

# Außerunterrichtliches Peer Tutoring mit deutsch-türkischsprachigen Grundschulkindern: Umsetzungsgenauigkeit und Umfang von bilingualer Kommunikation

Jasmin Decristan  · Martin Schastak · Valentina Reitenbach · Dominique Rauch

Eingegangen: 23. September 2020 / Überarbeitet: 25. Februar 2021 / Angenommen: 5. Mai 2021 / Online publiziert: 14. Juni 2021  
© Der/die Autor(en) 2021

**Zusammenfassung** Der vorliegende Beitrag fokussiert die Implementation von außerunterrichtlichen Peer Tutoring-Trainings im Lesen und Rechnen mit türkisch-deutschsprachigen Grundschulkindern. Hierbei bildeten jeweils zwei Grundschul-kinder (Peers) ein Tandem, das von einer geschulten studentischen Trainingsleitung angeleitet wurde. Zur Erfassung der Implementation von Peer Tutoring-Elementen unter diesen Bedingungen wurde die Umsetzungsgenauigkeit herangezogen. Eine Besonderheit der Studie war, dass es einem Teil der Trainingsgruppen erlaubt war, beide Sprachen während des Trainings zu sprechen. Für die bilinguale Kommunikation wurden zusätzliche Operationalisierungen geprüft, um die Umsetzung bilingualer Kommunikation zu erfassen. Darüber hinaus wird im Beitrag konsequent zwischen der Implementation durch Trainingsleitungen und durch Lernende unterschieden. Die Ergebnisse unterstreichen die hohe Umsetzungsgenauigkeit der außerunterrichtlichen Peer Tutoring-Trainings. Sie zeigen aber auch auf, dass die Lernenden in diesem Setting nur wenig miteinander in ihrer Herkunftssprache kommunizierten und mehr türkischsprachige Impulse durch Trainingsleitungen nicht mit mehr bilingualer Kommunikation durch die Lernenden einherging. Der Beitrag liefert somit

---

Prof. Dr. Jasmin Decristan (✉) · Valentina Reitenbach  
Bergische Universität Wuppertal, Rainer-Gruenter-Straße 21, 42119 Wuppertal, Deutschland  
E-Mail: [decristan@uni-wuppertal.de](mailto:decristan@uni-wuppertal.de)

Valentina Reitenbach  
E-Mail: [reitenbach@uni-wuppertal.de](mailto:reitenbach@uni-wuppertal.de)

Dr. Martin Schastak  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Rostocker  
Straße 6, 60323 Frankfurt, Deutschland  
E-Mail: [schastak@dipf.de](mailto:schastak@dipf.de)

Prof. Dr. Dominique Rauch  
Pädagogische Hochschule Ludwigsburg, Reuteallee 46, 71634 Ludwigsburg, Deutschland  
E-Mail: [dominique.rauch@ph-ludwigsburg.de](mailto:dominique.rauch@ph-ludwigsburg.de)

weitere Erkenntnisse zur Aktivierung und zum Umfang bilingualer Kommunikation beim kooperativen Lernen.

**Schlüsselwörter** Peer Tutoring · Mehrsprachigkeit · Implementation · Umsetzungsgenauigkeit

## **Extracurricular peer tutoring with Turkish-German primary school children: adherence and extent of bilingual communication**

**Abstract** The paper focusses on the implementation of extracurricular peer tutoring trainings in reading and arithmetic with Turkish-German primary school children. Two children formed a dyad and were accompanied by a trained undergraduate instructor. Adherence was used to measure the implementation of peer tutoring elements under these conditions. Some of the learners were allowed to speak both languages during the training. For bilingual communication, additional operationalizations were examined to capture its implementation. Furthermore, a consistent distinction is made between implementation by trainers and by peers. The results show high treatment adherence of extracurricular peer tutoring trainings. However, they also show that the learners communicated only little with each other in their language of origin and that more Turkish-language impulses by trainees did not correspond with more bilingual communication by the learners. The article thus provides further insights into the activation and extent of bilingual communication in cooperative learning.

**Keywords** Peer tutoring · Multilingualism · Implementation · Adherence

### **1 Einleitung und Vorstellung der Forschungsfrage**

Das aktive Einbeziehen der Mehrsprachigkeit von Lernenden wird seit einigen Jahren von Akteuren innerhalb und außerhalb der Wissenschaft gefordert und hat auch in der Aus- und Weiterbildung von Lehrkräften einen besonderen Stellenwert erhalten. Dennoch liegen für die Schulpraxis bislang kaum empirisch erprobte Konzepte vor. Einen vielversprechenden Ansatz stellt das Peer Tutoring (PT) dar, bei dem zwei Personen (Peers) nach vorgegebenen Abläufen unter Anleitung und Begleitung einer Trainingsleitung miteinander lernen. Die Methode hat sich nicht nur als lernförderlich erwiesen, sie bietet gleichzeitig theoretisch einen Rahmen, in dem Schülerinnen und Schüler ihre verschiedenen Sprachen für das Lernen nutzen können. Inwiefern solche Innovationen in der Praxis umsetzbar sind, ist bislang kaum untersucht worden. Der Beitrag untersucht deshalb die Umsetzungsgenauigkeit von außerunterrichtlichen PT-Trainings und exploriert, in welchem Umfang die Aktivierung der Herkunftssprache in diesem Setting gelingt. Dabei wird konsequent zwischen der Umsetzung durch Trainingsleitungen und durch Peers unterschieden.

## 2 Theoretischer Rahmen

Das Thema Mehrsprachigkeit und Bildung wird in Deutschland seit den 1980er-Jahren diskutiert (z. B. Gogolin 1994). In der Folge der ersten PISA-Erhebung um die Jahrtausendwende wurde die Diskussion weiter intensiviert, da im politisch-gesellschaftlichen Diskurs vermeintliche Sprachdefizite in Deutsch von Kindern und Jugendlichen mit Migrationshintergrund dominierten (Schastak 2020; Stošić 2017), wenngleich die Ergebnisse strukturelle Probleme des Bildungssystems anhand des hohen Zusammenhangs von sozioökonomischem Status und Bildungserfolg offenlegten (z. B. Gebhardt et al. 2013). Bis heute und nicht zuletzt durch globale Migrations- und Fluchtbewegungen hat das Thema Mehrsprachigkeit und Bildung große gesellschaftliche und bildungspraktische Relevanz (Gogolin et al. 2020). Dennoch zeichnet sich in Deutschland schulische Bildung weiterhin durch einen „monolingualen Habitus“ aus (Gogolin 1994): Unterrichtsmaterial liegt in der Regel ausschließlich in der Verkehrssprache Deutsch vor und es wird nahezu vollständig auf Deutsch kommuniziert. Damit einhergehen Annahmen zum Nutzen einer Fokussierung allein auf die Unterrichtssprache (z. B. Esser 2009) sowie Befürchtungen von Lehrkräften eines Kontrollverlustes oder eines negativeren Klassenklimas durch mehrsprachige Kommunikation (zusammenfassend Bredthauer und Engfer 2018).

Aus anerkennungstheoretischer Perspektive (Honneth 1992) gilt es jedoch, die (Sprach-)Potenziale mehrsprachig aufwachsender Lernender wertzuschätzen und es den Lernenden zu ermöglichen, ihre Kompetenzen in Bildungsprozesse einzubringen. Bezüglich einer Lernförderung wird argumentiert, dass der Rückgriff auf das gesamte Sprachrepertoire beim Lernen unter anderem einen Transfer akademischsprachlicher Fähigkeiten zwischen den Sprachen (Cummins 2000), ein Überwinden sprachlicher Hürden durch Code-Switching (Moschkovich 2007), ein tieferes Verständnis von Lerninhalten durch „Translanguaging“ (Gantefort und Sánchez Oroquieta 2015), eine Entlastung des Arbeitsgedächtnisses (Kern 1994), ein Anknüpfen an in der Herkunftssprache kodierte Wissens-elemente (Lyster 2019) und nicht zuletzt eine höhere Partizipation am Unterrichtsgespräch (Meyer und Prediger 2011) ermöglicht. Vor diesem Hintergrund wird gefordert, die Herkunftssprachen der Lernenden fruchtbar in den Regelunterricht zu integrieren (z. B. Rehbein 2011).

Als zwei zentrale Anknüpfungspunkte zur Aktivierung der Herkunftssprachen listen Meyer et al. (2016) Material auch in den Herkunftssprachen der Lernenden sowie Partnerinnen und Partner mit der gleichen Sprache. Beide Anknüpfungspunkte lassen sich in kooperativen Settings wie dem Peer Tutoring (PT) realisieren. Auch neuere Arbeiten verweisen auf das Potenzial kooperativen Lernens für mehrsprachige Kommunikation (BiSS-Trägerkonsortium 2020). Vor diesem Hintergrund wird im Folgenden näher auf PT eingegangen.

