich überarbeitet worden war. Ziel war es, das Wissen der Kursleitungen sowie ihre Akzeptanz in Bezug auf die Inhalte und Durchführung des Kurses zu erhöhen und gleichzeitig die persönlichen und zeitlichen Kosten zu reduzieren (vgl. Souvignier und Mokhesgerami 2006; Humphrey et al. 2016). Dementsprechend erhielten die Kursleitungen ein detailliertes Kursmanual (siehe Appendix B im Zusatzmaterial online) und nahmen an einer vorbereitenden Qualifizierungsveranstaltung teil. Die Kursleitungen, die den Kurs bereits in der Vergangenheit angeboten hatten, erhielten eine kürzere Schulung als die neue Kursleitung. Alle Schulungen wurden von der Kursentwicklerin durchgeführt. In der Veranstaltung für die neue Kursleitung wurde zunächst der theoretische Hintergrund zum Kurs und der Aufbau des Manuals vorgestellt. Dem folgte eine detaillierte Einführung in jede Kurseinheit, mit Vorstellung der Lernziele, Inhalte, Kurselemente und benötigten Materialien. Zuletzt wurde der Ablauf der Studie thematisiert. Die kürzere Schulung für die Kursleitungen, die den Kurs bereits angeboten hatten, startete mit der Vorstellung der inhaltlichen und methodischen Änderungen des Kurskonzeptes, die aufgrund der Studien und Rückmeldungen der Kursleitungen vorgenommen worden waren. Dem schloss sich der organisatorische Part, mit Informationen zum Studienablauf, an.

4.2.2 Studiendesign

Zur Überprüfung der Effektivität des Trainings war ein randomisiertes Wartekontrollgruppendesign mit Prä-Post-Messung geplant. Um die Kinder zufällig einer der beiden Gruppen zuzuteilen wurde innerhalb jeder Akademie blockweise, mittels computergenerierter Zufallszahlen, randomisiert. Damit wurde sichergestellt, dass innerhalb jeder Hector Kinderakademie gleich viele Kinder in beiden Gruppen waren (vgl. Moher et al. 2010). Hatten sich an einer Akademie weniger als acht Kinder für den Kurs angemeldet, wurde cluster-randomisiert und der ganze Kurs einer der Studienbedingungen zugeordnet, da die Mindestteilnehmerzahl pro Kurs vier war. Aufgrund der geringen Anmeldezahlen konnte lediglich für zwei Kurse eine Blockrandomisierung realisiert werden. Die anderen acht Kurse wurden cluster-randomisiert. Nach der Randomisierung waren 29 Kinder in der Interventionsgruppe (48% Mädchen, Durchschnittsalter 9,00 Jahre; $SD=0,58$) und 36 Kinder in der Kontrollgruppe (47% Mädchen, Durchschnittsalter 8,83 Jahre; $SD=0,70$).

Bevor die Kinder und Eltern darüber informiert wurden in welcher Gruppe sie waren, wurde der Prätest durchgeführt. Dieser fand gemeinsam für die Kinder beider Gruppen an der jeweiligen Akademie statt und bildete die erste Kurseinheit der Interventionsgruppe. Nach dem Prätest besuchte die Interventionsgruppe den Kurs über 10 Wochen. Der Posttest wurde erneut für alle Kinder einer Akademie zeitgleich durchgeführt und zwar am letzten Termin der Interventionsgruppe. Danach besuchte die Kontrollgruppe das Training.

4.3 Instrumente

Die Durchführungstreue wurde mittels Selbstbericht der Kursleitungen am Ende jeder Kurseinheit erfasst. Die abhängigen Variablen (Angemessenheit der Präsentationsfähigkeiten sowie Sprechangst) wurden beim Prä- und Posttest erfasst. Die Kontrollvariablen (figurale kognitive Fähigkeiten und verbale Intelligenz) wurden nur zum Prätest erfasst. Ergänzend und zur Erfassung der Reaktion der Kinder, schätzten die Kinder der Interventionsgruppe beim Posttest die Akzeptanz des Kurses ein.

4.3.1 Durchführungstreue

Jede Intervention besteht aus spezifischen Komponenten. Deshalb ist auch die Erfassung der Durchführungstreue jeweils verschieden (Abry et al. 2015). Da sowohl die Quantität als auch die Qualität der implementierten Kurselemente einen Einfluss auf die Entwicklung der an der Intervention teilnehmenden Kinder haben kann (Odom et al. 2010; Greene 2015), wurden für die Erfassung der Durchführungstreue des Präsentationstrainings zwei Dimensionen berücksichtigt. Erstens, die *Einhaltung* (d.h. inwieweit die einzelnen Kurselemente tatsächlich durchgeführt wurden) und zweitens, die *Qualität* (d.h. die Güte der Umsetzung für jedes Kurselement; vgl. Nelson et al. 2012; Humphrey et al. 2016).

Um die Durchführungstreue zu erfassen, füllten die Kursleitungen am Ende jeder 90-minütigen Kurseinheit einen Fragebogen aus. Für die Einhaltung gaben sie zunächst für jedes Kurselement an, ob sie es durchgeführt hatten (0 = *nein*, 1 = *ja*). Da die Kurselemente in ihrer Bedeutsamkeit für die Erreichung der Trainingsziele variierten, wurde jedes Element gewichtet (1 = *geringe Bedeutsamkeit* bis 3 = *hohe Bedeutsamkeit*). Die Gewichtung erfolgte durch die Kursentwicklerin in Abhängigkeit der umgesetzten Kernkomponenten. Dazu wurde für jedes Kurselement definiert, welche Kernkomponenten ihm zugrunde lagen. Ein Kurselement, das mehrere Komponenten wie z. B. Theorieinput, Modelllernen, praktisches Üben und Feedback verknüpft, wurde höher gewichtet als ein Element, das eine reine Wissensvermittlung beinhaltet. Ersteres trifft beispielsweise auf eine Übung zu, bei der die Kinder nacheinander vor der Gruppe sprechen mussten und dabei eine nonverbale Präsentationsfähigkeit, wie Gestik, in einer vorgegebenen Ausprägung verwenden mussten (praktisches Üben). Die anderen Kinder nahmen die Beobachterrolle ein, reflektierten was sie wahrgenommen hatten und welche Wirkung damit einherging (Modelllernen, Feedback und Theorie). Insgesamt bestand der Kurs aus 76 Kurselementen. Aufgrund der Gewichtung konnten maximal 161 Punkte für die Einhaltung erreicht

werden. Für die Einhaltung wurde der Prozentsatz der tatsächlich durchgeführten Elemente berichtet. Zur Erfassung der Qualität schätzten die Kursleitungen für jedes Kurselement die Umsetzungsgüte ein („Wie gut hat die Umsetzung geklappt?“; dreistufige Likert-Skala von 1 = *schlecht* bis 3 = *gut*).

4.3.2 Präsentationsfähigkeiten

Um die Präsentationsfähigkeiten der Kinder zu erfassen, erhielten sie beim Prä- und Posttest die gleiche, standardisierte Präsentationsaufgabe. Die resultierenden Präsentationen wurden jeweils auf Video aufgezeichnet. Die Kinder hatten die Aufgabe, einen Wissenschaftler oder eine Wissenschaftlerin zu malen (Chambers 1983) und zu überlegen, was er oder sie den ganzen Tag macht. Zudem sollten sie überlegen, was sie selbst gerne erforschen würden. Sie hatten 15 min Zeit, um auf einem vorstrukturierten Blatt Papier ihre Ideen aufzumalen und/oder etwas zu schreiben. Sie wurden außerdem darüber informiert, dass sie nach der Vorbereitung ihre Ideen in einem separaten Raum einem jungen Erwachsenen vorstellen würden und, dass sie dabei gefilmt werden. Nach der Vorbereitung im Gruppensetting wurden die Kinder zufällig einer von zwei Testleitungen zugeteilt. Dieser präsentierten sie dann einzeln und in randomisierter Reihenfolge ihr Bild.

Die gefilmten Präsentationen wurden genutzt, um mittels Videoratings die Angemessenheit der Präsentationsfähigkeiten, d. h. Passung der gezeigten Präsentationsfähigkeiten zum Kontext, zu erfassen. Dazu wurde eine leicht adaptierte Version des Beobachtungsbogens verwendet, der bereits in den vorherigen Studien zur Überprüfung der Wirksamkeit des Präsentationstrainings eingesetzt worden war (Herbein et al. 2018b). Die 35 Items des Bogens (z. B. „Die Gestik ist situativ angemessen.“) erfassen Präsentationsfähigkeiten aller vier Dimensionen der PK: nonverbal-visuelle (6 Items), nonverbal-auditive (15 Items) und organisatorische Fähigkeiten (8 Items), Sprachgebrauch (5 Items) und ein Item zum Gesamtauftreten. Die Angemessenheit wurde auf einer vierstufigen Likert-Skala (1 = *stimme nicht zu* bis 4 = *stimme zu*) eingeschätzt. Vor den Ratings wurden die drei Rater in der Anwendung des Bogens und der Einschätzung der Präsentationsfähigkeiten geschult. Danach schätzten alle Rater alle 111 Videos der Studie (60 Prätest-, 51 Posttestvideos, mittlere Dauer = 57 s, $SD = 22,44$) in einer unterschiedlichen, randomisierten Reihenfolge ein. Aufgrund von Abwesenheiten im Kurs lagen nicht für alle 65 Kinder Videos zum Prä- und zum Posttest vor.

Zur Berechnung der Interrater Reliabilität wurde die Intraklassenkorrelation $ICC_{unjustiert, MW}$ verwendet (Wirtz und Caspar 2002). Achtzehn Items hatten eine gute (zwischen 0,60 und 0,74) oder exzellente ICC ($\geq 0,75$; Cicchetti 1994) und wurden für die Analyse der Interventionseffekte verwendet. Vier der Items erfassen nonverbal-visuelle (Körperhaltung, Gestik, Blickkontakt, Raumverhalten), 5 nonverbal-auditive (Sprechatmung, Stimme, Artikulation, Lautstärke, Sprechflüssigkeit) und 6 organisatorische Fähigkeiten (Sprechdauer, Länge der Einleitung, Länge des Schlusses, Informationsmenge, Kommunikationsabsicht, Hörerbezug), 2 den Sprachgebrauch (persönliche Ansprache, Stilmittelverwendung) und eines

das Gesamtaufreten. Für eine Übersicht der ICCs siehe Tab. 1. Alle Items mit einer ICC < 0,60 wurden aus den Analysen ausgeschlossen.¹

4.3.3 Sprechangst

Zur Erfassung der Sprechangst wurde eine adaptierte Version des Bühnenangstfragebogens (Fehm und Hille 2005) eingesetzt. Dieser bestand aus 20 Items und erfasste körperliche (10 Items, z. B. „Ich habe Herzklopfen.“) und kognitive Symptome (10 Items, z. B. „Ich mache mir Sorgen über meinen Vortrag.“) der Sprechangst. Die Kinder beantworteten den Fragebogen auf einer fünfstufigen Likert-Skala (1 = *gar nicht* bis 5 = *sehr*). Dabei wurden sie gebeten, allgemein an Vortragssituationen zu denken. Sie füllten den Fragebogen vor der praktischen Präsentation aufgabe aus. Cronbachs Alpha reichte von 0,80–0,89 (Tab. 1).

