

Kathrin Fussangel, Gino Casale, Jacquelin Kluge,
Markus Spilles & Michael Grosche

Die Messung kokonstruktiver Kooperation: Entwicklung und Validierung eines Fragebogens für Lehrkräfte in der Inklusion

Zusammenfassung

Der Beitrag stellt die Entwicklung eines Fragebogens zur Messung von kokonstruktiver Kooperation im Rahmen schulischer Inklusion dar. Die theoretische Grundlage ist das umfassende Modell der Kokonstruktion von Grosche, Fussangel und Gräsel (2020a), auf dessen Basis eine große Anzahl an Items formuliert wurde. Die Entwicklung des Fragebogens wird anhand von vier aufeinander aufbauenden und unabhängigen Teilstudien dargestellt, in denen die Anzahl der Items immer weiter reduziert und die Anforderungen an die Messgüte des Fragebogens systematisch gesteigert wurden. Die finalen konfirmatorischen Faktorenanalysen zeigen, dass der Fragebogen die theoretischen Dimensionen des Kokonstruktionsmodells empirisch adäquat abbildet.¹

Schlagworte

Konstruktion, Fragebogenentwicklung, Inklusion, Kooperation

Prof. Dr. Kathrin Fussangel (Korrespondenzautorin) · Prof. Dr. Gino Casale · Jacquelin Kluge · Dr. Markus Spilles · Prof. Dr. Michael Grosche, Bergische Universität Wuppertal Institut für Bildungsforschung, Gaußstr. 20, 42119 Wuppertal
E-Mail: fussangel@uni-wuppertal.de
gcasale@uni-wuppertal.de
jkluge@uni-wuppertal.de
spilles@uni-wuppertal.de
grosche@uni-wuppertal.de

1 Der Fragebogen ist unter einer CC-BY-Lizenz zur Nutzung frei verfügbar.

Measuring Co-constructive Collaboration: Development and Validating of a Questionnaire for Teachers in Inclusion

Abstract

The article presents the development of a questionnaire to measure co-constructive cooperation in the context of inclusion in schools. The comprehensive model of co-construction by Grosche, Fussangel, and Gräsel (2020a) represents the theoretical framework for the development of a larger number of items. The development of the questionnaire is pictured on the basis of four consecutive and independent sub-studies in which the number of items was continuously decreased, while the requirements for the measurement quality of the questionnaire were systematically increased. The final confirmatory factor analyses show that the questionnaire supports the theoretical dimensions of the co-construction model empirically in an adequate manner.²

Keywords

Co-construction, questionnaire development, inclusion, cooperation

1. Einleitung

Die Kooperation von Regelschul- und sonderpädagogischen Lehrkräften zur Umsetzung schulischer Inklusion ist ein bedeutsames Thema der beruflichen Professionalisierung. Es sind v. a. diese beiden Gruppen, die im unterrichtlichen Alltag für Kinder und Jugendliche mit sonderpädagogischen Förderbedarfen verantwortlich sind (Jurkowski & Müller, 2018). Die Umsetzung von Inklusion scheint jedoch hohe Anforderungen an die Kooperation von Lehrkräften zu stellen (Grosche & Moser Opitz, 2023). Es genügt nicht, bloß einige schüler:innenbezogene Informationen auszutauschen. Vielmehr müssen alle Lehrkräfte gemeinsam für das Lernen aller Schüler:innen und die gemeinsame Konzipierung von Unterricht und Fördermaßnahmen verantwortlich sein (Kluge & Grosche, 2021; Lütje-Klose & Urban, 2014; Werning, 2013).

Vor diesem Hintergrund steht insbesondere die sogenannte Kokonstruktion als eine intensive Form der Kooperation im Fokus des Interesses vieler Forschenden geraten. Kokonstruktion meint, dass kooperierende Personen ihr Wissen aufeinander beziehen müssen, um neue Konzepte, Ideen und Maßnahmen zu entwickeln und zu reflektieren (Gräsel, Fussangel & Pröbstel, 2006). Die so gemeinsam konstruierten (ko-konstruierten) Wissensbestände oder Kompetenzen dürften auf einem Niveau liegen, das eine Person alleine nicht hätte erreichen können.

Das Konstrukt der Kokonstruktion wurde jüngst von Grosche, Fussangel und Gräsel (2020a) theoretisch geschärft und auf den Kontext der schulischen Inklusi-

² The questionnaire is freely available under a CC BY license.

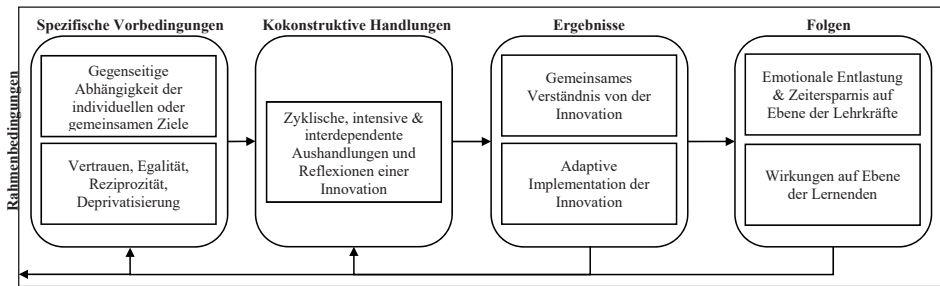
on übertragen. Demnach beschreibt Kokonstruktion die intensiven Aushandlungsprozesse, die nötig sind, um die regelschul- und die sonderpädagogische Perspektive aufeinander zu beziehen. Durch diesen Bezug zueinander entwickeln Lehrkräfte eine gemeinsame Sicht auf und eine gemeinsame Verantwortung für alle Lernenden mit ihren individuellen Bedürfnissen. Diese gemeinsame Verantwortung könnte dazu beitragen, Inklusion zielgerichteter und erfolgreicher umzusetzen (Kluge & Grosche, 2021). Damit liegt ein theoretisches Modell der Kokonstruktion vor, dessen empirische Überprüfung jedoch noch aussteht (Grosche et al., 2020a).

Um kokonstruktive Kooperation empirisch-quantitativ zu erforschen, bedarf es zunächst eines psychometrisch fundierten Messverfahrens, das das theoretisch postulierte Modell reliabel und valide abbildet. Im vorliegenden Beitrag werden die Entwicklung und psychometrische Überprüfung eines Instruments zur Messung der kokonstruktiven Kooperation beschrieben. Das neue Instrument geht dabei über die bisherigen Skalen zur Erfassung verschiedener Kooperationsformen hinaus und verfolgt das Ziel, Kokonstruktion mitsamt den spezifischen Vorbedingungen und Ergebnissen abzubilden. Zur Entwicklung und Validierung des Instruments wurden vier aufeinander aufbauende und unabhängige Teilstudien durchgeführt.

2. Kokonstruktive Kooperation

Das von Grosche et al. (2020a) ausgearbeitete theoretische Modell der Kokonstruktion definiert diese intensive Kooperationsform als „gemeinsame, zyklische, intensive und interdependente Aushandlungen und Reflexionen von Innovationen und ihren Konkretisierungen“ (S. 476). Dabei wird deutlich, dass Kokonstruktion nicht zu jeder Zeit im schulischen Alltag umgesetzt werden kann; vielmehr stellt die Kokonstruktion eine anspruchsvolle Kooperationsform dar, die bereits in der ersten Ausarbeitung des Kooperationsansatzes von Gräsel et al. (2006) als „high cost“ (S. 211) bezeichnet wurde und sich damit von den anderen beiden in diesem Ansatz beschriebenen Kooperationsformen, dem Austausch und der Arbeitsteilung, abgrenzt. Austausch meint beispielsweise die Weitergabe von Materialien oder Informationen. Arbeitsteilung beinhaltet die Aufteilung von Arbeitspaketen, die zwar auf ein übergeordnetes gemeinsames Ziel bezogen sind (z. B. Entwicklung einer Unterrichtsreihe), aber unabhängig voneinander bearbeitet werden (z. B. individuelle Ausarbeitung einzelner Stunden dieser Reihe). Das gemeinsame Ziel ist deutlich unverbindlicher als bei der Kokonstruktion und damit keine zwingend notwendige Bedingung für das Zustandekommen dieser Kooperationsform. Austausch und Arbeitsteilung wurden von den Autor:innen als weniger voraussetzungsreich und damit leichter im Alltag umsetzbar beschrieben (Gräsel et al., 2006). Die Definition der Kokonstruktion hingegen bringt zum Ausdruck, dass das Engagement der beteiligten Lehrkräfte höher sein muss, da diese Kooperationsform an spezifischere Bedingungen geknüpft ist. Die folgende Abbildung 1 zeigt das neu ausgearbeitete

Abbildung 1: Zyklisches Wirkmodell der kokonstruktiven Kooperation



theoretische Modell nach Grosche et al. (2020a), das in den folgenden Abschnitten überblicksartig beschrieben wird.

Das Modell besteht aus fünf Dimensionen (Rahmenbedingungen, spezifische Vorbedingungen, kokonstruktive Handlungen, Ergebnisse, Folgen), die in Abbildung 1 durch die vier inneren Kästen und den einen äußeren Rahmen repräsentiert sind. *Allgemeine Rahmenbedingungen* schulischer Kooperation führen dazu, dass Kooperation im schulischen Alltag gefördert und aufrechterhalten wird. Zu diesen allgemeinen Rahmenbedingungen zählen strukturelle (Zeitfenster, Räume), schulkulturelle (Kooperationsklima, Unterstützung der Schulleitung) sowie individuelle Aspekte (Reflexionsfähigkeit, grundsätzliche Autonomie, Sympathie, individuelle Erfahrungen und Einstellungen zur Kooperation). Die Bedeutung dieser allgemeinen Rahmenbedingungen für das Zustandekommen und die Ausprägung von Kooperation konnte in der Forschung vielfach gezeigt werden (Massenkeil & Rothland, 2016; Vangrieken, Dochy, Raes & Kyndt, 2015). Weil diese Rahmenbedingungen auch andere Formen der Kooperation ermöglichen, also auch relevant für die oben beschriebenen Formen des Austausches und der Arbeitsteilung sind, stellen sie keine spezifische Vorbedingung für die hier betrachtete Kokonstruktion dar.