### 2.1 Peer Tutoring

PT zeichnet sich nach Topping (2005, S. 632) durch „specific role-taking as tutor or tutee, with high focus on curriculum content and usually also on clear procedures for interaction“ aus. Beim PT lernen Schülerinnen und Schüler (Peers) gemeinsam in Zweier-Teams (Tandems) nach strukturierten Abläufen und nehmen die Rollen

als Lehrende und Lernende ein. Meta-Analysen belegen positive soziale, behaviorale und fachliche Wirkungen von PT (z. B. Ginsburg-Block et al. 2006; Rohrbeck et al. 2003). Als theoretische Grundlagen dienen sozialpsychologische (Wir-Gefühl; Cohen 1994) und motivationale Zugänge (Erleben von Kompetenz, Autonomie und sozialer Eingebundenheit; Deci und Ryan 2002). Hinsichtlich des Lernpotenzials von PT werden konstruktivistische (z. B. kognitive Konflikte beim gemeinsamen Lernen; Piaget und Inhelder 1983) und sozio-kulturelle Ansätze (z. B. Lernen in der Zone der proximalen Entwicklung; Wygotsky 1978) herangezogen und die Bedeutung qualitativ hochwertiger Dialoge betont (z. B. Littleton und Mercer 2010).

Beim PT werden in der Regel auch Strategien des selbstregulierten Lernens eingeübt (z. B. Fuchs et al. 1997; McMaster et al. 2006). Diese haben sich als effektiv für das Lernen herausgestellt (Dignath und Büttner 2008). Neben meta-kognitiven Strategien werden auch domänenspezifische Strategien erlernt und geübt. Im Lesen ist beispielsweise das Zusammenfassen von Textabschnitten eine bewährte Strategie (z. B. Spörer et al. 2008), im Rechnen zählen Strategien zu Grundrechenarten dazu (Heinze et al. 2009). Für die Schulpraxis existieren ausgearbeitete Konzepte (z. B. „Peer Assisted Learning Strategies“; Fuchs et al. 1997), die auch für den deutschsprachigen Raum adaptiert wurden (z. B. Munser-Kiefer und Kirschhock 2012). Aus den Vorteilen, die bei einer Anwendung von PT nachgewiesen sind, kann jedoch nicht ohne weiteres geschlossen werden, dass es in der Praxis auch umgesetzt wird.

## 2.2 Implementation von Peer Tutoring

Die Implementation von PT kann sowohl in ganzen Schulklassen unter Anleitung von Lehrkräften realisiert werden, als auch als außerunterrichtliches zusätzliches Training erfolgen, wenn es darum geht, spezifische Zielgruppen und Ziele zu erreichen. Wie und unter welchen Bedingungen spezifische Maßnahmen in der Praxis umgesetzt werden, ist zentraler Gegenstand der Implementationsforschung. Im Bildungsbereich sind diese Erkenntnisse entscheidend für eine Verknüpfung von Bildungsforschung und Praxis (z. B. Schrader et al. 2020). Um Einflussfaktoren und Wirkungen einer Implementation zu systematisieren, verwenden Schrader et al. (2020) sogenannte Angebots-Nutzungs-(Wirkungs)-Heuristiken (z. B. Fend 1998). Im Kern dieser Modelle steht das von Lehrenden implementierte Angebot, das getrennt von der Nutzung durch die Lernenden betrachtet wird. Vor diesem Hintergrund scheint es vielversprechend, die Implementation sowohl durch Lehrende als auch durch Lernende vergleichend in den Blick zu nehmen. Für PT wurde die Implementation ausschließlich über Lehrende (z. B. Maheady et al. 1991), nur über Lernende (z. B. Dufrene et al. 2005) oder in einem aus beiden Akteuren gemeinsam bestimmten Implementationsparameter (Adl-Amini et al. 2014; Topping et al. 2011) erfasst.

Wenn ein PT-Training möglichst genau – so wie es konzipiert wurde – umgesetzt werden soll (Top-down-Strategie; Gräsel und Parchmann 2004), steht die Fidelity-Perspektive im Fokus der Implementationsforschung (Dane und Schneider 1998; O'Donnell 2008). Dane und Schneider (1998) unterscheiden hierbei verschiedene Parameter (vgl. auch Sanetti und Kratochwill 2009), von denen die „adherence“ bzw. Umsetzungsgenauigkeit der gängigste Parameter ist (Dane und Schneider 1998;

O'Donnell 2008). Um die Wirksamkeit einer Maßnahme tatsächlich auf die Maßnahme selbst zurückführen zu können (interne Validität), sollte eine hinreichend hohe Umsetzungsgenauigkeit vorliegen. Dafür wird oft ein Cut-off-Wert von 70 % der vorgesehenen Elemente vorgeschlagen (z. B. Reinecker 2009).

Studien zur Implementation von PT im Klassenkontext zeigen, dass die Umsetzungsgenauigkeit, wenn Lehrkräfte geschult wurden und Trainings-Material vorliegt, im Mittel sehr hoch ist und weit über 70 % liegt (z. B. Adl-Amini et al. 2014; Dufrene et al. 2005; Maheady et al. 1991).

### 2.3 Implementation von mehrsprachiger Kommunikation beim Lernen

Bei der Implementation mehrsprachiger Kommunikation im institutionalisierten Bildungskontext müssen jedoch die spezifischen Bedingungen noch stärker berücksichtigt werden: Modelle mehrsprachiger Kommunikation (z. B. Grosjean 2008) weisen auf die Bedeutung des Kontexts, der Eigenschaften der interagierenden Personen sowie des Gesprächsthemas hinsichtlich der Quantität und Qualität bilingualen Sprachhandelns hin. So kommunizieren Schülerinnen und Schüler mit verschiedenen Herkunftssprachen im Unterricht im Submersionskontext nur in einer offiziellen Unterrichtssprache. Mit Ausnahme neu zugewanderter Schülerinnen und Schüler liegen in Deutschland in der Regel allerdings Sprachasymmetrien, insbesondere hinsichtlich bildungssprachlicher Fähigkeiten, zu Gunsten der Verkehrssprache vor (z. B. Reich 2009). Zudem können bilingual aufwachsende Schülerinnen und Schüler selbst negative Einstellungen zu ihrer Herkunftssprache (als Bildungsressource) in der Submersion entwickeln (vgl. Krumm 2009), was auch durch assimilatorische Sprachideologien (vgl. Bourhis 2001) und mediale Problematisierung von migrationsbedingter Mehrsprachigkeit (z. B. „Deutschpflicht auf dem Schulhof“) zu erklären ist. Diese Faktoren können mehrsprachige Kommunikation hemmen, auch wenn den Lernenden ein explizites Angebot gemacht wird.

Die bisherige Evidenz aus verschiedenen Bildungskontexten zum Umfang bilingualer Kommunikation beim kooperativen Lernen zeigt eine hohe Heterogenität bilingualer Kommunikation mit einem Anteil von 21 bis 85 % (Planas und Setati 2009; Schüler-Meyer et al. 2017; Storch und Wigglesworth 2003; Swain und Lapkin 2000). Für den Submersionskontext berichten Planas und Setati (2009) in ihrer Fallstudie bis zu 85 % herkunftssprachlicher Kommunikation bei zwei LI-dominanten Gruppen (Spanisch-Katalanisch). Jedoch ist das Katalanische zwar die offizielle Schulsprache, aber das Spanische außerhalb der Schule eine gleichberechtigte Landessprache. Zudem sind beide Sprachen linguistisch ähnlich. Schüler-Meyer et al. (2017) berichten in ihrer Interventionsstudie im Schnitt 31 % bilinguale Kommunikation bei türkisch-deutschsprachigen Schülerinnen und Schülern mit schriftsprachlichen Fähigkeiten im Türkischen, aber einer Sprachdominanz zu Gunsten des Deutschen. Die Kleingruppen bei Schüler-Meyer et al. (2017) wurden dabei von einer bilingual türkisch-deutschsprachigen Übungsleitung mit bilingualen Materialien durchgehend angeleitet, die im Schnitt zu 67 % bilingual mit den Schülerinnen und Schülern kommunizierte.

