



Sprachunterstützende Massnahmen in Schweizer Schulen.  
Studie zur Kooperation multiprofessioneller Teams in integrativen Settings.

# Statistischer Pretest

-

## Fragebogenevaluation

Dr. Christoph Till  
Projektleiter  
PHBern, Institut für Heilpädagogik  
Fabrikstrasse 8, 3012 Bern  
christoph.till@phbern.ch

Jasmin Kolb  
Hilfsassistentin  
PHBern, Institut für Forschung,  
Entwicklung und Evaluation  
Fabrikstrasse 2, 3012 Bern  
jasmin.kolb@phbern.ch



**Kanton Bern**  
Canton de Berne

## Inhalt

1 Einleitung .....	5
2 Kognitiver Pretest .....	5
3 Statistischer Pretest .....	6
3.1 Stichprobenbeschreibung .....	6
3.2 Methode .....	7
4 Ergebnisse .....	8
4.1 Itemschwierigkeiten und Standardabweichungen der abhängigen Variablen.....	8
4.1.1 unterrichtsintegrierte sprachunterstützende Maßnahmen (sM).....	8
4.1.2 Häufigkeit und Intensität der Zusammenarbeit in Bezug auf sprachunterstützende Maßnahmen (sM) .....	11
4.1.3 Einstellungen zur Zusammenarbeit in Bezug auf sprachunterstützende Maßnahmen (sM) .....	15
4.2 Analyse der Missings.....	17
4.2.1 Analyse der Missings in der Skala «Sprachunterstützende Maßnahmen im Allgemeinen»	18
4.2.2 Analyse der Missings in der Skala «Unterrichtsformen».....	19
4.2.3 Analyse der Missings in der Skala «Räumlichkeiten» .....	19
4.2.4 Analyse der Missings in der Skala «Austausch mit den Logopäd:innen» .....	20
5 Objektivität des Fragebogens .....	20
5.1 Durchführungsobjektivität.....	20
5.2 Auswertungsobjektivität.....	21
5.3 Interpretationsobjektivität .....	21
6 Reliabilität des Fragebogens .....	22
6.1 Reliabilität der Subskala «sprachunterstützende Maßnahmen im Allgemeinen» .....	23
6.2 Reliabilität der Subskala «sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten».....	25
6.3 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten».....	28
6.4 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Unterrichtsformen».....	28
6.5 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Fachbereichen» .....	30
6.6 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Erwerbsphasen».....	31
6.7 Reliabilität der Subskala «sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers».....	31
6.8 Reliabilität der Subskala «Austausch mit den Regellehrpersonen» .....	34
6.9 Reliabilität der Subskala «Austausch mit den Schulischen Heilpädagog:innen».....	35

6.10	Reliabilität der Subskala «Austausch mit den Logopäd:innen».....	36
6.11	Reliabilität der Subskala «Synchronisation mit den Regellehrpersonen».....	36
6.12	Reliabilität der Subskala «Synchronisation mit den Schulischen Heilpädagog:innen» .....	37
6.13	Reliabilität der Subskala «Synchronisation mit den Logopäd:innen» .....	38
6.14	Reliabilität der Subskala «Ko-Konstruktion mit den Regellehrpersonen».....	38
6.15	Reliabilität der Subskala «Ko-Konstruktion mit den Schulischen Heilpädagog:innen» .....	40
6.16	Reliabilität der Subskala «Ko-Konstruktion mit den Logopäd:innen» .....	41
6.17	Reliabilität der Subskala «Team-Teaching» .....	42
6.18	Reliabilität der Subskala «Intention».....	43
6.19	Reliabilität der Subskala «Interesse».....	44
6.20	Reliabilität der Subskala «Kosten».....	45
6.21	Reliabilität der Subskala «Nützlichkeit für die professionelle Entwicklung».....	45
6.22	Reliabilität der Subskala «Nützlichkeit für die Schüler und Schülerinnen sowie die Schule» ..	46
6.23	Reliabilität der Subskala «prozessorientierte Erfolgserwartung» .....	47
6.24	Reliabilität der Subskala «ergebnisorientierte Erfolgserwartung».....	48
6.25	Reliabilität der Subskala «Zufriedenheit».....	49
6.26	Zusammenfassung der Reliabilitätsanalyse.....	50
7	Validität des Fragebogens.....	52
7.1	Inhaltsvalidität .....	52
7.2	Kriteriumsvalidität.....	53
7.3	Konstruktvalidität.....	55
7.3.1	CFA der Skala «sprachunterstützende Maßnahmen im Allgemeinen» .....	56
7.3.2	CFA der Skala «sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten» .	57
7.3.3	CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten».....	59
7.3.4	CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Unterrichtsformen» .....	61
7.3.5	CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Fachbereichen».....	62
7.3.6	CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Erwerbsphasen» .....	64
7.3.7	CFA der Skala «sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers».....	65
7.3.8	CFA der Skala «Austausch» .....	65
7.3.9	CFA der Skala «Synchronisation».....	66
7.3.10	CFA der Skala «Ko-Konstruktion» .....	68
7.3.11	CFA der Skala «Team-Teaching».....	71

7.3.12 CFA der Skala «Intention».....	72
7.3.13 CFA der Skala «Interesse».....	74
7.3.13 CFA der Skala «Kosten» .....	76
7.3.14 CFA der Skala «Nützlichkeit für die professionelle Entwicklung».....	78
7.3.15 CFA der Skala «Nützlichkeit für die Schüler und Schülerinnen sowie die Schule» .....	80
7.3.16 CFA der Skala «prozessorientierte Erfolgserwartung» .....	82
7.3.17 CFA der Skala «ergebnisorientierte Erfolgserwartung» .....	83
7.3.18 CFA der Skala «Zufriedenheit».....	84
7.3.19 Zusammenfassung der CFAs .....	86
8 Analyse der Antworttendenzen .....	89
8.1 soziale Erwünschtheit .....	89
8.2 Akquieszenz .....	89
8.3 Tendenz zur Mitte und zu Milde/Härte .....	90
8.4 Effekte der Fragereihenfolge .....	90
8.5 Reihenfolge der Antwortvorgaben .....	91
9 Schlussfolgerungen und Implikationen.....	91
10 Verzeichnisse .....	94
10.1 Literaturverzeichnis .....	94
10.2 Abbildungsverzeichnis .....	95
10.3 Tabellenverzeichnis.....	96

# 1 Einleitung

Im Projekt SprICH - Sprachunterstützende Massnahmen an Schweizer Schulen: Studie zur Kooperation multiprofessioneller Teams in integrativen Settings wird die Zusammenarbeit von Regellehrpersonen, Schulischen Heilpädagog:innen und Logopäd:innen in der Unterstützung von sprachentwicklungsauffälligen Kindern in der Regelschule untersucht. Dabei sollten folgende Fragestellungen beantwortet werden:

- Unter welchen Bedingungen werden sprachunterstützende Maßnahmen bevorzugt integrativ umgesetzt?
- Wie oft und wie intensiv wird zu diesem Zweck zwischen den verschiedenen Fachkräften zusammengearbeitet?
- Welche Bedingungen werden dabei als förderlich oder hinderlich erlebt?

Die Beantwortung der Fragen soll mit der Skala zum unterrichtsbezogenen Kooperationsverhalten (Drossel, 2015) vorgenommen werden. Items mit Bezug zur Sprachunterstützung werden verschiedenen Fragebögen entnommen (v. a. Elksnin & Capilouto, 1994; Reber & Blechschmidt, 2014), gezielt zusammengestellt und durch Fragen zu soziodemografischen Angaben ergänzt. Die Items der genannten Fragebögen müssen jedoch adaptiert werden, damit sie zum beschriebenen Forschungsanliegen passen und von allen Berufsgruppen in gleicher Weise verstanden werden.

Um zu testen, ob die adaptierten Fragebogenitems die Daten in gewünschter Weise erheben, wurden kognitive und statistische Pretests durchgeführt (s. Kap. 2 und 3). Die Ergebnisse (Itemschwierigkeiten und Missings) werden in Kapitel 4 dargestellt und analysiert. In Kapitel 5, 6 und 7 werden die Gütekriterien des Messinstruments untersucht. In Kapitel 8 werden noch ein paar abschließende Überlegungen zu möglichen Antworttendenzen angestellt und in Kapitel 9 schließlich Konsequenzen für die Fragebogenüberarbeitung gezogen.

## 2 Kognitiver Pretest

Da die Originalitems der verwendeten Fragebögen für andere Forschungsfragen entwickelt wurden, ist die Formulierung für den vorliegenden Zweck nicht ganz angemessen. Es war nötig, die Items z.T. so umzuformulieren, dass die Thematik «unterrichtsintegrierte Sprachunterstützung in multiprofessionellen Teams» möglichst gut erfasst wird. Ein Problem dabei war, dass die Sprachunterstützung verschiedener Fachkräfte auch begrifflich und inhaltlich unterschieden werden musste, da ja von vornherein klar ist, dass Regellehrpersonen keine Sprachtherapie anbieten (das machen Logopäd:innen) und Logopäd:innen keinen sprachsensiblen Unterricht geben (das machen Regellehrpersonen und/oder Schulische Heilpädagog:innen).

Außerdem mussten Formulierungen gefunden werden, die von allen drei Berufsgruppen gleichermaßen gut verstanden und auch gleichermaßen gut beantwortet werden können. Da schon die Ausbildung dieser Fachkräfte und v.a. deren Arbeitsalltag vollkommen andersartig geprägt sind, war dies keine leichte Herausforderung.

In einer Reihe kognitiver Interviews (vgl. Prüfer & Rexroth 2005) wurden diese Formulierungen vorgetestet. Mittels verschiedener Techniken des Nachfragens, der Bewertung der Verlässlichkeit und des Paraphrasierens wurden insgesamt 21 Personen (5 Regellehrpersonen, 8 Schulische Heilpädagog:innen und 8 Logopäd:innen) befragt. Dabei gab es einzelne Fragen, die von Beginn an von Vertreter:innen aller drei Berufsgruppen schnell und korrekt verstanden. Andere Fragen hingegen

waren schwierig zu formulieren, da sie von verschiedenen Personen aus den drei Berufsgruppen sehr unterschiedlich interpretiert und dementsprechend auch beantwortet wurden. Die Rückmeldungen wurden fortlaufend in der Überarbeitung des Fragebogens berücksichtigt und in den nächsten Interviews erprobt. Nach und nach konnten die Formulierungen ermittelt werden, die mit grösster Wahrscheinlichkeit die gewünschte Antwort evozieren. Absolute Sicherheit gab es jedoch nie. Insgesamt wurden im Laufe der Zeit sieben Fragebogenversionen erstellt und in den Interviews getestet. Letzten Endes ist ein Fragebogen entstanden, dessen Items durch Personen aus allen drei Berufsgruppen mitgestaltet wurden.

### 3 Statistischer Pretest

Dieser in den kognitiven Interviews geprüfte Fragebogen sollte nun auch an einer grösseren Stichprobe unter realen Untersuchungsbedingungen getestet werden, sodass auch statistische Analysen angewendet werden konnten.

#### 3.1 Stichprobenbeschreibung

Um die Zielstichprobe (Regellehrpersonen, Schulische Heilpädagog:innen und Logopäd:innen an Regelschulen im Kanton Bern) nicht schon frühzeitig zu erfassen, wurden die drei Berufsgruppen gezielt ausserhalb des Kantons Bern angefragt. Ab Oktober 2021 wurden persönliche Kontakte gebeten, den Fragebogen unter ihren Mitarbeitenden zu streuen. Die Teilnahme erfolgte auf freiwilliger Basis. Als Gegenleistung wurde jedoch eine Zusammenfassung der Ergebnisse zur Zusammenarbeit in ihrem Schulhaus, ihrer Gemeinde oder sogar ihrem Kanton (je nach Datenlage) angeboten. Der Rücklauf fiel auch nach dem Versenden eines Reminders (nach zwei bis drei Wochen) zu gering aus, sodass ab Mitte November 2021 die kantonalen Berufsverbände für Lehrpersonen (inkl. den Fraktionen für Schulische Heilpädagogik) und Logopädie angeschrieben wurden. Das Interesse relativ gross und viele Personen, primär Logopäd:innen, füllten den Fragebogen aus. Anfangs Dezember 2021 wurden v.a. Regellehrpersonen und Schulische Heilpädagog:innen noch einmal animiert, an der Erhebung teilzunehmen, da diese beiden Berufsgruppen eher unterrepräsentiert waren. Am Ende konnten Daten von 58 Lehrpersonen, 65 Schulischen Heilpädagog:innen und 155 Logopäd:innen (gesamt: 278 Personen) gesammelt werden (s. Tab. 1).

Tabelle 1: Stichprobenbeschreibung

	Gesamt (N = 278)	Regellehrpersonen (n = 58)	Schulische Heilpädagog:innen (n = 65)	Logopäd:innen (n = 155)
Alter in Jahren (SD)	45,0 (10,7)	46,2 (11,2)	48,2 (9,1)	43,3 (10,9)
Geschlecht (♂:♀ in %)	7,7:92,3	12,7:87,3	10,8:89,2	4,6:95,4
Berufserfahrung in Jahren (SD)	15,7 (10,6)	20,4 (12,0)	12,4 (9,4)	15,0 (9,8)
Lektionen pro Woche (SD)	20,1 (6,2)	22,5 (6,6)	19,6 (6,5)	19,5 (5,7)
Anzahl Arbeitstage pro Woche (SD)	3,7 (1,0)	3,9 (1,0)	3,9 (1,1)	3,6 (1,0)
Anzahl Arbeitsorte (SD)	2,6 (2,5)	1,4 (1,4)	1,5 (1,1)	3,6 (2,8)
Kinder in Zuständigkeit (SD)	22,6 (21,7)	25,4 (16,5)	17,4 (11,6)	19,3 (7,8)
• Davon mit Schwierigkeiten der gesprochenen Sprache (in %)	59,7%	37,4%	50,6%	84,5%

Logopäd:innen sind in dieser Stichprobe die grösste Gruppe, was der tatsächlichen Verteilung widerspricht. Auch gibt es hinsichtlich weiterer Variablen (Alter in Jahren, Berufserfahrung in Jahren, Lektionen pro Woche und Anzahl Arbeitsorte) bedeutende Unterschiede zwischen den Gruppen, die Hypothesentests allenfalls konfundieren könnten und kontrolliert werden müssten. Die Unterschiede in der Menge Kinder mit Schwierigkeiten der gesprochenen Sprache pro Berufsgruppe sind zu erwarten und haben mit dem jeweiligen Aufgabenprofil der Berufsgruppen zu tun.

Geografisch gesehen verteilen sich die Teilnehmenden über die ganze Deutschschweiz (außer Kanton Bern), wobei die Anzahl Personen pro Kanton sehr stark schwanken kann (s. Abb. 1).