Daher beschreiben Grosche et al. (2020a) *spezifische Vorbedingungen* für Kokonstruktion, die sich in intra- und interpersonale Bedingungen aufteilen lassen und die als notwendig für das Zustandekommen von Kokonstruktion angesehen werden. Damit unterscheidet sich die Kokonstruktion vom Austausch und der Arbeitsteilung, die nicht auf diese Vorbedingungen angewiesen sind. Eine zentrale interpersonale Variable stellt die Zielinterdependenz dar. Ohne Zielinterdependenz würden Lehrkräfte keine kokonstruktiven Kooperationen eingehen können, weil sie ihre Abhängigkeit von der Expertise und den Kompetenzen ihrer Kooperationspartner:innen für ihre Zielerreichung nicht sehen. Dies bedeutet, dass nur dann von Kokonstruktion gesprochen werden kann, wenn die beteiligten Personen ein gemeinsames Ziel aushandeln und ihre (gemeinsame) Tätigkeit daraufhin ausrichten. Diese Vorbedingung spielt für inklusive Settings eine wichtige Rolle, gestaltet sich jedoch oft problematisch, da viele Lehrkräfte eher (noch) in getrennten Zuständigkeitsbereichen denken, statt eine gemeinsame Vision des inklusiven Unterrichts zu entwickeln (Gebhard et al., 2014). Zusätzlich sind weitere intrapersonale Bedingungen zentral für das Zustandekommen von Kokonstruktion; dazu gehören gegensei-

tiges Vertrauen, Egalität (die Annahme einer Gleichwertigkeit der Kooperationspartner:innen), Reziprozität (gewissermaßen das gegenseitige Geben und Nehmen innerhalb einer Kooperation) sowie die Bereitschaft zur Deprivatisierung des eigenen Unterrichts. Grosche et al. (2020a) beschreiben diese Variablen als intrapersonal, „weil sie von jedem Individuum selbst ausgehen“ (S. 470), auch wenn sie sich in kokonstruktiven Zusammenhängen auf die jeweiligen Kooperationspartner:innen beziehen. Diese Vorbedingungen lassen sich sowohl theoretisch herleiten als auch aus bisherigen Forschungsarbeiten zur Kooperation erschließen (Grosche et al., 2020a; Kluge & Grosche, 2021; Spieß, 2004).

Den Kern des theoretischen Modells stellen die *kokonstruktiven Handlungen* dar, die aus Aushandlungs- und Reflexionsprozessen bestehen, einen zyklischen Charakter aufweisen und als „interdependent, intensiv, langfristig, kontinuierlich“ (Grosche et al., 2020a, S. 471) beschrieben werden. Durch diese kokonstruktiven Handlungen können die individuellen aber voneinander abhängigen bzw. interdependenten Ziele der Lehrkräfte zu gemeinsam geteilten Zielen werden. Durch langfristige Aushandlungs- und Reflexionsprozesse kann es den Lehrkräften gelingen, eine gemeinsame Sicht und damit geteilte Ziele für ihr Handeln zu entwickeln. Kokonstruktion ist immer dann sinnvoll und zielführend, wenn es darum geht eine Innovation (wie z. B. schulische Inklusion) zu implementieren, da Lehrkräfte in dieser Situation aushandeln müssen, was diese Innovation für das Handeln als Lehrkraft, für die Kommunikation mit der Lerngruppe und für die Unterrichtsvorbereitung und -durchführung bedeutet (Coburn & Stein, 2006; Vescio, Ross & Adams, 2008).

Gelingen diese kokonstruktiven Handlungen, stellen sich bestimmte *proximale Ergebnisse* ein, die sich unmittelbar aus der Kokonstruktion ergeben. Hierzu gehören zum einen das gemeinsame Verständnis der Innovation, das sich in einer gleichen oder zumindest ähnlichen Sicht auf z. B. typische Probleme in einer Lerngruppe oder das Vorgehen bei der Förderung einzelner Schüler:innen zeigt. Zum anderen wird die Innovation insgesamt an die Situation der Handelnden adaptiert; sie erfährt eine Rekontextualisierung (Fend, 2006) an die individuelle Situation der Klasse bzw. Lerngruppe. Beides ist unmittelbar mit den kokonstruktiven Prozessen in der Gruppe verbunden, bei denen ausgehandelt wird, was die Innovation für die eigene Praxis bedeutet.

Die proximalen Ergebnisse der Kokonstruktion lassen sich von *distalen Folgen* sowohl auf der Lehrkräfteebene als auch auf der Ebene der Schüler:innen abgrenzen. Lehrkräfte profitieren langfristig von der Kokonstruktion, indem sie sich emotional und zeitlich entlastet fühlen, weil sie eine gemeinsame Verantwortung tragen und die Umsetzung einer Innovation auf verschiedene Schultern verteilt ist (Fussangel & Dizinger, 2014). Die Schüler:innen profitieren ebenfalls, weil Lehrkräfte in der Kokonstruktion die Innovation auf die spezifischen Lernbedürfnisse und den individuellen Lerngruppenkontext adaptieren (Lütje-Klose, Neumann, Gorges & Wild, 2018).

Konstruktion ist kein in sich abgeschlossener Prozess, sondern ist langfristig und durch *reflexive Zyklen* gekennzeichnet, d. h. die Lehrkräfte nehmen ihre Aushandlungs- und Reflexionsprozesse immer wieder auf, diskutieren neu und

entwickeln weiter. Die beschriebenen Ergebnisse und Folgen wirken dabei auf die Rahmenbedingungen und Vorbedingungen zurück und spiegeln sich langfristig in veränderten kokonstruktiven Handlungen wider (Grosche et al., 2020a). Auf diese Art und Weise bauen die beteiligten Lehrkräfte kontinuierlich eine gemeinsame Wissens- und Handlungsbasis auf. Für eine ausführlichere Beschreibung des theoretischen Modells sei noch einmal auf Grosche et al. (2020a) verwiesen.

3. Entwicklung eines Instruments zur Messung kokonstruktiver Kooperation

Der entwickelte Fragebogen zur Erfassung der kokonstruktiven Kooperation soll das beschriebene theoretische Modell abbilden und kokonstruktive Kooperation empirisch operationalisieren. Für jede einzelne der fünf theoretischen Dimensionen des Modells (Rahmenbedingungen, Vorbedingungen, kokonstruktive Handlungen, Ergebnisse, Folgen) wurde ein umfangreicher Itempool entwickelt, der vor der empirischen Überprüfung iterativ überarbeitet und eingegrenzt wurde. Als Orientierung für die Formulierung dienten sowohl das theoretische Modell von Grosche et al. (2020a) als auch verschiedene, bereits veröffentlichte Instrumente zur Erfassung von Kooperation bzw. Kokonstruktion, die auf dem Modell von Gräsel et al. (2006) basieren und die gesichtet wurden (z. B. Drossel, 2015; Fussangel, 2008).

Die Items wurden für die Zielgruppe von Lehrkräfteteams in inklusiven Settings entwickelt, d. h. für Teams bestehend aus Regelschul- und sonderpädagogischen Lehrkräften, da diese auch den Bezugskontext für die Entwicklung des theoretischen Modells darstellten (Grosche et al., 2020a). Der erste Entwurf des Fragebogens umfasste insgesamt 143 Items, die auf einer 7-stufigen Ratingskala beantwortet werden sollten, wobei nur die beiden Endpunkte mit den Beschreibungen 1 = „stimme überhaupt nicht zu“ und 7 = „stimme völlig zu“ ausformuliert waren. Die entwickelten Items verteilten sich folgendermaßen auf die fünf Dimensionen des theoretischen Modells:

1. Rahmenbedingungen (43 Items)
2. Vorbedingungen von Kokonstruktion (29 Items)³
3. kokonstruktive Handlungen (24 Items)
4. Ergebnisse der Kokonstruktion (20 Items)
5. Folgen der Kokonstruktion (27 Items)

3 Sowohl in Studie 1 als auch in Studie 2 enthielt der Fragebogen für die Dimension „Vorbedingungen von Kokonstruktion“ noch keine Items für den Faktor der „Deprivatisierung“. Erst in Studie 3 und 4 (in Studie 4 zudem nur in einer Teilstichprobe) wurde der Faktor „Deprivatisierung“ in den Fragebogen mit aufgenommen, weshalb die Anzahl an Faktoren und Items in dieser Dimension zwischen den Studien differiert. Der Grund der späten Aufnahme war, dass das theoretische Modell parallel zur Durchführung der empirischen Studien weiterentwickelt wurde; der Faktor „Deprivatisierung“ wurde erst später in die Theorie aufgenommen.

In insgesamt vier Studien entwickelten wir eine erste Version eines Instruments zur Messung von Kokonstruktion. Die im Folgenden beschriebenen Studien 1 (kognitive Pretests), 2 (Itemanalyse und explorative faktorielle Überprüfung) und 4 (Überprüfung der Faktorstruktur und Finalisierung) fanden Anfang 2019 bis zum Sommer 2020 im Rahmen des BMBF-Projekts „Partizipation in Response-To-Intervention“⁴ (PARTI) statt (Grosche et al., 2020b). Studie 3 (Kreuzvalidierung) fand im Rahmen eines mit Lehramtsstudierenden durchgeführten Forschungsprojekts im Wintersemester 2019/2020 statt.

Das Ziel der vorliegend beschriebenen Studien war die Entwicklung eines Instruments zur Messung von Kokonstruktion durch psychometrisch abgesicherte Selbsteinschätzungsskalen. Die relativ hohe Anzahl an auf der Grundlage des theoretischen Rahmenmodells entwickelten Items sollte dabei systematisch gekürzt werden, um ein Instrument zu entwickeln, das von der Länge her in empirischen Untersuchungen gut handhabbar ist. Über die vier Studien hinweg wählten wir ein iteratives Vorgehen, bei dem wir die Anforderungen bzw. die empirischen Bewährungsgrade an das Instrument stetig verschärften und so das Instrument weiterentwickelten. Angesichts der sehr hohen Anzahl an Items, die sämtlich neu entwickelt wurden, entschieden wir uns in Studie 2 zunächst explorative Faktorenanalysen zu rechnen und erst danach, entsprechend der Verschärfung der Bewährungskriterien, konfirmatorische Faktorenanalysen durchzuführen. Die explorativen Faktorenanalysen wurden mit SPSS (Vers. 27), die konfirmatorischen Faktorenanalysen mit R (Version 4.0.3) und den Paketen *lavaan* 0.6–7 (Rosseel, 2012) und *psych* 2.0.9 (Revelle, 2020) durchgeführt. Insgesamt möchten wir mit dem vorliegenden Beitrag die Entwicklung einer ersten Version des Instruments vorstellen, das in weiteren Replikationsstudien zu überprüfen ist.

4. Studie 1: Kognitive Pretests

In Studie 1 legten wir sechs Lehrkräften den Erstentwurf vor und evaluierten die Items mittels kognitiver Pretests (Prüfer & Rexroth, 2005). Die Lehrkräfte wurden im Rahmen des BMBF-Projekts PARTI rekrutiert. Drei Sonderpädagogen (alle männlich, zwischen 29 und 60 Jahre alt) und drei Grundschullehrkräfte (alle weiblich, zwischen 29 und 30 Jahre alt) bearbeiteten den Erstentwurf und schätzten Verständnis, Klarheit und Relevanz aller 143 Items einzeln ein. Dabei wurden mit Probing, Think-aloud-Techniken, Paraphrasing und Confidence Ratings verschiedene kognitive Interviewtechniken eingesetzt (Prüfer & Rexroth, 2005). Die Items wurden überwiegend als verständlich, klar und relevant eingeschätzt; zehn als etwas unverständlich eingestufte Items wurden umformuliert. Die Lehrkräfte schätzten insbesondere die Items zur Dimension der Zielinterdependenz als schwierig zu

4 Das diesen drei Studien zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) unter den Förderkennzeichen 01NV1733A und 01NV1733B gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen und Autoren.

erreichen ein. Da diese Dimension aus theoretischer Sicht jedoch eine zentrale Rolle für das Zustandekommen von Kokonstruktion spielt und theoretisch sogar anzunehmen ist, dass Zielinterdependenz aufgrund der notwendigen Bezugnahme auf die Kooperationspartner:innen im eigenen Handeln schwierig zu erreichen ist (Gräsel et al., 2006), wurden die Items beibehalten.