4.3.4 Akzeptanz

Ergänzend wurde zum Posttest die Kursakzeptanz (d.h. Erfahrung bezüglich der Inhalte und der Komplexität des Kurses; Petermann 2014) aus Sicht der Kinder erfasst. Dazu beantwortete die Interventionsgruppe auf jeweils einem Item Fragen zur Kursschwierigkeit („Im Kurs bin ich gut mitgekommen.“), zum Spaß an der Kursteilnahme („Der Kurs hat mir Spaß gemacht.“) und zum Neuigkeitsgehalt der Trainingsinhalte („Im Kurs habe ich viel gelernt.“). Zusätzlich wurde mittels 4 Items von Bos et al. (2004) die Mitarbeit im Kurs erfasst (Beispielitem „Im Kurs arbeitete ich meistens intensiv mit.“; $\alpha=0,50$). Alle Items wurden auf einer vierstufigen Likert-Skala (1 = *stimmt gar nicht* bis 4 = *stimmt genau*) beantwortet. Als letztes vergaben die Kinder noch eine Note für den Kurs (von 1–6).

4.3.5 Kontrollvariablen

Als Kontrollvariablen wurden verbale Intelligenz und figurale kognitive Fähigkeiten erfasst, da das Training als Teil eines Enrichment-Angebots für besonders begabte und hochbegabte Kinder angeboten wurde. Vorausgegangene Studien an den Hector Kinderakademien hatten gezeigt, dass die Intelligenz der Kinder einen Faktor für deren Nominierung darstellt (vgl. Rothenbusch et al. 2016). Verbale Intelligenz wurde mit den 30 Items des Wortschatztests des CFT 20-R (Weiß 2006) erfasst ($\alpha=0,78$). Zur Erfassung der figuralen kognitiven Fähigkeiten wurde eine Subskala der Kurzversion des Berliner Test zur Erfassung fluider und kristalliner Intelligenz verwendet (BEFKI; Schroeders et al. 2016), die aus 16 Items besteht ($\alpha=0,70$).

¹ Zusätzlich zur Interrater Reliabilität wurde mittels Fleiss Kappa auch die Übereinstimmung zwischen den Ratern überprüft. Außer für das Item „Der Grad der persönlichen Ansprache ist situativ angemessen.“ (Dimension Sprachgebrauch; $\kappa=0,63$) zeigten sich zwischen den Ratern keine substantiellen oder besseren Übereinstimmungswerte ($\kappa \geq 0,61$; Landis und Koch 1977).

Tab. 1 Mittelwerte und Standardabweichungen aller Variablen zum Prä- und Posttest für beide Gruppen

	Prätest		Posttest		Items	α t1/t2	ICC					
	Präsentationstraining		Präsentationstraining									
	M	SD	M	SD								
AV (Videorating)												
<i>Nonverbal-visuell</i>												
Körperhaltung	3,11	0,89	2,92	0,82	3,41	0,61	2,84	0,81	1	-	0,79	
Gestik	2,67	0,78	2,68	0,82	2,94	0,83	2,78	0,90	1	-	0,80	
Blickkontakt	3,10	0,67	2,67	0,63	3,06	0,67	2,93	0,71	1	-	0,74	
Raumverhalten	3,84	0,37	3,73	0,50	3,98	0,07	3,88	0,25	1	-	0,74	
<i>Nonverbal-auditiv</i>												
Sprechatmung	3,44	0,55	3,55	0,50	3,41	0,56	3,47	0,53	1	-	0,63	
Stimme	3,68	0,53	3,67	0,36	3,76	0,35	3,77	0,33	1	-	0,61	
Artikulation	3,73	0,49	3,74	0,43	3,79	0,41	3,77	0,31	1	-	0,68	
Lautstärke	3,57	0,55	3,69	0,37	3,70	0,39	3,60	0,58	1	-	0,69	
Sprechflussigkeit	3,29	0,50	3,12	0,54	3,25	0,52	3,33	0,52	1	-	0,71	
<i>Organisation</i>												
Sprechdauer	3,15	0,87	3,17	0,71	3,25	0,74	3,26	0,64	1	-	0,89	
Länge der Einleitung	1,27	0,38	1,30	0,61	1,83	1,05	1,33	0,61	1	-	0,85	
Länge des Schlusses	1,09	0,18	1,26	0,60	1,29	0,59	1,24	0,42	1	-	0,82	
Informationsmenge	2,78	0,72	2,96	0,65	3,21	0,54	3,17	0,52	1	-	0,71	
Kommunikationsabsicht	2,77	0,34	2,96	0,31	3,16	0,45	3,02	0,21	1	-	0,66	
Hörerbezug	3,26	0,52	3,29	0,55	3,48	0,43	3,27	0,48	1	-	0,61	
<i>Sprachgebrauch</i>												
Persönliche Ansprache	1,06	0,16	1,07	0,25	1,37	0,49	1,10	0,26	1	-	0,82	
Stilmittelverwendung	1,70	0,59	1,71	0,59	2,00	0,75	1,87	0,68	1	-	0,75	
<i>Gesamtaufreiten</i>	3,01	0,46	3,08	0,43	3,32	0,48	3,21	0,48	1	-	0,76	

Tab. 1 (Fortsetzung)

	Prätest		Posttest				Items	α t1/t2	ICC		
	Präsentationstraining		Kontrollgruppe		Präsentationstraining					Kontrollgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				<i>M</i>	<i>SD</i>
AV (Selbstberichte)											
Sprechangst kognitiv	2,24	0,69	2,24	0,65	2,19	0,74	2,07	0,71	10	0,85/0,89	-
Sprechangst körperlich	1,82	0,73	1,66	0,57	1,79	0,74	1,64	0,68	10	0,80/0,86	-
Kontrollvariablen											
Figurale kognitive Fähigkeiten	9,83	2,59	9,90	3,34	-	-	-	-	16	0,70	-
Verbale Intelligenz	111,85	14,86	112,61	12,46	-	-	-	-	30	0,78	-

α t1/t2 Cronbachs Alpha für Prä-/Posttest

ICC Intraklassenkorrelation ICC_{unjustiert,MW} (für alle 111 Videos und drei Rater)

4.4 Fehlende Werte

Fehlende Werte entstanden durch Abwesenheit der Kinder zum Prä- oder Posttest, z. B. wegen Erkrankung. Zusätzlich konnten die figuralen kognitiven Fähigkeiten an einer Akademie aus Zeitgründen nicht mehr erfasst werden. Insgesamt fehlen für den Prätest die Werte von 8 % der Kinder auf allen abhängigen Variablen, für insgesamt 18 % fehlen die Daten für die figuralen kognitiven Fähigkeiten. Zum Posttest fehlen die Werte von 22 % der Kinder. Das Ausmaß der fehlenden Werte unterschied sich nicht zwischen den Gruppen, weder für die Prätest- noch die Posttestwerte (alle p -Werte $> 0,05$). Vergleicht man die Gruppenmittelwerte der abhängigen Variablen zum Prätest zwischen den Kindern, die am Posttest teilnahmen, und denen, die am Posttest fehlten, dann zeigen sich keine signifikanten Mittelwertunterschiede (alle p -Werte $> 0,05$). Basierend auf diesen Ergebnissen ist die Annahme, dass die fehlenden Werte zufällig entstanden (missing at random), plausibel. Deshalb wurde für die Analysen das in Mplus integrierte Full Information Maximum Likelihood-Verfahren (FIML) zu Schätzung der fehlenden Werte (Enders 2010) verwendet.

4.5 Statistische Analysen

Trotz Randomisierung wurde zunächst überprüft, ob zum Prätest signifikante Unterschiede zwischen der Interventions- und der Kontrollgruppe auf den abhängigen Variablen sowie Geschlecht, Alter, Klassenstufe, verbale Intelligenz und figurale kognitive Fähigkeiten bestanden. Zur Überprüfung der Interventionseffekte wurden multiple lineare Regressionen in Mplus 7 (Muthén und Muthén 1998–2012), unter Verwendung des robusten Maximum-Likelihood-Schätzverfahrens (MLR), gerechnet. Die Effekte auf die Präsentationsfähigkeiten (18 Variablen; Videoratings) und Sprechangst (2 Variablen) wurden überprüft. Pro Zielvariable (Posttestwert) wurde ein separates Model gerechnet. Für jede Präsentationsfähigkeit wurde dazu der Mittelwert der Einschätzungen der Rater gebildet. Zusätzlich wurden die Prätestwerte als Prädiktor aufgenommen, um die Teststärke zu erhöhen (Cohen et al. 2003). Zur Überprüfung differentieller Effekte für Kinder mit hohen versus niedrigen Präsentationsfähigkeiten/Sprechangst vor dem Training, wurde die Interaktion zwischen Prätestwert und Treatment als weiterer Prädiktor im Model aufgenommen. Aufgrund der Charakteristika der Zielgruppe wurden ergänzend verbale Intelligenz und figurale kognitive Fähigkeiten als Kontrollvariablen aufgenommen. Vor den Analysen wurden alle kontinuierlichen Variablen (Prä- und Posttestwerte) standardisiert. Die Gruppenzugehörigkeit wurde dummy-kodiert (1 = *Interventionsgruppe*; 0 = *Kontrollgruppe*).

Insgesamt wurden 20 Modelle gerechnet, um die Hypothesen zu testen. Aufgrund der Anzahl der Tests wurde ein strengerer p -Wert verwendet. Dazu wurde die Benjamini-Hochberg Korrektur angewandt um die false discovery rate zu kontrollieren (Benjamini und Hochberg 1995). Es wurde einseitig getestet ($\alpha = 0,05$), da gerichtete Hypothesen formuliert worden waren. Da die abhängige Variable standardisiert wurde, kann der Regressionskoeffizient als Effektstärke interpretiert werden. Parallel zu den vorherigen Studien und da, soweit uns bekannt, keine vergleichbaren Interventionsstudien zur Effektivität von Präsentationstrainings für Grundschul Kinder

existieren, verwendeten wir zur Interpretation der Effektstärken die Klassifikation nach Cohen (1992). Dem folgend ist eine Effektstärke von $d=0,20$ als klein, $d=0,50$ als moderat und $d=0,80$ als groß zu bezeichnen.