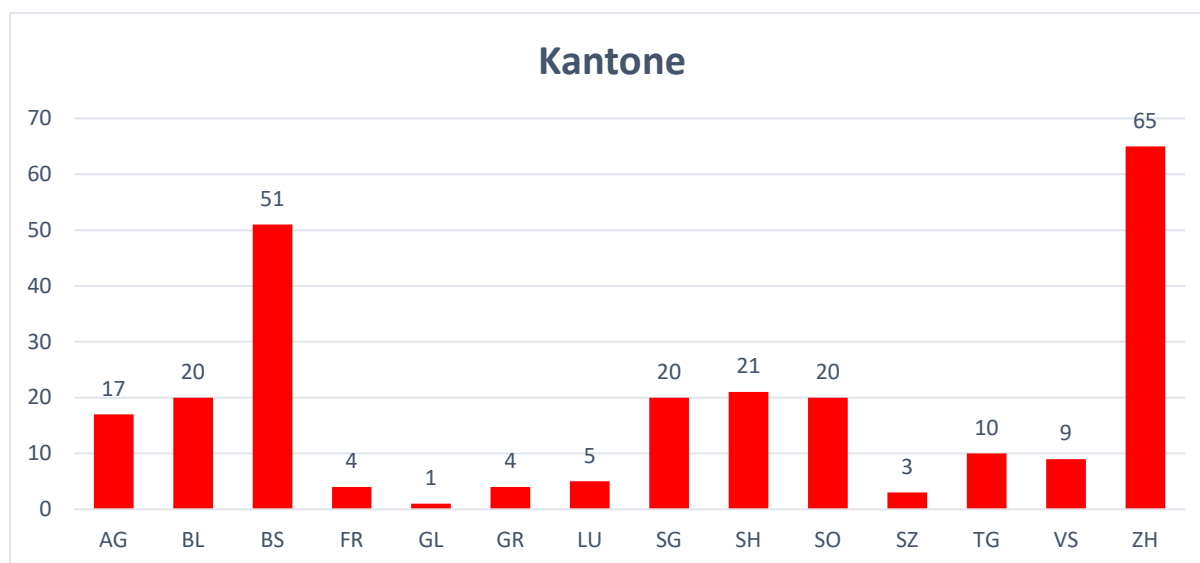


Abbildung 1: Anzahl Teilnehmende pro Kanton

Die Kantone Basel-Stadt und Zürich sind sehr stark repräsentiert. Da zwischen den Kantonen verschiedene (Zusammen-)Arbeitsbedingungen vorherrschen und auch unterschiedliche Vorgaben existieren, werden gerade jene aus den beiden genannten Kantonen besonders gut abgebildet. Zudem kann es sein, dass aus einigen Kantonen nur die Antworten von ein oder zwei, aber eben nicht allen drei Berufsgruppen erhoben werden konnten. Auch diesbezüglich gilt, dass die angefallene Stichprobe nicht repräsentativ ist. Auch sind die Teilnehmendenzahlen pro Kanton zu klein, um entsprechende Unterschiede statistisch überprüfen zu können.

### 3.2 Methode

Der Fragebogen wurde als Online-Umfrage mithilfe von Unipark per Link im Zeitraum Oktober bis Dezember 2021 in Umlauf gebracht. Er besteht aus vier Teilen: im ersten Teil werden wichtige Ausgangsvariablen erfasst (sozio-demografische Angaben, Angaben zur Ausbildung und Beschäftigung). Im zweiten Teil wird die Häufigkeit der Umsetzung unterrichtsintegrierter sprachunterstützender Maßnahmen erhoben, im dritten Teil die Häufigkeit und Intensität der multiprofessionellen Kooperation und im vierten Teil letztlich die Einstellungen zur multiprofessionellen Kooperation. Die Teile 2-4 des Fragebogens bestehen aus insgesamt 19 Subskalen bzw. 115 Items.

Drei Subskalen («Austausch», «Synchronisation» und «Ko-Konstruktion») sind aufgesplittet. So werden die Items je zweimal erfragt, jeweils in Bezug auf eine andere Berufsgruppe. Eine

Regellehrperson beispielsweise muss also einmal einschätzen, wie oft sie sich mit der Schulischen Heilpädagog:in austauscht, und dann noch einmal, wie oft sie sich mit der Logopäd:in austauscht. In den kognitiven Pretests hat sich nämlich gezeigt, dass es hier zu bedeutenden Unterschieden kommen kann. Diese werden aber nur ersichtlich, wenn man sie differenziert erfasst. Würde man hingegen fragen, wie oft sich jemand «mit den anderen Fachkräften» austauscht, neigen die befragten Personen dazu, einen Mittelwert zu generieren, der jedoch nicht erkennen lässt, wie oft sie sich tatsächlich und mit wem genau austauscht. Auch wenn der Fragebogen dadurch deutlich verlängert wurde, hat sich der Informationsgehalt in einer wichtigen Hinsicht gesteigert.

Die mittlere Bearbeitungsdauer für den Online-Fragebogen beträgt 28 Minuten, der Median liegt bei 21 Minuten. D.h. konkret, dass die Hälfte der Personen weniger als 21 Minuten gebraucht hat, was eine erstrebenswerte Zahl ist. Die andere Hälfte jedoch hat mehr als 21 Minuten benötigt, z.T. sogar deutlich mehr, sodass der Mittelwert relativ stark vom Median abweicht. Es könnte allerdings auch sein, dass einige Personen eine Pause gemacht haben, während der Fragebogen geöffnet und aktiv war. Dies würde in einigen Fällen den Mittelwert deutlich verzerren, was aber nicht überprüft werden kann. Letzten Endes wird eine Bearbeitungsdauer von 20 bis 25 Minuten angestrebt. Dazu sind einige Kürzungen notwendig.

## 4 Ergebnisse

### 4.1 Itemschwierigkeiten und Standardabweichungen der abhängigen Variablen

In den Tabellen 2-26 werden für alle Subskalen die Itemschwierigkeiten (Mittelwert), Standardabweichungen sowie die Anzahl fehlender Werte (Missings) angegeben.

#### 4.1.1 unterrichtsintegrierte sprachunterstützende Maßnahmen (sM)

*Tabelle 2: Häufigkeit sprachunterstützender Maßnahmen (sM) im Allgemeinen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings*

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
sM im Klassensetting (sm_kl)	2.95 (1.61)	14 (5.04)
sM im Gruppensetting (sm_gr)	3.67 (1.31)	14 (5.04)
sM im Einzelsetting (sm_ein)	4.38 (1.14)	11 (3.96)
Sprachbeobachtungen (sm_beo)	3.61 (1.18)	14 (5.04)
DaZ-Unterricht (sm_daz)	1.91 (1.48)	40 (14.39)
Kollegiale Beratung (sm_ber)	3.06 (1.03)	7 (2.52)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag



Tabelle 3: Häufigkeit sprachunterstützender Maßnahmen im Allgemeinen bei Auffälligkeiten ..., basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
des Redeflusses (rfl_all)	3.15 (1.38)	9 (3.24)
der Aussprache (aus_all)	4.54 (0.76)	6 (2.16)
des Wortschatzes (ws_all)	4.73 (0.51)	4 (1.44)
der Grammatik (gra_all)	4.50 (0.77)	5 (1.80)
des Sprachverstehens (sv_all)	4.65 (0.69)	4 (1.44)
des Sprachgebrauchs (kom_all)	3.88 (1.11)	7 (2.52)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 4: Häufigkeit unterrichtsintegrierter Maßnahmen bei Auffälligkeiten ..., basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
des Redeflusses (rfl_kl)	2.28 (1.49)	10 (3.60)
der Aussprache (aus_kl)	2.62 (1.62)	7 (2.52)
des Wortschatzes (ws_kl)	2.87 (1.66)	5 (1.80)
der Grammatik (gra_kl)	2.64 (1.59)	9 (3.24)
des Sprachverstehens (sv_kl)	2.90 (1.68)	5 (1.80)
des Sprachgebrauchs (kom_kl)	2.63 (1.57)	7 (2.52)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 5: Häufigkeit unterrichtsintegrierter sprachunterstützender Maßnahmen bei verschiedenen Unterrichtsformen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Gruppenarbeiten (uf_gru)	2.90 (1.54)	10 (3.60)
Einzelarbeiten (uf_ein)	3.16 (1.69)	8 (2.88)
offener Unterricht (uf_off)	2.78 (1.59)	9 (3.24)
Frontalunterricht (uf_fron)	2.26 (1.44)	23 (8.27)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 6: Häufigkeit unterrichtsintegrierter sprachunterstützender Maßnahmen in verschiedenen Fachbereichen, basierend auf den Antworten der Gesamtstichprobe unter Ausschluss der «trifft nicht zu»-Antworten

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	n <sup>b</sup>
Deutsch (fb_deu)	3.68 (1.52)	205
Fremdsprachen (fb_frem)	2.37 (1.55)	102
Mathematik (fb_mat)	3.17 (1.64)	156
NMG (fb_nmg)	3.23 (1.57)	159
Gestalten (fb_gest)	2.70 (1.61)	122
Musik (fb_mus)	2.42 (1.62)	111
Bewegung und Sport (fb_spo)	2.43 (1.58)	110
Medien und Informatik (fb_inf)	2.01 (1.35)	110
Berufliche Orientierung (fb_ber)	1.74 (1.14)	65

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag; <sup>b</sup> Das es auch eine «trifft nicht zu»-Option gab, wurden nicht die Missings, sondern die Anzahl verwertbarer Antworten angegeben

Tabelle 7: Häufigkeit unterrichtsintegrierter sprachunterstützender Maßnahmen in verschiedenen Erwerbsphasen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Einführung (pha_ein)	2.58 (1.47)	11 (3.96)
Übung (pha_ueb)	3.07 (1.63)	8 (2.88)
Transfer (pha_anw)	3.08 (1.59)	7 (2.52)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 8: Häufigkeit sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Gang (rau_gan)	2.39 (1.44)	24 (8.63)
Gruppenraum (rau_gru)	2.69 (1.59)	23 (8.27)
Separates Arbeitszimmer (rau_sep)	3.12 (1.77)	14 (5.04)
Außerhalb der Schule (rau_aus)	2.05 (1.64)	32 (11.51)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

#### 4.1.2 Häufigkeit und Intensität der Zusammenarbeit in Bezug auf sprachunterstützende Maßnahmen (sM)

Tabelle 9: Häufigkeit der Kooperationsform "Austausch" mit den Regellehrpersonen, basierend auf den Antworten der Schulischen Heilpädagog:innen und Logopäd:innen (n = 220) nach paarweisem Ausschluss der Missings

Item «Austausch ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> SD)	Missings (%)
von Materialien (aus1_lp)	2.73 (1.00)	7 (3.18)
über relevante Themen (aus2_lp)	3.60 (0.99)	3 (1.36)
über Inhalte der sM (aus3_lp)	3.39 (0.97)	3 (1.36)
über kindlichen Sprachentwicklungsstand (aus4_lp)	3.33 (0.90)	6 (2.73)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 10: Häufigkeit der Kooperationsform "Austausch" mit den Schulischen Heilpädagog:innen, basierend auf den Antworten der Regellehrpersonen und Logopäd:innen (n = 213) nach paarweisem Ausschluss der Missings

Item «Austausch ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
von Materialien (aus1_sh)	2.74 (0.99)	6 (2.82)
über relevante Themen (aus2_sh)	3.30 (0.98)	3 (1.41)
über Inhalte der sM (aus3_sh)	3.13 (0.98)	6 (2.82)
über kindlichen Sprachentwicklungsstand (aus4_sh)	3.22 (0.92)	7 (3.29)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 11: Häufigkeit der Kooperationsform "Austausch" mit den Logopäd:innen, basierend auf den Antworten der Regellehrpersonen und Schulischen Heilpädagog:innen (n = 123) nach paarweisem Ausschluss der Missings

Item «Austausch ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
von Materialien (aus1_lo)	2.07 (1.02)	4 (3.25)
über relevante Themen (aus2_lo)	2.38 (1.01)	3 (2.44)
über Inhalte der sM (aus3_lo)	2.29 (1.00)	3 (2.44)
über kindlichen Sprachentwicklungsstand (aus4_lo)	2.39 (0.96)	7 (5.69)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 12: Häufigkeit der Kooperationsform "Synchronisation" mit den Regellehrpersonen, basierend auf den Antworten der Schulischen Heilpädagog:innen und Logopäd:innen (n = 220) nach paarweisem Ausschluss der Missings

Item «Absprachen bezüglich ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Themen/Inhalte der sM (syn1_lp)	3.09 (0.91)	2 (0.91)
Anforderungen von Aufträgen (syn2_lp)	2.87 (0.99)	3 (1.36)
Termine von Aufträgen (syn3_lp)	2.39 (1.26)	5 (2.27)
Ziele und Schwerpunkte der sM (syn4_lp)	2.73 (0.89)	5 (2.27)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 13: Häufigkeit der Kooperationsform "Synchronisation" mit den Schulischen Heilpädagog:innen, basierend auf den Antworten der Regellehrpersonen und Logopäd:innen (n = 213) nach paarweisem Ausschluss der Missings

Item «Absprachen bezüglich ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Themen/Inhalte der sM (syn1_sh)	2.92 (0.89)	3 (1.41)
Anforderungen von Aufträgen (syn2_sh)	2.64 (1.01)	4 (1.88)
Termine von Aufträgen (syn3_sh)	2.17 (1.12)	7 (3.29)
Ziele und Schwerpunkte der sM (syn4_sh)	2.59 (0.89)	7 (3.29)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 14: Häufigkeit der Kooperationsform "Synchronisation" mit den Logopäd:innen, basierend auf den Antworten der Regellehrpersonen und Schulischen Heilpädagog:innen (n = 123) nach paarweisem Ausschluss der Missings

Item «Absprachen bezüglich ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Themen/Inhalte der sM (syn1_lo)	2.27 (0.98)	2 (1.63)
Anforderungen von Aufträgen (syn2_lo)	2.05 (1.02)	3 (2.44)
Termine von Aufträgen (syn3_lo)	1.69 (0.96)	5 (4.07)
Ziele und Schwerpunkte der sM (syn4_lo)	2.11 (0.90)	6 (4.88)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 15: Häufigkeit der Kooperationsform "Ko-Konstruktion" mit den Regellehrpersonen, basierend auf den Antworten der Schulischen Heilpädagog:innen und Logopäd:innen (n = 220) nach paarweisem Ausschluss der Missings