5. Studie 2: Itemanalyse und explorative faktorielle Überprüfung (1. Pilotierung)

In der ersten Pilotierungsstudie wurden die 143 Items, die anhand der Ergebnisse des kognitiven Pretests teilweise überarbeitet wurden, erstmalig einer größeren Stichprobe vorgelegt. Die Ziele dieser ersten Pilotierungsstudie waren die Analyse der Itemcharakteristiken, die Identifikation von problematischen Items und die erste Kürzung des Fragebogens. Die folgenden Analysen verstehen sich explorativ und modellgenerierend.

5.1 Stichprobe

Insgesamt füllten $N=151$ Lehrkräfte (76% Lehrerinnen) aus nordrhein-westfälischen Grundschulen den überarbeiteten Fragebogen aus. Darunter waren 34% Grundschullehrkräfte und 66% Sonderpädagog:innen mit einem Durchschnittsalter von 41.11 Jahren ($SD=9.88$); es lagen keine signifikanten Unterschiede im Alter vor ($t(147)=-.252, p=.802$). Die Lehrkräfte verfügten im Durchschnitt über eine Berufserfahrung von 13.36 Jahren ($SD=9.13$) und arbeiteten seit durchschnittlich 7.83 Jahren ($SD=7.26$) an ihrer aktuellen Schule (auch hier lagen keine signifikanten Unterschiede zwischen den Lehrkräftegruppen vor: Berufserfahrung $t(146)=.652, p=.802$; Jahre an aktueller Schule: $t(148)=-.270, p=.787$). Die Erhebung fand im Sommer 2019 statt und wurde als *paper-pencil*-Erhebung durchgeführt. Der Fragebogen wurde entlang der fünf theoretischen Dimensionen und den dazugehörigen Items in fünf Abschnitte unterteilt.

5.2 Methode

Die Analyse wurde in zwei Schritten vorgenommen: Zunächst wurden die empirischen Verteilungen analysiert. In Bezug auf Schiefe und Kurtosis gibt es zwar keinen Konsens für spezifische Grenzwerte, in Simulationsstudien konnte jedoch gezeigt werden, dass eine Schiefe ab einem Wert von 2 und eine Kurtosis ab einem Wert von 7 (bzw. 4 bei der im Folgenden berichteten excess kurtosis) problematisch werden (Ryu, 2011). Für beide Angaben legten wir etwas konservativer einen Wert von 2 als Grenzwert fest. Boden- bzw. Deckeneffekte wurden durch Itemschwierigkeiten von unter 0.1 bzw. über 0.9 festgelegt. Items, die diese Grenzwerte über-

trafen, wurden gestrichen. Die Werte wurden pro Item in den Substichproben der Grundschullehrkräfte sowie der Sonderpädagog:innen geprüft.

Der zweite Kürzungsschritt umfasste eine Reihe von explorativen Faktorenanalysen für jede der fünf theoretisch postulierten Dimensionen (Hauptachsenfaktorenanalyse mit Oblimin-Rotation). Die Prüfungen der Voraussetzungen zur Durchführung von Faktorenanalysen ergab Folgendes: Die Korrelationsmatrizen zeigten für fast alle Items keine Anzeichen von Multikollinearität (d.h. keine Inter-Item-Korrelationen $> .80$). Eine Ausnahme stellten drei als hoch (zwischen $r = .80$ und $r = .86$) einzustufende Inter-Item-Korrelationen dar. Die betroffenen Items wurden jedoch aus theoretischen Gründen zunächst beibehalten. Das Maß der Stichprobenadäquanz (*measure of sampling adequacy; MSA*) lag für alle Items über $.60$ und somit im unproblematischen Bereich (Pett, Lackey & Sullivan, 2003).

Die Entscheidung über die Anzahl der zu extrahierenden Faktoren wurde auf Grundlage mehrerer Aspekte getroffen. Zunächst wurden die Eigenwerte sowie die *Scree Plots* für eine Beurteilung der Güte herangezogen, wobei Eigenwerte > 1.0 als Kriterium angelegt wurden. Neben den *Scree Plots* wurden zusätzlich die Parallelanalysen als Entscheidungsgrundlage genutzt und geprüft, ob diese dieselbe Anzahl an Faktoren nahelegen. Nach der Faktorextraktion wurden die Muster der Faktorladungen überprüft, um Items zu identifizieren, die entweder (a) schlecht auf den Primärfaktor luden (d.h. Ladungen $< .50$) oder (b) zu stark ($> .32$) auf mehr als einem Faktor luden (Tabachnick & Fidell, 2012). Diese Items wurden entfernt. Für die auf diese Weise extrahierten Faktoren wurde die interne Konsistenz mittels Cronbachs α berechnet.

5.3 Ergebnisse

Es zeigte sich, dass sehr viele Items sehr leicht waren ($M = 4.70$, $SD = 0.39$, $Min = 2.51$, $Max = 5.34$). Insgesamt wiesen 34 Items eine Schwierigkeit von über 90% und/oder eine Schiefe oder Kurtosis > 2 auf. Diese Items wurden im Autor:inenteam dahingehend diskutiert, ob sie von zentraler theoretischer Bedeutung sind. Da dies nicht der Fall war, wurden sie gestrichen, sodass nach dem ersten Kürzungsschritt 109 der 143 Items verblieben, mit denen im zweiten Schritt für jede der fünf theoretischen Dimensionen eine explorative Faktorenanalyse durchgeführt wurde.

5.3.1 Rahmenbedingungen der kokonstruktiven Kooperation

Laut theoretischem Modell umfasst die Dimension der Rahmenbedingungen drei Faktoren: schulstrukturelle, schulkulturelle sowie individuelle Rahmenbedingungen. Nach Eigenwerten hätten 7 Faktoren, nach *Scree-Plot* 2 bis 5 Faktoren extrahiert werden können. Eine dreifaktorielle Lösung war gut interpretierbar. Es zeigte sich jedoch, dass – anders als theoretisch angenommen – ein gemeinsamer

Schulfaktor vorlag und sich die individuellen Rahmenbedingungen in zwei Faktoren aufteilten. Insgesamt sieben der 27 Items wurden aufgrund von zu geringen oder Doppelladungen gestrichen. Die Faktorenstruktur der verbleibenden 20 Items prüften wir nochmals mittels eines dreifaktoriellen Modells. Der Kaiser-Meyer-Olkin (KMO)-Wert lag bei .806, der Bartlett-Test war signifikant ($p < .001$). Dieses dreifaktorielle Modell war sehr gut interpretierbar, enthielt keine Doppelladungen und wies nur schwache Korrelationen zwischen den Faktoren auf (r zwischen .06 und .28). Die drei extrahierten Faktoren sind: (a) Individuelle Rahmenbedingungen (8 Items, $\alpha = .87$, Bsp.: „Ich erkenne Meinungen meiner Teammitglieder an.“), (b) Individuelle Erfahrungen (7 Items, $\alpha = .75$, Bsp.: „Ich habe bereits intensiv in Teams zusammengearbeitet.“) und (c) Schulkulturelle und schulstrukturelle Rahmenbedingungen (5 Items, $\alpha = .71$, Bsp.: „Mein Arbeitsalltag sieht feste Zeiten für die Teamarbeit vor.“).

5.3.2 Vorbedingungen der kokonstruktiven Kooperation

Theoretisch wurde eine vierfaktorielle Struktur erwartet: Zielinterdependenz, Vertrauen, Egalität und Reziprozität. Die 22 Items konnten in einer solchen vierfaktoriellen Struktur sinnvoll interpretiert werden. Drei Items wurden allerdings aufgrund von zu schwachen ($< .50$) oder Doppelladungen ($< .32$) gestrichen. Die verbleibenden 19 Items wurden nochmals mittels einer 4-Faktor-Struktur überprüft. Der KMO-Wert war .874, der Bartlett-Test war signifikant ($p < .001$). Die Lösung war gut interpretierbar, enthielt keine Doppelladungen und die Faktor-Interkorrelationen waren theoriekonform (Zielinterdependenz und Vertrauen korrelierten mit $r = .19$ am niedrigsten, Reziprozität und Vertrauen mit $r = .52$ am höchsten). Die vier extrahierten Faktoren sind: (a) Reziprozität (6 Items, $\alpha = .88$, Bsp.: „Im Team herrscht ein ausgewogenes Geben und Nehmen.“), (b) Zielinterdependenz (6 Items, $\alpha = .85$, Bsp.: „Nur durch die verschiedenen Kompetenzen im Team bin ich neuen Anforderungen gewachsen.“), (c) Egalität (4 Items, $\alpha = .82$, Bsp.: „Alle Teammitglieder sind gleichberechtigt.“) und (d) Vertrauen (3 Items, $\alpha = .78$, Bsp.: „Ich traue mich im Team, auch unbequeme Meinungen zu äußern.“).

5.3.3 Kokonstruktive Handlungen

Laut theoretischem Modell umfasst die Dimension der kokonstruktiven Handlungen zwei Faktoren: Aushandlungsprozesse und Reflexionsprozesse. Die empirische Überprüfung dieser Dimension mit insgesamt 21 Items zeigte zunächst keine sinnvoll interpretierbaren Faktormodelle, was darauf hindeutete, dass es noch nicht hinreichend gelungen war, die theoretische Dimension gut abzubilden. Es ergaben sich inhaltliche Faktoren (z. B. Reflexion der Teamprozesse), die nicht dem Kern der Theorie zur Kokonstruktion bzw. den kokonstruktiven Handlungen entsprachen. In den verschiedenen Faktorenlösungen konnten 10 problematische Items iden-