Insgesamt erscheint die Implementation bilingualer Kommunikation beim kooperativen Lernen mit deutsch-dominanten bilingualen Schülerinnen und Schülern im

Submersionskontext durch bilinguale Materialien sowie permanenter Anleitung und bilingualen Inputs mehrsprachiger Übungsleitungen durchaus möglich. Die bisherigen Ergebnisse beziehen sich aber auf ganz unterschiedliche Bedingungen und erlauben keine allgemeine Verortung zum Umfang bilingualer Kommunikation im Submersionskontext. Es lässt sich jedoch grundsätzlich ableiten, dass die Fidelity-Perspektive mit der Annahme, dass eine Intervention möglichst genau wie intendiert umgesetzt wird, bei der Erfassung mehrsprachiger Kommunikation im Submersionskontext weder ein realistischer noch ein inhaltlich sinnvoller Maßstab sein kann. Vielmehr muss eine Auswertung anhand unterschiedlicher Kriterien erfolgen, um neue Erkenntnisse zur Aktivierung bilingualer Kommunikation beim Lernen zu gewinnen. Im Kontext der Implementationsforschung kann zusätzlich der Parameter der Trainings-Differenziertheit („program differentiation“; z. B. Dane und Schneider 1998) herangezogen werden. Diese ist gegeben, wenn der Umfang bilingualer Äußerungen in Gruppen mit Mehrsprachigkeitsangebot über den Umfang in Gruppen ohne Sprachangebot als reguläres Setting hinausgeht. Hierdurch kann die Bedeutung des bilingualen Angebots jenseits individueller sprachlicher Voraussetzungen der Tandems aufgezeigt werden. Als Drittes gilt es im Kontext der Mehrsprachigkeitsforschung auch stets mit zu untersuchen, inwieweit eine Aktivierung bilingualer Kommunikation in den einzelnen Trainingsgruppen überhaupt gelingt (nominales Kriterium).

### 3 Fragestellungen und Hypothesen

Der Beitrag zielt darauf ab, neue Erkenntnisse (a) zur Implementation außerunterrichtlicher PT-Trainings, (b) zum Umfang bilingualer Kommunikation beim PT und zu deren Erfassung im Kontext der Implementationsforschung sowie (c) zur Übereinstimmung der Implementation durch Trainingsleitungen und durch Lernende zu gewinnen. Dafür werden folgende Fragestellungen verfolgt und Hypothesen geprüft:

(1) Inwieweit gelingt die Implementation von Elementen außerunterrichtlicher PT-Trainings?

(1a) Wir erwarten bei geschulten Trainingsleitungen mit manualgestützten Trainingsabläufen eine im Mittel hohe Umsetzungsgenauigkeit der PT-Trainings.

(1b) Wir erwarten bei den Tandems eine im Mittel hohe Umsetzungsgenauigkeit der PT-Trainings.

(2) Inwieweit wird bei den außerunterrichtlichen PT-Trainings bilingual kommuniziert?

(2a) Wir erwarten bei geschulten Trainingsleitungen mit manualgestützten Trainingsabläufen im Mittel eine hohe Umsetzungsgenauigkeit der vorgesehenen bilingualen Kommunikation und eine hohe Trainingsdifferenziertheit sowie eine Erfüllung des nominalen Kriteriums bei allen Trainingsleitungen.

(2b) Bei den Tandems erwarten wir auf Basis bisheriger Befunde, dass die Aktivierung bilingualer Kommunikation im Mittel gelingt, aber nur in moderatem Umfang, sodass im Mittel keine hohe Umsetzungsgenauigkeit vorliegt, aber eine hohe Trainingsdifferenziertheit gegeben ist und die Tandems mehrheitlich das nominale Kriterium erfüllen.

(3) Welche Übereinstimmung zwischen Trainingsleitungen und Tandems zeigt sich bei der Implementation der außerunterrichtlichen PT-Trainings sowie bei der bilingualen Kommunikation?

(3a) Wir erwarten im Sinne eines Angebots-Nutzungs-Modells, dass die Umsetzung der PT-Elemente durch die Trainingsleitungen mit der Umsetzung durch die Tandems zusammenhängt.

(3b) Inwieweit die bilinguale Kommunikation der Trainingsleitungen mit der bilingualen Kommunikation durch die Tandems zusammenhängt, bleibt zu explorieren.

## 4 Methode

### 4.1 Interventionsdesign

Die Daten stammen aus einer Interventionsstudie in der dritten und vierten Jahrgangsstufe im Zeitraum 2015 bis 2017 (Rauch et al. [in Vorb.](#)). Hierbei fanden am Nachmittag nach dem Unterricht PT-Trainings in den jeweiligen Grundschulen der teilnehmenden Kinder statt. Immer zwei Peers (ein Tandem) wurden von einer Trainingsleitung instruiert und begleitet. Die Trainingsleitungen (mehrheitlich Studierende des Lehramts, der Psychologie und Erziehungswissenschaft) wurden in den Abläufen zweitägig geschult und nutzten für die Durchführung der Sitzungen ein standardisiertes Manual sowie vorgegebene Arbeitsmaterialien. Tandems der Gruppen mit bilingualer Kommunikation wurden von türkisch-deutschsprachigen Trainingsleitungen begleitet. Der Interventionsstudie lag ein besonderes Design zugrunde (vgl. [Tab. 1](#)): In den Treatment-Gruppen fand eine Förderung von Lesestrategien, in den Treatment-Kontrollgruppen eine Förderung von Rechenstrategien statt. Zu jedem Inhalt gab es jeweils drei Gruppen, sodass die für das Lernen verwendeten Sprachen (nur Deutsch vs. Türkisch und Deutsch) sowie die Sprachhintergründe eines Tandems (monolingual-bilingual vs. bilingual-bilingual) variiert wurden. Aufgrund des Designs ist es möglich, Gruppen mit bilingualen Lernenden, die Türkisch für das Lernen nutzen durften (LG3 + RG3), mit Gruppen mit bilingualen Lernenden, die nur Deutsch als Lehr-Lernsprache nutzen sollten (LG2 + RG2), zu vergleichen und hieraus Aussagen zur Implementation von bilingualer Kommunikation durch das Training abzuleiten. Für die Umsetzungsgenauigkeit der PT-Elemente wurden die drei Lese- sowie die drei Rechengruppen aufgrund paralleler PT-Konzeption und zur Erhöhung der statistischen Power zusammengefasst. Vor, während und nach der Intervention fanden quantitative Erhebungen statt. Während des Trainings wurden

**Tab. 1** Interventionsdesign

Trainingsinhalt	Lesen			Rechnen		
	LG1	LG2	LG3	RG1	RG2	RG3
Gruppenname	D	D	T-D	D	D	T-D
Sprachangebot	m + b	b + b	b + b	m + b	b + b	b + b
Sprachhintergrund (Peer 1 + Peer 2)						

Anmerkungen: *LG* Lesegruppe, *RG* Rechengruppe, *D* Deutsch, *T-D* Türkisch-Deutsch, *m* monolingual Deutsch, *b* bilingual Türkisch-Deutsch

die Gespräche der Sitzungen sechs, neun und zwölf über Diktiergeräte aufgezeichnet. Sitzung neun wurde zusätzlich transkribiert.

## 4.2 Trainingsinhalte

Im Lesetraining wurden die Strategien Wörter klären, Zusammenfassen und Vorhersagen (vgl. Spörer et al. 2009), im Rechentraining halbschriftliche Strategien des Addierens und Subtrahierens (vgl. Heinze et al. 2009) eingeübt. In Anlehnung an etablierte PT-Programme (vgl. Fuchs et al. 1997; Munser-Kiefer und Kirschhock 2012) nahmen die Peers zur Einübung der Strategien die Rollen „Spielerin“ bzw. „Spieler“ (wendet Strategien an, kann um Hilfe bitten) und „Trainerin“ bzw. „Trainer“ (leitet an, evaluiert, gibt Hilfestellungen) ein. Unterstützend konnten die Peers Hilfskarten einbeziehen, die die Teilschritte der Strategien darstellten sowie die Handlungsrouitinen in den jeweiligen Rollen verdeutlichten. In den Gruppen LG3 und RG3 durften die Kinder auch auf Türkisch oder Türkisch-Deutsch miteinander kommunizieren. Zur Anregung wurden Material (z. B. türkisch-deutschsprachige Hilfskarten, eine Fallvignette zu mehrsprachigem Lernen; siehe Schastak 2020) und bilinguale Spiele eingesetzt. Zudem gaben die türkisch-deutschsprachigen Trainingsleitungen Instruktionen auf Türkisch, die in den Manualen zur Durchführung der Instruktionssitzungen ausformuliert waren.

Die Trainings umfassten je 12 Sitzungen à 45 min. In den ersten sechs Sitzungen (Instruktionsphase) wurden die Strategien sowie die Rollen von einer geschulten Trainingsleitung eingeführt und modelliert. Jede Sitzung wurde durch ein kurzes Spiel eröffnet, Inhalte der vorherigen Sitzung wurden von den Peers wiederholt und anschließend wurden die Strategien eingeführt und erklärt sowie von den Peers unter Supervision und Feedback der Trainingsleitung erprobt. In den anschließenden sechs Sitzungen (Übungsphase) wurden die Inhalte von den Peers weiter eingeübt, während die Trainingsleitung die Abläufe vorrangig supervidierte, sodass die Verantwortung zunehmend bei den Peers lag. Die Tandems erhielten zu Beginn der Sitzung die benötigten Arbeitsmaterialien und sollten diese nach den vorgegebenen Abläufen bearbeiten und am Ende gemeinsam ihre Zusammenarbeit reflektieren. Im Lesetraining arbeiteten die Tandems durchgängig in reziproken Rollen, wobei die Rollenhandlungen und deren Wechsel an die Strategiebearbeitung gekoppelt waren, sodass die Rollen vom Fortschritt bei der Strategie-Anwendung abhingen. Im Rechentraining arbeiteten die Tandems nur in der ersten Hälfte der Sitzung („Teamrechnen“, ca. 20 min) in reziproken Rollen, in der anschließenden „Rechenkonferenz“ (ca. 20 min) bearbeiteten die Peers die Aufgaben individuell, verglichen anschließend ihre Ergebnisse und kontrollierten diese abschließend mit einem Lösungsblatt. Im Gegensatz zum Lesetraining war der Rollenwechsel im Rechentraining somit nicht durch die Bearbeitungsgeschwindigkeit der Peers bestimmt, sondern wurde zeitlich durch die Trainingsleitung angeleitet: Nach 10 min sollten die Rollen in der Teamrechenphase wechseln.