Item «gemeinsame/r ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Durchführung von Hospitationen (kok1_lp)	1.71 (0.83)	3 (1.36)
Durchführung von Elterngesprächen (kok2_lp)	2.41 (0.68)	4 (1.82)
Besuch von Weiterbildungen (kok3_lp)	1.87 (0.51)	5 (2.27)
Planung von sM (kok4_lp)	1.92 (0.98)	6 (2.73)
Erprobung von sM (kok5_lp)	1.73 (0.85)	5 (2.27)
Erstellung von Arbeitsblättern (kok6_lp)	1.55 (0.94)	4 (1.82)
Erarbeitung von Konzepten für sM (kok7_lp)	1.80 (0.87)	5 (2.27)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 16: Häufigkeit der Kooperationsform "Ko-Konstruktion" mit den Schulischen Heilpädagog:innen, basierend auf den Antworten der Regellehrpersonen und Logopäd:innen (n = 213) nach paarweisem Ausschluss der Missings

Item «gemeinsame/r ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Durchführung von Hospitationen (kok1_sh)	1.50 (0.62)	6 (2.82)
Durchführung von Elterngesprächen (kok2_sh)	2.31 (0.74)	5 (2.35)
Besuch von Weiterbildungen (kok3_sh)	1.87 (0.59)	4 (1.88)
Planung von sM (kok4_sh)	1.85 (0.89)	6 (2.82)
Erprobung von sM (kok5_sh)	1.69 (0.83)	6 (2.82)
Erstellung von Arbeitsblättern (kok6_sh)	1.36 (0.68)	7 (3.29)
Erarbeitung von Konzepten für sM (kok7_sh)	1.72 (0.80)	5 (2.35)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 17: Häufigkeit der Kooperationsform "Ko-Konstruktion" mit den Logopäd:innen, basierend auf den Antworten der Regellehrpersonen und Schulischen Heilpädagog:innen (n = 123) nach paarweisem Ausschluss der Missings

Item «gemeinsame/r ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Durchführung von Hospitationen (kok1_lo)	1.17 (0.50)	5 (4.07)
Durchführung von Elterngesprächen (kok2_lo)	1.49 (0.57)	4 (3.25)
Besuch von Weiterbildungen (kok3_lo)	1.56 (0.52)	6 (4.88)
Planung von sM (kok4_lo)	1.57 (0.78)	4 (3.25)
Erprobung von sM (kok5_lo)	1.36 (0.65)	5 (4.07)
Erstellung von Arbeitsblättern (kok6_lo)	1.09 (0.31)	6 (4.88)
Erarbeitung von Konzepten für sM (kok7_lo)	1.49 (0.77)	3 (2.44)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 18: Häufigkeit der Kooperationsform "Team-Teaching" im Allgemeinen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
Unterricht und Einzelunterstützung (tt_ein)	2.95 (1.63)	9 (3.24)
Unterricht in Stationen (tt_sta)	2.15 (1.41)	9 (3.24)
Paralleles Unterrichten (tt_par)	2.01 (1.38)	11 (3.96)
Klassen- und Förderunterricht (tt_foe)	2.07 (1.45)	7 (2.52)
Interaktives Unterrichten (tt_int)	1.78 (1.26)	10 (3.60)

<sup>a</sup> 1 = nie, 2 = 1-2 Mal pro Schuljahr, 3 = 1-2 Mal pro Monat, 4 = 1-2 Mal pro Woche, 5 = fast jeden Tag

Tabelle 19: Intention, die Häufigkeit der sprachunterstützenden Maßnahmen (sM) und die Zusammenarbeit künftig zu intensivieren, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «Intensivierung ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
der unterrichtsintegrierten sM (int1)	2.21 (1.08)	3 (1.08)
des Austauschs (int2)	2.68 (0.93)	3 (1.08)
der Synchronisation (int3)	2.18 (0.89)	3 (1.08)
der Ko-Konstruktion (int4)	2.33 (0.89)	8 (2.88)
des Team-Teachings (int5)	1.91 (0.90)	3 (1.08)

<sup>a</sup> 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft zu

#### 4.1.3 Einstellungen zur Zusammenarbeit in Bezug auf sprachunterstützende Maßnahmen (sM)

Tabelle 20: Interesse an multiprofessioneller Kooperation, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «multiprofessionelle Kooperation ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
ist sinnvoll (intr1)	3.82 (0.41)	1 (0.36)
hat einen hohen Stellenwert (intr2)	3.47 (0.66)	3 (1.08)
macht Spaß (intr3)	3.49 (0.64)	2 (0.72)
wird gerne durchgeführt (intr4)	3.58 (0.59)	3 (1.08)
ist ein wichtiger Teil der Arbeit (intr5)	3.25 (0.79)	1 (0.36)
ist wichtig (intr6)	3.61 (0.58)	2 (0.72)
fördert die Teamzugehörigkeit (intr7)	3.31 (0.88)	1 (0.36)

<sup>a</sup> 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft zu

Tabelle 21: Empfundene Kosten der multiprofessionellen Kooperation, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «multiprofessionelle Kooperation ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
kostet unnötig viel Zeit (kost1)	1.96 (0.77)	3 (1.08)
kostet unnötig viel Energie (kost2)	1.88 (0.74)	4 (1.44)
kostet Überwindung (kost3)	1.41 (0.63)	4 (1.44)
kostet Zeit für wichtigere Dinge (kost4)	2.25 (1.01)	5 (1.80)
beeinträchtigt private Verhältnisse (kost5)	1.79 (0.91)	5 (1.80)
lohnt sich mittel- und langfristig nicht (kost6)	1.39 (0.69)	3 (1.08)
ist verlorene Zeit (kost7)	1.61 (0.74)	6 (2.16)
ist mit schlechten Erfahrungen verknüpft (kost8)	1.57 (0.82)	3 (1.08)

<sup>a</sup> 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft zu

Tabelle 22: Wahrgenommene Nützlichkeit der multiprofessionellen Kooperation für die eigene, professionelle Entwicklung, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «multiprofessionelle Kooperation ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
motiviert mich bei meiner Arbeit (nue_fk1)	3.25 (0.70)	3 (1.08)
ist für meine Karriere vorteilhaft (nue_fk2)	2.81 (0.92)	5 (1.80)
führt zur Verbesserung meiner sM (nue_fk3)	3.29 (0.71)	3 (1.08)
stärkt das Teamzugehörigkeitsgefühl (nue_fk4)	3.42 (0.73)	4 (1.44)
spart langfristig gesehen Ressourcen ein (nue_fk5)	2.79 (0.88)	4 (1.44)
liefert wertvolles Feedback zur Arbeit (nue_fk6)	3.20 (0.75)	6 (2.16)
entlastet emotional (nue_fk7)	2.92 (0.84)	5 (1.80)
offenbart gemeinsamen Problemen (nue_fk8)	3.32 (0.64)	5 (1.80)
Vermindert das Gefühl des Alleinseins (nue_fk9)	3.11 (0.80)	4 (1.44)
ermöglicht gemeinsame Problemlösungen (nue_fk10)	3.03 (0.82)	3 (1.08)
stärkt den Rückhalt (nue_fk11)	3.27 (0.68)	3 (1.08)

<sup>a</sup> 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft zu

Tabelle 23: Wahrgenommene Nützlichkeit der multiprofessionellen Kooperation für die Schüler und Schülerinnen sowie die Schule, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «multiprofessionelle Kooperation ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
steigert die Qualität der sM (nue_sus1)	3.44 (0.67)	3 (1.08)
führt zur Entwicklung gemeinsamer Strategien (nue_sus2)	3.34 (0.65)	4 (1.44)
schafft ein besseres Arbeitsklima (nue_sus3)	3.27 (0.75)	4 (1.44)
steigert die Fortschritte in der Sprachentwicklung (nue_sus4)	3.35 (0.68)	8 (2.88)
unterstützt die sM in allen Fachbereichen (nue_sus5)	3.32 (0.71)	6 (2.16)
führt zur Entwicklung gemeinsamer Handlungsmuster (nue_sus6)	3.24 (0.67)	8 (2.88)
etabliert gemeinsame Vorgehensweisen (nue_sus7)	3.36 (0.66)	5 (1.80)
führt zur Entwicklung gemeinsamer Materialien (nue_sus8)	2.83 (0.87)	7 (2.52)
trägt zum fachlichen Austausch bei (nue_sus9)	3.57 (0.62)	6 (2.16)
führt zur Weitergabe von Erfahrungen (nue_sus10)	3.57 (0.59)	6 (2.16)
stärkt das Zusammengehörigkeitsgefühl (nue_sus11)	3.47 (0.67)	5 (1.80)

<sup>a</sup> 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft zu



Tabelle 24: Einschätzung der prozessorientierten Erfolgserwartung, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «Ich ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
kann gut mit anderen kooperieren (erf_pro1)	3.70 (0.47)	2 (0.72)
werde als Kooperationspartner:in geschätzt (erf_pro2)	3.52 (0.53)	4 (1.44)
komme gut mit anderen aus (erf_pro2)	3.55 (0.51)	2 (0.72)

<sup>a</sup> 1 = gar nicht wahrscheinlich, 2 = eher unwahrscheinlich, 3 = eher wahrscheinlich, 4 = sehr wahrscheinlich

Tabelle 25: Einschätzung der ergebnisorientierten Erfolgserwartung, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «Ich ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
kann anderen bei Problemen helfen (erf_erg1)	3.41 (0.59)	3 (1.08)
werde oft zu Rate gezogen (erf_erg2)	3.25 (0.64)	3 (1.08)
bringe die Schule voran (erf_erg3)	3.01 (0.70)	6 (2.16)
habe Kompetenzen, von denen andere profitieren (erf_erg4)	3.48 (0.57)	2 (0.72)

<sup>a</sup> 1 = gar nicht wahrscheinlich, 2 = eher unwahrscheinlich, 3 = eher wahrscheinlich, 4 = sehr wahrscheinlich

Tabelle 26: Zufriedenheit mit der Arbeitssituation, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings

Item «Zufriedenheit mit ...» (Abkürzung)	Itemschwierigkeit <sup>a</sup> (SD)	Missings (%)
der Rolle im Team (zufr1)	3.47 (0.62)	1 (0.36)
der Wertschätzung durch andere (zufr2)	3.37 (0.66)	5 (1.80)
dem Gehalt (zufr3)	3.30 (0.78)	2 (0.72)
den Möglichkeiten zur Zusammenarbeit (zufr4)	3.12 (0.89)	2 (0.72)
der ausgewiesenen Vor- und Nachbereitungszeit (zufr5)	2.83 (0.91)	2 (0.72)
den räumlichen Ressourcen (zufr6)	3.13 (1.00)	3 (1.08)
dem Arbeitsmaterial (zufr7)	3.42 (0.79)	5 (1.80)
der Qualifikation zur Zusammenarbeit (zufr8)	3.01 (0.95)	3 (1.08)

<sup>a</sup> 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft zu

## 4.2 Analyse der Missings

Mithilfe des MCAR-Tests (MCAR = missing completely at random) nach Little kann überprüft werden, ob fehlende Werte vollkommen zufällig oder systematisch (MNAR = missing not at random) fehlen. Hier sollten alle Skalen untersucht werden, bei denen der Anteil fehlender Werte 5% an der untersuchten (Teil-)Stichprobe übersteigt, da ein entsprechender oder gar größerer Anteil fehlender Werte die Untersuchungsergebnisse bedeutend verzerren kann (Leonhart 2013, S. 105). Nur wenn das Fehlen der Werte zufällig zustande kommt, dürfen sie durch entsprechende Verfahren ersetzt werden.

Gemäß Dong und Peng (2013) sind die Ersetzungsverfahren MI (Multiple Imputation), FIML (Full Information-Maximum-Likelihood) und EM (Expectation Maximization) gleichermaßen gut dafür geeignet.

Da MNAR nie endgültig bestätigt werden kann (Leonhart 2013, S. 677), wird im Falle des Nichtvorliegens von MCAR automatisch von MAR (missing at random) ausgegangen und entsprechende Ersetzungsverfahren eingesetzt. Dies gilt auch für einen nicht-signifikanten MCAR-Test, da hiermit nur ermittelt werden kann, dass MCAR nicht verworfen werden darf, MCAR kann aber nicht endgültig bestätigt werden (Leonhart 2013, S. 681). Demnach wird auch in diesen Fällen von MAR ausgegangen.

Die Skalen in den Tabellen 2, 5, 8 und 11 weisen Werte von mehr als 5% Missings auf und sollen hier näher untersucht werden.

#### 4.2.1 Analyse der Missings in der Skala «Sprachunterstützende Maßnahmen im Allgemeinen»

Der MCAR-Test nach Little wird nicht signifikant ( $\chi^2 = 71.55$ ,  $df = 70$ ,  $p = .426$ ), was bedeutet, dass sich die Fälle mit fehlenden Werten nicht von denen ohne fehlende Werte unterscheiden und die Missings rein zufällig verteilt sind. Deswegen dürfen die fehlenden Werte im Rahmen weiterer Analysen durch geeignete Verfahren ersetzt werden.

Dennoch können einige bedeutsame, möglicherweise systematische Zusammenhänge bei der Analyse der Missings entdeckt werden. Gemäß der T-Tests bei unterschiedlicher Varianz unterscheiden sich Personen, die bei

- sm\_kl fehlende Werte haben bezüglich sm\_ein und sm\_daz,
- sm\_beo fehlende Werte haben bezüglich sm\_kl und
- sm\_daz fehlende Werte haben bezüglich sm\_kl, sm\_beo und sm\_ber

signifikant von denjenigen, die dort keine fehlenden Werte haben. V.a. das Item sm\_daz («DaZ-Unterricht») weist eine hohe Anzahl Missings auf und führt zu systematischen Unterschieden in der Verteilung weiterer Variablenausprägungen.

Betrachtet man die Verteilung der Missings über die verschiedenen Berufsgruppen hinweg, kann Folgendes festgestellt werden:

- sm\_ein und sm\_beo weisen v.a. in der Gruppe der Regellehrpersonen einen hohen Anteil Missings auf (je 12.1%)
- sm\_daz weist in allen Gruppen eine hohe Anzahl Missings auf (Regellehrpersonen = 10.3%, Schulische Heilpädagog:innen = 16.9%, Logopäd:innen = 14.8%).

Auch hier erweist sich das Item sm\_daz («DaZ-Unterricht») als schwierig. So scheinen die befragten Personen lieber gar nichts zu antworten, als die Option «nie» auszuwählen, da diese Option für sie nicht zutreffend sein mag. Es muss überlegt werden, ob das Item umformuliert werden kann oder lieber weggelassen wird.

Die Items sm\_ein («sprachunterstützende Maßnahmen im Einzelsetting») und sm\_beo («Sprachbeobachtungen») hingegen können nicht weggelassen werden, da sie von zwei Berufsgruppen sicher beantwortet werden können. Eine Umformulierung der Items kommt jedoch der Gruppe der Regellehrpersonen evtl. zugute.

#### 4.2.2 Analyse der Missings in der Skala «Unterrichtsformen»

Der MCAR-Test nach Little wird nicht signifikant ( $\chi^2 = 17.59$ ,  $df = 14$ ,  $p = .226$ ), was bedeutet, dass sich die Fälle mit fehlenden Werten nicht von denen ohne fehlende Werte unterscheiden und die Missings rein zufällig verteilt sind. Deswegen dürfen die fehlenden Werte im Rahmen weiterer Analysen durch geeignete Verfahren ersetzt werden.