5 Ergebnisse

Die deskriptiven Statistiken sind in Tab. 1 dargestellt. Für die abhängigen Variablen fanden sich zum Prätest keine signifikanten Unterschiede zwischen den Gruppen, außer für Blickkontakt ($t(58)=-2,54$; $p=0,014$) und Kommunikationsabsicht ($t(58)=2,30$; $p=0,025$). Die Korrelationen der abhängigen Variablen zum Prä- und Posttest finden sich in Tab. 2. Die Rückmeldungen der Kinder zur Akzeptanz des Kurses waren $M=3,95$ ($SD=0,22$) für Neuigkeitsgehalt, $M=3,90$ ($SD=0,30$) für Schwierigkeit und $M=3,86$ ($SD=0,36$) für Spaß an der Kursteilnahme. Die Mitarbeit lag im Mittel bei $3,56$ ($SD=0,40$). Als Note vergaben 81 % der Kinder eine 1, 19 % eine 2. Für die weiteren Ergebnisse wird zunächst die Durchführungstreue berichtet. Danach folgen die Interventionseffekte auf die Präsentationsfähigkeiten und die Sprechangst.

5.1 Durchführungstreue

Zur Überprüfung der Hypothese 1, dass der Kurs mit hoher Durchführungstreue implementierbar ist, wurden Einhaltung und Qualität erfasst. Für die Einhaltung zeigte sich ein hoher Wert. Über alle Kursleitungen und Kurseinheiten hinweg wurden im Mittel 86,71 % ($SD=9,18$) der Kurselemente durchgeführt. Die mittlere Qualität lag bei $2,73$ ($SD=0,50$; Tab. 3) über alle Kursleitungen und Elemente. Die Einhaltung und Qualität der Durchführung waren demnach hoch ausgeprägt, was eine Analyse der Interventionseffekte im nächsten Schritt erlaubte.

5.2 Effekte auf die Angemessenheit der Präsentationsfähigkeiten

Für Hypothese 2 wurde angenommen, dass Kinder, die am Präsentationstraining teilgenommen hatten, angemessenere, das heißt zum Kontext passendere, Präsentationsfähigkeiten zeigen als die Kinder der Kontrollgruppe. Für die nonverbalen Präsentationsfähigkeiten zeigte sich, nach Adjustierung der p -Werte, ein signifikanter Interventionseffekt auf Körperspannung ($B=0,66$; $p=0,035$). Keine signifikant angemesseneren Präsentationsfähigkeiten zeigten die Kinder der Interventionsgruppe, im Vergleich zur Kontrollgruppe, für die nonverbal-visuellen Fähigkeiten Gestik, Blickkontakt und Raumverhalten (Tab. 4) sowie für die nonverbal-auditiven Fähigkeiten Sprechatmung, Stimme, Artikulation, Lautstärke und Sprechflüssigkeit (Tab. 5; alle p -Werte $> 0,05$).

Für die organisatorischen Präsentationsfähigkeiten fanden sich keine signifikanten Interventionseffekte (Sprechdauer, Länge der Einleitung, Länge des Schlusses, Informationsmenge, Kommunikationsabsicht und Hörerbezug; alle p -Werte $> 0,05$; Tab. 6). Für die beiden Items zum Sprachgebrauch fand sich ein signifikanter Effekt auf persönliche Ansprache ($B=0,68$; $p=0,020$) und kein signifikanter Effekt

Tab. 2 Korrelationen zwischen den abhängigen Variablen zum Prätest (unterhalb der Diagonalen) und Posttest (oberhalb der Diagonalen)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1) Körperhaltung	-	0,32*	0,40*	0,25	0,25	0,11	0,03	0,07	0,13	-0,06	0,31*	0,10	0,04	0,23	0,19	0,34*	0,21	0,32*	-0,01	0,06
(2) Gestik		-	0,13	0,18	0,39*	0,48*	0,36*	0,38*	0,26	0,38*	0,27	0,06	0,37*	0,18	0,37*	0,14	0,35*	0,67*	-0,11	0,01
(3) Blickkontakt			-	0,27	0,13	0,31*	0,14	0,13	0,46*	-0,08	0,23	0,20	0,19	0,19	0,30*	0,23	0,24	0,41*	-0,08	-0,09
(4) Raumverhalten				-	-0,01	0,18	0,09	0,11	0,10	0,03	0,14	0,05	0,16	0,19	0,14	0,12	0,20	0,23	0,12	0,21
(5) Sprechatmung					-	0,44*	0,20	0,22	0,38*	-0,03	-0,06	-0,16	0,07	-0,17	0,02	-0,02	0,15	0,31*	-0,05	0,02
(6) Stimme						-	0,68*	0,78*	0,37*	0,31*	0,29*	0,02	0,29*	0,28*	0,29*	0,19	0,36*	0,51*	-0,03	0,11
(7) Artikulation							-	0,57*	0,24	0,16	0,28*	0,14	0,09	0,28*	0,22	0,17	0,21	0,36*	0,14	0,22
(8) Lautstärke								-	0,23	0,26	0,26	-0,03	0,18	0,27	0,21	0,23	0,28*	0,44*	-0,06	0,01
(9) Sprechflüssigkeit									-	0,01	0,14	0,11	0,22	0,16	0,24	0,19	0,38*	0,57*	0,09	0,10
(10) Sprechdauer										-	0,13	0,16	0,61*	0,33*	0,40*	0,09	0,39*	0,43*	-0,06	0,13

Tab. 2 (Fortsetzung)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(11) Länge der Einleitung	0,03	0,34*	-0,18	-0,27*	0,08	0,13	-0,27*	0,08	-0,11	0,20	-	0,34*	0,16	0,70*	0,20	0,79*	0,33*	0,33*	0,09	0,01
(12) Länge des Schlusses	0,12	0,24	-0,07	-0,07	0,03	-0,01	-0,24	0,07	0,14	0,12	0,26*	-	0,37*	0,37*	0,26	0,34*	0,34*	0,40*	0,23	0,13
(13) Informationsmenge	0,05	0,33*	0,13	0,05	-0,08	0,21	0,29*	0,45*	0,18	0,71*	0,14	0,10	-	0,45*	0,67*	0,03	0,69*	0,64*	0,01	0,01
(14) Kommunikationsabsicht	-0,05	0,39*	-0,04	-0,19	-0,12	0,12	-0,02	0,24	0,06	0,40*	0,36*	0,26*	0,57*	-	0,42*	0,48*	0,45*	0,43*	0,01	-0,02
(15) Hörerbezug	0,22	0,34*	0,26*	0,11	0,01	0,17	0,36*	0,43*	0,25	0,57*	0,11	0,18	0,61*	0,49*	-	0,11	0,61*	0,63*	-0,07	0,03
(16) Persönliche Ansprache	0,12	0,31*	-0,04	-0,04	0,14	0,02	-0,24	0,04	-0,04	0,03	0,60*	0,68*	-0,05	0,20	0,06	-	0,24	0,29*	0,23	0,17
(17) Stimmittelverwendung	0,14	0,38*	0,18	0,03	0,03	0,15	0,08	0,38*	0,22	0,47*	0,25	0,52*	0,55*	0,45*	0,54*	0,28*	-	0,70*	0,05	0,08
(18) Gesamtaufreten	0,20	0,58*	0,33*	0,09	0,17	0,45*	0,28*	0,56*	0,43*	0,53*	0,22	0,38*	0,66*	0,54*	0,67*	0,23	0,73*	-	0,02	0,08
(19) Sprechangst kognitiv	0,07	0,04	0,08	0,05	-0,14	-0,12	-0,01	-0,02	0,08	0,01	-0,08	0,19	-0,07	0,15	0,03	0,21	0,06	0,06	-	0,79*
(20) Sprechangst körperlich	0,09	0,11	0,27*	0,02	-0,16	-0,01	-0,01	-0,10	0,07	0,03	-0,09	0,03	-0,03	0,09	0,02	0,07	-0,01	0,04	0,71*	-

Die Variablen 1–18 wurden mittels Videoratings erfasst, Variablen 19 und 20 basieren auf Selbstberichten der Kinder

* $p < 0,05$

Tab. 3 Durchführungstreue: Einhaltung und Qualität der Durchführung

Kurseinheit	Kursleitung																		
	A			B			C			D			E			Gesamt			
	Einh	Qualität	Einh	Qualität	Einh	Qualität	Einh	Qualität	Einh	Qualität	Einh	Qualität	Einh	Qualität	Einh	Qualität	Einh	Qualität	
%	(SD)	M	(SD)	%	(SD)	M	(SD)	%	(SD)	M	(SD)	%	(SD)	M	(SD)	%	(SD)	M	(SD)
2	100	2,64	(0,67)	67	2,75	(0,46)	92	2,40	(1,26)	75	3,00	(0)	67	2,33	(0,82)	80,00	(15,14)	2,62	(0,27)
3	100	2,50	(0,53)	100	2,88	(0,35)	89	3,00	(0)	100	2,88	(0,35)	68	3,00	(0)	91,58	(13,72)	2,86	(0,17)
4	92	3,00	(0)	100	2,71	(0,76)	100	3,00	(0)	85	2,80	(0,45)	92	2,33	(0,52)	93,85	(6,44)	2,77	(0,27)
5	84	2,71	(0,49)	100	3,00	(0)	100	3,00	(0)	89	2,71	(0,49)	68	3,00	(0)	88,42	(13,11)	2,89	(0,16)
6	88	2,86	(0,38)	88	2,71	(0,76)	100	3,00	(0)	81	3,00	(0)	50	2,00	(0)	81,25	(18,75)	2,71	(0,42)

Tab. 3 (Fortsetzung)

Kurseinheit	Kursleitung																			
	A			B			C			D			E			Gesamt				
	Einh	Qualität																		
%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	
7	100	2,67 (0,50)	45	2,43 (0,53)	100	3,00 (0)	85	2,86 (0,35)	95	2,43 (0,53)	85,00	2,68 (0,26)	100	2,90 (0,32)	100	3,00 (0)	90	2,38 (0,74)	95,86	2,82 (0,26)
8	100	2,90 (0,32)	100	3,00 (0)	100	2,82 (0,40)	90	3,00 (0)	90	2,38 (0,74)	95,86	2,82 (0,26)	100	2,40 (0,55)	77	3,00 (0)	62	1,50 (0,71)	69,23	2,34 (0,61)
9	92	2,40 (0,55)	77	3,00 (0)	77	2,80 (0,45)	62	2,00 (0)	38	1,50 (0,71)	69,23	2,34 (0,61)	100	3,00 (0)	100	3,00 (0)	88	1,67 (0,58)	90,00	2,73 (0,60)
10	100	3,00 (0)	75	3,00 (0)	100	3,00 (0)	88	3,00 (0)	88	3,00 (0)	20,35	2,73 (0,60)	96	2,73 (0,48)	84	2,82 (0,47)	84	2,4 (0,66)	86,71	2,73 (0,50)
Gesamt	96	2,73 (0,48)	84	2,82 (0,47)	96	2,87 (0,55)	84	2,85 (0,34)	74	2,4 (0,66)	86,71	2,73 (0,50)	—	—	—	—	—	—	—	—