### 4.3 Erfassung der Implementation der Trainings

Die Erfassung der Implementation erfolgte für die Trainingsleitungen auf Basis der Audioaufnahmen der kompletten sechsten Sitzung (Ende Instruktionsphase), während für die Peers inhaltlich zentrale Phasen der neunten Sitzung herangezogen wurden: Im Rechentraining wurden die beiden Teamrechenphasen berücksichtigt, im Lesetraining waren es die Phasen der Anwendung der Strategien „Zusammenfassen“ und „Vorhersagen“, da diese alle zentralen Elemente des PT enthielten. Die Rechenkonferenz wurde aufgrund nicht vorgesehener Rollenhandlungen ausgeklammert. Die Lesestrategie „Wörter klären“ wurde ebenso nicht einbezogen, da diese nur eingesetzt werden sollte, wenn Wörter unbekannt waren. Ein Drittel des Materials wurde von einer zweiten Person kodiert. Für die Rolleneinnahme durch die Tandems im Rechentraining ist die Interrater-Reliabilität als gut zu bewerten ( $\kappa > 0,66$ ); alle anderen Interventionselemente wurden mit einer sehr guten Übereinstimmung kodiert ( $\kappa > 0,80$ ).

Für die Umsetzung von PT wurden die beiden Kernelemente „Strategien“ und „Rollen“ einbezogen und der Implementationsparameter der Umsetzungsgenauigkeit bestimmt. Für die Trainingsleitungen wurden dazu Indikatoren anhand der in den Manualen vorgesehenen Instruktionen bestimmt und anhand des Audiomaterials geprüft, ob diese realisiert wurden (= 1) oder nicht (= 0). Hierbei war es unerheblich, ob die Übungsleitungen den genauen Wortlaut aus dem Manual verwendeten oder die Inhalte sinngemäß vermittelten. Für die Tandems wurden Indikatoren auf Basis der im Manual vorgesehenen Strategie- und Rollenhandlungen gebildet und bestimmt, ob die Tandems diese realisiert hatten (= 1) oder nicht (= 0). Jedes Kernelement wurde mit mehreren solcher dichotomen Indikatoren erfasst. Die Anzahl der Indikatoren hing dabei von den in den Manualen vorgesehenen übergeordneten Trainingsabläufen, den Handlungen sowie Handlungsspielräumen der Strategieausführung und der Rollen ab, sodass sowohl für Trainingsleitungen und Tandems als auch für Lesen und Rechnen teils unterschiedlich viele Indikatoren definiert wurden. Für die Umsetzung bilingualer Kommunikation wurden die Gruppen mit ausschließlich türkisch-deutsch bilingualen Kindern (LG2 + 3 bzw. RG2 + 3) einbezogen. Bei den Trainingsleitungen wurden die im Manual vorgesehenen türkischen Instruktionen identifiziert und geprüft, ob diese auf türkisch umgesetzt wurden (= 1) oder nicht (= 0). Bezüglich der Lese- und Rechen-Tandems wurden deutschsprachige sowie türkischsprachige und türkisch-deutsche Gesprächsbeiträge bestimmt und der Anteil türkisch(-deutsch-)sprachiger Turns an allen Gesprächsbeiträgen berechnet (Schastak 2020). Tab. 2 gibt einen Überblick über die einbezogenen Interventionselemente und deren Erfassung.

### 4.4 Stichprobe und Fehlende Werte

Insgesamt nahmen 170 Kinder (85 Tandems) an mindestens 9 Sitzungen der Interventionsstudie teil, davon 80 (40 Tandems) am Lesetraining und 90 (45 Tandems) am Rechentraining. Das durchschnittliche Alter lag bei 9;4 Jahren (identische Werte in den Lese- und Rechengruppen). Der Mädchenanteil betrug 55,9% ( $n=95$ ) und war im Lesen (57,5% bzw.  $n=46$ ) ähnlich hoch wie im Rechnen (54,4% bzw.  $n=49$ ).

**Tab. 2** Erfassung der Implementation der Interventionselemente

	Trainingsleitung	Tandems
Strategien	Instruiert die Trainingsleitung die Strategie-Anwendung? <i>Lesen:</i> 4 Indikatoren <i>Rechnen:</i> 8 Indikatoren <i>Beispiel-Item:</i> Die Trainingsleitung fordert das Tandem auf, die Lesestrategie „Vorhersagen“/die Rechenstrategie „Ergänzen“ zu wiederholen	Wenden die Peers die Strategien an? <i>Lesen:</i> 2 Indikatoren <i>Rechnen:</i> 2 Indikatoren <i>Beispiel-Item:</i> Peers fassen einen Abschnitt zusammen und machen eine Vorhersage zum nächsten Abschnitt/Peers rechnen mind. eine Rechenaufgabe je Teamrechenphase
Rollen	Instruiert die Trainingsleitung die Rollen? <i>Lesen:</i> 3 Indikatoren <i>Rechnen:</i> 3 Indikatoren <i>Beispiel-Item:</i> Die Trainingsleitung instruiert das Tandem, ihre Rollen zu tauschen	Arbeiten die Peers in den vorgesehenen Rollen? <i>Lesen:</i> 4–5 Indikatoren je Rolle und Strategie <i>Rechnen:</i> 5–6 Indikatoren je Rolle in jeder Teamrechenphase <i>Beispiel-Item:</i> Spieler/Spielerin verschriftlicht Zusammenfassung oder verbalisiert Rechenweg bzw. Trainer/Trainerin evaluiert Zusammenfassung oder Lösung
Bilinguale Kommunikation (nur LG2 + LG3 bzw. RG2 + RG3)	Spricht die Trainingsleitung türkisch? <i>Lesen:</i> 12 Indikatoren <i>Rechnen:</i> 12 Indikatoren <i>Beispiel-Item:</i> Die Trainingsleitung setzt einen lt. Manual vorgesehenen Impuls in türkischer Sprache bei der Instruktion der Strategien und Abläufe	Sprechen die Peers türkisch miteinander? <i>Lesen und Rechnen:</i> Kodierung von türkisch-(deutsch-)sprachigen sowie deutschsprachigen Gesprächsbeiträgen (Turns)

*Anmerkung:* Bei der Arbeit in Rollen wurde festgelegt, dass für jede Lesestrategie bzw. für jede Teamrechenphase mindestens zwei (Lesen) bzw. vier (Rechnen) typische Rollenhandlungen sowohl als Spielerin bzw. Spieler als auch als Trainerin bzw. Trainer umgesetzt werden mussten. Damit wurde inhaltlich festgesetzt, dass die Rollenhandlungen hinreichend umfangreich stattgefunden haben, da einige Elemente, wie ein kleinschrittiges Ausführen und Auffordern der Strategieausübung beim Zusammenfassen oder nach Hilfe fragen und Hilfestellungen geben, bei routinierten Strategieanwendung nicht zwingend realisiert werden mussten.

Von den 85 Tandems setzten sich 26 (30,6 %) aus zwei Jungen und 36 (42,4 %) aus zwei Mädchen zusammen; 23 Tandems (27,1 %) waren geschlechterheterogen zusammengesetzt. In den Lesegruppen gab es 12 Jungen-Tandems (30,0 %), 19 Mädchen-Tandems (47,5 %) und 9 geschlechtergemischte Tandems (22,5 %).

Aufgrund eines technischen Defekts/fehlerhafter Bedienung des Audioaufzeichnungsgeräts fehlen zwei Aufzeichnungen (1 × LG1, 1 × RG1) aus der sechsten Sitzung (Implementation durch die Trainingsleitungen) sowie eine Aufzeichnung (LG1) aus der neunten Sitzung (Implementation durch die Tandems). In der LG3 konnte nur für 11 von 13 Trainingsleitungen die bilinguale Kommunikation kodiert werden, da aufgrund verlangsamer Sitzungsabläufe nicht ausreichend Implementationsindikatoren zur Berechnung eines Scores vorlagen.