tifiziert werden (unterschiedliches Ladungsverhalten, Doppelladungen, inhaltliche Nähe zu anderen Dimensionen), die ausgeschlossen wurden. Eine daraufhin identifizierte zweifaktorielle Lösung mit den verbleibenden 11 Items ließ sich gut interpretieren (KMO-Wert: .903, Bartlett-Test: $p < .001$): (a) Kokonstruktive Handlungen (8 Items, $\alpha = .87$, Bsp.: „Im Team handeln wir die zu bearbeitenden Inhalte solange aus, bis wir einen gemeinsamen Plan erarbeitet haben.“) und (b) Kokonstruktive Reflexionen (3 Items, $\alpha = .94$, Bsp.: „Wir reflektieren im Team regelmäßig die Erfolge und Misserfolge unserer Arbeit.“). Beide Faktoren korrelierten recht hoch mit $r = .76$. Das zweifaktorielle Modell wurde dennoch beibehalten, weil das Erzwingen eines einfaktoriellen Modells schlechter theoretisch interpretierbar war. Da die kokonstruktiven Handlungen im theoretischen Modell das zentrale Element darstellen, war es unbefriedigend, dass sich die Überprüfung dieser Dimension schwieriger gestaltete. Zur weiteren Entwicklung dieser zentralen Skalen führten wir deshalb zusätzliche konfirmatorische Faktorenanalysen mit ML-Schätzer und FIML-Ersetzung (Savalei & Rhemtulla, 2012) in R durch. Diese verfolgten im vorliegenden Stadium der Instrumentenentwicklung jedoch weniger das Ziel der Prüfung, sondern eher der Weiterentwicklung der Faktorenlösung. Das in der vorherigen explorativen Faktorenanalyse gefundene zweifaktorielle Modell konnte entlang der empfohlenen Kriterien (CFI/TLI $> .95$, RMSEA $< .06$ und SRMR $< .08$) von Hu und Bentler (1999) mit folgenden Fit-Indizes nur teilweise bestätigt werden: $\chi^2 = 105$, $p < .001$, CFI = .94, TLI = .92, RMSEA = .099, $p = .001$, SRMR = .047. Die Analyse der Modifikation-Indizes zeigte, dass korrelierte Fehlerterme vorlagen (bei den Items, die zum einen den Umgang mit Meinungsverschiedenheiten und zum anderen das iterative Überarbeiten erfassen). Ein daraufhin spezifiziertes vierfaktorielles Modell zeigte zufriedenstellende Fit-Indizes ($\chi^2 = 67$, $p = .036$, CFI = .98, TLI = .98, RMSEA = .052, $p = .438$, SRMR = .029). Um bei der Skala *Umgang mit Meinungsverschiedenheiten* auf eine Mindestanzahl von drei Items zu kommen, wurde das negativ formulierte Item „Meinungsverschiedenheiten werden im Team übergangen.“, das im ersten Kürzungsschritt aus inhaltlichen Erwägungen zunächst ausgeschlossen wurde, an dieser Stelle mitberücksichtigt. Es lagen gut interpretierbare Faktoren vor: (a) Kokonstruktive Aushandlungen (4 Items, $\alpha = .80$, Bsp.: „Im Team nutzen wir unterschiedliche pädagogische Perspektiven für die Weiterentwicklung von Konzepten oder Methoden.“), (b) Kokonstruktive Reflexionen (3 Items, $\alpha = .94$, Bsp.: „Wir reflektieren regelmäßig die Umsetzbarkeit unserer Ideen, Konzepte und Methoden.“), (c) Meinungsverschiedenheiten (3 Items, $\alpha = .82$, Bsp.: „Im Team versuchen wir, auch Missverständnisse und Meinungsverschiedenheiten produktiv zu nutzen.“) und (d) Iterative Überarbeitungen (2 Items, $\alpha = .77$, Bsp.: „Im Team feilen wir manchmal lange am selben Gegenstand, um ihn zu optimieren.“). Auch für die letztgenannte Skala wurde die Itemanzahl auf drei erhöht und ein vollständig neues und damit nicht evaluiertes Item formuliert: „Im Team arbeiten wir solange an Modellen und Konzepten, bis wir mit dem Ergebnis zufrieden sind.“ Insgesamt liegen bei dieser Lösung recht hohe Faktorinterkorrelationen zwischen min. $r = .66$ (Reflexionen – iterative Überarbeitung) und max. $r = .84$ (Kokonstruktive Aushand-

lungen – Meinungsverschiedenheiten) vor. An dieser Stelle sei noch einmal darauf hingewiesen, dass diese Analysen explorativ erfolgten und keinen Nachweis der Messgüte liefern.

5.3.4 Ergebnisse der kokonstruktiven Kooperation

Die ursprünglich in dieser Dimension theoretisch formulierten vier Faktoren (gemeinsame Verantwortung, geteilte Kognition, inklusives Rollenverständnis sowie adaptive Implementation) ließen sich anhand der 13 Items empirisch zunächst nicht gut abbilden. Insbesondere die zwei verbliebenen Items der *adaptiven Implementation* mussten aufgrund problematischen Ladungsverhaltens ausgeschlossen werden. Nach einer faktoriellen Überprüfung und Identifikation der zwei problematischen Items lag eine zufriedenstellende dreifaktorielle Lösung vor (KMO-Wert: .789, Bartlett-Test: $p < .001$): (a) Geteilte Kognition (4 Items, $\alpha = .83$, Bsp.: „Wir haben im Team eine gemeinsame Wissensbasis über die zu bearbeitenden Inhalte.“), (b) Inklusives Rollenverständnis (4 Items, $\alpha = .70$, Bsp.: „In unserem Team werden keine Grenzen zwischen den einzelnen Professionen gezogen.“), (c) Gemeinsame Verantwortung (3 Items, $\alpha = .83$, Bsp.: „Wir fühlen uns im Team für alle Schüler:innen verantwortlich.“). Die drei Faktoren sind mit maximal $r = .50$ korreliert. Da die *adaptive Implementation* im theoretischen Modell eine zentrale Rolle innerhalb der Ergebnisse von kokonstruktiven Prozessen spielt, entschieden wir uns die zwei Items aus theoretischen Gründen zunächst beizubehalten. Zudem formulierten wir eines der gestrichenen Items um, um hier ebenfalls drei Items vorliegen zu haben.

5.3.5 Folgen der kokonstruktiven Kooperation

Die drei theoretisch erwarteten Faktoren (emotionale sowie zeitliche Entlastung und Wirkungen auf die Lernenden), erfasst über 24 Items, fanden sich in einer gut interpretierbaren Faktorlösung empirisch wieder. Es zeigte sich jedoch, dass vier Items des Faktors *Wirkungen auf Lernende* hohe Doppelladungen hatten. Da in dieser Dimension ohnehin viele redundante Items vorhanden waren, strichen wir hier insgesamt neun Items. Mit den verbleibenden 11 Items zeigte sich eine gut interpretierbare dreifaktorielle Lösung (KMO-Wert: .869, Bartlett-Test: $p < .001$): (a) zeitliche Entlastung (4 Items, $\alpha = .90$, Bsp.: „Durch die Kooperation im Team habe ich mehr Zeit für andere Dinge.“), (b) Wirkung auf Ebene der Lernenden (3 Items, $\alpha = .91$, Bsp.: „Wenn die im Team gemeinsam erarbeiteten Konzepte oder Methoden umgesetzt werden, können alle Schüler:innen der Klasse besser lernen.“) und (c) emotionale Entlastung (4 Items, $\alpha = .78$, Bsp.: „Durch die Arbeit im Team lasse ich mich durch Misserfolge nicht mehr so schnell entmutigen.“).

Zusammenfassend lässt sich festhalten, dass nach der Auswertung der zweiten Studie von den ursprünglich 143 Items 77 Items im Itempool verblieben.

6. Studie 3: Kreuzvalidierung (2. Pilotierung)

In Studie 3 wurde der gekürzte Fragebogen nochmals anhand einer unabhängigen und größeren Stichprobe kreuzvalidiert. Ziel dieser Studie war neben einer weiteren Kürzung des Fragebogens die erstmalige konfirmatorische Analyse der Faktorenstruktur. Zur Übersicht der Items sowie zur besseren Nachvollziehbarkeit befindet sich im Anhang (Tabelle A2: Ursachen für Item-Ausschluss) eine Tabelle der in Studie 3 ausgeschlossenen Items mit der jeweiligen Ursache für den Ausschluss.

6.1 Stichprobe

Die Stichprobe bestand aus $N=914$ Lehrkräften (76 % Lehrerinnen), darunter 71.4 % Regelschullehrkräfte und 27.8 % Sonderpädagog:innen im Alter von $M=41.7$ Jahre ($SD=10.59$). Insgesamt verfügten sie über eine durchschnittliche Berufserfahrung von 12.24 Jahren ($SD=9.47$). Die meisten der Befragten unterrichteten an Grund- (39.9 %), Förder- (20.6 %) und Gesamtschulen (12.7 %). Die weiteren Schulformen waren jeweils im einstelligen Prozentbereich vertreten (Hauptschulen: 5.4 %, Realschulen: 3.6 %, Gymnasien: 9.6 %, Berufskolleg: 5 %, sonstige: 2.5 %). Die Befragung wurde von Lehramtsstudierenden im Rahmen eines Forschungsprojektes im Wintersemester 2019/2020 durchgeführt. Der Fragebogen wurde als Online-Fragebogen eingesetzt.

6.2 Methode

Die Berechnungen wurden mit dem Statistikprogramm R (Version 4.0.3) und den Paketen *lavaan* (Rosseel, 2012) und *psych* (Revelle, 2020) durchgeführt. Aufgrund der ausreichenden Größe der Stichprobe teilten wir den Datensatz randomisiert in zwei gleich große Teildatensätze, um im ersten sogenannten Training-Datensatz die Faktorenstruktur weiter zu verbessern und im zweiten sogenannten Test-Datensatz die Faktorenstruktur unabhängig zu überprüfen (Bortz & Schuster, 2010; De Rooij & Weeda, 2020). Die Analysen basieren also entweder auf dem Training-Datensatz ($N=457$) oder auf dem Test-Datensatz ($N=457$). Für die weitere Kürzung des Fragebogens verfolgten wir folgende Strategie: Am Training-Datensatz gingen wir analog zu Studie 2 vor und (a) eliminierten zunächst Items mit einer Schiefe oder Kurtosis > 2 . Dann führten wir (b) explorative Faktorenanalysen durch und strichen Items mit einer Faktorladung $\leq .50$ oder einer Kross-Ladung $> .32$ (Tabachnick & Fidell, 2007). Da wir insgesamt das Ziel verfolgten, einen handhabbaren Fragebogen zu entwickeln, kürzten wir die Anzahl der Items auf max. vier pro Faktor und strichen (c) Items aufgrund von schwächeren Ladungen oder höheren Doppelladungen; wenn noch zu viele Items vorlagen, strichen wir auch in dieser Phase noch einmal, wenn inhaltliche Redundanzen bei den Items vorlagen. Im Anschluss an diese explorative Entwicklung wurde die Faktorstruktur anhand des Test-Daten-

satzes erstmalig konfirmatorisch getestet (ML-Schätzer mit FIML-Ersetzung) und die Fit-Indizes anhand der zuvor genannten Kriterien von Hu und Bentler (1999) beurteilt. Bei nicht zufriedenstellenden Fit-Indizes ($CFI/TLI < .95$, $RMSEA > .06$ und $SRMR > .08$) wurden die Items auf der Grundlage der Modifikation-Indizes einer weiteren inhaltlichen Analyse unterzogen sowie die standardisierten Faktorladungen und Itemcharakteristika beurteilt. Dieses Vorgehen wurde für jede der fünf theoretisch postulierten Dimensionen durchgeführt.