Dennoch können einige bedeutsame, möglicherweise systematische Zusammenhänge bei der Analyse der Missings entdeckt werden. Gemäß der T-Tests bei unterschiedlicher Varianz unterscheiden sich Personen, die bei

- uf\_fron fehlende Werte haben bezüglich der Items uf\_gru, uf\_ein und uf\_off

signifikant von denjenigen, die dort keine fehlenden Werte haben. Das Item uf\_fron («Frontalunterricht») weist also eine hohe Anzahl Missings auf und führt zu systematischen Unterschieden in der Verteilung weiterer Variablenausprägungen.

Betrachtet man die Muster fehlender Werte, fällt auf, dass es v.a. Schulische Heilpädagog:innen sind (18.5% Missings), die das Item uf\_fron («Frontalunterricht») nicht angewählt haben. Die Vermutung liegt nahe, dass die genannte Berufsgruppe sehr wohl weiß, was Frontalunterricht ist, sich aber aus anderen Gründen entschieden hat, dieses Item lieber nicht zu beantworten als die Option «nie» anzuwählen. Eine Überarbeitung der Itemformulierung in einer Weise, die es auch Schulischen Heilpädagog:innen ermöglicht, hier eine Antwort auszuwählen, ist notwendig.

Eine weitere berufsgruppenspezifische Häufung von Fehlermustern findet sich bei den Logopäd:innen – hier finden sich gleich vier Personen, die keines der Items der Skala beantwortet haben. Evtl. besteht hier eine gewisse Unsicherheit bezüglich der Auslegung der Items.

#### 4.2.3 Analyse der Missings in der Skala «Räumlichkeiten»

Der MCAR-Test nach Little wird signifikant ( $\chi^2 = 61.62$ ,  $df = 23$ ,  $p < .001$ ), was bedeutet, dass die fehlenden Werte nicht zufällig, sondern wahrscheinlich systematisch verteilt sind. Es scheint einen bedeutenden Zusammenhang zwischen den verschiedenen Missings zu geben, weswegen die fehlenden Werte nicht einfach durch die genannten Verfahren ersetzt werden dürfen.

Gemäß der T-Tests bei unterschiedlicher Varianz unterscheiden sich Personen, die bei

- rau\_gan fehlende Werte haben bezüglich des Items rau\_sep,
- rau\_gru fehlende Werte haben bezüglich der Items rau\_gan, rau\_sep und rau\_aus,
- rau\_sep fehlende Werte haben bezüglich der Items rau\_gan und rau\_aus und
- rau\_aus fehlende Werte haben bezüglich der Items rau\_gru und rau\_sep

signifikant von denjenigen, die dort keine fehlenden Werte haben.

Betrachtet man die Muster fehlender Werte, fällt auf, dass das Item rau\_aus («ausserhalb der Schule») in 11 Fällen die einzige fehlende Angabe auf dieser Skala ist und alle anderen Items beantwortet werden konnten. Anscheinend ist nicht eindeutig geworden, was mit «ausserhalb der Schule» gemeint ist, obwohl ein Beispiel angegeben wurde. Die anderen Missings hingegen treten eher in Kombination auf. In sieben Fällen wurden keine Angaben bei rau\_gan («im Gang»), rau\_gru («im Gruppenraum») und rau\_aus («ausserhalb der Schule») gemacht, sodass lediglich eine Aussage zum rau\_sep («separaten Arbeitszimmer») vorliegt. Es ist unklar, ob die anderen Antwortoptionen nicht

voneinander unterschieden und deswegen nicht ausgewählt werden konnten, aber ob hier eine «ökonomische Strategie» zu erkennen ist, das Item zu beantworten, wo eine einfache und schnelle Beantwortung möglich war und die anderen auszulassen. Ausserdem fällt auf, dass es bei denjenigen, die zwar bei rau\_aus («ausserhalb der Schule») Angaben gemacht haben, aber sonst diverse Missings aufzeigen, ausschließlich um Logopäd:innen handelt. Evtl. sind diese bei einem logopädischen Dienst angestellt und deswegen per se außerhalb der Schule tätig und haben die anderen Optionen als für sich «nicht zutreffend» beurteilt. Darüber hinaus gibt es keine berufsgruppenspezifische Häufung von Missings in dieser Skala.

Ein systematischer Zusammenhang (MNAR) kann, wie eingangs beschrieben, jedoch nicht nachhaltig bestätigt werden, sodass von MAR ausgegangen wird.

#### 4.2.4 Analyse der Missings in der Skala «Austausch mit den Logopäd:innen»

Der MCAR-Test nach Little wird nicht signifikant ( $\chi^2 = 5.68$ ,  $df = 7$ ,  $p = .578$ ), was bedeutet, dass sich die Fälle mit fehlenden Werten nicht von denen ohne fehlende Werte unterscheiden und die Missings rein zufällig verteilt sind. Deswegen dürfen die fehlenden Werte im Rahmen weiterer Analysen durch geeignete Verfahren ersetzt werden.

Dennoch können einige bedeutsame, möglicherweise systematische Zusammenhänge bei der Analyse der Missings entdeckt werden. Gemäß der T-Tests bei unterschiedlicher Varianz unterscheiden sich Personen, die bei

- aus4\_lo fehlende Werte haben bezüglich des Items aus2\_lo

signifikant von denjenigen, die dort keine fehlenden Werte haben.

Von den vier Fällen, die lediglich bei aus4\_lo («Austausch über kindlichen Sprachentwicklungsstand») Missings haben, sind drei Lehrpersonen. Man darf vermuten, dass der Begriff «kindlicher Sprachentwicklungsstand» allenfalls nicht geläufig ist, da ansonsten ja auch eine Option «nie» ausgewählt werden könnte, wenn ein entsprechendes Gespräch mit der logopädischen Fachkraft eben nicht stattfindet.

Ansonsten häufen sich Fälle, bei denen alle Angaben auf der Skala fehlen. Aus den kognitiven Pretests ist bekannt, dass einzelne Personen lieber nichts auswählen als die Option «nie», wenn kein Kontakt zu einer logopädischen Fachkraft (aus welchen Gründen auch immer) besteht. Die Skala könnte durch eine Option «trifft nicht zu» ergänzt werden, wenn aktuell keine logopädische Fachkraft zum Team gehört, um derartige Missings zu vermeiden.

## 5 Objektivität des Fragebogens

### 5.1 Durchführungsobjektivität

Der Fragebogen wird online via Unipark (Tivian XI GmbH, 2021) zur Verfügung gestellt. Die Teilnehmenden werden der Reihe nach durch die Seiten des Fragebogens geführt. Dabei erfolgen alle Angaben auf freiwilliger Basis. Lediglich die Zugehörigkeit zu einer Berufsgruppe (Regellehrpersonen, Schulische Heilpädagog:innen, Logopäd:innen) ist eine Pflichtangabe, worauf teilnehmende Personen hingewiesen werden, sofern hier keine Option ausgewählt wird.

Bei den meisten Fragen müssen beschriftete Antwortoptionen bzw. konkrete Skalenwerte (1-5 bzw. 1-4) ausgewählt werden. Wenn mehrere Optionen ausgewählt werden können, wird in einem Hilfetext darauf hingewiesen.

Bei insgesamt acht Fragen müssen die Antworten frei eingegeben werden (das Alter in Jahren, das Jahr der Diplomierung, die Anzahl der Arbeitsorte, die Anzahl der Lektionen pro Woche, die Anzahl der Arbeitstage pro Woche, die Anzahl der Kinder, für die eine Person verantwortlich ist, sowie die Anzahl der Kinder mit Auffälligkeiten der gesprochenen Sprache, für die eine Person verantwortlich ist). Bei sieben dieser acht Fragen werden somit Zahleneingaben verlangt. Nur bei einer Frage (Name und Postleitzahl des Arbeitsortes) können auch Wörter eingegeben werden. Die Teilnehmenden erhalten Hilfestellungen und Beispiele, wie eine mögliche Antwort aussehen sollte. Dennoch konnte festgestellt werden, dass es immer wieder Teilnehmende gibt, die von diesen Vorgaben abweichen und ggf. eine Interpretation oder gar ein Verwerfen der Eingabe notwendig machen. Hier ist allenfalls eine Präzisierung bzw. Umformulierung der entsprechenden Items notwendig.

Darüber hinaus konnte festgestellt werden, dass der Umgang mit den Skalenwerten keine Probleme bereitet und beibehalten werden kann. Die Durchführungsobjektivität wird – abgesehen von den erwähnten Einschränkungen – als gegeben betrachtet.

## 5.2 Auswertungsobjektivität

Die meisten Items werden auf einer vier- bzw. fünfstufigen Likertskala beantwortet. Die Items, die zu einer Subskala des Fragebogens gehören, können als Skalenmittelwert zusammengefasst werden.

Auch die Antworten, die per freier Eingabe erfolgen, sind in der Regel leicht auszuwerten, da hier sehr konkrete Sachverhalte (wie z.B. die Anzahl Arbeitstage pro Woche) erfragt werden. Schwierig wird es mitunter dann, wenn das vorgegebene Antwortformat verletzt wird. So antworteten mehrere Personen auf die Frage, in welchem Jahr sie ihr berufsqualifizierendes Diplom erworben haben, mit einer zweistelligen Zahl, z.B. «10». Es ist unklar, ob das Diplom also «im Jahr 2010» (vorgegebenes vierstelliges Antwortformat) oder «vor zehn Jahren» erworben wurde. In Einzelfällen bot das Alter der Person Aufschluss, sodass die eine oder die andere Auswertungsart wahrscheinlicher schien.

Weniger gut aufzulösen sind unvollständige Angaben bei der Frage nach dem aktuellen Arbeitsort (Kindergarten bzw. Schulhaus inkl. Postleitzahl), bei der einige Personen entweder nur den Ort oder nur die Postleitzahl nannten. Gewisse Ortsnamen gibt es mehrmals in der Schweiz, sodass kein eindeutiger Ort ermittelt werden konnte. Oder der Ort lässt sich eindeutig identifizieren, aber es gibt dort mehrere Kindergärten bzw. Schulhäuser, sodass die Zugehörigkeit zu einem bestimmten Schulteam nicht bestimmt werden kann. Gleiches gilt für die alleinige Angabe der Postleitzahl.

Alles in allem werden aber auch diese Items dank der angebotenen Eingabehilfen in den meisten Fällen eindeutig beantwortet. Dies und der erfolgreiche Umgang der Teilnehmenden mit den Likertskalen spricht hier für eine höchst objektive Auswertung.

## 5.3 Interpretationsobjektivität

Da der Fragebogen aus verschiedenen existierenden Fragebogenteilen gezielt zusammengesetzt und erstmalig bei den genannten Berufsgruppen eingesetzt wurde, ist eine Orientierung an Normwerten zur Interpretation der Skalenwerte nicht möglich. Auch ist die vorliegende Stichprobe zu klein, um eigene Normwerte zu ermitteln. Dennoch können die Veröffentlichungen von Reber und Blechschmidt (2014) und Drossel (2015) Anhaltspunkte liefern, wie die hier eingegangenen Antworten einzuordnen sind.

Des Weiteren muss darauf hingewiesen werden, dass die Logopäd:innen in der angefallenen Stichprobe den größten Anteil der Teilnehmenden ausmachen, was einer realistischen und damit repräsentativen Verteilung widerspricht. Treffen die hypothetischen Annahmen zu, dass gerade Logopäd:innen bezüglich bestimmter Subskalen deutlich niedrigere Werte erzielen als die beiden anderen Berufsgruppen, führt dies zu einer Verzerrung der Itemschwierigkeiten in Richtung der

Logopäd:innen. Eine ergänzende Itemanalyse mit einer zufälligen Auswahl von Logopäd:innen, die zu einer gleichmäßigen Verteilung der Teilnehmenden über die drei Berufsgruppen führt, könnte hier Aufschluss geben. Dies ist jedoch nur als Annäherung an «objektive» Werte zu sehen und kann nicht als definitive Interpretationsgrundlage herangezogen werden.

Die Interpretation der Ergebnisse wird dadurch vereinfacht, dass die Stufen der verwendeten Likertskalen konkrete Bezeichnungen erhalten haben (s. Tab. 27). Die Skalentypen II und III wurden unverändert der Vorlage von Drossel (2015) entnommen. Auch der Skalentyp I entstammt dem Fragebogen von Drossel, wurde aber leicht adaptiert. So wurden auch bei den Stufen 3 und 4 konkrete Angaben (1-2 Mal) ergänzt (im Original nur bei Stufe 2), da sich in den kognitiven Pretests herausstellte, dass es den Teilnehmenden die Auswahl erleichterte.

*Tabelle 27: Verwendete Likertskalen und deren Beschriftung*

Anzahl Stufen	Skalentyp I	Skalentyp II	Skalentyp III
1	nie	trifft nicht zu	gar nicht wahrscheinlich
2	1-2 Mal pro Schuljahr	trifft eher nicht zu	eher unwahrscheinlich
3	1-2 Mal pro Monat	trifft eher zu	eher wahrscheinlich
4	1-2 Mal pro Woche	trifft zu	sehr wahrscheinlich
5	fast jeden Tag	-	-

Die verwendeten Skalen erfüllen außerdem mehrere Empfehlungen von Krosnick und Presser (2010) zur Gestaltung von Likertskalen: sie sollten eher wenige Stufen (mind. drei, max. neun) umfassen, da eine Person ihren persönlichen Wert erst in eine Skalenstufe übersetzen und dann übertragen können muss – zu breite Skalen erschweren diesen Prozess; jede Skalenstufe sollte konkret benannt und eindeutig interpretiert werden können; die Skalenstufen sollten kontinuierlich ansteigend sein und sich inhaltlich nicht überlappen. Einzig das Anbieten einer neutralen Skalenstufe («weder-noch», «weiß ich nicht») wurde hier nicht berücksichtigt.

Bezüglich des Skalentyps I ist fraglich, ob dieser ein Kontinuum darstellt. Offensichtlich gibt es Lücken bzw. Brüche in den dargestellten Zeiträumen – so ist z.B. fraglich ob «3 Mal pro Monat» schon zur nächst höheren Stufe zählen darf oder nicht. In einer ähnlich angelegten Skala übersetzt Theisel (2014) diese Stufen in 1 = «nie», 2 = «selten», 3 = «gelegentlich», 4 = «oft» und 5 = «sehr oft». Hier wird zwar eher ein Kontinuum dargestellt, dies aber zu Lasten der eindeutigen Interpretierbarkeit, da keine klaren Abstufungen von «selten» zu «gelegentlich» bzw. von «gelegentlich» zu «oft» möglich sind. In den kognitiven Pretests hat sich die in Tabelle 27 angegebene Stufenbezeichnung bewährt und wird unter den genannten Einschränkungen beibehalten.