### 4.5 Datenanalyse

Die Prüfung der Hypothesen erfolgte unter Berücksichtigung der hierarchischen Datenstruktur (zwei Kinder in einem Tandem unter Supervision einer Trainingsleitung). Entsprechend gingen in die Analysen ausschließlich Werte auf Ebene der Trainingsleitungen bzw. Tandems ein. Zur Prüfung der ersten Hypothese wurde aus den jeweiligen Indikatoren die Umsetzungsgenauigkeit bestimmt, indem pro PT-Element („Strategien“ und „Rollen“) ein Summenwert und hieraus wiederum ein Prozentwert berechnet wurde. Anhand des Cut-off-Kriteriums von 70 % (Reinecker 2009) wurden die Prozentwerte in dichotome Variablen (1 = umgesetzt; 0 = nicht umgesetzt) überführt (Ausnahmen: bei max. 3 Indikatoren (Trainingsleitung Rollen) wurde bereits bei 2 erfüllten Indikatoren (66,7 %) eine 1 kodiert; bei max. 2 Indikatoren (Peers Strategien) mussten beide Indikatoren erfüllt sein (100 %), um eine 1 zu kodieren). Zur Prüfung der zweiten Hypothese (Umsetzung bilingualer Kommunikation) wurden bei den Trainingsleitungen ebenso auf Basis der Einzelindikatoren Summen- und Prozentwerte gebildet. Für die Tandems wurden auf Basis des Prozentsatzes türkisch-(deutsch-)sprachiger Kommunikation an der Gesamtkommunikation verschiedene Umsetzungskriterien gebildet: (a) die Umsetzungsgenauigkeit anhand eines Cut-off-Werts von 70 % (Reinecker 2009), (b) die Trainings-Differenziertheit als Anzahl der Tandems mit bilinguaem Kommunikationsangebot (LG3 bzw. RG3), deren Prozentsatz bilingualer Kommunikation über dem mittleren Prozentsatz bilingualer Kommunikation der Gruppen ohne bilinguaem Sprachangebot (LG2 bzw. RG2) lag, sowie (c) ein nominales Kriterium, bei dem die bilinguale Kommunikation als umgesetzt galt, wenn mindestens ein türkisch-(deutsch-)sprachiger Gesprächsbeitrag erfasst wurde.

Zur Prüfung der dritten Hypothese und Fragestellung wurde die Übereinstimmung der PT-Elemente zwischen der Implementation durch Trainingsleitungen (Angebot) und durch Tandems (Nutzung) in einem Vier-Felder-Schema dargestellt und als Übereinstimmungsmaß das korrigierte Kappa nach Brennan und Prediger (1981) berechnet. Die Übereinstimmung zwischen Trainingsleitungen und Tandems hin-

**Tab. 3** Umsetzungsgenauigkeit der PT-Elemente durch Trainingsleitungen (L) und Tandems (T)

	Strategien		Rollen		Gesamt (Strategien + Rollen)	
	L	T	L	T	L	T
<i>Lesen</i>	<i>n</i> = 39	<i>n</i> = 39	<i>n</i> = 39	<i>n</i> = 39	<i>n</i> = 39	<i>n</i> = 39
<i>M%</i> ( <i>SD%</i> )	93,2 (15,5)	94,9 (19,2)	86,3 (21,9)	84,5 (24,0)	89,7 (16,1)	89,7 (18,5)
Min-max (%)	50–100	0–100	33–100	0–100	42–100	25–100
Q1 Q2 Q3 Q4 ( <i>n</i> )	0 31 4 32	1 21 0 36	0 6 6 27	2 31 5 29	0 21 7 30	1 31 1 34
UG ( <i>n</i> ≥ 70 %)	36	36	33	33	34	34
<i>Rechnen</i>	<i>n</i> = 44	<i>n</i> = 45	<i>n</i> = 44	<i>n</i> = 45	<i>n</i> = 44	<i>n</i> = 45
<i>M%</i> ( <i>SD%</i> )	99,7 (1,8)	97,8 (10,4)	92,4 (17,4)	99,4 (3,7)	96,1 (9,0)	98,6 (6,6)
Min-max (%)	88–100	50–100	33–100	75–100	67–100	63–100
Q1 Q2 Q3 Q4 ( <i>n</i> )	0 0 0 44	0 21 0 43	0 21 6 36	0 0 11 44	0 0 2 42	0 0 2 43
UG ( <i>n</i> ≥ 70 %)	44	43	42	45	42	44

Anmerkung: Q1 bis Q4 Quartile, UG Umsetzungsgenauigkeit

sichtlich der bilingualen Kommunikation in den Gruppen LG3 bzw. RG3 wurde über eine Korrelation erfasst.

## 5 Ergebnisse

### 5.1 Umsetzungsgenauigkeit von PT-Elementen im außerunterrichtlichen Setting

Entsprechend der ersten Hypothese zeigt sich eine sehr hohe Umsetzungsgenauigkeit der beiden PT-Elemente im außerunterrichtlichen Paar-Setting, und zwar sowohl für die Lese- und Rechentrainings als auch für Trainingsleitungen und Tandems (vgl. Tab. 3). Deskriptiv lässt sich eine höhere Varianz im Lese- im Vergleich zum Rechentraining feststellen. Anhand der Quartile lässt sich aufzeigen, dass diese auf einzelne Trainingsleitungen bzw. Tandems zurückzuführen ist.

### 5.2 Aktivierung und Implementation von bilingualer Kommunikation beim außerunterrichtlichen PT

Die Ergebnisse der realisierten bilingualen Kommunikation mit (LG3 + RG3) und ohne (LG2 + RG2) bilinguaalem Sprachangebot sind in Tab. 4 dargestellt. Insgesamt zeigen die deskriptiven Daten, dass während der herangezogenen Audioaufnahmen in Gruppen ohne Sprachangebot kein nennenswerter Anteil bilingualer Kommunika-

**Tab. 4** Deskriptive Daten zur bilingualen Kommunikation von Trainingsleitungen (L) und Tandems (T) in Gruppen ohne und mit vorgesehenem bilingualen Sprachangebot

	Trainingssprache deutsch (LG2 bzw. RG2)		Trainingssprache türkisch-deutsch (LG3 bzw. RG3)	
	L	T	L	T
<i>Lesen</i>	<i>n</i> = 13	<i>n</i> = 13	<i>n</i> = 11	<i>n</i> = 13
<i>M%</i> ( <i>SD%</i> )	0 (0)	0,3 (0,7)	72,1 (16,4)	7,9 (10,8)
Min-max (%)	0	0–2,7	40,0–88,9	0–38,2
Umsetzungsgenauigkeit ( <i>n</i> ≥ 70 %)	0	0	6	0
Trainings-Differenziertheit ( <i>n</i> > <i>M</i> <sub>LG2</sub> )	–	–	11	10
Nominales Kriterium ( <i>n</i> > 0 %)	0	3	11	10
<i>Rechnen</i>	<i>n</i> = 17	<i>n</i> = 17	<i>n</i> = 15	<i>n</i> = 14
<i>M%</i> ( <i>SD%</i> )	0 (0)	1,0 (2,6)	69,4 (20,4)	12,9 (17,1)
Min-max (%)	0	0–10,6	38,0–100	0–66,0
Umsetzungsgenauigkeit ( <i>n</i> ≥ 70 %)	0	0	7	0
Trainings-Differenziertheit ( <i>n</i> > <i>M</i> <sub>RG2</sub> )	–	–	15	13
Nominales Kriterium ( <i>n</i> > 0 %)	0	7	15	13

tion stattfand – weder von Trainingsleitungen noch von den Tandems. In den Gruppen mit angebotener bilingualer Kommunikation (LG3 + RG3) setzten die Trainingsleitungen in etwa 70% der vorgesehenen türkischsprachigen Impulse um, während in den Tandems im Mittel zu etwa 10% auf türkisch-deutsch kommuniziert wurde (Lesen: 7,9%; Rechnen: 12,9%).

Um die Implementation von bilingualer Kommunikation von Trainingsleitungen und Tandems beim außerunterrichtlichen PT zu prüfen (Hypothesen 2a und 2b), wurden drei verschiedene Kriterien herangezogen: (1) ein Cut-off-Wert zur Erfassung der Umsetzungsgenauigkeit, (2) der Vergleich zwischen Gruppen mit und ohne bilingualem Sprachangebot zur Erfassung der Trainingsdifferenziertheit sowie (3) die Setzung von mindestens einem bilingualen Gesprächsanteil als nominales Kriterium.

Für die Trainingsleitungen lässt sich die angenommene hohe Umsetzung nicht bestätigen: Legt man den Cut-off-Wert von 70% zugrunde, implementierten nur jeweils knapp die Hälfte in den Lese- (6/11) bzw. Rechengruppen (7/15) die bilingualen Impulse. Die beiden weiteren Kriterien, Trainings-Differenziertheit und nominales Kriterium, zeigen dagegen eine Implementation in allen Gruppen mit bilingualem Angebot und keine Implementation in allen Gruppen ohne bilingualem Sprachangebot.