6.3 Ergebnisse

6.3.1 Rahmenbedingungen

Die Ergebnisse von Studie 2 legten für die Dimension der Rahmenbedingungen 3 Skalen mit insgesamt 20 Items nahe: Individuelle Rahmenbedingungen, Individuelle Erfahrungen und Schulkulturelle und schulstrukturelle Rahmenbedingungen. Die Analysen am Training-Datensatz führten zum Ausschluss von vier Items (Schiefe und/oder Kurtosis > 2). Die aufgrund einer Parallel-Analyse empfohlene explorative Faktorenanalyse mit drei Faktoren führte zum Ausschluss weiterer vier Items aufgrund von zu geringer Ladung oder inhaltlicher Nähe zu anderen Items. Die in Studie 2 identifizierten Skalen veränderten sich folgendermaßen: Die *schulculturellen* und *schulstrukturellen Rahmenbedingungen* ließen sich identifizieren, die Skala der *individuellen Rahmenbedingungen* entfiel bzw. verteilten sich die verbliebenen Items auf *Erfahrungen mit Teamarbeit* sowie *Einstellungen zur Teamarbeit*. Die anschließend am Test-Datensatz durchgeführte konfirmatorische Faktorenanalyse ergab zunächst noch nicht zufriedenstellende Fit-Indizes. Die Analyse der Modifikation-Indizes zeigte korrelierte Fehlerterme zwischen verschiedenen Items; dies führte zum Ausschluss zwei weiterer Items. Die anschließende konfirmatorische Faktorenanalyse zeigte einen guten Fit dieses 3-Faktor-Modells (vgl. auch Tab. 1): $\chi^2 = 77.677$, $df = 32$, $p < .001$, $CFI = .978$, $TLI = .969$, $SRMR = .036$, $RMSEA = .056$. Die drei identifizierten Dimensionen sind: (a) Schulische Rahmenbedingungen (3 Items, $\alpha = .76$), (b) Einstellungen zur Kooperation (4 Items, $\alpha = .86$) sowie (c) Erfahrungen mit Kooperation (3 Items, $\alpha = .82$). Aufgrund der Verwendung der Modifikation-Indizes im Test-Datensatz gilt das Modell jedoch nicht im Sinne einer CFA als empirisch getestet, da das Modell durch die Verwendung der Indizes post hoc verändert wurde.

6.3.2 Vorbedingungen der kokonstruktiven Kooperation

Die Ergebnisse von Studie 2 legten für die Dimension der kokonstruktions-spezifischen Vorbedingungen vier Skalen mit insgesamt 19 Items nahe: Zielinterdependenz, Reziprozität, Egalität und Vertrauen. In der vorliegenden Studie 3 sind zudem neun Items zur *Deprivatisierung* Bestandteil der Analysen, die in Studie 1 und

2 nicht enthalten waren. Der Grund für die nachträgliche Aufnahme war, dass das theoretische Modell parallel zur Durchführung der empirischen Studien weiterentwickelt wurde und die Deprivatisierung erst später in die Theorie aufgenommen wurde. Die deskriptiven Analysen führten zum Ausschluss zweier Items (Schiefe/Kurtosis > 2). Um das Ziel von maximal vier Items pro Skala zu erreichen, wurden in der gesamten theoretischen Dimension auf der Basis der explorativen Faktorenanalysen am Training-Datensatz insgesamt fünfzehn Items gestrichen (schwächste Ladungen, Doppelladungen auf anderen Faktoren, inhaltliche Redundanz). Eine abschließende Parallelanalyse legte drei Faktoren nahe, die die Zielinterdependenz, die Reziprozität sowie die Deprivatisierung repräsentieren (d. h., dass Egalität und Vertrauen empirisch nicht abgebildet werden konnten). Dieses 3-Faktor-Modell wurde mittels konfirmatorischer Faktorenanalyse am Test-Datensatz geprüft und konnte mit den folgenden Fit-Indizes bestätigt werden (vgl. Tab.1): $\chi^2 = 112.640$, $df = 51$, $p = .00$, CFI = .964, TLI = .960, SRMR = .046, RMSEA = .052. Die drei Faktoren sind: (a) Zielinterdependenz (4 Items, $\alpha = .88$), (b) Reziprozität (4 Items, $\alpha = .84$) sowie (c) Deprivatisierung (4 Items, $\alpha = .85$).

6.3.3 Kokonstruktive Handlungen

Die Ergebnisse von Studie 2 legten für die zentrale Dimension der kokonstruktiven Handlungen vier Skalen mit insgesamt 12 Items nahe: kokonstruktive Aushandlungen, Reflexionen, Umgang mit Meinungsverschiedenheiten sowie iterative Überarbeitung von Materialien und Konzepten (hier wurde noch ein zusätzliches Item formuliert, s. Studie 2). Die deskriptiven Analysen in der vorliegenden Studie wiesen für alle Items zufriedenstellende Werte auf. Eine Parallelanalyse wies vier Faktoren auf und die daraufhin durchgeführte explorative Faktorenanalyse am Training-Datensatz zeigte, dass lediglich zwei Items eine zu geringe Ladung aufwiesen. Auch wenn diese Streichung dazu führte, dass die Skala *Aushandlungen* aus nur noch zwei Items bestand, behielten wir diese Skala bei, da sie im theoretischen Modell eine Kerndimension darstellt. Die anschließende konfirmatorische Überprüfung am Test-Datensatz ergab zufriedenstellende Fit-Indizes für die 4-Faktorlösung (vgl. Tab. 1): $\chi^2 = 95.475$, $df = 38$, $p < .001$, CFI = .977, TLI = .967, SRMR = .027, RMSEA = .059. Die vier Faktoren lassen sich folgendermaßen beschreiben: (a) Kokonstruktive Aushandlungen (2 Items, $\alpha = .79$), (b) Reflexionsprozesse (3 Items, $\alpha = .89$), (c) Iterative Überarbeitung von Konzepten und Materialien (3 Items, $\alpha = .83$) sowie (d) Aushalten von Meinungsverschiedenheiten (3 Items, $\alpha = .76$).

6.3.4 Ergebnisse der Kokonstruktion

Die Ergebnisse von Studie 2 legten für die Dimension der Ergebnisse der Kokonstruktion drei Skalen mit insgesamt 11 Items nahe: geteilte Kognitionen, inklusives Rollenverständnis und gemeinsame Verantwortung. Zudem wurde aus theoretischen Überlegungen eine Skala für die Dimension der gemeinsamen Verantwortung

Tabelle 1: Fit-Indizes der Faktormodelle pro Dimension

	Studie 3 (N = 914)										Studie 4 (N = 351)									
	χ^2	df	CFI	TLI	SRMR	RMSEA	90 % KI RMSEA	χ^2	df	CFI	TLI	SRMR	RMSEA	90 % KI RMSEA						
Rahmenbedingungen	77.677	32	0.978	0.969	0.036	0.056	[0.040, 0.072]	28.088 ^a	32	1.000	1.003	0.020	0.000	[0.000, 0.033]						
Vorbedingungen	112.640	51	0.964	0.960	0.046	0.052	[0.039, 0.065]	67.004 ^b	51	0.966	0.956	0.069	0.062	[0.000, 0.100]						
Konstruktive Handlungen	95.475	38	0.977	0.967	0.027	0.059	[0.044, 0.074]	72.394 ^d	38	0.984	0.976	0.031	0.052	[0.033, 0.070]						
Ergebnisse	91.383	38	0.977	0.967	0.031	0.057	[0.042, 0.072]	141.363 ^e	38	0.946	0.922	0.050	0.089	[0.074, 0.105]						
Folgen	55.073	24	0.991	0.986	0.028	0.055	[0.036, 0.074]	84.337 ^f	29	0.965	0.945	0.045	0.075	[0.057, 0.094]						

Anmerkungen. CFI: Comparative Fit Index, TLI: Tucker-Lewis Index, RMSEA: Root-Mean-Square-Error-of-Approximation, 90 % KI RMSEA: 90 % Konfidenzintervall für RMSEA. Alle χ^2 sind mit drei Ausnahmen (Studie 4 – Rahmenbedingungen, Vorbedingungen, konstruktive Handlungen) mit $p < .001$ signifikant. ^a $p = .665$. ^b Werte aus der Teilstichprobe 2 der Studie 4 (N = 83), bei denen die Items zur Deprivatisierung eingesetzt wurden. Analog zu dem Modell bei Studie 3 umfasst der Faktor Deprivatisierung vier Items. $p = .066$. ^c Werte aus der Teilstichprobe 1 der Studie 4 (N = 351) ohne den Faktor Deprivatisierung. ^d $p = .001$. ^e Gleiches Modell wie bei Studie 3. ^f Modellfit ohne das Item GK4, welches im Rahmen der Replikation ausgeschlossen wurde.

tischen Gründen die Skala zur *adaptiven Implementation* beibehalten, die jedoch in Studie 2 empirisch nicht abgebildet werden konnte. Die deskriptiven Werte der Items der vier Skalen waren zufriedenstellend. Die explorative Faktorenanalyse am Training-Datensatz zeigte, dass drei Items eine zu geringe Ladung aufwiesen und gestrichen werden mussten. Eine abschließende Parallelanalyse wies vier Faktoren aus. Dieses 4-Faktor-Modell wurde am Test-Datensatz konfirmatorisch überprüft und wies zufriedenstellende Fit-Indizes auf (vgl. Tab. 1): $\chi^2 = 91.383$, $df = 38$, $p < .001$, CFI = .977, TLI = .967, SRMR = .031, RMSEA = .057. Die vier Skalen lassen sich folgendermaßen beschreiben: (a) Geteilte Kognitionen (4 Items, $\alpha = .87$), (b) Adaptive Implementation (3 Items, $\alpha = .82$), (c) Gemeinsame Verantwortung (2 Items, $\alpha = .83$) sowie (d) Inklusives Rollenverständnis (2 Items, $\alpha = .75$).

6.3.5 Folgen der Kokonstruktion

Die Ergebnisse der Studie 2 legten für die Dimension der langfristigen Folgen von kokonstruktiven Prozessen drei Skalen mit insgesamt 11 Items nahe: zeitliche Entlastung, emotionale Entlastung und Wirkungen auf Ebene der Lernenden. Die deskriptiven Werte, die anhand des Training-Datensatzes geprüft wurden, waren zufriedenstellend. Im Rahmen der explorativen Faktorenanalyse am Training-Datensatz wurden zwei Items eliminiert, es verblieben drei Faktoren, die im Anschluss am Test-Datensatz konfirmatorisch überprüft wurden. Die Fit-Indizes dieses 3-Faktor-Modells waren zufriedenstellend (vgl. Tab. 1): $\chi^2 = 55.073$, $df = 24$, $p < .001$, CFI = .991, TLI = .986, SRMR = .028, RMSEA = .055. Die drei Faktoren sind (a) Zeitliche Entlastung (4 Items, $\alpha = .95$), (b) Emotionale Entlastung (2 Items, $\alpha = .83$) sowie (c) Wirkungen auf die Lernenden (3 Items, $\alpha = .92$).

Eine Übersicht über alle Skalen der Dimensionen von Kokonstruktion befindet sich auf der linken Seite in Tabelle 2. Insgesamt zeigt sich nach Studie 3, dass ein Fragebogen mit 53 Items die 17 Skalen auf fünf Dimensionen zufriedenstellend abbildet. Gleichwohl ist zu konstatieren, dass die konfirmatorischen Faktorenanalysen bezüglich der *Rahmenbedingungen* auch zur Modellweiterentwicklung genutzt wurden. Die Fit-Indizes dürfen daher nicht überinterpretiert werden. Vielmehr muss eine nochmalige unabhängige Replikation in Studie 4 erfolgen.

7. Studie 4: Überprüfung der Faktorstruktur und Finalisierung

Studie 4 stellt die letzte Teilstudie dar, mit der die erste Version des Fragebogens zur Kokonstruktion finalisiert wurde, indem die Faktorstruktur anhand eines weiteren unabhängigen Datensatzes konfirmatorisch überprüft wurde.