Insgesamt wird die Interpretationsobjektivität als «eingeschränkt» eingeschätzt, was vor allem auf fehlende Normwerte und eine mögliche Verzerrung der Mittelwerte durch die nicht-repräsentative Stichprobe zurückzuführen ist. Die eindeutigen Skalenbezeichnungen ermöglichen dennoch eine inhaltliche Interpretation der Antworten.

## 6 Reliabilität des Fragebogens

Da eine Wiederholung der Erhebung oder eine Aufteilung des Fragebogens in gleichwertige Teile nicht möglich war, wird zur Überprüfung der Reliabilität die interne Konsistenz überprüft. Da es sich bei der Gesamtstichprobe um eine recht heterogene Gruppe bestehend aus Personen aus drei

unterschiedlichen Berufsgruppen handelt, werden die Subskalen getrennt für die Gesamtstichprobe wie für die verschiedenen Berufsgruppen untersucht, sollte dies nötig sein. Fälle mit fehlenden Werten werden, wenn nicht anders erwähnt, listenweise ausgeschlossen. Das bedeutet auch, dass sich die hier dargestellten Itemschwierigkeiten von denjenigen in den Tabellen 2-26 unterscheiden können.

Als Entscheidungskriterien für oder gegen das Beibehalten von Skalen und/oder einzelnen Items dienen folgende Kennwerte: für Cronbachs Alpha  $> .9$  exzellent,  $> .8$  gut,  $> .7$  akzeptabel,  $> .6$  fragwürdig,  $> .5$  schlecht und  $< .5$  inakzeptabel (Blanz, 2021, S. 250); für die Itemtrennschärfe  $> .50$  hoch,  $.50 - .30$  mittel und  $< .30$  niedrig (Bortz & Döring, 2016, S. 478). Items mit Trennschärfen kleiner als  $.30$  sollten gemäß Blanz (2021, S. 239) aus dem Itempool entfernt werden.

### 6.1 Reliabilität der Subskala «sprachunterstützende Maßnahmen im Allgemeinen»

Da diese Subskala überdurchschnittlich viele Missings aufweist, werden die fehlenden Werte durch das EM-Verfahren geschätzt.

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit  $.476$  als inakzeptabel zu bewerten (s. Tabelle 28). Drei der sechs Items sind mit Werten zwischen  $0.3$  und  $0.5$  als «mittelmäßig» trennscharf und zwei Items mit kleineren Werten als nicht trennscharf einzuordnen. Das Item *sm\_ein* («sprachunterstützende Maßnahmen im Einzelsetting») weist sogar eine negative Itemtrennschärfe auf, was bedeutet, dass dieses Item negativ mit dem übrigen Summenwert dieser Subskala korreliert. Personen, die hier einen hohen Wert angeben, erreichen im Rest der Skala tiefere Werte und umgekehrt. Da die zugrunde liegende Stichprobe mit Teilnehmenden aus drei unterschiedlichen Berufsgruppen in dieser Hinsicht sehr heterogen ist und auch sein darf, überrascht dies nicht unbedingt. Eine Reliabilitätsanalyse der Subskala mit den Teilstichproben kann aufschlussreich sein.

Tabelle 28: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
sm_kl	2.95 (1.57)	.354	.354
sm_gr	3.68 (1.28)	.365	.359
sm_ein	4.37 (1.12)	-.144	.596
sm_beo	3.62 (1.15)	.241	.431
sm_daz	1.96 (1.40)	.389	.336
sm_ber	3.05 (1.02)	.253	.428

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Regellehrpersonen ist mit  $.717$  als akzeptabel einzustufen. Alle Items weisen mittlere bis hohe Trennschärfen auf (s. Tab. 29, S. 24). Das Cronbachs Alpha kann durch das Weglassen eines Items nicht verbessert werden.

Tabelle 29: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Stichprobe der Regellehrpersonen

<b>Regellehrpersonen</b>			
<b>Item</b>	<b>Itemschwierigkeit<sup>a</sup> (SD)</b>	<b>Korrigierte Itemtrennschärfe</b>	<b>Cronbachs Alpha, wenn Item weggelassen</b>
sm_kl	4.49 (0.85)	.464	.683
sm_gr	4.31 (0.81)	.445	.688
sm_ein	3.38 (1.46)	.424	.693
sm_beo	3.43 (0.99)	.508	.666
sm_daz	2.96 (1.64)	.409	.711
sm_ber	2.73 (1.20)	.602	.630

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Schulischen Heilpädagog:innen ist mit .616 als fragwürdig einzustufen (s. Tab. 30). Zwei Items weisen mittlere bis hohe Trennschärfen auf, die restlichen Items korrelieren nur schwach mit der Restskala. Das Cronbachs Alpha könnte durch Weglassen des Items smdaz («DaZ-Unterricht») leicht, aber nicht wesentlich verbessert werden.

Tabelle 30: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Stichprobe der Schulischen Heilpädagog:innen

<b>Schulische Heilpädagog:innen</b>			
<b>Item</b>	<b>Itemschwierigkeit<sup>a</sup> (SD)</b>	<b>Korrigierte Itemtrennschärfe</b>	<b>Cronbachs Alpha, wenn Item weggelassen</b>
sm_kl	3.89 (1.29)	.512	.495
sm_gr	4.42 (0.68)	.291	.599
sm_ein	3.84 (1.19)	.272	.603
sm_beo	3.55 (1.02)	.361	.569
sm_daz	2.65 (1.52)	.271	.624
sm_ber	2.80 (1.04)	.468	.529

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Logopäd:innen ist mit .516 als schlecht einzustufen (s. Tab. 31, S. 25). Drei Items weisen mittlere Trennschärfen auf, die übrigen Items korrelieren nur schwach mit der Restskala. Das Cronbachs Alpha könnte durch Weglassen des Items sm\_ein leicht, aber nicht wesentlich verbessert werden.



Tabelle 31: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Stichprobe der Logopäd:innen

Item	Logopäd:innen		
	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
sm_kl	1.94 (1.06)	.370	.413
sm_gr	3.13 (1.35)	.294	.467
sm_ein	4.96 (0.23)	.116	.529
sm_beo	3.71 (1.25)	.236	.498
sm_daz	1.29 (0.74)	.347	.446
sm_ber	3.28 (0.87)	.310	.453

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Es fällt auf, dass die Reliabilitätswerte der Subskala gemessen mit den Teilstichproben besser sind als die gemessen mit der Gesamtstichprobe, und auch die negative Itemtrennschärfe fällt weg. Aber die hier untersuchte Subskala weist alles in allem eher schlechte Reliabilitätswerte auf, sodass der Einsatz dieser Skala in unveränderter Form fragwürdig erscheint.

## 6.2 Reliabilität der Subskala «sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .632 als fragwürdig einzustufen (s. Tab. 32, S. 26). Die meisten Items weisen mittlere bis hohe Trennschärfen auf, außer dem Item rfl\_all («des Redeflusses»). Mit Blick auf die berufsgruppenspezifischen Mittelwerte begründet sich der Verdacht, dass die Regellehrpersonen und Schulische Heilpädagog:innen nicht genau wissen, um was es sich bei Auffälligkeiten des Redeflusses genau handelt, obwohl ein Beispiel (Stottern) angegeben wurde. So geben Regellehrpersonen und Schulische Heilpädagog:innen an, im Durchschnitt beinahe wöchentlich damit zu tun zu haben (M = 3.76 bzw. 3.33), Logopäd:innen, als Fachpersonen für Redeflussstörungen, jedoch noch nicht einmal monatlich (M = 2.85). Dies spricht dafür, dieses Item umzuformulieren oder einfach wegzulassen. Das Cronbachs Alpha würde sich auf .710 erhöhen und wäre somit akzeptabel.

Tabelle 32: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rfl_all	3.15 (1.38)	.199	.710
aus_all	4.54 (0.77)	.476	.553
ws_all	4.74 (0.51)	.553	.565
gra_all	4.52 (0.74)	.436	.568
sv_all	4.65 (0.70)	.488	.556
kom_all	3.89 (1.11)	.366	.591

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Regellehrpersonen ist mit .775 als akzeptabel einzustufen (s. Tab. 33). Alle Items weisen hohe Trennschärfen auf. Das Cronbachs Alpha kann durch das Weglassen eines Items nicht verbessert werden.

Tabelle 33: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Regellehrpersonen

Item	Regellehrpersonen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rfl_all	3.81 (1.32)	.542	.754
aus_all	4.25 (0.83)	.589	.727
ws_all	4.60 (0.63)	.614	.734
gra_all	4.11 (0.93)	.469	.754
sv_all	4.53 (0.87)	.582	.727
kom_all	4.32 (0.85)	.465	.755

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Schulischen Heilpädagog:innen ist mit .755 als akzeptabel einzustufen (s. Tab. 34, S. 27). Ein Item weist eine mittlere Trennschärfe auf, die restlichen Items korrelieren hoch mit der Restskala. Das Cronbachs Alpha könnte durch Weglassen des Items rfl\_all verbessert werden.

Tabelle 34: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Schulischen Heilpädagog:innen

<b>Schulische Heilpädagog:innen</b>			
Item	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rfl_all	3.33 (1.53)	.376	.790
aus_all	4.23 (1.05)	.668	.670
ws_all	4.58 (0.56)	.628	.718
gra_all	4.22 (0.86)	.550	.710
sv_all	4.47 (0.78)	.500	.724
kom_all	3.83 (1.20)	.537	.708

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Logopäd:innen ist mit .460 als inakzeptabel einzustufen (s. Tab. 35). Vier Items weisen mittlere Trennschärfen auf, die übrigen Items korrelieren nur schwach mit der Restskala. Auch hier könnte das Cronbachs Alpha durch Weglassen des Items rfl\_all wesentlich verbessert werden, sodass das Cronbachs Alpha bei .604 läge, was allerdings immer noch als fragwürdig einzuordnen wäre. Allgemein muss überlegt werden, ob die Bezeichnung der logopädischen Störungsbilder als Auffälligkeiten «der Aussprache», «des Wortschatzes» usw. den Gegenstand ungebührlich vereinfacht. Logopäd:innen operieren eher mit Begriffen, die diagnostizierte Störungsbilder bezeichnen, wie z.B. «semantisch-lexikalische» oder «morpho-syntaktische» Störungen. Es gilt zu überlegen, ob eine Anpassung der Itemformulierungen in der Subgruppe der Logopäd:innen die Reliabilität erhöhen könnte.

Tabelle 35: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Logopäd:innen

<b>Logopäd:innen</b>			
Item	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rfl_all	2.85 (1.25)	.058	.604
aus_all	4.77 (0.47)	.210	.433
ws_all	4.85 (0.39)	.441	.375
gra_all	4.79 (0.44)	.424	.368
sv_all	4.77 (0.56)	.398	.353
kom_all	3.76 (1.13)	.303	.371

### 6.3 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .966 als exzellent einzustufen (s. Tab. 36). Die meisten Items weisen durchweg hohe Trennschärfen auf. Dennoch zeigt sich auch hier noch einmal, dass sich das Cronbachs Alpha verbessern würde, wenn man das Item rfl\_kl («des Redeflusses») weglässt.

*Tabelle 36: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten" für die Gesamtstichprobe*

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rfl_kl	2.29 (1.50)	.783	.970
aus_kl	2.60 (1.63)	.892	.960
ws_kl	2.84 (1.67)	.930	.955
gra_kl	2.61 (1.59)	.909	.958
sv_kl	2.86 (1.68)	.942	.954
kom_kl	2.60 (1.60)	.890	.960

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .882 (gut)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .922 (exzellent)
- Cronbachs Alpha für Logopäd:innen: .924 (exzellent)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen, außer beim Item rfl\_kl bei den Logopäd:innen, welches mit .460 aber immer noch als «mittel» einzustufen wäre. Dennoch bestätigt sich in allen berufsgruppenspezifischen Analysen noch einmal, dass das Cronbachs Alpha für diese Subskala erhöht werden kann, wenn auf das Item rfl\_kl verzichtet wird.

### 6.4 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Unterrichtsformen»

Da diese Subskala überdurchschnittlich viele Missings aufweist, werden die fehlenden Werte durch das EM-Verfahren geschätzt.

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .894 als gut zu bewerten (s. Tab. 37, S. 29). Alle Items weisen hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines der Items nicht verbessert werden.

Tabelle 37: Schwierigkeiten, Trennschärpen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Unterrichtsformen" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
uf_gru	2.91 (1.52)	.813	.845
uf_ein	3.16 (1.67)	.741	.874
uf_off	2.78 (1.57)	.795	.852
uf_fron	2.30 (1.40)	.721	.879

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Die berufsgruppenspezifischen Reliabilitätsanalysen belegen die Reliabilität für die Regellehrpersonen und die Logopäd:innen

- Cronbachs Alpha für Regellehrpersonen: .705 (akzeptabel)
- Cronbachs Alpha für Logopäd:innen: .803 (gut)

Auch die Itemtrennschärpen sind bei allen Items für die Regellehrpersonen und die Logopäd:innen als mittel bis hoch einzuschätzen. Eine Verbesserung des Cronbachs Alpha kann für diese Teilstichproben nicht erreicht werden.

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Schulischen Heilpädagog:innen ist mit .525 als schlecht einzustufen (s. Tab. 38). Das Item uf\_ein («Einzelarbeiten») weist eine schwache Trennschärfe auf, die restlichen Items korrelieren mittelmäßig mit der Restskala. Das Cronbachs Alpha könnte durch Weglassen des Items uf\_ein verbessert werden.

Tabelle 38: Schwierigkeiten, Trennschärpen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Unterrichtsformen" für die Schulischen Heilpädagog:innen

Item	Schulische Heilpädagog:innen		
	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
uf_gru	4.22 (0.43)	.344	.425
uf_ein	4.17 (0.98)	.209	.546
uf_off	4.12 (0.97)	.313	.453
uf_fron	2.83 (0.83)	.409	.380

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Schon bei der Analyse der fehlenden Werte hat sich gezeigt, dass die große Anzahl Missings beim Item uf\_fron («Frontalunterricht») in einem bedeutenden Zusammenhang mit den Items uf\_gru

(«Gruppenarbeiten») und uf\_ein steht. Diese Subskala scheint zumindest in der Subgruppe der Schulischen Heilpädagog:innen problembehaftet zu sein. Es gilt zu überlegen, ob hier allenfalls Hilfestellungen, Begriffsklärungen und ähnliches zu einer verbesserten Reliabilität führen könnten. Alles in allem sind die Reliabilitätswerte für diese Subskala doch als zufriedenstellend zu betrachten.