Kurseinheiten 1 und 11 fehlen, da diese den Prä- und Posttest bildeten. Für fünf Kursleitungen liegen keine Daten vor, da aufgrund zu geringer Anmeldezahlen zum Kurs, der gesamte Kurs zufällig der Kontrollbedingung zugeordnet worden war

Einh., Einhaltung (d. h. Prozent der durchgeführten Kurselemente), *Qualität* Güte der Umsetzung

Tab. 4 Interventionseffekte auf die nonverbal-visuellen Präsentationsfähigkeiten (Posttest)

	Nonverbal-visuell											
	Körperspannung		Gestik		Blickkontakt		Raumverhalten					
	<i>B</i>	(<i>SE</i>)	<i>p</i>	<i>B</i>	(<i>SE</i>)	<i>p</i>	<i>B</i>	(<i>SE</i>)	<i>p</i>			
Training	0,66	0,24	0,035	0,21	0,23	0,345	-0,05	0,25	0,569	0,43	0,20	0,085
Prätest Wert	0,40	0,19	0,038	0,73	0,13	<0,001	0,35	0,24	0,135	0,10	0,17	0,570
Training × Prätest Wert	-0,30	0,31	0,321	-0,29	0,24	0,243	0,01	0,35	0,969	-0,17	0,21	0,418
Figurale kognitive Fähigkeiten	0,07	0,15	0,669	0,02	0,08	0,764	0,01	0,17	0,978	0,38	0,23	0,096
Verbale Intelligenz	0,18	0,12	0,119	0,04	0,12	0,706	0,15	0,17	0,377	-0,03	0,12	0,813
Erklärte Varianz (R^2)	0,25	-	-	0,48	-	-	0,13	-	-	0,20	-	-

Vor den Analysen wurden alle kontinuierlichen Variablen standardisiert

Training wurde dummy-kodiert (1 = Interventionsgruppe, 0 = Kontrollgruppe)

Für Training wurde einseitig getestet

p adjustierter *p*-Wert

Tab. 5 Interventionseffekte auf die nonverbal-auditiven Präsentationsfähigkeiten (Posttest)

	Nonverbal-auditiv														
	Sprechatmung			Stimme			Artikulation			Lautstärke			Sprechflüssigkeit		
	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p
Training	-0,04	0,23	0,544	-0,03	0,22	0,519	0,03	0,24	0,471	0,23	0,21	0,341	-0,37	0,23	0,167
Prätest Wert	0,33	0,19	0,089	0,93	0,20	<0,001	0,23	0,13	0,068	0,75	0,28	0,007	0,34	0,14	0,017
Training × Prätest Wert	0,06	0,28	0,827	-0,29	0,28	0,307	0,57	0,33	0,082	-0,26	0,28	0,355	0,26	0,22	0,235
Figurale kognitive Fähigkeiten	0,19	0,12	0,118	0,14	0,11	0,179	0,03	0,11	0,804	0,23	0,10	0,019	0,17	0,10	0,080
Verbale Intelligenz	0,14	0,14	0,328	0,08	0,10	0,405	-0,04	0,10	0,710	-0,01	0,11	0,936	0,40	0,11	<0,001
Erklärte Varianz (R ²)	0,25	-	-	0,53	-	-	0,34	-	-	0,45	-	-	0,40	-	-

Vor den Analysen wurden alle kontinuierlichen Variablen standardisiert
 Training wurde dummy-kodiert (1 = Interventionsgruppe, 0 = Kontrollgruppe)
 Für Training wurde einseitig getestet
 p adjustierter p-Wert

Tab. 6 Interventionseffekte auf die organisatorischen Präsentationsfähigkeiten (Posttest)

	Organisatorische Präsentationsfähigkeiten																	
	Sprechdauer			Länge der Einleitung			Länge des Schlusses			Informationsmenge			Kommunikationsabsicht			Hörerbezug		
	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p
Training	0,03	0,27	0,464	0,69	0,29	0,053	0,17	0,26	0,477	0,13	0,25	0,496	0,61	0,32	0,114	0,42	0,27	0,173
Prätest Wert	0,43	0,17	0,011	0,43	0,07	<0,001	0,33	0,16	0,037	0,36	0,15	0,016	0,42	0,10	<0,001	0,29	0,16	0,063
Training x Prätest Wert	-0,16	0,32	0,606	-0,36	0,66	0,583	-0,35	0,54	0,510	0,03	0,22	0,895	-0,12	0,23	0,609	-0,18	0,29	0,534
Figurale kognitive Fähigkeiten	-0,34	0,13	0,008	-0,30	0,11	0,006	-0,08	0,14	0,589	-0,09	0,15	0,524	-0,05	0,10	0,585	-0,05	0,12	0,690
Verbale Intel- ligenz	0,32	0,13	0,013	0,41	0,13	0,001	0,27	0,19	0,155	0,25	0,14	0,077	0,39	0,13	0,003	0,20	0,13	0,117
Erklärte Varianz (R ²)	0,30	-	-	0,43	-	-	0,17	-	-	0,22	-	-	0,32	-	-	0,14	-	-

Vor den Analysen wurden alle kontinuierlichen Variablen standardisiert
 Training wurde dummy-kodiert (1 = *Interventionsgruppe*, 0 = *Kontrollgruppe*)
 Für Training wurde einseitig getestet
 p adjustierter p-Wert

auf Stilmittelverwendung ($p=0,531$; Tab. 7). Auch für das Gesamtaufreten fand sich kein signifikanter Effekt ($p=0,344$; Tab. 7). Insgesamt waren die Kinder der Interventionsgruppe nach dem Besuch des Trainings besser in der Lage, ihre Körperspannung angemessen einzusetzen. Zudem verwendeten sie eine zum Kontext passendere persönliche Ansprache der Zuhörer als die Kinder der Kontrollgruppe. Auf allen anderen Präsentationsfähigkeiten zeigten die Kinder der Interventionsgruppe kein signifikant angemesseneres Verhalten. Der Klassifikation von Cohen folgend ist die Stärke der Interventionseffekte bei zehn der 18 Variablen als klein oder moderat einzuordnen. Moderate Effekte fanden sich für die vier Präsentationsfähigkeiten Körperspannung ($d=0,66$), Länge der Einleitung ($d=0,69$), Kommunikationsabsicht ($d=0,61$) und persönliche Ansprache ($d=0,68$). Für die sechs Fähigkeiten Gestik ($d=0,21$), Raumverhalten ($d=0,43$), Lautstärke ($d=0,23$), Sprechflüssigkeit ($d=-0,37$), Hörerbezug ($d=0,42$) und Gesamtaufreten ($d=0,22$) fanden sich kleine Effekte.

Bezüglich der differentiellen Effekte des Trainings in Abhängigkeit der Prätestwerte der Kinder, fand sich eine signifikante Interaktionen zwischen Training und Prätestwert für persönliche Ansprache ($B=-0,85$; $p=0,003$). Das heißt Kinder, die vor dem Training eine weniger angemessene persönliche Ansprache in ihrer Präsentation einsetzten, profitierten stärker vom Training.

5.3 Effekte auf Sprechangst

Hypothese 3 beinhaltete, dass Kinder, die am Training teilgenommen hatten, geringere Sprechangst berichten als Kinder der Kontrollgruppe. Es zeigten sich keine signifikanten Effekte auf Sprechangst, weder auf die kognitiven Symptome ($B=0,08$; $p=0,488$) noch auf die körperlichen ($B=-0,03$; $p=0,492$; Tab. 7). Kinder der Interventionsgruppe berichteten keine geringere Sprechangst als Kinder der Kontrollgruppe. Für Sprechangst wurde keine signifikante Interaktion zwischen Prätestwert und Training gefunden.

6 Diskussion

Die Studie ist die vierte eines Forschungsprogramms in der die Durchführbarkeit und Effektivität eines Präsentationstrainings untersucht wurde. Sie ist eine Replikationsstudie der vorausgegangenen Effektivitätsstudie, d. h. das Training wurde erneut von geschulten Kursleitungen aus der Praxis angeboten. Diese erhielten eine Schulung sowie ein Kursmanual um die Durchführungstreue zu erhöhen. Um sich den Bedingungen einer direkten Replikation möglichst gut anzunähern wurden zudem die Zielgruppe, der Trainingskontext sowie das Design der Studie und die Instrumente zur Erfassung der Zielvariablen parallel gehalten (Erdfelder und Ulrich 2018). Zur Überprüfung der Interventionseffekte wurde ein cluster-randomisiertes Design mit Wartekontrollgruppe und Prä- und Posttest verwendet. Die Angemessenheit der Präsentationsfähigkeiten wurde mittels Videoratings und Sprechangst mit Fragebögen erfasst.

Tab. 7 Interventionseffekte auf Sprachgebrauch und Gesamtaufreten sowie Sprechangst (Posttest)

	Sprachgebrauch			Gesamtaufreten			Sprechangst								
	Persönliche Ansprache			Stilmittelverwendung			Kognitive Symptome			Körperliche Symptome					
	<i>B</i>	(<i>SE</i>)	<i>p</i>	<i>B</i>	(<i>SE</i>)	<i>p</i>	<i>B</i>	(<i>SE</i>)	<i>p</i>	<i>B</i>	(<i>SE</i>)	<i>p</i>			
Training	0,68	0,22	0,020	0,07	0,22	0,531	0,22	0,22	0,344	0,08	0,17	0,488	-0,03	0,20	0,492
Prätest Wert	0,43	0,05	<0,001	0,36	0,13	0,007	0,42	0,10	<0,001	0,86	0,10	<0,001	0,80	0,13	<0,001
Training x Prätest Wert	-0,85	0,29	0,003	0,19	0,20	0,341	0,16	0,23	0,471	0,02	0,15	0,915	-0,10	0,18	0,585
Figurale kognitive Fähigkeiten	-0,12	0,08	0,159	0,02	0,12	0,889	0,17	0,10	0,076	-0,00	0,10	0,966	-0,06	0,07	0,430
Verbale Intelligenz	0,37	0,12	0,003	0,33	0,12	0,004	0,30	0,11	0,008	0,05	0,09	0,609	0,17	0,13	0,188
Erklärte Varianz (<i>R</i> ²)	0,43	-	-	0,38	-	-	0,47	-	-	0,67	-	-	0,61	-	-

Sprechangst wurde mittels Selbstberichte der Kinder erfasst
 Vor den Analysen wurden alle kontinuierlichen Variablen standardisiert
 Training wurde dummy-kodiert (1 = *Interventionsgruppe*, 0 = *Kontrollgruppe*)
 Für Training wurde einseitig getestet
p adjustierter *p*-Wert

6.1 Durchführungstreue

Um die Effektivität einer Intervention überprüfen zu können ist es notwendig, dass Kursleitungen die Intervention wie intendiert durchführen (Petermann 2014). In der vorliegenden Studie wurden die Quantität (Einhaltung) und die Qualität der eingesetzten Kurselemente erfasst. Für die Einhaltung berichteten die Kursleitungen, dass sie im Mittel 86,71 % ($SD=9,18$) der Kurselemente eingesetzt haben. Setzt man das Ergebnis mit der Einhaltung in anderen Studien in Bezug, dann ist es als hoch einzuschätzen (vgl. Durlak und DuPre 2008). Zudem ist das Ergebnis vergleichbar zur Einhaltung in der vorherigen Effektivitätsstudie ($M=85,33\%$; Herbein et al. 2018b). Für die Qualität fanden sich mittlere bis hohe Werte in der aktuellen Studie. Dabei sind sie, im Vergleich zur Effektivitätsstudie 1, tendenziell etwas höher.