7.1 Stichprobe

Aus 21 nordrhein-westfälischen Grundschulen wurden zu einem ersten Messzeitpunkt (t₁) insgesamt 376 Regelschul- und sonderpädagogische Lehrkräfte befragt. Von diesen Lehrkräften nahmen 322, darunter 84.8 % Regelschul- und 15.2 % sonderpädagogische Lehrkräfte, an der Befragung zu ihrer Kooperation teil. Zusätzlich wurden 7 Referendar:innen und 22 Lehrkräfte ohne Angabe des Lehramts, die ebenfalls an der genannten Befragung teilgenommen haben, in die Stichprobe mit aufgenommen. Daraus ergibt sich für die vorliegenden Analysen eine Stichprobengröße für t₁ von $n_{t_1} = 351$ (angehenden) Lehrkräften (89.3 % weiblich, 11.4 % männlich, 0.3 % divers). Das durchschnittliche Alter beträgt 42.41 Jahre ($SD = 9.48$). Im Durchschnitt hatten die Befragten 7.89 Jahre ($SD = 6.41$) Erfahrungen mit Schüler:innen mit sonderpädagogischem Förderbedarf. Die Befragung fand online im Winter 2019/2020 statt.

Zu einem zweiten Messzeitpunkt (t₂) im Sommer 2020 wurden 231 Regelschul- und sonderpädagogische Lehrkräfte befragt. Für die vorliegenden Analysen wurde auf eine Teilstichprobe von $n_{t_2} = 83$ Lehrkräften der Kontrollgruppe (91.6 % weiblich, 8.4 % männlich) zurückgegriffen, die bereits zu t₁ an der Befragung teilgenommen hatten (80.7 % Regelschul-, 18.1 % sonderpädagogische Lehrkräfte und 1.2 % Lehrkräfte ohne Angabe des Lehramts). Das durchschnittliche Alter beträgt 44.41 Jahre ($SD = 7.24$). Im Durchschnitt hatten die Befragten 9.16 Jahre ($SD = 6.40$) Erfahrungen mit Schüler:innen mit sonderpädagogischem Förderbedarf.

7.2 Methode

In Studie 4 wurden zur faktoriellen Überprüfung an der Teilstichprobe zu t₁ die gleichen konfirmatorischen Faktorenanalysen wie am Test-Datensatz von Studie 3 durchgeführt.

7.3 Ergebnisse

Die Faktorstruktur wurde in der vorliegenden Studie in vier von fünf Dimensionen vollständig bestätigt (vgl. rechte Seite in Tabelle 1). Die *Rahmenbedingungen*, die *Vorbedingungen der Kokonstruktion*, die *kokonstruktiven Handlungen* und die *Folgen der Kokonstruktion* ließen sich über die finale Faktorstruktur aus Studie 3 vollständig replizieren. In Tabelle 1 finden sich für die Dimension der Vorbedingungen zwei Modelle, was dem Umstand geschuldet ist, dass in Studie 4 die Items zur *Deprivatisierung* lediglich in der Teilstichprobe zu t₂ eingesetzt wurden. Aus diesem Grund wurde das Modell je mit (t₂) und ohne (t₁) diesen Faktor gerechnet, wobei beide Modelle grundsätzlich gute Fit-Indizes aufwiesen (in der Stichprobe ohne Deprivatisierungsitems mit einer minimalen Abweichung des TLI von .945). In der Dimension *Ergebnisse* verfehlten CFI = .946, TLI = .922 und RMSEA = .089 die de-

Tabelle 2: Skalenübersicht und -werte

	Studie 3						Studie 4							
	Items	N	M	SD	Skew	Kurt	α	Items	N	M	SD	Skew	Kurt	α
Rahmenbedingungen	3	875	4.94	1.49	-0.40	-0.74	.76	3	341	5.65	1.26	-1.01	0.62	.71
Schulische Rahmenbedingungen	4	847	5.52	1.16	-1.00	0.86	.86	4	341	5.68	1.16	-1.14	1.03	.89
Einstellungen	3	860	5.69	1.26	-1.21	1.19	.82	3	345	5.83	1.10	-1.22	1.43	.77
Erfahrungen														
Vorbedingungen														
Zielinterdependenz	4	863	3.85	1.54	0.02	-0.85	.88	4	343	4.44	1.32	-0.20	-0.53	.82
Reziprozität	4	829	4.89	1.47	-0.39	-0.61	.84	4	338	5.38	1.30	-0.66	-0.34	.79
Deprivatisierung	4	398	2.76	1.65	0.84	-0.27	.85	4	79 ^a	2.66	1.36	0.98	0.54	.80
Kokonstruktive Handlungen														
Aushandlungen	2	840	4.74	1.43	-0.57	-0.27	.79	2	337	4.92	1.36	-0.69	-0.06	.80
Reflexion	3	851	4.99	1.44	-0.61	-0.36	.89	3	337	5.12	1.41	-0.80	0.27	.91
Iterative Überarbeitung	3	835	4.65	1.40	-0.47	-0.26	.83	3	337	4.60	1.35	-0.37	-0.26	.82
Meinungsverschiedenheiten	3	834	5.48	1.11	-0.81	0.36	.76	3	336	5.55	1.12	-0.97	0.77	.78
Ergebnisse														
Geteilte Kognitionen	4	830	5.07	1.24	-0.69	0.07	.87	3 ^b	337	4.96	1.27	-0.50	-0.29	.82
Implementation	3	826	5.29	1.14	-0.72	0.59	.82	3	336	5.28	1.13	-0.77	0.64	.83
Gemeinsame Verantwortung	2	836	5.65	1.42	-1.09	0.46	.83	2	337	5.52	1.55	-1.08	0.24	.83
Inklusives Rollenverständnis	2	808	5.57	1.40	-0.82	-0.13	.75	2	337	5.56	1.50	-1.01	0.15	.80
Folgen														
Zeitliche Entlastung	4	838	4.41	1.78	-0.34	-0.93	.95	4	336	4.74	1.62	-0.48	-0.65	.94
Emotionale Entlastung	2	833	4.74	1.63	-0.59	-0.52	.83	2	338	5.09	1.36	-0.80	0.40	.81
Schüler:innen	3	782	4.93	1.29	-0.53	0.10	.92	3	339	5.37	1.04	-0.33	-0.45	.91

Anmerkungen. ^a Die Skalenwerte Deprivatisierung stammen ausschließlich aus der Teilstichprobe zu Messzeitpunkt 2 des Projekts (da zum ersten Messzeitpunkt nicht erhoben). ^b Im Rahmen der Replikation wurde ein Item zusätzlich ausgeschlossen.

finierten Grenzen knapp. Daher wurde auf der Basis der Modifikation-Indizes ein Item gestrichen, was den Modellfit verbesserte. Tabelle 1 (rechte Seite) zeigt einen Überblick der Fit-Indizes für alle Dimensionen der Kokonstruktion. Die Items des finalen Fragebogens sowie die zugehörigen Faktoren mit Faktorladungen befinden sich im Anhang (Tabelle A1: Übersicht Finale Items und Faktorladungen). Die Skalenwerte der in Studie 4 bestätigten Skalen befinden sich in Tabelle 2 (rechte Seite).

8. Diskussion

Ziel des Beitrags war die Darstellung der Entwicklung eines Instruments zur Erfassung der kokonstruktiven Kooperation im Rahmen der Umsetzung von schulischer Inklusion. Die Entwicklung einer ersten Version dieses Instruments wurde anhand von vier aufeinander aufbauenden und unabhängigen Teilstudien vorgestellt, in denen der Fragebogen systematisch gekürzt und stetig strengeren Bewährungskriterien unterzogen wurde. Insgesamt ist es gelungen, eine erste Fassung eines Instruments zur Messung von kokonstruktiver Kooperation zu entwickeln, durch das sich das theoretische Modell von Grosche et al. (2020a) empirisch abbilden lässt. Damit liegt erstmalig ein Instrument vor, das nicht nur die kokonstruktiven Handlungen, sondern auch Rahmenbedingungen, Vorbedingungen sowie Ergebnisse und Folgen von Kokonstruktion erfasst. Es enthält 52 Items in 17 Skalen auf fünf Dimensionen und erweitert damit die bisherigen Instrumente, die auf der Grundlage des ursprünglichen Ansatzes von Gräsel et al. (2006) zur Kokonstruktion entwickelt wurden (Fussangel, 2008; Keller-Schneider & Albisser, 2013; Pröbstel, 2008). Darüber hinaus stellt es in der aktuellen Debatte um die Entwicklung von Inklusion in Unterricht und Schule ein wichtiges Instrument dar, um die Kooperation in den Lehrkräfteteams abbilden zu können.

Bezüglich der vier vorgestellten Teilstudien lässt sich zusammenfassen, dass die Faktorstruktur der theoretischen Dimensionen des Kokonstruktionsmodells überwiegend durch die Daten abgebildet werden und lediglich zwei der ursprünglich entwickelten Skalen nicht beibehalten werden konnten: *Vertrauen* und *Egalität* als Faktoren der Vorbedingungen von Kokonstruktion konnten mit den Daten nicht bestätigt werden, obwohl sie in der Kooperationsforschung durchaus als wichtige Bedingungen von Kooperation betrachtet werden (Spieß, 2004) und insbesondere die Egalität im Rahmen der Zusammenarbeit von sonderpädagogischen und Regelschullehrkräften eine wichtige Rolle spielt (Lütje-Klose et al., 2018). Die beiden Faktoren ließen sich jedoch in Studie 3 nicht empirisch abbilden und konnten mit den konfirmatorischen Faktorenanalysen nicht bestätigt werden, wobei sich nicht klar sagen lässt, woran dies liegt. Es ist möglich, dass die formulierten Items nicht trennscharf genug formuliert waren oder die Konstrukte auch auf theoretischer Ebene als weniger trennscharf zu bewerten sind, als dies durch das theoretische Modell nahegelegt wird. Es bedarf hier in jedem Fall einer Replikationsstudie, mit

der überprüft werden kann, ob es sich im vorliegenden Fall um ein Artefakt handelt oder ob die theoretischen Annahmen einer Überarbeitung bedürfen.

Die einzelnen Teilstudien fokussierten verschiedene Ziele und zeigten Hürden des Entwicklungsprozesses auf unterschiedlichen Ebenen auf. Der Kern der Kokonstruktion – die kokonstruktiven Handlungen – ließ sich in Studie 3 durch vier Faktoren gut abbilden und wurde auch in Studie 4 bestätigt, obwohl sich die ersten Analysen zu dieser Dimension in Studie 2 zunächst schwieriger gestalteten. Auch die *adaptive Implementation* der Innovation innerhalb der Dimension Ergebnisse von Kokonstruktion ließ sich in Studie 2 nicht gut durch die Daten abbilden; wurde jedoch aus theoretischen Gründen beibehalten. Die konfirmatorische Überprüfung dieses Faktors in Studie 3 sowie Studie 4 gestaltete sich dann unproblematisch.