## 6.5 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Fachbereichen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .924 als exzellent einzustufen (s. Tab. 39). Die meisten Items weisen durchweg hohe Trennschärfen auf. Nur das Item fb\_ber («Berufliche Orientierung») hat eine als mittelmäßig einzustufende Trennschärfe. Dieses Item wurde wahrscheinlich nur in den seltensten Fällen ausgewählt, da es lediglich im Zyklus 3 der Schweizer Volksschule vorkommt, welcher in dieser Befragung nicht so gut repräsentiert ist. Zwar könnte das Cronbachs Alpha durch Weglassen dieses Items etwas verbessert werden, aber durch die vorliegende Auswahl von Items werden sämtliche Fachbereiche des Lehrplans 21 abgedeckt. Eine Eliminierung dieses Items würde eine künstliche Reduktion der Antwortoptionen bedeuten.

*Tabelle 39: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Fachbereichen" für die Gesamtstichprobe*

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
fb_deu	2.54 (1.67)	.900	.904
fb_frem	1.65 (1.14)	.571	.925
fb_mat	2.04 (1.58)	.900	.903
fb_nmg	2.00 (1.45)	.876	.905
fb_gest	1.76 (1.32)	.792	.911
fb_mus	1.48 (1.03)	.678	.919
fb_spo	1.57 (1.13)	.742	.915
fb_inf	1.43 (1.07)	.603	.923
fb_ber	1.30 (0.76)	.497	.928

Auf eine berufsgruppenspezifische Analyse wird hier verzichtet, da bei einer großen Anzahl von «trifft nicht zu»-Antworten die Stichprobengröße in den Berufsgruppen auf ein Minimum reduziert würde (Regellehrpersonen: n = 5; Schulische Heilpädagog:innen: n = 7). Lediglich für die Logopäd:innen lässt sich bestätigen, dass die Subskala mit einem Cronbachs Alpha von .898 und durchweg guten Itemtrennschärfen nach wie vor gute Reliabilitätswerte aufweist.

## 6.6 Reliabilität der Subskala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Erwerbsphasen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .943 als exzellent einzustufen (s. Tab. 40). Die Items weisen durchweg hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

*Tabelle 40: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Erwerbsphasen" für die Gesamtstichprobe*

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
pha_ein	2.58 (1.47)	.849	.942
pha_ueb	3.08 (1.64)	.928	.879
pha_anw)	3.10 (1.60)	.872	.923

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .795 (akzeptabel)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .861 (gut)
- Cronbachs Alpha für Logopäd:innen: .910 (exzellent)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Das Cronbachs Alpha könnte durch das Weglassen eines Items nur minimal verbessert werden.

## 6.7 Reliabilität der Subskala «sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers»

Da diese Subskala überdurchschnittlich viele Missings aufweist, werden die fehlenden Werte durch das EM-Verfahren geschätzt.

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .270 als schlecht zu bewerten (s. Tab. 41, S. 32). Alle Items sind unzureichend trennscharf. Es fällt ein negativer Cronbachs Alpha auf, der als Nullwert interpretiert werden darf (Bühner, 2011, S. 168). Ein negativer Cronbachs Alpha kann auf ein falsch gepoltes Item hindeuten, was hier ausgeschlossen werden kann. Von daher ist eher davon auszugehen, dass in dieser Subskala inhaltlich Verschiedenes erfasst wird. Das Item rau\_aus («außerhalb der Schule») weist eine negative Itemtrennschärfe auf, was bedeutet, dass dieses Item negativ mit dem übrigen Summenwert dieser Subskala korreliert. Personen, die hier einen hohen Wert angeben, erreichen im Rest der Skala tiefere Werte und umgekehrt. Da die zugrunde liegende Stichprobe mit Teilnehmenden aus drei unterschiedlichen Berufsgruppen in dieser Hinsicht sehr heterogen ist und auch sein darf, überrascht dies nicht unbedingt. Eine Reliabilitätsanalyse der Subskala mit den Teilstichproben kann aufschlussreich sein.

Tabelle 41: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rau_gan	2.40 (1.38)	.235	.103
rau_gru	2.71 (1.52)	.298	-.002
rau_sep	3.13 (1.73)	.102	.268
rau_aus	2.05 (1.54)	-.045	.427

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Regellehrpersonen ist mit .628 als fragwürdig einzustufen (s. Tab. 42). Alle Items weisen mittlere Trennschärfen auf. Das Cronbachs Alpha kann durch das Weglassen eines Items nicht verbessert werden.

Tabelle 42: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Regellehrpersonen

Item	Regellehrpersonen		
	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rau_gan	2.86 (1.44)	.369	.586
rau_gru	3.45 (1.40)	.468	.513
rau_sep	2.32 (1.52)	.412	.556
rau_aus	1.95 (1.29)	.385	.574

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Schulischen Heilpädagog:innen ist mit .308 als inakzeptabel einzustufen (s. Tab. 43, S. 33). Alle Items weisen schwache Trennschärfen auf. Das Cronbachs Alpha könnte durch Weglassen des Items rau\_sep («separates Arbeitszimmer») leicht, aber nicht wesentlich verbessert werden.



Tabelle 43: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Schulische Heilpädagog:innen

<b>Schulische Heilpädagog:innen</b>			
Item	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rau_gan	3.18 (1.41)	.215	.178
rau_gru	3.83 (1.39)	.143	.268
rau_sep	3.27 (1.62)	.068	.384
rau_aus	1.53 (1.03)	.253	.175

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Logopäd:innen ist mit .206 ebenfalls als inakzeptabel einzustufen (s. Tab. 44). Ein Item weist eine mittlere Trennschärfe auf, die übrigen Items korrelieren nur schwach mit der Restskala. Das Item rau\_aus weist eine negative Itemtrennschärfe auf, was bedeutet, dass dieses Item negativ mit dem übrigen Summenwert dieser Subskala korreliert. Personen, die hier einen hohen Wert angeben, erreichen im Rest der Skala tiefere Werte und umgekehrt. Das Cronbachs Alpha könnte durch Weglassen dieses Items wesentlich verbessert werden, läge aber immer noch im inakzeptablen Bereich.

Tabelle 44: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Logopäd:innen

<b>Logopäd:innen</b>			
Item	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
rau_gan	1.91 (1.12)	.124	.142
rau_gru	1.96 (1.16)	.394	-.173
rau_sep	3.37 (1.76)	.131	.110
rau_aus	2.30 (1.74)	-.111	.483

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Schon im Kontext der Missing-Analyse wurde deutlich, dass die verschiedenen Items dieser Subskala nicht eindeutig interpretiert werden konnten und es bei der Antwortanalyse zu komplexen und nur schwer interpretierbaren bedeutenden Zusammenhängen kommt. Es besteht der Verdacht, dass mit den hier beschriebenen Räumlichkeiten keine realen örtlichen Bedingungen verknüpft werden konnten. Oder, und das ist wesentlich grundlegender, der Kontrast aus diversen Fragen zur «unterrichtsintegrierten Sprachunterstützung» rührt eher an prinzipielle Einstellungen und Haltungen und wird mit dem Einbezug einfacher räumlicher Situationen nur unzureichend kontrastiert,

weswegen eine gewisse Unsicherheit beim Ausfüllen des Fragebogens an dieser Stelle erzeugt wurde. Diese Skala wird in dieser Form künftig nicht mehr eingesetzt werden.

### 6.8 Reliabilität der Subskala «Austausch mit den Regellehrpersonen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Schulischen Heilpädagog:innen und Logopäd:innen ist mit .784 als akzeptabel einzustufen (s. Tab. 45). Die Items weisen durchweg mittlere bis hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen des Items aus1\_lp («von Materialien») deutlich verbessert werden. aus1 ist hier (wie auch in den anderen Subskalen zum Austausch immer) das schwierigste der vier Items und korreliert deswegen weniger gut mit dem Rest der Skala. Dies gilt auch für die Reliabilitätsanalyse der Subgruppen.

Tabelle 45: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Regellehrpersonen" für die Schulische Heilpädagog:innen und Logopäd:innen

<b>Schulische Heilpädagog:innen &amp; Logopäd:innen</b>			
<b>Item</b>	<b>Itemschwierigkeit (SD)</b>	<b>Korrigierte Itemtrennschärfe</b>	<b>Cronbachs Alpha, wenn Item weggelassen</b>
aus1_lp	2.75 (0.99)	.427	.814
aus2_lp	3.63 (0.97)	.614	.719
aus3_lp	3.41 (0.95)	.735	.654
aus4_lp	3.32 (0.90)	.607	.724

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Schulischen Heilpädagog:innen ist mit .555 als schlecht einzustufen (s. Tab. 46). Zwei Items weisen schwache Trennschärfen auf, zwei mittelmäßige. Das Cronbachs Alpha könnte durch Weglassen des Items aus1\_lp leicht, aber nicht wesentlich verbessert werden.

Tabelle 46: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Regellehrpersonen" für die Schulische Heilpädagog:innen

<b>Schulische Heilpädagog:innen</b>			
<b>Item</b>	<b>Itemschwierigkeit (SD)</b>	<b>Korrigierte Itemtrennschärfe</b>	<b>Cronbachs Alpha, wenn Item weggelassen</b>
aus1_lp	3.32 (0.97)	.272	.557
aus2_lp	4.21 (0.68)	.285	.526
aus3_lp	3.71 (0.82)	.527	.319
aus4_lp	3.56 (0.82)	.310	.508

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Logopäd:innen ist mit .801 als gut einzustufen (s. Tab. 47, S. 35). Ein Item weist eine mittlere Trennschärfe auf, die übrigen Items

korrelieren hoch mit der Restskala. Durch Weglassen des Items aus1\_lp könnte das Cronbachs Alpha verbessert werden.

*Tabelle 47: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Regellehrpersonen" für die Logopäd:innen*

Item	Logopäd:innen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
aus1_lp	2.51 (0.90)	.390	.850
aus2_lp	3.39 (0.97)	.629	.744
aus3_lp	3.29 (0.97)	.785	.661
aus4_lp	3.22 (0.92)	.679	.720

Für die Subskala «Austausch mit den Regellehrpersonen» scheint es sich zu lohnen, das Item aus1\_lp zu streichen oder umzuformulieren, um die Reliabilität der Subskala zu steigern. Ob dies auch für die anderen Austausch-Skalen so zeigt, wird im Folgenden untersucht.

### 6.9 Reliabilität der Subskala «Austausch mit den Schulischen Heilpädagog:innen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Regellehrpersonen und den Logopäd:innen ist mit .871 als gut einzustufen (s. Tab. 48). Die Items weisen durchweg hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

*Tabelle 48: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Schulischen Heilpädagog:innen" für die Regellehrpersonen und Logopäd:innen*

Item	Regellehrpersonen & Logopäd:innen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
aus1_sh	2.74 (0.97)	.636	.870
aus2_sh	3.30 (0.95)	.732	.832
aus3_sh	3.15 (0.97)	.782	.811
aus4_sh	3.22 (0.92)	.752	.824

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .871 (gut)
- Cronbachs Alpha für Logopäd:innen: .926 (exzellent)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Das Cronbachs Alpha könnte durch das Weglassen keines der Items verbessert werden.

### 6.10 Reliabilität der Subskala «Austausch mit den Logopäd:innen»

Da diese Subskala überdurchschnittlich viele Missings aufweist, werden die fehlenden Werte durch das EM-Verfahren geschätzt. Die hier angegebenen Itemschwierigkeiten unterscheiden sich daher von denen in Tabelle 11.

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Regellehrpersonen und den Schulischen Heilpädagog:innen ist mit .936 als exzellent einzustufen (s. Tab. 49). Die Items weisen durchweg hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

*Tabelle 49: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Logopäd:innen" für die Regellehrpersonen und Schulischen Heilpädagog:innen*

Item	Regellehrpersonen & Schulische Heilpädagog:innen		
	Itemschwierigkeit <sup>a</sup> (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
aus1_lo	2.07 (1.01)	.793	.934
aus2_lo	2.38 (1.00)	.886	.904
aus3_lo	2.28 (0.99)	.897	.900
aus4_lo	2.36 (0.94)	.819	.925

<sup>a</sup> Missings wurden im EM-Verfahren geschätzt

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .929 (exzellent)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .942 (exzellent)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Das Cronbachs Alpha könnte durch das Weglassen eines der Items bei den Regellehrpersonen (aus1\_lp) leicht verbessert werden. Da dies für die anderen Austauschskalen jedoch nicht zutrifft, wird dieses Item beibehalten.

### 6.11 Reliabilität der Subskala «Synchronisation mit den Regellehrpersonen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Schulischen Heilpädagog:innen und Logopäd:innen ist mit .890 als gut einzustufen (s. Tab. 50, S. 37). Die Items weisen durchweg hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

Tabelle 50: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Synchronisation mit den Regellehrpersonen" für die Schulische Heilpädagog:innen und Logopäd:innen

<b>Schulische Heilpädagog:innen &amp; Logopäd:innen</b>			
<b>Item</b>	<b>Itemschwierigkeit (SD)</b>	<b>Korrigierte Itemtrennschärfe</b>	<b>Cronbachs Alpha, wenn Item weggelassen</b>
syn1_lp	3.09 (0.90)	.747	.866
syn2_lp	2.86 (0.99)	.808	.841
syn3_lp	2.40 (1.26)	.760	.875
syn4_lp	2.73 (0.89)	.780	.856

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Schulische Heilpädagog:innen: .784 (akzeptabel)
- Cronbachs Alpha für Logopäd:innen: .848 (gut)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Das Cronbachs Alpha könnte durch das Weglassen keines der Items verbessert werden.

## 6.12 Reliabilität der Subskala «Synchronisation mit den Schulischen Heilpädagog:innen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Regellehrpersonen und Logopäd:innen ist mit .872 als gut einzustufen (s. Tab. 51). Die Items weisen durchweg hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

Tabelle 51: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Synchronisation mit den Schulischen Heilpädagog:innen" für die Regellehrpersonen und Logopäd:innen

<b>Regellehrpersonen &amp; Logopäd:innen</b>			
<b>Item</b>	<b>Itemschwierigkeit (SD)</b>	<b>Korrigierte Itemtrennschärfe</b>	<b>Cronbachs Alpha, wenn Item weggelassen</b>
syn1_sh	2.92 (0.87)	.698	.849
syn2_sh	2.65 (1.00)	.802	.805
syn3_sh	2.18 (1.12)	.687	.861
syn4_sh	2.59 (0.88)	.750	.830

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .943 (exzellent)
- Cronbachs Alpha für Logopäd:innen: .821 (gut)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Das Cronbachs Alpha könnte durch das Weglassen keines der Items verbessert werden.

### 6.13 Reliabilität der Subskala «Synchronisation mit den Logopäd:innen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Regellehrpersonen und Schulischen Heilpädagog:innen ist mit .909 als exzellent einzustufen (s. Tab. 52). Die Items weisen durchweg hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen des Items syn3\_lo («Termine von Aufträgen») leicht verbessert werden.