Um nach der ersten Effektivitätsstudie 2015 die Durchführungstreue für einzelne Kurselemente zu erhöhen, wurde das Kurskonzept, basierend auf den Rückmeldungen der Kursleitungen, leicht adaptiert. Die Rückmeldungen betrafen beispielsweise Aspekte der Machbarkeit, d. h. die Umsetzbarkeit im spezifischen Kontext und unter den gegebenen Rahmenbedingungen (z. B. zeitliche Probleme), oder die Akzeptanz (z. B. Einschätzung der Komplexität des Trainings; Petermann 2014; Hebbecke und Souvignier 2018). Ziel war es, durch eine erhöhte Akzeptanz und Machbarkeit, eine höhere Durchführungstreue zu erreichen. Da in der Effektivitätsstudie 1 vor allem die Kurselemente zur Präsentationsgliederung, zum Videofeedback und zur Vorbereitung der Abschlusspräsentation in reduziertem Umfang oder mit geringerer Qualität eingesetzt worden waren, wurden der Aufbau der Übungen, das Material und die zur Verfügung stehenden Zeitfenster angepasst. Die Kursleitungen berichten hierfür nun eine tendenziell höhere Einhaltung und Qualität. Schwieriger scheint die Durchführung der 9. Kurssitzung zu sein, die ebenfalls für die Effektivitätsstudie 2 umstrukturiert worden war. Hier ist als nächstes zu prüfen, welche Faktoren die Implementation erschwert haben könnten (z. B. zeitliche Umsetzbarkeit, Interesse der Kinder, Schwierigkeit der Aufgaben).

Zusätzlich zur Einschätzung der Durchführungstreue durch die Kursleitungen wurden die Kinder gebeten, Rückmeldungen zum Kurs zu geben, um ihre Akzeptanz des Kurses einzuschätzen und um potentielle Änderungen in den Kursinhalten und Methoden abzuleiten. Die Kinder hatten Spaß am Kurs, berichteten in hohem Maße Neues gelernt zu haben und gaben dem Kurs nur sehr gute oder gute Noten. Damit scheint der Kurs Inhalte zu adressieren, die für die Zielgruppe ansprechend und passend sind.

6.2 Effektivität des Trainings

6.2.1 Replikation der Ergebnisse auf Präsentationsfähigkeiten

Für die Interventionseffekte konnten in dieser Studie Präsentationsfähigkeiten aller vier Dimensionen, nonverbal-visuell, nonverbal-auditiv, Organisation und Sprachgebrauch, erhoben werden. Erfasst wurde dabei die Angemessenheit, d. h. Passung der gezeigten Fähigkeit zum Kontext. Signifikante Effekte zeigten sich für Körperspannung (nonverbal-visuell) und persönliche Ansprache (Sprachgebrauch). Auf allen

Tab. 8 Übersicht über die Interventionseffekte der Wirksamkeitsstudie und der zwei Effektivitätsstudien

	Wirksamkeitsstudie	Effektivitätsstudie (I)	Effektivitätsstudie (II)
<i>Nonverbal-visuell</i>			
Körperspannung	0,69*	0,50	0,66*
Gestik	0,77*	0,24	0,21
Mimik	0,03	/	/
Blickkontakt	0,45*	0,26	-0,05
Raumverhalten	0,71*	/	0,43
<i>Nonverbal-auditiv</i>			
Sprechatmung	-0,70*	0,10	-0,04
Stimme	/	/	-0,03
Artikulation	/	/	0,03
Lautstärke	/	/	0,23
Pausen	0,64*	0,03	/
Sprechflüssigkeit	0,54*	0,42	-0,37
<i>Organisation</i>			
Sprechdauer	0,54*	0,61*	0,03
Länge der Einleitung	/	0,81*	0,69
Länge des Schlusses	0,59*	0,89*	0,17
Informationsmenge	/	/	0,13
Kommunikationsabsicht	/	/	0,61
Hörerbezug	0,54*	0,64*	0,42
<i>Sprachgebrauch</i>			
Persönliche Ansprache	/	/	0,68*
Stilmittelverwendung	/	/	0,07
<i>Gesamtauftreten</i>	0,62*	0,56*	0,22
<i>Sprechangst</i>	-0,15	-0,48* und -0,68*	0,08 und -0,03

Zwischen den Studien variierte die Art der Kontrollgruppe. In der Wirksamkeitsstudie wurde eine Kontrollgruppe mit Behandlung, in den beiden Effektivitätsstudien eine Wartekontrollgruppe verwendet. Kontrollvariablen – zusätzlich zu verbaler Intelligenz und figuralen kognitiven Fähigkeiten – wurden dann in die Analysen aufgenommen, wenn sich die Gruppen zum Prätest auf bedeutsamen Variablen unterschieden hatten. Die Interventionseffekte sind als Cohen's d berichtet. Da für die Videoratings nicht auf allen Items eine ICC $>0,60$ erreicht werden konnte, konnten nicht alle Präsentationsfähigkeiten in den Analysen berücksichtigt werden (/)

* $p < 0,05$

anderen Variablen waren die Effekte nicht signifikant. Ziel der vorliegenden Studie war es, die Ergebnisse der ersten Effektivitätsstudie zu replizieren, in welcher sich signifikante Effekte auf die Angemessenheit der organisatorischen Fähigkeiten (Sprechdauer, Länge der Einleitung, Länge des Schlusses und Hörerbezug) sowie Gesamtauftreten gezeigt hatten (Herbein et al. 2018b; siehe Tab. 8). In der vorliegenden Studie fanden sich auf diesen Variablen geringere, nicht signifikante Effekte (d zwischen 0,03 und 0,69). Für die nonverbalen Fähigkeiten, visuell und auditiv, fanden sich – bis auf Körperspannung in Effektivitätsstudie 2 – in keiner der beiden Studien signifikante Effekte (d für Effektivitätsstudie 1 zwischen 0,03 und 0,50, für Effektivitätsstudie 2 zwischen -0,37 und 0,66; Tab. 8).

In beiden Effektivitätsstudien fanden sich geringe Effekte auf nonverbale Fähigkeiten. Ein ähnliches Muster berichten Studien zur Förderung von PK im Hochschulkontext, nämlich signifikante Effekte auf die organisatorischen Fähigkeiten und keine signifikanten Effekte auf nonverbale Fähigkeiten (vgl. De Grez et al. 2009a, b). Mögliche Gründe für fehlende Interventionseffekte können

- a) die fehlende Wirksamkeit des Trainings,
- b) Mängel in der Implementation oder
- c) eine fehlerhafte Studiendurchführung sein (Gräsel und Parchmann 2004; Greene 2015).

Besteht die Annahme, dass die fehlenden Effekte auf die Unwirksamkeit der gewählten Kernkomponenten oder auf eine ungeeignete Operationalisierung der Trainingsaktivitäten zurückzuführen ist, dann ist ein weiterer Einsatz des Trainings ohne grundlegende Überarbeitung nicht gerechtfertigt. Da jedoch in der ersten Wirksamkeitsstudie des Präsentationstrainings unter kontrollierten Bedingungen positive Effekte auf die nonverbalen Fähigkeiten gefunden wurden (Herbein et al. 2018a) und da die Effektstärken in den Effektivitätsstudien zwar nicht mehr signifikant, aber auf einem Teil immer noch klein ausfielen, ist es sinnvoll, noch weitere Faktoren zu diskutieren, die Ursache für die fehlenden Effekte sein könnten.

Die Effektivität einer Intervention kann auch durch die Art der Implementation sowie Charakteristika der Trainerperson beeinflusst werden (Humphrey et al. 2016). Für die Implementation können implementationsnahe und -ferne Faktoren unterschieden werden (Souvignier und Mokhlesgerami 2005; Humphrey et al. 2016), wobei erstere die Dimensionen der Durchführungstreue umfassen (vgl. Humphrey et al. 2016). Für die Einhaltung und Qualität berichteten die Kursleitungen hohe Werte in beiden Studien. Ergänzend ist deshalb als nächstes genauer zu überprüfen, inwieweit die Kursleitungen die eingesetzten Kurselemente adaptierten. Studien haben gezeigt, dass Interventionen dann am effektivsten sind, wenn sie mit hoher Durchführungstreue implementiert werden (Durlak und DuPre 2008; O'Donnell 2008; Nelson et al. 2012). Dennoch müssen Adaptionen nicht automatisch zu einer Reduktion der Effektivität des Trainings führen. Gezielte Adaptionen können beispielsweise eingesetzt werden um die Praktikabilität oder die Passung zur Zielgruppe zu erhöhen (z.B. durch die Verwendung von Beispielen aus den Interessensbereichen der Kinder; Blase und Fixsen 2013). Die zugrundeliegenden Kernkomponenten werden dabei jedoch weiterhin praktisch umgesetzt. Bei dem vorliegenden Training handelt es sich um eine Multikomponenten-Intervention. Um für den Einsatz des Trainings in der Praxis festzulegen, wo gezielte Adaptionen möglich sind und wo Lehrpersonen Prioritäten setzen können, muss genauer überprüft werden, welche (Kombination der) Komponenten die größten Interventionseffekte hervorrufen (Nelson et al. 2012; Blase und Fixsen 2013; Abry et al. 2015). Dazu können verschiedene Vorgehensweisen genutzt werden. Die Wirksamkeit einzelner Komponenten lassen sich beispielsweise in mehrarmigen, randomisierten kontrollierten Studien überprüfen (Montgomery et al. 2003), welche es erlauben, Kombinationen von bestimmten Kernkomponenten oder einzelne Komponenten gegeneinander zu testen (vgl. Spörer et al. 2009; Supanc et al. 2017).

Zu den implementationsfernen Aspekten zählt die Unterrichtsqualität (Souvignier und Mokhlesgerami 2005). Um Kursleitungen, v. a. mit wenig Lehrerfahrung, bestmöglich zu unterstützen das Training flüssig und strukturiert durchzuführen, ist der Ablauf jeder Kurseinheiten detailliert vorstrukturiert (vgl. Grossman und Thompson 2008). Dennoch kann es die Klassenführung negativ beeinflussen, wenn Kursleitungen mit dem Ablauf der Einheit nicht vertraut sind oder sich zu eng am Manual orientieren. Zukünftige Studien sollten deshalb, über die Durchführungstreue hinaus, die Unterrichtsqualität mit berücksichtigen und deren Einfluss auf die Wirksamkeit der Intervention untersuchen.