Kritisch ist anzumerken, dass die Variable der *Deprivatisierung* nicht in allen Teilstudien berücksichtigt wurde; sie konnte in Studie 3 jedoch gut durch die Daten abgebildet werden. Ebenso ist zu kritisieren, dass nicht der gesamte Fragebogen einer einzigen konfirmatorischen Analyse unterzogen wurde, sondern getrennte Faktorenanalysen pro Dimension berechnet wurden. Dieses Vorgehen erscheint gerechtfertigt, weil die schiere Anzahl an Dimensionen, Skalen und Items eine extrem große Stichprobe erfordern würde. Dies hat jedoch zur Folge, dass die Analysen lediglich die Testgüte pro Dimension aber nicht für den gesamten Fragebogen nachweisen. Zudem konnten einige der theoretisch postulierten Skalen innerhalb der Dimensionen nicht faktorenanalytisch getrennt werden. Im vorliegenden Beitrag wurde dieses Problem empirisch (z. B. durch den Ausschluss von Items oder durch das Zusammenlegen von Skalen) bearbeitet. Gleichwohl besitzen gerade diese Befunde Implikationen für die Theorie von Grosche et al. (2020a), die vermutlich präzisiert werden müsste. Dennoch liegt mit dem vorliegenden Fragenbogen ein Instrument vor, das die Forschungstätigkeiten im Rahmen der Umsetzung von Inklusion sinnvoll erweitert und Hinweise zum Ausbau von kokonstruktiven Prozessen, die im Rahmen der Inklusion als wichtige Gelingensbedingung erachtet werden, geben kann. Es steht jedoch außer Frage, dass es sich mit dem vorliegenden Instrument lediglich um eine erste Version handelt, die sich in weiteren Replikationen bewähren muss. Das Ziel war es, die empirischen Bewährungsgrade stetig zu steigern und so die relativ hohe Anzahl an anfangs erstmalig formulierten Items zu kürzen. In einem zukünftigen Entwicklungs- und Überarbeitungsschritt werden wir auf der einen Seite die theoretischen Annahmen der Kokonstruktionstheorie intensiver mit *Item Parcels* prüfen (Little et al., 2013) und die psychometrische Qualität der Items mit *exploratory structural equation modeling* (ESEM) testen (Marsh et al., 2014). Zusätzlich werden wir die Messinvarianz der Faktormodelle für Regelschul- und sonderpädagogische Lehrkräfte evaluieren. Der Einsatz dieser drei Verfahren erschien uns in den vorliegenden vier Teilstudien noch zu früh, da die Verfahren teilweise eine bereits sehr hohe Messgüte erfordern, die in unserem frühen Entwicklungsstadium noch nicht vorausgesetzt werden konnte. Zudem hat ESEM ein großes Problem, sparsame Modelle zu identifizieren (Marsh et al., 2014), weshalb ESEM nicht zu unserem Ziel der Itemreduktion passte. Eine weitere empiri-

sche Bewährung der jetzigen Version des Instruments steht demnach noch aus. Darüber hinaus stehen zur weiteren Validierung im nächsten Schritt die Überprüfung von Zusammenhängen mit externen Variablen im Sinne der diskriminanten Validität an.

Literatur

- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler (7., vollst. überarb. und erw. Aufl.)*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-12770-0>
- Coburn, C. E., & Stein, M. K. (2006). Communities of Practice Theory and the Role of Teacher Professional Community in Policy Implementation. In M. Honig (Hrsg.), *Confronting complexity: Defining the field of education policy implementation* (S. 25–46). Albany: Suny Press.
- De Rooij, M., & Weeda, W. (2020). Cross-Validation: A Method Every Psychologist Should Know. *Advances in Methods and Practices in Psychological Science*, 3(2), 248–263. <https://doi.org/10.1177/2515245919898466>
- Drossel, K. (2015). *Motivationale Bedingungen von Lehrerkooperation. Eine empirische Analyse der Zusammenarbeit im Projekt Ganz* in. Münster: Waxmann.
- Fend, H. (2006). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fussangel, K. (2008). *Subjektive Theorien von Lehrkräften zur Kooperation. Eine Analyse der Zusammenarbeit von Lehrerinnen und Lehrern in Lerngemeinschaften*. Wuppertal. <http://elpub.bib.uni-wuppertal.de/edocs/dokumente/fbg/paedagogik/diss2008/fussangel/index.html>.
- Fussangel, K., & Dizinger, V. (2014). The challenge of change? The development of all-day schools and its implications for teacher stress. *Journal for Educational Research Online*, 6(3), 115–133. <https://doi.org/10.25656/01:9690>
- Gebhard, S., Happe, C., Paape, M., Riestenpatt, J., Vögler, A., Wollenweber, K. U., & Castello, A. (2014). Merkmale und Bewertung der Kooperation von Sonderpädagoginnen und Regelschullehrkräften in inklusiven Unterrichtssettings. *Empirische Sonderpädagogik* (1), 17–32. <https://doi.org/10.25656/01:9243>
- Gräsel, C., Fußangel, K., & Pröbstel, C. (2006). Lehrkräfte zur Kooperation anregen – eine Aufgabe für Sisyphos? *Zeitschrift für Pädagogik*, 52(2), 205–219. <https://doi.org/10.25656/01:4453>
- Grosche, M., Fussangel, K., & Gräsel, C. (2020a). Kokonstruktive Kooperation zwischen Lehrkräften: Aktualisierung und Erweiterung der Kokonstruktionstheorie sowie deren Anwendung am Beispiel schulischer Inklusion. *Zeitschrift für Pädagogik*, 66(4), 461–479. <https://doi.org/10.25656/01:25803>
- Grosche, M., König, J., Huber, C., Hennemann, T., Fussangel, K., Gräsel, C., Kaspar, K., Melzer, C., Strauß, S., Lüke, T., Krull, J., Neroznikova, K., Spilles, M., Casale, G. & Bartling, A. (2020b). Das Forschungsprojekt PARTI: Evaluation einer Fortbildungsreihe zur kokonstruktiven Umsetzung eines um Partizipation ergänzten Response-To-Intervention-Modells im Förderschwerpunkt Emotional-soziale Entwicklung. In G. Ricken & S. Degenhardt (Hrsg.), *Vernetzung, Kooperation, Sozialer Raum – Inklusion als Querschnittsaufgabe* (S. 116–121). Bad Heilbrunn: Klinkhardt.
- Grosche, M., & Moser Opitz, E. (2023). Kooperation von Lehrkräften zur Umsetzung von inklusivem Unterricht – notwendige Bedingung, zu einfach gedacht oder überbewerteter Faktor? *Unterrichtswissenschaft*, 51(2), 245–263. <https://doi.org/10.1007/s42010-023-00172-3>

- Hu, L.-T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jurkowski, S., & Müller, B. (2018). Co-teaching in inclusive classes: The development of multi-professional cooperation in teaching dyads. *Teaching and Teacher Education*, 75, 224–231. <https://doi.org/10.1016/j.tate.2018.06.017>
- Keller-Schneider, M., & Albisser, S. (2013). Kooperation von Lehrpersonen und die Bedeutung von individuellen und kooperativen Ressourcen. In M. Keller-Schneider, S. Albisser & J. Wissinger (Hrsg.), *Professionalität und Kooperation in Schulen. Beiträge zur Diskussion über Schulqualität* (S. 33–56). Bad Heilbrunn: Klinkhardt.
- Kluge, J., & Grosche, M. (2021). Hängen die disziplinären Selbstverständnisse von Lehrkräften in inklusiven Schulen von ihrer kokonstruktiven Kooperation ab? *Empirische Pädagogik*, 35(4), 356–377.
- Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, 18(3), 285–300. <https://doi.org/10.1037/a0033266>
- Lütje-Klose, B., Neumann, P., Gorges, J., & Wild, E. (2018). Die Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements (Bi-LieF) – Zentrale Befunde. *Die Deutsche Schule*, 110(2), 109–123. <https://doi.org/10.31244/dd.2018.02.02>
- Lütje-Klose, B., & Urban, M. (2014). Professionelle Kooperation als wesentliche Bedingung inklusiver Schul- und Unterrichtsentwicklung. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 83(2), 112–123. <https://doi.org/10.2378/vhn2014.art09d>
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. *Annual Review of Clinical Psychology*, 10(1), 85–110. <https://doi.org/10.1146/annurev-clinpsy-032813-153700>
- Massenkeil, J., & Rothland, M. (2016). Kollegiale Kooperation im Lehrerberuf. Überblick und Systematisierung aktueller Forschung. *Schulpädagogik heute*, 7(13). <https://doi.org/10.2307/j.ctvss3xtt.8>
- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. Thousand Oaks: SAGE Publications. <https://doi.org/10.4135/9781412984898>.
- Pröbstel, C. (2008). *Lehrkooperation und die Umsetzung von Innovationen. Eine Analyse der Zusammenarbeit von Lehrkräften aus Perspektive der Bildungsforschung und der Arbeits- und Organisationspsychologie*. Berlin: Logos Verlag.
- Prüfer, P., & Rexroth, M. (2005). *Kognitive Interviews*. Mannheim: Zentrum für Umfragen, Methoden und Analysen.
- Revelle, W. (2020). *psych: Procedures for Personality and Psychological Research*. Evanston, Illinois: Northwestern University. <https://CRAN.R-project.org/package=psych> Version = 2.0.9
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal for Statistical Software*, 48(2), 1–36. <http://www.jstatsoft.org/v48/i02/>
- Ryu, E. (2011). Effects of skewness and kurtosis on normal-theory based maximum likelihood test statistics in multilevel structural equation modeling. *Behavior Research Methods*, 43, 1066–1074. <https://doi.org/10.3758/s13428-011-0115-7>
- Savalei, V., & Rhemtulla, M. (2012). On Obtaining Estimates of the Fraction of Missing Information From Full Information Maximum Likelihood. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(3), 477–494. <https://doi.org/10.1080/10705511.2012.687669>

- Spieß, E. (2004). Kooperation und Konflikt. In H. Schuler (Hrsg.), *Organisationspsychologie* (3. Aufl., S. 193–247). Göttingen: Hogrefe.
- Tabachnick, B.G., & Fidell, L.S. (2007). *Using Multivariate Statistics* (5 ed.). Boston: Pearson International.
- Vangrieken, K., Dochy, F., Raes, E., & Kynndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, *15*, 17–40. <https://doi.org/10.1016/j.edurev.2015.04.002>
- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, *24*, 80–91. <https://doi.org/10.1016/j.tate.2007.01.004>
- Werning, R. (2013). Inklusive Schulentwicklung. In V. Moser (Hrsg.), *Die inklusive Schule* (S. 51–63). Stuttgart: Kohlhammer.