Für die Tandems resultiert ein vergleichbares Ergebnis: Unter Verwendung des 70%-Cut-off-Werts zeigt sich in keinem Tandem eine hinreichend hohe Umsetzungsgenauigkeit. Die Annahmen zu Trainings-Differenziertheit und nominalem Kriterium ließen sich bestätigen. In 10 von 13 Lese-Tandems (76,9%) sowie in 13 von 14 Rechen-Tandems (92,9%) waren beide Kriterien erfüllt. Wenngleich die Lernenden auch ohne bilinguales Sprachangebot in 3 der 13 Lese-Tandems bzw. 7 der 17 Rechen-Tandems vereinzelt auch auf türkisch kommunizierten, war der Anteil an Tandems mit bilingualer Kommunikation in den Gruppen mit Sprachangebot signifikant und bedeutsam höher (Lesen:  $\chi^2=7,54$ ,  $p=0,006$ ,  $\phi=0,538$ ; Rechnen:  $\chi^2=8,96$ ,  $p=0,003$ ,  $\phi=0,916$ ). Verwendet man für den Vergleich beider Gruppen die prozentualen Gesprächsanteile, resultiert ebenso ein signifikanter und bedeutsamer Unterschied (Lesen: Mann-Whitney-U-Test=27,00,  $p=0,002$ ,  $\eta^2=0,334$ ; Rechnen: Mann-Whitney-U-Test=23,00,  $p<0,001$ ,  $\eta^2=0,468$ ).

### 5.3 Übereinstimmung der Implementation durch Trainingsleitungen und Tandems

Als Drittes wurde geprüft, inwiefern die Umsetzung der Elemente der außerunterrichtlichen PT-Trainings sowie die bilinguale Kommunikation durch Trainingsleitungen und durch Tandems übereinstimmen (s. Tab. 5). Für die Umsetzungsgenauigkeit der PT-Elemente zeigen sich überwiegend gute bis sehr gute Übereinstimmungen. Nur in den Lesegruppen gibt es eine moderate Übereinstimmung in der Umsetzungsgenauigkeit der PT-Rollen durch Trainingsleitungen und durch Tandems. Für die bilinguale Kommunikation konnte mit den drei Kriterien (Umsetzungsgenauigkeit, Trainings-Differenziertheit, nominales Kriterium) keine Übereinstimmung berechnet werden, da entweder die Umsetzung der Tandems (beim Kriterium Umsetzungsgenauigkeit) oder die der Trainingsleitungen (bei Trainings-Differenziertheit und nominales Kriterium) keine Varianz aufwiesen. Bezogen auf den prozentualen

**Tab. 5** Übereinstimmung der Umsetzungsgenauigkeit durch Trainingsleitungen (L) und Tandems (T)

	Strategien		Rollen		Gesamt (Strategien+ Rollen)	
<i>Lesen</i>	T=0	T=1	T=0	T=1	T=0	T=1
L=0	1	2	1	5	1	3
L=1	2	33	5	27	4	30
$\kappa$	0,79		0,47		0,63	
<i>Rechnen</i>	T=0	T=1	T=0	T=1	T=0	T=1
L=0	0	0	0	2	0	1
L=1	2	42	0	42	1	41
$\kappa$	0,91		0,87		0,86	

Anmerkung:  $\kappa$  korrigiertes Kappa nach Brennan und Prediger (1981)

türkisch-(deutsch-)sprachigen Gesprächsanteil zeigen sich wider Erwarten negative Zusammenhänge zwischen Trainingsleitungen und Tandems (Lesen:  $r = -0,38$ ,  $p = 0,251$ ; Rechnen:  $r = -0,32$ ,  $p = 0,272$ ).

## 6 Diskussion

Der vorliegende Beitrag untersucht die Umsetzungsgenauigkeit von außerunterrichtlichen PT-Trainings im Lesen und Rechnen mit türkisch-deutschsprachigen Grundschulkindern. Die Trainings fanden in einem spezifischen Setting statt, indem jeweils zwei Lernende von einer Trainingsleitung angeleitet und begleitet wurden. Eine Besonderheit der Studie war es, dass in einigen Trainingsgruppen bilingual kommuniziert werden sollte. Die Erfassung und der Umfang bilingualer Kommunikation beim PT war ein weiterer Forschungsfokus dieses Beitrags.

### 6.1 Umsetzungsgenauigkeit der außerunterrichtlichen PT-Trainings

Bisherige Studien im klassenweiten PT belegen bereits eine hohe Umsetzungsgenauigkeit, wenn die Trainings manualbasiert mit vorstrukturierten Abläufen und Materialien erfolgen (z. B. Adl-Amini et al. 2014; Dufrene et al. 2005; Maheady et al. 1991). Wie vor diesem Hintergrund erwartet, zeigt sich auch in dem in dieser Studie realisierten außerunterrichtlichen Setting eine hohe Umsetzungsgenauigkeit der Lese- und Rechentrainings. Trotz des spezifischen Settings sind die Ergebnisse zur Umsetzung von PT anschlussfähig an bisherige Befunde, denn die Trainings wurden in Anlehnung an bewährte Programme (z. B. Munser-Kiefer und Kirschhock 2012) konzipiert und beinhalteten zwei PT-Kernelemente, nämlich das Einüben von Strategien und typische Rollenhandlungen der Lernenden (Fuchs et al. 1997; McMaster et al. 2006; Topping 2005).

Anknüpfend an Angebots-Nutzungs-(Wirkungs)-Heuristiken (vgl. Schrader et al. 2020) wurde in diesem Beitrag systematisch zwischen der Umsetzung durch Trainingsleitungen und durch Tandems unterschieden sowie deren Übereinstimmung geprüft. Die Ergebnisse zeigen, dass die vorgesehenen Strategien und Rollenhandlungen sowohl von den Trainingsleitungen als auch von Lernenden im Mittel mit hoher

Genauigkeit umgesetzt wurden und die Umsetzung von Lehrenden und Lernenden insgesamt in hohem Maße übereinstimmte. Dabei lagen die Werte im Rechentraining deskriptiv höher als im Lesetraining. Dieses Ergebnis könnte auch mit dem höheren Grad an Strukturierung im Rechentraining zusammenhängen: Im Rechentraining wechselten die Rollen nach 10 min, während im Lesetraining die Rollenhandlungen vom Fortschritt der Strategiebearbeitung durch die Peers abhingen. Vor diesem Hintergrund lässt sich ableiten, dass vor allem bei Trainings mit vielen Freiheitsgraden auf Seiten der Lernenden eine getrennte Erfassung der Umsetzungsgenauigkeit durch Lehrende und Lernende sinnvoll ist.

## 6.2 Umfang bilingualer Kommunikation beim außerunterrichtlichen PT

Die Umsetzungsgenauigkeit ist ein etablierter Parameter der Implementationsforschung und ermöglicht Aussagen dazu, wie gut eine Maßnahme in der Praxis umgesetzt werden kann. Für eine hinreichende bilinguale Kommunikation scheint dieser Parameter jedoch nur eingeschränkt geeignet zu sein. Dies begründet sich theoretisch durch Annahmen zur Bedeutung des Kontextes und der individuellen Merkmale der Lernenden für mehrsprachige Kommunikation im institutionalisierten Bildungskontext (Grosjean 2008). Bereits bei Grundschulkindern zeigen sich negativere Einstellungen zur eigenen Mehrsprachigkeit (Krumm 2009) und oft auch geringere (bildungssprachliche) Fähigkeiten in den Herkunftssprachen mit einer Asymmetrie zu Gunsten des Deutschen (Reich 2009). Empirische Befunde zum Umfang bilingualer Kommunikation beim kooperativen Lernen variieren substanziell, abhängig vom realisierten Kontext und von linguistischen Ähnlichkeiten zwischen den Sprachen (z. B. Planas und Setati 2009; Schüler-Meyer et al. 2017). Zum gegenwärtigen Zeitpunkt erlauben die wenigen Studien mit großen Unterschieden in ihren kontextuellen und individuellen Bedingungen noch keine Orientierung dazu, welcher Umfang an bilingualer Kommunikation erwartbar ist.

In dieser Studie kommunizierten die Lernenden zu 10 % des erfassten Gesprächs auf Türkisch-Deutsch, wenn sie von Trainingsleitungen dazu angeregt wurden und bilinguales Material vorlag. Der „monolinguale Habitus“ (Gogolin 1994) an deutschen Schulen trägt vermutlich maßgeblich dazu bei, dass die Lernenden Schwierigkeiten haben, von der habituierten Unterrichtssprache abzuweichen. Das Dilemma aus grundsätzlicher Freiwilligkeit eines Mehrsprachigkeitsangebots und geringer Aktivierung von bilingualer Kommunikation lässt sich nur über einen zunehmenden und regelmäßigen Einbezug von Mehrsprachigkeit im Unterricht bearbeiten.