Weiterhin könnte die Effektivität des Trainings durch Eigenschaften der Kursleitungen beeinflusst werden (Humphrey et al. 2016), z. B. durch deren fachliche Expertise und praktische Erfahrung im Präsentieren. Die Kursleitungen sind ein Vorbild, an dem sich die Kinder orientieren und von dem sie lernen können (Staton und Tomlinson 2001). Das könnte besonders relevant für den Erwerb nonverbaler Fähigkeiten sein. Während des Trainings ist die Kursleitung zum einen ein positives Beispiel für den angemessenen Gebrauch nonverbaler Fähigkeiten, zum anderen kann sie gezielt verschiedene Extreme im Verhalten demonstrieren und anschließend die Wirkung diskutieren. Dadurch können die Kinder ihr Verhaltensrepertoire erweitern und den angemessenen Einsatz der Fähigkeiten reflektieren. Zukünftige Studien sollten demnach auch die Eigenschaften der Kursleitungen und deren Einfluss auf die Interventionseffekte berücksichtigen.

Mit Bezug zu Einschränkungen, die aus der Art der Studiendurchführung resultieren, muss hier die Erfassung der Durchführungstreue diskutiert werden. Da in den bestehenden Studien lediglich Selbsteinschätzungen der Lehrpersonen verwendet wurden, um die Einhaltung und die Qualität einzuschätzen, muss diskutiert werden, inwieweit soziale Erwünschtheit die Beantwortung der Fragebögen beeinflusst haben könnte. Zukünftige Studien sollten bei der Erfassung der Durchführungstreue weitere Perspektiven, wie externe Beobachter oder Videoratings, berücksichtigen.

Für die Interventionseffekte auf die organisatorischen Fähigkeiten fanden sich in der aktuellen Studie, im Vergleich zur Effektivitätsstudie 1 und der Wirksamkeitsstudie, geringere Effekte. Die Gründe könnten hier ebenfalls in der Art der Implementation, der Unterrichtsqualität und in Unterschieden in den Kursleitercharakteristika liegen.

Zusammengefasst lässt sich sagen: Von der kontrollierten Wirksamkeitsstudie, über die Effektivitätsstudie 1 zur Effektivitätsstudie 2 nahmen die Interventionseffekte kontinuierlich ab. (Dabei ist jedoch bei der Diskussion der Ergebnisse zu berücksichtigen, dass sich das Design der Wirksamkeitsstudie vom Design der anderen beiden Studien unterschied. Die Kinder der Kontrollgruppe besuchten hier einen alternativen Kurs zum Thema Wissenschaftsverständnis.) Obwohl in der vorliegenden Studie nur zwei der Interventionseffekte auf die Präsentationsfähigkeiten signifikant wurden, sind die Effektstärken dennoch auf sechs von 18 Zielvariablen klein – wobei ein negativer Effekt für Sprechflüssigkeit gefunden wurde – und auf vier moderat ausgeprägt. Zusammen mit dem Ergebnismuster der Effektivitätsstudie 1 lassen sich deshalb Hinweise ableiten, dass das Training, als extracurriculares Förderprogramm im Rahmen der Hector Kinderakademien, tendenziell wirksam ist. Aufgrund von methodischen Einschränkungen in der vorliegenden Studie sind jedoch weitere Stu-

dien zur Überprüfung der Wirksamkeit notwendig, um aussagekräftigere Ergebnisse zu erhalten.

6.2.2 Replikation der Ergebnisse auf Sprechangst

In der vorliegenden Effektivitätsstudie fanden sich keine Effekte auf Sprechangst, weder auf die kognitiven noch auf die körperlichen Symptome. Diese Ergebnisse sind parallel zur Wirksamkeitsstudie. Im Vergleich dazu hatten sich in der Effektivitätsstudie 1 signifikante Effekte gezeigt. Dabei war bei keiner der Studien ein Bodeneffekt für die Ausprägung der Sprechangst zum Prätest zu finden. Ein Grund für die reduzierten Interventionseffekte könnte sein, dass die Kurseinheit zur Sprechangst angepasst worden war. Die Kursleitungen hatten zurückgemeldet, dass Sprechangst für die Kinder keine große Relevanz hat, eher auf Desinteresse stößt und deshalb schwierig umzusetzen war. Aus Sicht der Forschung und sprecherzieherischen Praxis stellt Sprechangst jedoch eine Kernvariable dar, die in einem Präsentationstraining nicht zu vernachlässigen ist. Obwohl Kinder das Sprechen vor anderen noch positiver wahrnehmen oder beim Halten einer Spontanrede eine geringere biologischen Stressempfindlichkeit zeigen als Jugendliche (vgl. Sumter et al. 2010) ist es sinnvoll, bereits Kinder im Grundschulalter weiter in ihrer Sicherheit im Umgang mit der Präsentationssituation zu stärken, um zukünftiger Sprechangst entgegenzuwirken (Boyce et al. 2007; Klippert 2010). Zudem ist es das Ziel, Kinder die von Sprechangst betroffen sind früh zu erkennen, um eine Verstärkung der Angst zu verhindern. Aus diesem Grund wurden trotz der Rückmeldungen der Kursleitungen die Kernübungen aus der Effektivitätsstudie 1 beibehalten, jedoch der zeitliche Umfang von 90 auf 45 min gekürzt. Aufgrund der vorliegenden Ergebnisse ist zu überlegen, wie der Umgang mit Lampenfieber wieder stärker in den Kurs integriert werden kann, dabei jedoch besser zum Bedarf der Zielgruppe passt.

6.3 Stärken und Grenzen der Studie

Ziel der Studie war es, die vorausgegangene Effektivitätsstudie zu replizieren und den Einsatz des Präsentationstrainings in der Praxis sowie dessen Effektivität erneut zu überprüfen. Ein Scaling-up des Trainings, im Sinne eines flächendeckenden Einsatzes oder der Übertragung in einen anderen Kontext, wurde noch nicht in Betracht gezogen, da sich die vorherige Effektivitätsstudie nur auf eine kleine Stichprobe gestützt hatte. Einschränkend ist für die vorliegende Studie jedoch zu sagen, dass sich trotz sorgfältiger Planung, verschiedene Herausforderungen in der Studiendurchführung ergaben. Infolgedessen ist die Diskussion der Ergebnisse in der vorliegenden Studie erneut nur im selben Maß möglich, wie in der vorausgegangenen Studie.

Erstens war es nicht möglich, die geplante Stichprobengröße zu erreichen, die in einer Poweranalyse ermittelt worden war. Schwierigkeiten bei der Rekrutierung von Teilnehmer/inne/n für Studien im Feld sind nicht untypisch (vgl. Friedman et al. 2010). Beispielsweise erfordert die Teilnahme an randomisierten kontrollierten Studien, bei denen die Teilnehmenden erst nach dem Prätest erfahren, ob sie die Intervention sofort oder erst zu einem späteren Zeitpunkt erhalten, viel zeitliche Flexibilität und kurzfristige Organisation. Da dies in der Praxis, im Zusammen-

hang mit der Nachmittagsbetreuung von Kindern im Grundschulalter, für Eltern nicht immer gut zu ermöglichen ist, könnte das ein Grund für die geringe Teilnehmerzahl in der vorliegenden Studie sein ($N_{\text{Kinder}}=65$). Als Konsequenz auf die geringen Anmeldezahlen zu den einzelnen Kursen konnte nur an zwei Akademien eine Blockrandomisierung, und damit eine zufällige Zuteilung der einzelnen Kinder zu Interventions- und Kontrollgruppe, durchgeführt werden. Die anderen acht Kurse wurden cluster-randomisiert, das heißt alle Kinder wurden komplett, entweder der einen oder anderen Gruppe, zugeordnet (vgl. Spybrook et al. 2014). Folglich liegen nur für fünf der zehn Kursleitungen Rückmeldungen zur Durchführungstreue vor, denn alle anderen boten den Kurs lediglich für die Wartekontrollgruppe an. Dazu kommt die relativ hohe Anzahl an Missings zum Posttest, was als Einschränkung der Studie zu diskutieren ist. Da jedoch kein differentieller Ausfall vorlag (What Works Clearinghouse 2014) wurde angenommen, dass die fehlenden Werte zufällig entstanden. Dementsprechend wurde FIML verwendet, um diese zu schätzen.

Zweitens ist bezüglich des Designs zu diskutieren, dass keine Langzeiteffekte des Trainings erfasst wurden. Dies ist jedoch notwendig, um Aussagen über die Langfristigkeit der Effekte treffen zu können und um die Relevanz der Intervention zu stützen (Gottfredson et al. 2015). Da im vorliegenden Fall ein Wartekontrollgruppendesign verwendet wurde, bei dem die Kontrollgruppe direkt nach dem Posttest das Training erhielt, war ein längerfristiges Follow-up nicht möglich. Ein Design, bei dem das Training der Kontrollgruppe erst im nächsten Schulhalbjahr angeboten wird, schied aus, da die an der Studie teilnehmenden Viertklässler/innen dann bereits an eine weiterführende Schule gewechselt hätten.

Drittens basierte die Erfassung der Durchführungstreue auf Selbstberichten der Kursleitungen. Dieses Vorgehen wird häufig eingesetzt (vgl. z. B. Spörer et al. 2009; Hebecker und Souvignier 2018). Ein Vorteil von Selbstberichten ist, dass sie eine relativ ökonomische Erfassungsmethode darstellen, die kontinuierlich und über den ganzen Kurs hinweg möglich ist. Alternative Erfassungen der Durchführungstreue, über Beobachter oder Videoaufnahmen, sind hingegen relativ zeit- und kostenaufwändig (Petermann 2014). Zudem könnten diese das Commitment und die Akzeptanz der Lehrpersonen gegenüber der Studie reduzieren (James Bell Associates 2009) und/oder ein Gefühl der Kontrolle entstehen lassen (Krainer et al. 2012). Dazu kommt, dass wiederholte Beobachtungen notwendig sind, um valide Informationen zu erhalten. Trotz der Nachteile von Beobachtungen und Videoaufzeichnungen ist die alleinige Konzentration auf Selbstberichte in dieser Studie zu diskutieren, da diese beispielsweise anfällig für sozial erwünschte Antworten sind (Petermann 2014). Zudem wurden nur zwei Dimensionen der Durchführungstreue erfasst. Mit dem Schwerpunkt auf Quantität und Qualität sind zwar zwei relevante Dimensionen des Verhaltens der Kursleitung berücksichtigt (Odom et al. 2010; Greene 2015), die zusätzliche Erfassung der Adaption würde jedoch einen vertieften Einblick in die Implementation erlauben. Dazu kommt, dass neben der in dieser Studie erfassten Durchführungstreue, verschiedene weitere Faktoren zur Implementation gehören. Das sind beispielsweise die von den Lehrpersonen wahrgenommene Akzeptanz und Machbarkeit sowie die Angemessenheit einer Intervention (Petermann 2014; Hebecker und Souvignier 2018). Um detaillierter zu erfassen, in welchem Ausmaß die Implementation des Präsentationstrainings gelingt, sollten zukünftige Studien neben

der Durchführbarkeit auch Akzeptanz, Machbarkeit und Angemessenheit erheben. Diese bilden zusammen mit der Durchführungstreue, die Basis für eine gelungene Implementation (Souvignier und Philipp 2016). Ausgehend von den Ergebnissen kann das Trainingsmaterial weiter angepasst und optimiert werden.