Anhang

Tabelle A1: Items des finalen Fragebogens sowie zugehörige Faktoren und Faktorladungen (Studie 4)

Dimension	Faktoren		
Rahmenbedingungen			
Items	Schulische Rahmenbedingungen	Einstellungen	Erfahrungen
Mein Arbeitsalltag sieht feste Zeiten für die Teamarbeit vor.	.429		
Kooperation wird an unserer Schule „groß geschrieben“.	.951		
Unsere Schulleitung unterstützt die Kooperationsbemühungen der Mitarbeiter/innen.	.760		
Ich fühle mich meinem Team eng verbunden.		.840	
Meine Teamarbeit entspricht meinen Normen, Werten und Überzeugungen.		.811	
Ich freue mich auf bevorstehende Teamsitzungen.		.795	
Wenn ich im Team arbeite, fühle ich mich wohl.		.831	
Ich habe bereits intensiv in Teams zusammengearbeitet.			.509
Ich habe im Team Lösungen für praktische Herausforderungen erarbeitet.			.911
Ich habe bereits intensiv im Team zusammengearbeitet, um Konzepte oder Handlungsmöglichkeiten zu entwickeln.			.784
Vorbedingungen ^a			
Items	Zielinterdependenz	Reziprozität	Deprivatisierung
Ich bin auf mein Team angewiesen.	.793 ^a /.906 ^b		

Vorbedingungen ^a			
Items	Zielinterde- pendenz	Reziprozität	Deprivatisie- rung
Nur durch die verschiedenen Kompetenzen im Team bin ich neuen Anforderungen gewachsen.	.722 ^a /.798 ^b		
Ich bin von meinem Team abhängig.	.749 ^a /.747 ^b		
Ich kann meine Ziele nur durch mein Team erreichen.	.690 ^a /.774 ^b		
Im Team gibt es auch Trittbrettfahrer/innen.		.506 ^a /.661 ^b	
Im Team bekommt jede/r etwas von seinem/ihrem Engagement zurück.		.751 ^a /.851 ^b	
Das Engagement im Team ist ungleich verteilt.		.658 ^a /.678 ^b	
Im Team herrscht ein ausgewogenes Geben und Nehmen.		.937 ^a /.907 ^b	
In meinem Unterricht hospitieren regelmäßig Kolleg*innen.			.685 ^b
Ich führe systematische Beobachtungen im Unterricht meiner Kolleg*innen durch.			.755 ^b
Andere Lehrkräfte führen systematische Beobachtungen meines Unterrichts durch.			.949 ^b
Ich bitte meine Kolleg*innen um aussagekräftiges und gezieltes Feedback zu meinem Unterricht.			.517 ^b

Kokonstruktive Handlungen				
Items	Aushand- lungen	Reflexion	Meinun- gen	Überar- beitung
Im Team handeln wir die zu bearbeitenden Inhalte solange aus, bis wir einen gemeinsamen Plan erarbeitet haben.	.875			
Von außen herangetragene Ziele handeln wir im Team so lange aus, bis wir uns damit identifizieren können.	.759			
Wir reflektieren regelmäßig die Umsetzbarkeit unserer Ideen, Konzepte oder Methoden.		.835		
Wir reflektieren im Team regelmäßig die Erfolge und Misserfolge in unserer Arbeit.		.894		
Wir nutzen die Reflexionsergebnisse zur Weiterentwicklung unserer Ideen, Konzepte oder Methoden.		.901		
Meinungsverschiedenheiten werden im Team über- gangen.			.596	
Im Team versuchen wir, auch Missverständnisse und Meinungsverschiedenheiten produktiv zu nutzen.			.816	
Bei Meinungsverschiedenheiten im Team versuchen wir, den Standpunkt des/r anderen zu verstehen.			.831	
Im Team feilen wir manchmal lange am selben Ge- genstand, um ihn zu optimieren.				.652

Kokonstruktive Handlungen				
Items	Aushandlungen	Reflexion	Meinungen	Überarbeitung
Konzepte oder Methoden, die wir gerade entwickeln, überarbeiten wir im Team mehrfach.				.781
Im Team arbeiten wir solange an Modellen und Konzepten, bis wir mit dem Ergebnis zufrieden sind.				.879

Ergebnisse				
Items	Geteilte Kognitionen	Implementation	Gemeinsame Verantwortung	Inklusives Rollenverständnis
Jede/r im Team kennt unsere gemeinsam erarbeiteten Konzepte oder Methoden.	.779			
Wir haben im Team eine gemeinsame Wissensbasis über die zu bearbeitenden Inhalte.	.828			
Jede/r im Team weiß, wie die anderen bestimmte Dinge interpretieren.	.721			
Ich passe die erarbeiteten Konzepte oder Methoden fortlaufend an den Schulalltag an.		.808		
Ich setze unsere im Team entwickelten Konzepte oder Methoden für alle Schülerinnen und Schüler um.		.616		
Für eine optimale Umsetzung gleiche ich die im Team erarbeiteten Konzepte oder Methoden fortlaufend an die Praxisbedingungen an.		.966		
Wir fühlen uns als Team gemeinsam für alle Schülerinnen und Schüler verantwortlich.			.868	
Wir fühlen uns im Team unabhängig von unserer Profession für alle Schülerinnen und Schüler verantwortlich.			.820	
In unserem Team werden keine Grenzen zwischen den einzelnen Professionen gezogen.				.873
Unsere Teammitglieder legen großen Wert auf eine Abgrenzung zwischen den einzelnen Professionen.				.764

Folgen			
Items	Zeitliche Entlastung	Emotionale Entlastung	Schüler:innen
Durch die Kooperation im Team habe ich mehr Zeit für andere Dinge.	.881		
Durch die intensive Zusammenarbeit im Team bin ich bei der Unterrichtsvorbereitung effizienter.	.834		
Unsere Teamarbeit gibt mir anderswo zeitliche Freiräume.	.924		

Folgen			
Items	Zeitliche Entlastung	Emotionale Entlastung	Schüler:innen
Weil wir die Methoden so gut im Team entwickelt haben, spare ich Zeit.	.917		
Durch die Umsetzung unserer gemeinsam erarbeiteten Maßnahmen bin ich im Unterricht weniger gestresst.		.945	
Durch die Arbeit im Team lasse ich mich durch Misserfolge nicht mehr so schnell entmutigen.		.715	
Wenn die im Team gemeinsam erarbeiteten Konzepte oder Methoden umgesetzt werden ...			
... arbeiten die Schülerinnen und Schüler der Klasse gern miteinander.			.859
... haben alle Schülerinnen und Schüler der Klasse mehr Spaß am Lernen.			.920
... können alle Schülerinnen und Schüler der Klasse besser lernen.			.873

Anmerkungen. ^a Faktorladungen stammen aus den Analysen mit der Teilstichprobe 1 ohne den Faktor Deprivatisierung. ^b Faktorladungen stammen aus den Analysen mit der Teilstichprobe 2 mit dem Faktor Deprivatisierung.

Tabelle A2: Ursache der ausgeschlossenen Items sowie zugehörige Faktoren aus den Studien 3 und 4

Dimension und Skala	Item	Teildatensatz ^a	Ursache des Ausschlusses
Studie 3			
Rahmenbedingungen			
Schulische Rahmenbedingungen	Im Schulgebäude werden Räume zur Verfügung gestellt, in denen wir als Team ungestört arbeiten können.	Train	geringe Ladung
Schulische Rahmenbedingungen	An meiner Schule gibt es keine Strukturen für die Teamarbeit.	Test	korrelierter Fehlerterm
Erfahrungen	Ich habe in der Vergangenheit bisher wenige positive Erfahrungen in Teams sammeln können.	Train	zu geringe Ladung
Erfahrungen	Ich habe bereits über einen längeren Zeitraum in gut funktionierenden Teams gearbeitet.	Test	korrelierter Fehlerterm
Einstellungen	Meine Teammitglieder sind mit sympathisch.	Train	nicht normalverteilt
Einstellungen	Ich respektiere auch Arbeitsweisen und Haltungen im Team, die mit meinen nicht vereinbar sind.	Train	nicht normalverteilt
Einstellungen	Ich erkenne Meinungen meiner Teammitglieder an.	Train	nicht normalverteilt

Dimension und Skala	Item	Teildatensatz ^a	Ursache des Ausschlusses
Einstellungen	Ich würde auch freiwillig in meinem Team mitarbeiten	Train	nicht normalverteilt
	Ich kann meine Ideen und Vorstellungen in meinem Team umsetzen.	Train	geringe Ladung, inhaltliche Passung
Einstellungen	Ich identifiziere mich vollständig mit den Zielen unseres Teams.	Train	inhaltliche Nähe
Vorbedingungen			
Zielinterdependenz	Ich brauche mein Team nicht.	Train	zu geringe Ladung
Zielinterdependenz	Ich kann durch mein Team Dinge umsetzen, die ich alleine nicht schaffe.	Train	geringe Ladung
Vertrauen	Im Team können Fehler nicht offen angesprochen werden.	Train	geringe Ladung, inhaltliche Passung
Vertrauen	Ich traue mich im Team, auch unbequeme Meinungen zu äußern.	Train	zu geringe Ladung
Vertrauen	Im Team kann Kritik nicht offen geäußert werden.	Train	zu geringe Ladung
Egalität	Ich begegne meinen Teammitgliedern auf Augenhöhe.	Train	nicht normalverteilt
Egalität	Alle Teammitglieder sind gleichberechtigt.	Train	zu geringe Ladung
Egalität	In unserem Team bestimmt eine Person, wo es lang geht.	Train	zu geringe Ladung
Egalität	Alle Teammitglieder bringen eigene aber gleich wichtige Perspektiven ein.	Train	zu geringe Ladung
Reziprozität	Im Team nutzt keiner den anderen aus.	Train	geringe Ladung
Reziprozität	Im Team profitieren alle voneinander.	Train	inhaltliche Passung
Deprivatisierung	Ich finde, Unterricht ist Privatsache jeder einzelnen Lehrperson.	Train	nicht normalverteilt
Deprivatisierung	Ich lege meine Arbeitsweisen im Team offen.	Train	zu geringe Ladung
Deprivatisierung	Von meinen Kolleginnen und Kollegen lasse ich mir nur ungern in meinen Unterricht hineinreden.	Train	zu geringe Ladung
Deprivatisierung	Ich fühle mich unwohl, wenn Kolleginnen und Kollegen meinen Unterricht beobachten.	Train	Ladung auf falschem Faktor
Deprivatisierung	In meinem Unterricht können jederzeit andere Personen zuschauen.	Train	Ladung auf falschem Faktor

Dimension und Skala	Item	Teildatensatz ^a	Ursache des Ausschlusses
Kokonstruktive Handlungen			
Aushandlung	Im Team nutzen wir unterschiedliche pädagogische Perspektiven für die Weiterentwicklung von Konzepten oder Methoden.	Train	zu geringe Ladung
Aushandlung	Vor- und Nachteile verschiedener Handlungsalternativen werden im Team nicht diskutiert.	Train	zu geringe Ladung
Ergebnisse			
Gemeinsame Verantwortung	Bei den einzelnen Teammitgliedern ist das Verantwortungsgefühl für alle Schülerinnen und Schüler ungleich verteilt.	Train	zu geringe Ladung
Inklusives Rollenverständnis	Teammitglieder mit ursprünglich anderer Profession übernehmen auch Aufgaben aus anderen Bereichen.	Train	zu geringe Ladung
Inklusives Rollenverständnis	In unserem Team spielt es keine Rolle, für welche Schulform jemand ausgebildet wurde.	Train	zu geringe Ladung
Folgen			
Emotionale Entlastung	Auch zu Hause muss ich an Schwierigkeiten in der Teamarbeit denken.	Train	zu geringe Ladung
Emotionale Entlastung	Ich finde es gut, dass wir so intensiv im Team gearbeitet haben.	Train	geringe Ladung, inhaltliche Passung
Studie 4			
Ergebnisse			
Gemeinsame Kognitionen	Wir haben im Team ein gemeinsames Verständnis unserer Arbeit entwickelt.		Krossladung

Anmerkungen. ^a Der Datensatz wurde randomisiert in zwei gleich große Teildatensätze geteilt. Am Trainingsdatensatz wurden die Items zunächst deskriptiv analysiert (Schiefe und Kurtosis) und eine explorative Faktorenanalyse wurde durchgeführt. Im Anschluss wurde die Faktorstruktur anhand des Test-Datensatzes konfirmatorisch getestet.