*Tabelle 52: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Synchronisation mit den Logopäd:innen" für die Regellehrpersonen und Schulischen Heilpädagog:innen*

Item	Regellehrpersonen & Schulische Heilpädagog:innen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
syn1_lo	2.28 (0.97)	.839	.866
syn2_lo	2.05 (1.02)	.861	.858
syn3_lo	1.70 (0.96)	.683	.920
syn4_lo	2.11 (0.90)	.805	.880

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .908 (exzellent)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .912 (exzellent)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Wie auch bei der übergeordneten Stichprobe ließe sich das Cronbachs Alpha durch das Weglassen eines der Items (syn3) bei den Logopäd:innen leicht verbessern. Da dies für die anderen Synchronisationsskalen jedoch nicht zutrifft, kommt ein Weglassen des genannten Items nicht in Frage.

### 6.14 Reliabilität der Subskala «Ko-Konstruktion mit den Regellehrpersonen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Schulischen Heilpädagog:innen und Logopäd:innen ist mit .747 als akzeptabel einzustufen (s. Tab. 53, S. 39). Ein Item hat eine mittlere Trennschärfe, vier Items eine hohe und zwei Items sogar ungenügende Trennschärfen (kok2\_lp «Durchführung von Elterngesprächen» und kok3\_lp «Besuch von Weiterbildungen»). Das Weglassen dieser beiden Items würde zu einer Steigerung des Cronbachs Alpha führen.

Tabelle 53: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Ko-Konstruktion mit den Regellehrpersonen" für die Schulische Heilpädagog:innen und Logopäd:innen

<b>Schulische Heilpädagog:innen &amp; Logopäd:innen</b>			
<b>Item</b>	Itemschwierigkeit (SD)	Korrigierte	Cronbachs Alpha, wenn
		Itemtrennschärfe	Item weggelassen
kok1_lp	1.72 (0.83)	.312	.750
kok2_lp	2.42 (0.68)	.162	.771
kok3_lp	1.87 (0.51)	.272	.751
kok4_lp	1.91 (0.97)	.625	.675
kok5_lp	1.72 (0.85)	.638	.675
kok6_lp	1.55 (0.94)	.559	.693
kok7_lp	1.81 (0.87)	.652	.670

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Schulischen Heilpädagog:innen ist mit .635 als fragwürdig einzustufen (s. Tab. 54). Vier Items weisen mittlere bis hohe Trennschärfen auf, drei Items schwache (kok1\_lp «Durchführung von Hospitationen», kok2\_lp und kok3\_lp). Das Cronbachs Alpha könnte durch Weglassen dieser Items leicht, aber nicht wesentlich verbessert werden.

Tabelle 54: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Ko-Konstruktion mit den Regellehrpersonen" für die Schulische Heilpädagog:innen

<b>Schulische Heilpädagog:innen</b>			
<b>Item</b>	Itemschwierigkeit (SD)	Korrigierte	Cronbachs Alpha, wenn
		Itemtrennschärfe	Item weggelassen
kok1_lp	2.03 (1.13)	.266	.636
kok2_lp	2.38 (0.64)	.184	.639
kok3_lp	1.95 (0.39)	.055	.655
kok4_lp	2.78 (0.90)	.484	.553
kok5_lp	2.37 (0.88)	.478	.557
kok6_lp	2.42 (1.18)	.410	.580
kok7_lp	2.48 (0.89)	.524	.540

Das Cronbachs Alpha für diese Subskala gemessen mit Stichprobe der Logopäd:innen ist mit .659 als fragwürdig einzustufen (s. Tab. 55, S. 40). Drei Items weisen mittlere und zwei Items hohe

Trennschärfen auf, zwei Items schwache (kok1\_lp und kok2\_lp). Das Cronbachs Alpha könnte durch Weglassen dieser Items leicht, aber nicht wesentlich verbessert werden.

*Tabelle 55: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Ko-Konstruktion mit den Regellehrpersonen" für die Logopäd:innen*

Item	Logopäd:innen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
kok1_lp	1.59 (0.63)	.171	.677
kok2_lp	2.43 (0.69)	.273	.653
kok3_lp	1.84 (0.55)	.382	.621
kok4_lp	1.56 (0.75)	.428	.604
kok5_lp	1.46 (0.68)	.514	.576
kok6_lp	1.20 (0.49)	.328	.636
kok7_lp	1.53 (0.70)	.505	.578

### 6.15 Reliabilität der Subskala «Ko-Konstruktion mit den Schulischen Heilpädagog:innen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Regellehrpersonen und Logopäd:innen ist mit .786 als akzeptabel einzustufen (s. Tab. 56). Drei Items weisen mittlere und vier Items hohe Trennschärfen auf. Wiederum handelt es sich um die ersten drei Items auf der Skala, wie bei der vorherigen Untersuchung auch schon. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

*Tabelle 56: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Ko-Konstruktion mit den Schulischen Heilpädagog:innen" für die Regellehrpersonen und Logopäd:innen*

Item	Regellehrpersonen & Logopäd:innen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
kok1_sh	1.50 (0.62)	.374	.782
kok2_sh	2.32 (0.74)	.366	.786
kok3_sh	1.88 (0.60)	.414	.776
kok4_sh	1.85 (0.89)	.658	.727
kok5_sh	1.69 (0.83)	.679	.722
kok6_sh	1.36 (0.68)	.477	.765
kok7_sh	1.72 (0.80)	.613	.738



Dieses Bild bestätigt sich auch in den berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .830 (gut)
- Cronbachs Alpha für Logopäd:innen: .774 (akzeptabel)

In beiden Subgruppen weisen wiederum die ersten drei Items «lediglich» mittlere bzw. schlechte Trennschärfe auf. In der Subgruppe der Regellehrpersonen ließe sich das Cronbachs Alpha verbessern, wenn man die Items kok1\_sh («Durchführung von Hospitationen») und kok3\_sh («Besuch von Weiterbildungen») wegließe.

### 6.16 Reliabilität der Subskala «Ko-Konstruktion mit den Logopäd:innen»

Das Cronbachs Alpha für diese Subskala gemessen mit der Stichprobe der Regellehrpersonen und Schulischen Heilpädagog:innen ist mit .791 als akzeptabel einzustufen (s. Tab. 57). Vier Items weisen mittlere und drei Items hohe Trennschärfe auf. Das Cronbachs Alpha kann durch Weglassen des Items kok1\_lo («Durchführung von Hospitationen») leicht verbessert werden.

*Tabelle 57: Schwierigkeiten, Trennschärfe und Cronbachs Alpha der Items der Subskala "Ko-Konstruktion mit den Logopäd:innen" für die Regellehrpersonen und Schulische Heilpädagog:innen*

Item	Regellehrpersonen & Schulische Heilpädagog:innen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
kok1_lo	1.18 (0.51)	.328	.794
kok2_lo	1.48 (0.57)	.492	.769
kok3_lo	1.56 (0.52)	.370	.788
kok4_lo	1.60 (0.79)	.694	.725
kok5_lo	1.38 (0.66)	.691	.727
kok6_lo	1.09 (0.32)	.467	.782
kok7_lo	1.51 (0.79)	.648	.737

Dieses Bild bestätigt sich wiederum in den berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .817 (gut)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .773 (akzeptabel)

In der Subgruppe der Regellehrpersonen ließe sich das Cronbachs Alpha verbessern, wenn man die Items kok1\_lo und kok3\_lo wegließe. In der Gruppe der Schulischen Heilpädagog:innen könnte man das Item kok1\_lo weglassen, um ein leicht höheres Cronbachs Alpha zu erhalten.

Die Items kok2 und kok3 sind in der übergeordneten Analyse wiederholt mit ungenügenden Trennschärfen aufgefallen. Gemäß den Empfehlungen von Blanz (2021) werden diese also künftig weggelassen. So kann auch das Cronbachs Alpha in allen Subgruppen verbessert werden.

### 6.17 Reliabilität der Subskala «Team-Teaching»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .862 als gut einzustufen (s. Tab. 58). Alle Items weisen hohe Trennschärfen auf. Zwar könnte das Cronbachs Alpha durch Weglassen des Items tt\_ein («Unterricht und Einzelunterstützung») etwas verbessert werden, aber dann würde eine wesentliche Team-Teaching-Form fehlen. Eine Eliminierung dieses Items würde ein unvollständiges Bild hinterlassen.

*Tabelle 58: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Team-Teaching" für die Gesamtstichprobe*

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
tt_ein	2.92 (1.63)	.547	.873
tt_sta	2.14 (1.41)	.704	.827
tt_par	1.99 (1.38)	.762	.813
tt_foe	2.02 (1.43)	.703	.827
tt_int	1.77 (1.26)	.726	.825

In der Gruppe der Regellehrpersonen bestätigt sich das an der Gesamtstichprobe ermittelte Ergebnis (Cronbachs Alpha: .850). Jedoch fallen die Werte für die Schulischen Heilpädagog:innen (Cronbachs Alpha: .540) und für die Logopäd:innen (Cronbachs Alpha: .659) niedriger aus. In beiden Gruppen könnte das Cronbachs Alpha gesteigert werden, wenn das Item tt\_ein weggelassen würde – bei den Logopäd:innen sogar auf .800. Die Vermutung liegt nahe, dass «Unterricht und Einzelunterstützung» nicht als Team-Teaching-Form im engeren Sinne verstanden wird. Besonders bei den Schulischen Heilpädagog:innen zeichnet sich ab, dass vor allem das Item tt\_int («Interaktives Unterrichten») als Team-Teaching aufgefasst wird und die anderen Formen nicht dazugehören würden (s. Tab. 59, S. 43).

Tabelle 59: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Team-Teaching" für die Schulischen Heilpädagog:innen

Item	Schulische Heilpädagog:innen		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
tt_ein	4.15 (0.89)	.087	.578
tt_sta	3.56 (1.34)	.273	.503
tt_par	3.16 (1.38)	.351	.454
tt_foe	3.46 (1.36)	.284	.497
tt_int	2.79 (1.54)	.513	.326

Wie hier deutlich wird, besteht die Inkonsistenz dieser Skala darin, dass «Unterricht und Einzelunterstützung» am häufigsten durchgeführt wird (auch bei den Logopäd:innen), die anderen Team-Teaching-Formen werden mit einer geringeren Häufigkeit angegeben. Die Interpretation der Reliabilitätswerte geht nun dahin, dass «Unterricht und Einzelunterstützung» nicht als vollwertige Team-Teaching-Form anerkannt wird, aber dass diese Form deutlich einfacher umzusetzen ist und deswegen häufiger praktiziert wird. Infolgedessen korreliert dieses Item weniger gut mit dem Rest der Skala. Ein Weglassen dieses Items kann inhaltlich jedoch nicht gerechtfertigt werden. Allenfalls könnte in einem ergänzenden Text zu dieser Skala ausgeführt werden, was «Unterricht und Einzelunterstützung» alles umfasst (eine didaktische Begründung, ein geplantes Vorgehen usw.), anstatt dass damit einfach nur das unabhängige, nicht miteinander abgesprochene «assistieren» einzelner Kinder gemeint ist.

### 6.18 Reliabilität der Subskala «Intention»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .819 als gut einzustufen (s. Tab. 60). Fast alle Items weisen hohe Trennschärfen auf, nur eines hat eine mittlere Trennschärfe. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

Tabelle 60: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Intention" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
int1	2.20 (1.08)	.627	.781
int2	2.70 (0.93)	.499	.815
int3	2.18 (0.90)	.689	.762
int4	2.33 (0.89)	.651	.773
int5	1.91 (0.90)	.606	.785

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .842 (gut)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .875 (gut)
- Cronbachs Alpha für Logopäd:innen: .736 (akzeptabel)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Lediglich bei den Logopäd:innen gibt es drei Items mit «nur» mittlerer Trennschärfe. Das Cronbachs Alpha könnte durch das Weglassen eines Items jedoch in keiner Subgruppe verbessert werden.

### 6.19 Reliabilität der Subskala «Interesse»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .885 als gut einzustufen (s. Tab. 61). Alle Items weisen hohe Trennschärfen auf, das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

*Tabelle 61: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Interesse" für die Gesamtstichprobe*

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
intr1	3.82 (0.42)	.553	.885
intr2	3.47 (0.67)	.767	.857
intr3	3.49 (0.64)	.741	.861
intr4	3.58 (0.59)	.765	.859
intr5)	3.27 (0.78)	.624	.878
intr6	3.62 (0.58)	.804	.856
intr7	3.31 (0.89)	.626	.884

Auch die berufsgruppenspezifischen Reliabilitätsanalysen dieser Subskala weisen durchweg gute Ergebnisse auf, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .872 (gut)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .862 (gut)
- Cronbachs Alpha für Logopäd:innen: .892 (gut)

Auch die Itemtrennschärfen sind bei allen Items und in allen Teilstichproben als hoch einzuschätzen. Lediglich bei den Schulischen Heilpädagog:innen gibt es ein Item mit «nur» mittlerer Trennschärfe (intr1 «ist sinnvoll»). Das Cronbachs Alpha könnte durch das Weglassen dieses Items in dieser Subgruppe leicht verbessert werden.

## 6.20 Reliabilität der Subskala «Kosten»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .821 als gut einzustufen (s. Tab. 62). Fünf Items weisen hohe Trennschärfen auf, drei Items mittlere. Das Cronbachs Alpha könnte durch Weglassen der Items kost6 («lohnt sich mittel- und langfristig nicht») und kost8 («ist mit schlechten Erfahrungen verknüpft») leicht verbessert werden.

Tabelle 62: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Kosten" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
kost1	1.96 (0.78)	.712	.776
kost2	1.87 (0.75)	.696	.780
kost3	1.40 (0.63)	.498	.807
kost4	2.23 (1.00)	.638	.786
kost5	1.78 (0.90)	.505	.807
kost6	1.40 (0.69)	.349	.824
kost7	1.62 (0.74)	.660	.785
kost8	1.57 (0.82)	.329	.829

Dieses Bild zeigt sich ebenso in den berufsgruppenspezifischen Analysen, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .804 (gut)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .806 (gut)
- Cronbachs Alpha für Logopäd:innen: .831 (gut)

In der Gruppe der Regellehrpersonen weist ein Item eine ungenügende Trennschärfe auf (kost8). Das Cronbachs Alpha könnte durch Weglassen dieses Items deutlich verbessert werden (.840). In der Gruppe der Schulischen Heilpädagog:innen ist einzig das Item kost6 mit einer ungenügenden Trennschärfe verknüpft. Auch hier ließe sich das Cronbachs Alpha leicht verbessern, ließe man dieses Item weg. In der Gruppe der Logopäd:innen schließlich gibt es kein Item mit ungenügender Trennschärfe, allerdings ließe sich auch hier das Cronbachs Alpha leicht steigern, wenn man das Item kost6 wegließe.