Eine Stärke der Studie ist die Erfassung der Präsentationskompetenz mit Videoratings, da diese zuverlässiger sind als Selbstberichte der Sprecherin/des Sprechers (Carrell und Willmington 1996). Die Erfassung der PK über Videoratings gelang in der aktuellen Effektivitätsstudie besser als in den vorherigen Studien. Um die Überprüfung der Effektivität über alle Studien hinweg möglichst parallel zu halten, wurde der Beobachtungsbogen nur in geringem Ausmaß adaptiert. Da sich jedoch in der Vergangenheit Schwierigkeiten, v. a. bei der Erfassung der nonverbal-auditiven Präsentationsfähigkeiten und des Sprachgebrauchs gezeigt hatten – nicht nur im Rahmen dieses Forschungsprogramms, sondern auch in anderen Studien zu PK (vgl. Cheng und Warren 2005; Joe et al. 2015) – wurde die Raterschulung optimiert. Final konnte in der vorliegenden Studie auf 18 Items eine ICC > 0,60 erlangt werden. In Effektivitätsstudie 1 waren es nur 11 (Herbein et al. 2018b). Außerdem konnten in der momentanen Studie das erste Mal auch Effekte auf Items zum Sprachgebrauch erfasst und diskutiert werden. Trotz dieser Verbesserungen in der Erfassung ist es weiterhin notwendig, das Instrument zu optimieren, um langfristig alle Präsentationsfähigkeiten erheben zu können. Zudem ist die Überprüfung der Validität des Instrumentes ein erforderlicher weiterer Schritt.

6.4 Fazit

Die vorliegende Studie ist die vierte in einem systematischen Forschungsprogramm, in dem ein Training zur Förderung der Präsentationskompetenz schrittweise entwickelt, in der Praxis implementiert und dessen Wirksamkeit überprüft wurde. Ziel der Studie war eine Replikation der Effektivitätsstudie 1, da deren Ergebnisse auf einer relativ kleinen Stichprobe beruhten (vgl. Erdfelder und Ulrich 2018). Das Präsentationstraining richtet sich an Kinder im Grundschulalter, die für die Teilnahme an einem extracurricularen Enrichment-Programm nominiert sind. Die Ergebnisse der drei vorliegenden Studien zum Präsentationstraining geben Hinweise darauf, dass das Training für diese spezifische Zielgruppe und die Förderung im Rahmen der Hector Kinderakademien durchführbar und tendenziell wirksam ist.

Obwohl auch die Ergebnisse der zweiten Effektivitätsstudie auf einer kleinen Stichprobe basieren, lassen sich Schlüsse für die weitere Forschung und Implementation in der Praxis ableiten. Unter stark kontrollierten Bedingungen (Wirksamkeitsstudie) zeigten sich Hinweise auf die Effektivität des Trainings, sowohl auf nonverbale und organisatorische Präsentationsfähigkeiten als auch auf das Gesamtaufreten. Beim Schritt weiter in die Praxis (Effektivitätsstudien) reduzierten sich die Interventionseffekte. Gleichzeitig berichteten die Kursleitungen jedoch eine hohe Durchführbarkeit. Im Sinne der Implementationsforschung sind nun beispielsweise die Qualität der Implementation sowie Faktoren, die die Implementation beeinflussen, weiter zu untersuchen (Century und Cassata 2016; Humphrey et al. 2016). Daran orientiert können Anpassungen für das Unterstützungssystem für die Kursleitungen abgeleitet werden. Ziel könnte dabei die Erhöhung des fachlichen oder

praktischen Wissens zu Präsentation sein, um so die Durchführbarkeit des Kurses weiter zu optimieren und die Effektivität des Kurses in der Praxis zu steigern. Zudem kann auch überlegt werden, den flächendeckenden Einsatz des Trainings durch die Ausbildung von Multiplikatoren zu unterstützen, die das Training an weitere Kursleitungen aus der Praxis weitergeben. Dabei wäre jedoch erneut zu überprüfen, inwieweit das veränderte Unterstützungssystem die erfolgreiche Implementation des Trainings beeinflusst.

Basierend auf Ergebnissen weiterer, intensiver Forschung zur Implementation und Effektivität des Trainings, kann langfristig ein Scaling-up in einen anderen Kontext angestrebt werden. Um beispielsweise die Präsentationskompetenz besonders begabter und hochbegabter Kinder im Regelunterricht zu fördern, könnten die Kursmaterialien für die veränderten Rahmenbedingungen des Lernkontexts angepasst und erneut, in mehrschrittigen Verfahren, überprüft werden. Ein ähnliches Vorgehen wäre zu wählen, wenn eine Ausweitung auf eine andere Zielgruppe angestrebt wird, beispielsweise die Förderung älterer Schüler/innen in Arbeitsgruppen oder an Projekttagen.

Förderung Das Forschungsprojekt wurde gefördert durch die Hector Stiftung II und unterstützt von der Postdoktorandinnen- und Postdoktorandenakademie des Hector-Instituts für Empirische Bildungsforschung, die durch das Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg gefördert wird.

Funding Open Access funding provided by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Abry, T., Hulleman, C. S., & Rimm-Kaufman, S. E. (2015). Using indices of fidelity to intervention core components to identify program active ingredients. *American Journal of Evaluation*, 36(3), 320–338. <https://doi.org/10.1177/1098214014557009>.
- Backlund, P. M., & Morreale, S. P. (2015). Communication competence: Historical synopsis, definitions, applications, and looking to the future. In A. F. Hannawa & B. H. Spitzberg (Hrsg.), *Communication competence*. Handbooks of communication science, (Bd. 22, S. 11–38). Berlin: De Gruyter.
- Baker, E. L. (2006). Principles for scaling up: Choosing, measuring effects, and promoting the widespread use of educational innovation. In B. Schneider & S.-K. McDonald (Hrsg.), *Scale-up in education: Ideas in principle* (S. 37–54). Lanham: Rowman & Littlefield.
- Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, Ref. Öffentlichkeitsarbeit (2014). *LehrplanPLUS Grundschule: Lehrplan für die bayerische Grundschule*. <https://www.>

- lehrplanplus.bayern.de/sixcms/media.php/107/LehrplanPLUS%20Grundschule%20StMBW%20-%20Mai%202014.1749837.pdf. Zugegriffen: 19. Dez. 2019.
- Behrens, U., & Eriksson, B. (2011). Sprechen und Zuhören. In A. Bremerich-Vos, D. Granzer, U. Behrens, & O. Köller (Hrsg.), *Lehrer-Bücherei Grundschule: Kompakt. Bildungsstandards für die Grundschule: Deutsch konkret.: Aufgabenbeispiele, Unterrichtsarrangements, Fortbildungsideen; mit CD-ROM* (3. Aufl., S. 43–74). Berlin: Cornelsen.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300.
- Blase, K., & Fixsen, D.L. (2013). *Core intervention components: Identifying and operationalizing what makes programs work. ASPE research brief*. Washington: US Department of Health and Human Services.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R., & Walther, G. (Hrsg.) (2004). *IGLU: Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. Zusammenfassung ausgewählter Ergebnisse* (1. Aufl.). Münster: Waxmann.
- Boyce, J.S., Alber-Morgan, S.R., & Riley, J.G. (2007). Fearless public speaking: Oral presentation activities for the elementary classroom. *Childhood Education*, 83(3), 142–150. <https://doi.org/10.1080/00094056.2007.10522899>.
- Carrell, L.J., & Willmington, S.C. (1996). A comparison of self-report and performance data in assessing speaking and listening competence. *Communication Reports*, 9(2), 185–191.
- Century, J., & Cassata, A. (2016). Implementation research: Finding common ground on what, how, why, where, and who. *Review of Research in Education*, 40(1), 169–215. <https://doi.org/10.3102/0091732X16665332>.
- Chambers, D.W. (1983). Stereotypic images of the scientist: The Draw-A-Scientist Test. *Science Education*, 67(2), 255–265. <https://doi.org/10.1002/sce.3730670213>.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93–121. <https://doi.org/10.1191/0265532205lr298oa>.
- Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>.
- Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3. Aufl.). Mahwah: Erlbaum.
- Coldwell, M., & Maxwell, B. (2018). Using evidence-informed logic models to bridge methods in educational evaluation. *Review of Education*, 6(3), 267–300. <https://doi.org/10.1002/rev3.3151>.
- Common Core State Standards Initiative (2010). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington D.C.
- Croucher, S.M. (2013). Communication apprehension, self-perceived communication competence, and willingness to communicate: A French analysis. *Journal of International and Intercultural Communication*, 6(4), 298–316. <https://doi.org/10.1080/17513057.2013.769615>.
- De Grez, L., & Valcke, M. (2010). Learning and instruction of oral presentation skills. In L.E. Kattington (Hrsg.), *Education in a competitive and globalizing world. Handbook of curriculum development* (S. 177–208). New York: Nova Science.
- De Grez, L., Valcke, M., & Roozen, I. (2009a). The impact of goal orientation, self-reflection and personal characteristics on the acquisition of oral presentation skills. *European Journal of Psychology of Education*, 24(3), 293–306. <https://doi.org/10.1007/BF03174762>.
- De Grez, L., Valcke, M., & Roozen, I. (2009b). The impact of an innovative instructional intervention on the acquisition of oral presentation skills in higher education. *Computers & Education*, 53(1), 112–120. <https://doi.org/10.1016/j.compedu.2009.01.005>.
- Durlak, J.A., & DuPre, E.P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350. <https://doi.org/10.1007/s10464-008-9165-0>.
- EDK (2010). *SCHULSPRACHE: Wissenschaftlicher Kurzbericht und Kompetenzmodell: Provisorische Fassung (vor Verabschiedung der Standards)*. http://www.edudoc.ch/static/web/arbeiten/harmos/L1_wissB_25_1_10_d.pdf. Zugegriffen: 19. Dez. 2019.
- Enders, C.K. (2010). *Applied missing data analysis. Methodology in the social sciences*. New York: Guilford.