Um weitere Erkenntnisse zum Umfang bilingualer Kommunikation beim PT zu erlangen, wurden die Trainings-Differenziertheit und ein nominales Kriterium als ergänzende Indikatoren verwendet. Die Trainings-Differenziertheit vergleicht die bilinguale Kommunikation mit und ohne bilingualem Sprachangebot und ermöglicht es somit, die Bedeutung eines bilingualen Angebots weitgehend unabhängig von den individuellen sprachlichen Voraussetzungen der Lernenden aufzuzeigen. Das Kriterium lässt sich allerdings nur dann inhaltlich sinnvoll anwenden, wenn es Gruppen mit und ohne Mehrsprachigkeitsangebot gibt, in denen idealerweise jeweils bilinguale Lernende sind. Die Ergebnisse dieses Beitrags zeigen nicht nur, dass mehrsprachige Lernende ohne Sprachangebot in diesem Setting nicht in ihrer

Herkunftssprache kommunizierten, sondern sie zeigen auch, dass das zusätzliche Angebot mit mehrsprachigem Material sowie mit Aufforderungen und Modellierungen durch die Trainingsleitungen mit einem signifikanten Anstieg bilingualer Kommunikation einherging.

Das nominale Kriterium (mind. ein bilingualer Gesprächsbeitrag) bietet einen weiteren Erkenntnisgewinn. Hierbei interessiert weniger der Umfang der Kommunikation, sondern es wird geprüft, ob eine Aktivierung bilingualer Kommunikation in jeder einzelnen Trainingsgruppe gelingt. Die Ergebnisse dieser Studie zeigen, dass eine Aktivierung in dreiviertel der Lesegruppen und über 90 % der Rechengruppen gelang. Zusammengenommen weisen die Ergebnisse zu Umfang und Aktivierbarkeit bilingualer Kommunikation auf eine prinzipielle Nutzung des Angebots hin, welche durch eine längerfristige, fächerübergreifende Implementation im Schulkontext samt sich damit entwickelnder bildungssprachlicher Fähigkeiten in der Herkunftssprache durchaus steigen könnte.

Vor dem Hintergrund der bisherigen Ausführungen erscheint es umso wichtiger, bei der Aktivierung bilingualer Kommunikation die Lehrenden und Lernenden separat in den Blick zu nehmen und deren Umsetzung vergleichend zu betrachten. In der vorliegenden Studie zeigten sich keine bzw. sogar negative Zusammenhänge zwischen dem Umfang bilingualer Kommunikation durch die Trainingsleitungen und in den Tandems. Dieser Befund widerspricht der Annahme, dass mehrsprachige Initiierungen und Modellierungen mit dem Ziel der Etablierung eines bilingualen Sprachmodus dazu beitragen können, dass Lernende mehrsprachig kommunizieren. Andererseits ließe sich argumentieren, dass wiederholte freiwillige Angebote nicht zwingend zur Nutzung führen, sondern, dass weitere Faktoren, wie die wahrgenommene Nützlichkeit der bilingualen Kommunikation eine Rolle spielen. Deshalb sollten diese Zusammenhänge in zukünftigen Studien weiter erforscht werden, um nicht zuletzt das spezifische Trainingssetting als Grund für den negativen Zusammenhang auszuschließen. So könnte die negative Übereinstimmung auch darauf hindeuten, dass die Trainingsleitungen am Ende der Instruktionsphase auf die vorgesehenen türkischsprachigen Impulse verzichteten, wenn in den Tandems bereits bilingual kommuniziert wurde.

### 6.3 Limitationen und Ausblick

Limitierend ist zunächst die mangelnde Generalisierbarkeit der Ergebnisse aufgrund der geringen Stichprobengröße anzuführen. Darüber hinaus ist einschränkend anzumerken, dass aufgrund der hohen Umsetzungsgenauigkeit Erklärungsfaktoren der Implementation als „Schlüsselaspekt der Implementationsforschung“ (Schrader et al. 2020, S. 13) nicht analysiert werden konnten. Hier liegt ein klassischer Zielkonflikt vor: Das erreichte Ziel einer hohen Umsetzung führt dazu, dass das Ziel der Analyse von Erklärungsfaktoren der Implementation kaum erreichbar ist. Als proximale Einflussfaktoren gelten der Heuristik von Schrader et al. (2020) zufolge Merkmale der Durchführenden sowie Merkmale der Lernenden und deren Wahrnehmung der Maßnahme. Vor allem die Wahrnehmung bzw. Rezeption der Maßnahme stellt gegenwärtig noch eine Black-Box dar, einzig eine grundsätzlich ablehnende Haltung gegenüber dem Lerngegenstand wird als hemmender Einflussfaktor genannt.



Auch die Trennung der Implementation durch Trainingsleitungen und Tandems erfolgte unter Verweis von Angebots-Nutzungs-Modellen und somit mit Bezug auf die Argumentation, dass Angebote von den Lernenden auch genutzt werden müssen, damit sie ihre Wirkung entfalten können. Die Gleichsetzung von Nutzung mit der Implementation durch die Lernenden ist dabei vor allem als argumentative Brücke für eine getrennte Erfassung der Implementation durch Trainingsleitungen und Tandems zu verstehen. Was jeweils unter Nutzung in Angebots-Nutzungs-Modellen gefasst wird, unterscheidet sich je nach Modell. Bei Schrader et al. (2020) werden unter Nutzung Involvement, Beteiligung und Engagement der Lernenden verortet. Hierbei zeigen sich deutliche Bezüge zum Implementationsparameter „participant responsiveness“ (vgl. Dane und Schneider 1998; Sanetti und Kratochwill 2009).

Darüber hinaus fokussierte die Studie die Umsetzungsgenauigkeit als Kernaspekt der Implementation (Dane und Schneider 1998; O'Donnell 2008). Greift man Forschung zur zentralen Bedeutung von Unterrichtsqualität auf (z.B. Klieme 2018), sollten zukünftige Studien die Unterrichtsqualität als weiteren Implementationsparameter („quality of delivery“) einbeziehen. Dieser Parameter könnte auch für die bilinguale Kommunikation operationalisiert werden, indem die Qualität der Diskurse als entscheidender Faktor für das Lernen erfasst wird (Littleton und Mercer 2010).

Schließlich ist zu diskutieren, inwiefern auf Basis der herangezogenen Zeitabschnitte Schlussfolgerungen für die Implementation des gesamten Trainings möglich sind. Für die Trainingsleitungen wurde die letzte Instruktionssitzung herangezogen, in der sie alle zuvor instruierten Abläufe und Strategien noch mal anleiten sollte. Somit sollten die Inhalte vorab implementiert worden sein. Für die Peers wurde die Sitzung neun einbezogen, in der die Peers das eigenständige Arbeiten bereits mehrfach angewendet haben sollten. Eine Umsetzung durch die Peers zu diesem Zeitpunkt sollte somit deren bisherige Anwendung gut abbilden können.

**Danksagung** Wir bedanken und ausdrücklich für das detaillierte und hilfreiche Feedback der Herausgeber\*innen und der Gutachter\*innen.

**Förderung** Die Studie wurde gefördert vom Bundesministerium für Bildung und Forschung (FKZ 01JM1404).

**Funding** Open Access funding enabled and organized 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

- Adl-Amini, K., Decristan, J., Hondrich, A. L., & Hardy, I. (2014). Umsetzung von peer-gestütztem Lernen durch Lehrkräfte im naturwissenschaftlichen Sachunterricht der Grundschule. *Zeitschrift für Grundschulforschung*, 7, 74–87.
- BiSS-Trägerkonsortium (Hrsg.). (2020). *Leseverstehen kennt keine Sprachgrenzen. Kooperativ und mehrsprachig Texte verstehen*. Köln: Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache.
- Bourhis, R. (2001). Acculturation, language maintenance and language shift. In J. Klatte-Folmer & P. van Avermaet (Hrsg.), *Theories on maintenance and loss of minority languages. Towards a more integrated explanatory framework* (S. 5–37). Münster: Waxmann.
- Bredthauer, S., & Engfer, H. (2018). Natürlich ist Mehrsprachigkeit toll! Aber was hat das mit meinem Unterricht zu tun? [https://kups.uni-koeln.de/8092/1/Manuskript\\_BredthauerEngfer-2.pdf](https://kups.uni-koeln.de/8092/1/Manuskript_BredthauerEngfer-2.pdf). Zugriffen: 4. Mai 2021.
- Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41, 687–699.
- Cohen, E. G. (1994). Restructuring the classroom: conditions for productive small groups. *Review of Educational Research*, 64, 1–35.
- Cummins, J. (2000). The threshold and interdependence hypothesis revisited. In J. Cummins (Hrsg.), *Language, power, and pedagogy. Bilingual children in the crossfire* (S. 173–200). Clevedon: Multilingual Matters.
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review*, 18, 23–45.
- Deci, E. L., & Ryan, R. M. (2002). Overview of self-determination theory: an organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 3–33). Rochester: University Rochester Press.
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3, 231–264.
- Dufrene, B. A., Noell, G. H., Gilbertson, D. N., & Duhon, G. J. (2005). Monitoring implementation of reciprocal peer tutoring: Identifying and intervening with students who do not maintain accurate implementation. *School Psychology Review*, 34, 74–86.
- Esser, H. (2009). Der Streit um die Zweisprachigkeit: Was bringt die Bilingualität? In I. Gogolin & U. Neumann (Hrsg.), *Streitfall Zweisprachigkeit—The Bilingualism Controversy* (S. 69–90). Wiesbaden: VS.
- Fend, H. (1998). *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistungen*. Weinheim: Juventa.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174–206.
- Gantefort, C., & Oroquieta, S. M. (2015). Translanguaging-Strategien im Sachunterricht der Primarstufe: Förderung des Leseverstehens auf Basis der Gesamtsprachigkeit. *Transfer Forschung ↔ Schule*, 1, 24–37.
- Gebhardt, M., Rauch, D., Mang, J., Sälzer, C., & Stanat, P. (2013). Mathematische Kompetenzen von Schülerinnen und Schülern mit Zuwanderungshintergrund. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 275–308). Münster: Waxmann.
- Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, 98, 732–749.
- Gogolin, I. (1994). Allgemeine sprachliche Bildung als Bildung zur Mehrsprachigkeit. In K.-R. Bausch, H. Christ & H.-J. Krumm (Hrsg.), *Interkulturelles Lernen im Fremdsprachenunterricht Arbeitspapiere der 14. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts* (S. 73–84). Tübingen: Gunter Narr.
- Gogolin, I., Hansen, A., McMonagle, S., & Rauch, D. (Hrsg.). (2020). *Handbuch Mehrsprachigkeit und Bildung*. Berlin: Springer.
- Gräsel, C., & Parchmann, I. (2004). Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 32, 238–256.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford: Oxford University Press.

- Heinze, A., Marschick, F., & Lipowsky, F. (2009). Addition and subtraction of three-digit numbers: adaptive strategy use and the influence of instruction in German third grade. *ZDM Mathematics Education*, *41*, 591–604.
- Honneth, A. (1992). *Kampf um Anerkennung. Zur moralischen Grammatik sozialer Konflikte*. Frankfurt a. M.: Suhrkamp.
- Kern, R. (1994). The role of mental translation in second language reading. *Studies in Second Language Acquisition*, *16*, 441–460.
- Klieme, E. (2018). Unterrichtsqualität. In M. Harring, C. Rohlfes & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 393–408). Münster: Waxmann.
- Krumm, H.-J. (2009). Die Bedeutung der Mehrsprachigkeit in den Identitätskonzepten von Migrantinnen und Migranten. In I. Gogolin & U. Neumann (Hrsg.), *Streitfall Zweisprachigkeit—The Bilingualism Controversy* (S. 233–247). Wiesbaden: VS.
- Littleton, K., & Mercer, N. (2010). The significance of educational dialogues between primary school children. In K. Littleton & C. Howe (Hrsg.), *Educational dialogues. Understanding and promoting productive interaction* (S. 272–288). London: Routledge.
- Lyster, R. (2019). Translanguaging in immersion: cognitive support or social prestige? *The Canadian Modern Language Review*, *75*, 340–352.
- Maheady, L., Harper, G.F., Malette, B., & Winstanley, N. (1991). Training and implementation requirements associated with the use of a classwide peer tutoring system. *Education & Treatment of Children*, *14*, 177–198.
- McMaster, K.N., Fuchs, D., & Fuchs, L.S. (2006). Research on peer-assisted learning strategies: The promise and limitations of peer-mediated instruction. *Reading & Writing Quarterly*, *22*, 5–25.
- Meyer, M., & Prediger, S. (2011). The use of first language Turkish as a resource. A German case study on chances and limits for building conceptual understanding. In M. Setati, T. Nkambule & L. Goosen (Hrsg.), *Proceedings of the ICMI study 21 mathematics and language diversity* (S. 225–234). Sao Paulo: University Press.
- Meyer, M., Prediger, S., César, M., & Norén, E. (2016). Making use of multiple (non-shared) first languages: state of and need for research and development in the European language context. In R. Barwell, P. Clarkson, A. Halai, M. Kazima, J. Moschkovich, N. Planas, M. Setati-Phakeng, P. Valero & M. Villavicencio (Hrsg.), *Mathematics education and language diversity: The 21st ICMI Study* (S. 47–66). Heidelberg: Springer.
- Moschkovich, J. (2007). Using two languages when learning mathematics. *Educational Studies in Mathematics*, *64*, 121–144.
- Munser-Kiefer, M., & Kirschhock, E. (2012). *Lesestrategien im Leseteam trainieren. Lehrmanual und Unterrichtsmaterialien*. Donauwörth: Auer.
- 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*, 33–84.
- Piaget, J., & Inhelder, B. (1983). *Die Psychologie des Kindes*. Frankfurt a. M.: Fischer.
- Planas, N., & Setati, M. (2009). Bilingual students using their languages in the learning of mathematics. *Mathematics Education Research Journal*, *21*, 36–59.
- Rauch, D., Decristan, J., Schastak, M., & Reitenbach, V. (in Vorbereitung). Effects of an intervention fostering German reading skills among bilingual primary school children through peer learning.
- Rehbein, J. (2011). „Arbeitsprache“ Türkisch im mathematisch-naturwissenschaftlichen Unterricht der deutschen Schule – ein Plädoyer. In S. Prediger & E. Özdil (Hrsg.), *Mathematiklernen unter den Bedingungen der Mehrsprachigkeit. Stand und Perspektiven der Forschung und Entwicklung in Deutschland* (S. 205–232). Münster: Waxmann.
- Reich, H.H. (2009). Entwicklungswege türkisch-deutscher Zweisprachigkeit. In U. Neumann & H.H. Reich (Hrsg.), *Erwerb des Türkischen in einsprachigen und mehrsprachigen Situationen* (S. 63–90). Münster: Waxmann.
- Reinecker, H. (2009). Therapieforschung. In J. Margraf & S. Schneider (Hrsg.), *Lehrbuch der Verhaltenstherapie. Grundlagen, Diagnostik, Verfahren, Rahmenbedingungen* (S. 84–99). Heidelberg: Springer.
- Rohrbeck, C.A., Ginsburg-Block, M.D., Fantuzzo, J.W., & Miller, T.R. (2003). Peer-assisted learning interventions in elementary school students: a meta-analytic review. *Journal of Educational Psychology*, *95*, 240–257.
- Sanetti, L.M.H., & Kratochwill, T.R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, *38*, 445–459.
- Schastak, M. (2020). *Bilinguale Interaktion beim Peer-Learning in der Grundschule. Eine Mixed-Methods Studie mit bilingual türkisch-deutschsprachig aufwachsenden Schüler\*innen Reihe Mehrsprachigkeit*

- und Bildung (Bd. 4). Leverkusen: Barbara Budrich. Open Access: <https://shop.budrich-academic.de/produkt/bilinguale-interaktion-beim-peer-learning-in-der-grundschule/>
- Schrader, J., Hasselhorn, M., Hetfleisch, P., & Goeze, A. (2020). Stichwortbeitrag Implementationsforschung: Wie Wissenschaft zu Verbesserungen im Bildungssystem beitragen kann. *Zeitschrift für Erziehungswissenschaft*, 23(1), 9–59.
- Schüler-Meyer, A., Prediger, S., Kuzu, T., Wessel, L., & Redder, A. (2017). Is formal language proficiency in the home language required to profit from a bilingual teaching intervention in Mathematics? A mixed methods study on fostering multilingual students' conceptual understanding. *International Journal of Science and Mathematics Education*, 34, 1–23.
- Spörer, N., Brunstein, J.C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction*, 19, 272–286.
- Spörer, N., Seuring, V., Schünemann, N., & Brunstein, J.C. (2008). Förderung des Leseverständnisses von Schülern der 7. Klasse. *Zeitschrift für Pädagogische Psychologie*, 22, 247–259.
- Storch, N., & Wigglesworth, G. (2003). Is there a role for the use of the L1 in an L2 setting? *Teachers of English to Speakers of Other Languages*, 37, 760–770.
- Stošić, P. (2017). Kinder mit ‚Migrationshintergrund‘ – Reflexionen einer (erziehungs-)wissenschaftlichen Differenzkategorie. In I. Diehm, M. Kuhn & C. Machold (Hrsg.), *Differenz – Ungleichheit – Erziehungswissenschaft. Verhältnisbestimmungen im (Inter-)Disziplinären* (S. 81–99). Wiesbaden: Springer.
- Swain, M., & Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research*, 4, 251–274.
- Topping, K.J. (2005). Trends in peer learning. *Educational Psychology*, 25, 631–645.
- Topping, K., Miller, D., Murray, P., & Conlin, N. (2011). Implementation integrity in peer tutoring mathematics. *Educational Psychology*, 31, 575–593.
- Wygotsky, L.S. (1978). *Mind in society. The development of higher psychological processes*. Cambridge: MS: Harvard University Press.