- Erdfelder, E., & Ulrich, R. (2018). Zur Methodologie von Replikationsstudien. *Psychologische Rundschau*, 69(1), 3–21. <https://doi.org/10.1026/0033-3042/a000387>.
- Fehm, L., & Hille, C. (2005). Bühnenangst bei Musikstudierenden. *Verhaltenstherapie und Verhaltensmedizin*, 2, 199–212.
- Friedman, L. M., Furberg, C., & DeMets, D. L. (2010). *Fundamentals of clinical trials* (4. Aufl.). Heidelberg: Springer.
- van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2015). Towards a set of design principles for developing oral presentation competence: A synthesis of research in higher education. *Educational Research Review*, 14, 62–80. <https://doi.org/10.1016/j.edurev.2015.02.002>.
- Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. N., & Zafft, K. M. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science: the Official Journal of the Society for Prevention Research*, 16(7), 893–926. <https://doi.org/10.1007/s11121-015-0555-x>.
- Gräsel, C., & Parchmann, I. (2004). Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 32(3), 196–214.
- Greene, J. A. (2015). Serious challenges require serious scholarship: Integrating implementation science into the scholarly discourse. *Contemporary Educational Psychology*, 40, 112–120. <https://doi.org/10.1016/j.cedpsych.2014.10.007>.
- Grossman, P., & Thompson, C. (2008). Learning from curriculum materials: Scaffolds for new teachers? *Teaching and Teacher Education*, 24(8), 2014–2026. <https://doi.org/10.1016/j.tate.2008.05.002>.
- Hebbecker, K., & Souvignier, E. (2018). Formatives Assessment im Leseunterricht der Grundschule – Implementation und Wirksamkeit eines modularen, materialgestützten Konzepts. *Zeitschrift für Erziehungswissenschaft*, 21(4), 735–765. <https://doi.org/10.1007/s11618-018-0834-y>.
- Herbein, E., Golle, J., Tibus, M., Schiefer, J., Trautwein, U., & Zettler, I. (2018a). Fostering elementary school children's public speaking skills: A randomized controlled trial. *Learning and Instruction*, 55, 158–168. <https://doi.org/10.1016/j.learninstruc.2017.10.008>.
- Herbein, E., Golle, J., Tibus, M., Zettler, I., & Trautwein, U. (2018b). Putting a speech training program into practice: Its implementation and effects on elementary school children's public speaking skills and levels of speech anxiety. *Contemporary Educational Psychology*, 55, 176–188. <https://doi.org/10.1016/j.cedpsych.2018.09.003>.
- Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016). *Implementation and process evaluation (IPE) for interventions in education settings: A synthesis of the literature*. https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Review_Final.pdf. Zugegriffen: 4. März 2019.
- Hunt, S., Wright, A., & Simonds, C. (2014). Securing the future of communication education: Advancing an advocacy and research agenda for the 21st century. *Communication Education*, 63(4), 449–461. <https://doi.org/10.1080/03634523.2014.926016>.
- James Bell Associates (2009). *Evaluation brief: Measuring implementation fidelity*. <http://www.jbassoc.com/ReportsPublications/Evaluation%20Brief%20-%20Measuring%20Implementation%20Fidelity%20Octob%E2%80%A6.pdf>. Zugegriffen: 4. März 2019.
- Joe, J., Kitchen, C., Chen, L., & Feng, G. (2015). A prototype public speaking skills assessment: An evaluation of human-scoring quality. *ETS Research Report Series*, 2015(2), 1–21. <https://doi.org/10.1002/ets2.12083>.
- Klippert, H. (2010). *Kommunikations-Training: Übungsbausteine für den Unterricht* (12. Aufl.). Weinheim: Beltz.
- Krainer, L., Lerchster, R., & Goldmann, H. (2012). Interventionsforschung in der Praxis. In L. Krainer & R. E. Lerchster (Hrsg.), *Interventionsforschung* (Bd. 1, S. 175–243). Wiesbaden: Springer. https://doi.org/10.1007/978-3-531-19113-3_6.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Lee, S.-Y., Olszewski-Kubilius, P., & Thomson, D. T. (2012). Academically gifted students' perceived interpersonal competence and peer relationships. *Gifted Child Quarterly*, 56(2), 90–104. <https://doi.org/10.1177/0016986212442568>.
- Lee, O., Quinn, H., & Valdes, G. (2013). Science and language for English language learners in relation to next generation science standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher*, 42(4), 223–233. <https://doi.org/10.3102/0013189X13480524>.

- McCroskey, J. C., Beatty, M. J., Kearney, P., & Plax, T. G. (2009). The content validity of the PRCA-24 as a measure of communication apprehension across communication contexts. *Communication Quarterly*, 33(3), 165–173. <https://doi.org/10.1080/01463378509369595>.
- McCroskey, J. C., & McCroskey, L. L. (1988). Self-report as an approach to measuring communication competence. *Communication Research Reports*, 5(2), 108–113. <https://doi.org/10.1080/08824098809359810>.
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg (Hrsg.) (2016). *Deutsch. Bildungsplan der Grundschule* (Reihe A Nr. 10: Vol. 8). Villingen-Schwenningen: Neckar.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotsche, P. C., Devereaux, P. J., ... & Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *British Medical Journal BMJ*, 340(mar23 1), 1–28. <https://doi.org/10.1136/bmj.c869>.
- Montgomery, A. A., Peters, T. J., & Little, P. (2003). Design, analysis and presentation of factorial randomised controlled trials. *BMC Medical Research Methodology*, 3, 26. <https://doi.org/10.1186/1471-2288-3-26>.
- Murrah, W. M., Kosovich, J., & Hulleman, C. S. (2017). A framework for incorporating intervention fidelity in educational evaluation studies. In I. G. Roberts, S. Vaughn, S. N. Beretvas, & V. Wong (Hrsg.), *Treatment fidelity in studies of educational intervention* (S. 39–60). New York: Routledge.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7. Aufl.). Los Angeles: Muthén & Muthén.
- Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *The Journal of Behavioral Health Services & Research*, 39(4), 374–396. <https://doi.org/10.1007/s11414-012-9295-x>.
- Odom, S. L., Fleming, K., Diamond, K., Lieber, J., Hanson, M., Butera, G., ... & Marquis, J. (2010). Examining different forms of implementation and in early childhood curriculum research. *Early Childhood Research Quarterly*, 25(3), 314–328. <https://doi.org/10.1016/j.ecresq.2010.03.001>.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84. <https://doi.org/10.3102/0034654307313793>.
- Petermann, F. (2014). Implementationsforschung: Grundbegriffe und Konzepte. *Psychologische Rundschau*, 65(3), 122–128. <https://doi.org/10.1026/0033-3042/a000214>.
- Richter, T., Souvignier, E., Hertel, S., Heyder, A., & Kunina-Habenicht, O. (2019). Positionspapier zur Lage der Pädagogischen Psychologie in Forschung und Lehre. *Psychologische Rundschau*, 70(2), 109–118. <https://doi.org/10.1026/0033-3042/a000437>.
- Rimm-Kaufman, S. E., Larsen, R. A. A., Baroody, A. E., Curby, T. W., Ko, M., Thomas, J. B., & DeCoster, J. (2014). Efficacy of the Responsive Classroom Approach: Results from a 3-year, longitudinal randomized controlled trial. *American Educational Research Journal*, 51(3), 567–603. <https://doi.org/10.3102/0002831214523821>.
- Ringeisen, T., Rohrmann, S., Bürgermeister, A., & Tibubos, A. N. (2017). Assessing self-efficacy in presentation and moderation skills. *European Journal of Psychological Assessment*, 1–13. <https://doi.org/10.1027/1015-5759/a000423>.
- Rothembusch, S., Zettler, I., Voss, T., Löscher, T., & Trautwein, U. (2016). Exploring reference group effects on teachers' nominations of gifted students. *Journal of Educational Psychology*, 108(6), 883–897. <https://doi.org/10.1037/edu0000085>.
- Schleppegrell, M. J. (2012). Academic language in teaching and learning. *The Elementary School Journal*, 112(3), 409–418. <https://doi.org/10.1086/663297>.
- Schroeders, U., Schipolowski, S., Zettler, I., Golle, J., & Wilhelm, O. (2016). Do the smart get smarter? Development of fluid and crystallized intelligence in 3rd grade. *Intelligence*, 59, 84–95. <https://doi.org/10.1016/j.intell.2016.08.003>.
- Siddiqui, N., Gorard, S., & See, B. H. (2018). The importance of process evaluation for randomised control trials in education. *Educational Research*, 60(3), 357–370. <https://doi.org/10.1080/00131881.2018.1493349>.
- Souvignier, E., & Mokhlesgerami, J. (2005). Implementation eines Programms zur Vermittlung von Lese-strategien im Deutschunterricht. *Zeitschrift Für Pädagogische Psychologie*, 19(4), 249–261. <https://doi.org/10.1024/1010-0652.19.4.249>.
- Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction*, 16(1), 57–71. <https://doi.org/10.1016/j.learninstruc.2005.12.006>.

- Souvignier, E., & Philipp, M. (2016). Implementation – Begrifflichkeiten, Befunde und Herausforderungen. In M. Philipp & E. Souvignier (Hrsg.), *Implementation von Lesefördermaßnahmen: Perspektiven auf Gelingensbedingungen und Hindernisse* (1. Aufl. S. 9–22). Münster: Waxmann.
- Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction*, 19(3), 272–286. <https://doi.org/10.1016/j.learninstruc.2008.05.003>.
- Spybrook, J., Hedges, L., & Borenstein, M. (2014). Understanding statistical power in cluster randomized trials: Challenges posed by differences in notation and terminology. *Journal of Research on Educational Effectiveness*, 7(4), 384–406. <https://doi.org/10.1080/19345747.2013.848963>.
- Staton, A. Q., & Tomlinson, S. D. (2001). Communication education outreach in elementary school classrooms. *Southern Communication Journal*, 66(3), 211–223. <https://doi.org/10.1080/10417940109373200>.
- Sumter, S. R., Bokhorst, C. L., Miers, A. C., van Pelt, J., & Westenberg, P. M. (2010). Age and puberty differences in stress responses during a public speaking task: Do adolescents grow more sensitive to social evaluation? *Psychoneuroendocrinology*, 35(10), 1510–1516. <https://doi.org/10.1016/j.psneuen.2010.05.004>.
- Supanc, M., Völlinger, V. A., & Brunstein, J. C. (2017). High-structure versus low-structure cooperative learning in introductory psychology classes for student teachers: Effects on conceptual knowledge, self-perceived competence, and subjective task values. *Learning and Instruction*, 50, 75–84. <https://doi.org/10.1016/j.learninstruc.2017.03.006>.
- What Works Clearinghouse (2014). *What Works Clearinghouse procedures and standards handbook, version 3.0*. http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf. Zugegriffen: 4. März 2019.
- Weiß, R. H. (2006). *CFT 20-R mit WS/ZF-R: Grundintelligenztest Skala 2 – Revision (CFT 20-R) mit Wortschatztest und Zahlenfolgentest – Revision (WS/ZF-R)*. Göttingen: Hogrefe.
- Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.