## 6.21 Reliabilität der Subskala «Nützlichkeit für die professionelle Entwicklung»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .908 als exzellent einzustufen (s. Tab. 63, S. 46). Alle Items weisen hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

Tabelle 63: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala " Nützlichkeit für die professionelle Entwicklung" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
nue_fk1	3.26 (0.70)	.644	.900
nue_fk2	2.82 (0.92)	.586	.905
nue_fk3	3.29 (0.71)	.662	.899
nue_fk4	3.42 (0.73)	.670	.899
nue_fk5	2.78 (0.89)	.623	.902
nue_fk6	3.19 (0.75)	.660	.899
nue_fk7	2.91 (0.84)	.677	.898
nue_fk8	3.31 (0.65)	.621	.902
nue_fk9	3.10 (0.80)	.707	.897
nue_fk10	3.01 (0.83)	.629	.901
nue_fk11	3.26 (0.69)	.769	.894

Dieses Bild zeigt sich auch in den berufsgruppenspezifischen Analysen, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .925 (exzellent)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .890 (gut)
- Cronbachs Alpha für Logopäd:innen: .891 (gut)

In der Gruppe der Schulischen Heilpädagog:innen weist ein Item eine «nur» mittlere Trennschärfe auf (nue\_fk10 «ermöglicht gemeinsame Problemlösungen»). Das Cronbachs Alpha könnte durch Weglassen dieses Items jedoch nicht verbessert werden. Alle übrigen Itemtrennschärfen in allen Berufsgruppen sind hoch.

## 6.22 Reliabilität der Subskala «Nützlichkeit für die Schüler und Schülerinnen sowie die Schule»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .922 als exzellent einzustufen (s. Tab. 64, S. 47). Alle Items weisen hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

Tabelle 64: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala " Nützlichkeit für die Schüler und Schülerinnen sowie die Schule" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
nue_sus1	3.45 (0.67)	.740	.912
nue_sus2	3.35 (0.65)	.743	.912
nue_sus3	3.27 (0.76)	.590	.920
nue_sus4	3.36 (0.68)	.751	.912
nue_sus5	3.33 (0.71)	.733	.913
nue_sus6	3.25 (0.67)	.745	.912
nue_sus7	3.37 (0.66)	.761	.911
nue_sus8	2.85 (0.87)	.651	.918
nue_sus9	3.57 (0.63)	.654	.916
nue_sus10	3.57 (0.59)	.668	.916
nue_sus11	3.47 (0.67)	.603	.919

Auch in den berufsgruppenspezifischen Analysen bestätigt sich die hohe Reliabilität der Skala:

- Cronbachs Alpha für Regellehrpersonen: .885 (gut)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .947 (exzellent)
- Cronbachs Alpha für Logopäd:innen: .914 (exzellent)

Die Items weisen in allen Gruppen zumeist hohe, teils auch mittlere Trennschärfen auf. In der Gruppe der Regellehrpersonen ließe sich das Cronbachs Alpha durch Weglassen des Items nue\_sus11 («stärkt das Zusammengehörigkeitsgefühl») minimal verbessern. Auch in der Gruppe der Schulischen Heilpädagog:innen ließe sich eine minimale Verbesserung des Cronbachs Alpha durch das Weglassen der Items nue\_sus3 («schafft ein besseres Arbeitsklima») und nue\_sus8 («führt zur Entwicklung gemeinsamer Materialien») bewirken. In der Gruppe der Logopäd:innen sind keine Verbesserungen möglich.

### 6.23 Reliabilität der Subskala «prozessorientierte Erfolgserwartung»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .755 als akzeptabel einzustufen (s. Tab. 65, S. 48). Alle Items weisen hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

Tabelle 65: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "prozessorientierte Erfolgserwartung" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
erf_pro1	3.70 (0.47)	.562	.699
erf_pro2	3.52 (0.53)	.628	.620
erf_pro2	3.55 (0.51)	.567	.692

Auch in den berufsgruppenspezifischen Analysen bestätigt sich dieses Ergebnis der Reliabilitätsanalyse:

- Cronbachs Alpha für Regellehrpersonen: .717 (akzeptabel)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .754 (akzeptabel)
- Cronbachs Alpha für Logopäd:innen: .770 (akzeptabel)

Die Items weisen in allen Gruppen zumeist hohe, teils auch mittlere Trennschärfen auf. In der Gruppe der Schulischen Heilpädagog:innen ließe sich das Cronbachs Alpha durch Weglassen des Items «kann gut mit anderen kooperieren» leicht verbessern. Ansonsten sind keine Verbesserungen des Cronbachs Alpha möglich.

#### 6.24 Reliabilität der Subskala «ergebnisorientierte Erfolgserwartung»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .751 als akzeptabel einzustufen (s. Tab. 66). Die Items weisen mittlere bis hohe Trennschärfen auf. Das Cronbachs Alpha kann durch Weglassen eines Items nicht verbessert werden.

Tabelle 66: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "ergebnisorientierte Erfolgserwartung" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
erf_erg1	3.42 (0.60)	.458	.739
erf_erg2	3.25 (0.64)	.597	.665
erf_erg3	3.01 (0.70)	.500	.726
erf_erg4	3.48 (0.57)	.655	.640

Auch in den berufsgruppenspezifischen Analysen bestätigt sich dieses Ergebnis der Reliabilitätsanalyse:



- Cronbachs Alpha für Regellehrpersonen: .700 (akzeptabel)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .816 (gut)
- Cronbachs Alpha für Logopäd:innen: .747 (akzeptabel)

Die Items weisen in allen Gruppen zumeist hohe, teils auch mittlere Trennschärfen auf. In den Gruppen der Regellehrpersonen und der Schulischen Heilpädagog:innen ließe sich das Cronbachs Alpha durch Weglassen des Items erf\_erg1 («kann anderen bei Problemen helfen») (leicht) verbessern. Ansonsten sind keine Verbesserungen des Cronbachs Alpha möglich.

### 6.25 Reliabilität der Subskala «Zufriedenheit»

Das Cronbachs Alpha für diese Subskala gemessen mit der Gesamtstichprobe ist mit .721 als akzeptabel einzustufen (s. Tab. 67). Zwei Items weisen hohe, fünf Items mittlere und ein Item eine ungenügende Trennschärfe auf. Das Cronbachs Alpha kann durch Weglassen dieses Items zufr8 («Qualifikation zur Zusammenarbeit») verbessert werden.

Tabelle 67: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Zufriedenheit" für die Gesamtstichprobe

Item	Gesamtstichprobe		
	Itemschwierigkeit (SD)	Korrigierte Itemtrennschärfe	Cronbachs Alpha, wenn Item weggelassen
zufr1	3.48 (0.61)	.507	.683
zufr2	3.38 (0.66)	.455	.689
zufr3	3.29 (0.79)	.387	.699
zufr4	3.15 (0.87)	.479	.679
zufr5	2.84 (0.91)	.453	.685
zufr6	3.13 (1.01)	.537	.664
zufr7	3.43 (0.78)	.447	.687
zufr8	3.02 (0.94)	.142	.754

Dieses Bild zeigt sich auch in den berufsgruppenspezifischen Analysen, weswegen hier nur eine Zusammenfassung der Analysen präsentiert wird:

- Cronbachs Alpha für Regellehrpersonen: .801 (gut)
- Cronbachs Alpha für Schulische Heilpädagog:innen: .755 (akzeptabel)
- Cronbachs Alpha für Logopäd:innen: .711 (akzeptabel)

Auch in den Gruppen der Regellehrpersonen und der Logopäd:innen ist das Item zufr8 das einzige mit einer ungenügenden Trennschärfe (die übrigen Items weisen mittlere bis hohe Trennschärfen auf). Das Weglassen dieses Items würde bei den Regellehrpersonen zu einem Cronbachs Alpha von .818

und bei den Logopäd:innen zu einem Cronbachs Alpha von .725 führen. Da das Item auch auf Ebene der Gesamtstichprobe nur ungenügend trennscharf war, ist es ratsam, dieses künftig wegzulassen.

## 6.26 Zusammenfassung der Reliabilitätsanalyse

In der folgenden Zusammenfassung sind noch einmal alle Subskalen, deren Cronbachs Alpha und ein Verweis auf niedrige Itemtrennschärfen aufgeführt. Subskalen mit ungenügenden Cronbachs Alpha-Werten sind fett markiert, Subskalen mit ungeeigneten Itemtrennschärfen sind kursiv verschriftet (s. Tab. 68, S. 51).

Zwei Subskalen («sprachunterstützende Maßnahmen (sM) im Allgemeinen» und «sM in Räumlichkeiten außerhalb des Klassenzimmers») werden in der überarbeiteten Version des Fragebogens nicht mehr eingesetzt. Die Items sind insgesamt nicht brauchbar und können auch nur schwerlich umformuliert werden. Um trotzdem an entsprechende Informationen zu gelangen, werden künftig Mehrfachauswahl-Fragen zu den diversen Maßnahmen und Räumlichkeiten gestellt, ohne deren Häufigkeit auf einer Likert-Skala einschätzen zu müssen. Es wird sich zeigen, ob die so gewonnenen Daten brauchbarer als die aktuell vorliegenden sind.

Zwar weist auch die Skala «sM im Allgemeinen bei Auffälligkeiten» ein ungenügendes Cronbachs Alpha auf, doch kann dieses genügend gesteigert werden, wenn man das Item «des Redeflusses» weglässt. Dieses Item wies in allen Subgruppenanalysen eher ungenügende oder geringe Trennschärfen auf, weswegen das Weglassen des Items gut gerechtfertigt werden kann. Konsequenterweise muss es dann aber auch in der Skala «unterrichtsintegrierte sM bei Auffälligkeiten» weggelassen werden, was dieser Skala aber nicht zum Nachteil gereicht.

Die Subskala «sM in verschiedenen Unterrichtsformen» hat v.a. in der Gruppe der Schulischen Heilpädagog:innen schlechte Reliabilitätswerte. Hier müssen allenfalls Erläuterungen ergänzt werden, damit die Items eindeutig und über alle Gruppen hinweg gleichwertig interpretiert werden. In ähnlicher Weise muss auch in der Subskala «Team-Teaching» das Item «Unterricht und Einzelunterstützung» differenziert erläutert werden, da dieses Item ebenfalls bei den Schulischen Heilpädagog:innen eine ungenügende Trennschärfe aufweist. Es muss deutlich gemacht werden, was diese Team-Teaching-Form umfasst, damit diese nicht fehlinterpretiert wird.

In der Skala «Ko-Konstruktion mit Regellehrpersonen» weisen zwei Items («Durchführung von Elterngesprächen» und «Besuch von Weiterbildungen») ungenügende Trennschärfen auf. Die Eliminierung dieser zwei Items trägt maßgeblich zur Verbesserung des Cronbachs Alpha in dieser Skala bei. Konsequenterweise müssen die genannten Items aber auch in den anderen Subskalen zur Ko-Konstruktion weggelassen werden, was aber auch dort zu einer leichten Verbesserung des Cronbachs Alpha beiträgt oder diesen zumindest nicht verschlechtert.

Zu guter Letzt wird künftig das Item «Qualifikation zur Zusammenarbeit» in der Subskala «Zufriedenheit» weggelassen, da es eine geringe Itemtrennschärfe aufweist und so das Cronbachs Alpha deutlich gesteigert werden kann.

Tabelle 68: Subskalen des Fragebogens, Angaben zur internen Konsistenz und zu den Itemtrennschärfen

Subskala	Cronbachs Alpha für Gesamtstichprobe	Enthält Items mit niedriger Trennschärfe
<b>sprachunterstützende Maßnahmen im Allgemeinen</b>	.476 (inakzeptabel)	ja
<b>sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten</b>	.632 (fragwürdig)	ja
unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten	.966 (exzellent)	nein
unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Unterrichtsformen	.894 (gut)	nein
unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Fachbereichen	.924 (exzellent)	nein
unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Erwerbsphasen	.943 (exzellent)	nein
<b>sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers</b>	.270 (inakzeptabel)	ja
Austausch mit den Regellehrpersonen	.784 (akzeptabel)	nein
Austausch mit den Schulischen Heilpädagog:innen	.871 (gut)	nein
Austausch mit den Logopäd:innen	.936 (exzellent)	nein
Synchronisation mit den Regellehrpersonen	.890 (gut)	nein
Synchronisation mit den Schulischen Heilpädagog:innen	.872 (gut)	nein
Synchronisation mit den Logopäd:innen	.909 (exzellent)	nein
<i>Ko-Konstruktion mit den Regellehrpersonen</i>	.747 (akzeptabel)	ja
Ko-Konstruktion mit den Schulischen Heilpädagog:innen	.786 (akzeptabel)	nein
Ko-Konstruktion mit den Logopäd:innen	.791 (akzeptabel)	nein
Team-Teaching	.862 (gut)	nein
Intention	.819 (gut)	nein
Interesse	.885 (gut)	nein
Kosten	.821 (gut)	nein
Nützlichkeit für die professionelle Entwicklung	.908 (exzellent)	nein
Nützlichkeit für die Schüler und Schülerinnen sowie die Schule	.922 (exzellent)	nein
prozessorientierte Erfolgserwartung	.755 (akzeptabel)	nein
ergebnisorientierte Erfolgserwartung	.751 (akzeptabel)	nein
<i>Zufriedenheit</i>	.721 (akzeptabel)	ja

## 7 Validität des Fragebogens

### 7.1 Inhaltsvalidität

Mit der Inhaltsvalidität wird ausgedrückt, wie präzise die verschiedenen Skalenitems das zu messende Konstrukt wiedergeben (Bühner, 2011, S. 62). So muss gut begründet werden, warum von allen möglichen Items bzw. Itemformulierungen gerade die vorliegenden ausgewählt wurden.

Der hier zu analysierende Fragebogen besteht inhaltlich gesehen aus drei Teilen: ein Teil, in dem die Umsetzung unterrichtsintegrierter sprachunterstützender Maßnahmen erfragt wird, ein Teil zur Häufigkeit und Intensität der multiprofessionellen Zusammenarbeit in diesem Kontext und ein Teil zu den Einstellungen zur multiprofessionellen Zusammenarbeit. Dabei wurden Items aus bereits existierenden Fragebögen gezielt zusammengestellt. Der erste Teil beruht zum Großteil auf einem Fragebogen von Karin Reber, welcher für das Projekt «Inklusion und Sprachtherapie/ Logopädie/ Sprachheilpädagogik: Angebote für Kinder und Jugendliche mit Sprachbehinderungen an europäischen Schulen» entwickelt wurde (vgl. Reber, 2012; Reber & Blechschmidt, 2014). Der Fragebogen wurde in kognitiven und statistischen Pretests voruntersucht, Details hierzu sind jedoch nicht bekannt. Der zweite und dritte Teil beruht auf einem Fragebogen zur Lehrpersonenkooperation von Drossel (2015), der wiederum zum Teil auf bewährten Subskalen, zum Teil aber auch auf Eigenentwicklungen beruht. Dieser Fragebogen wurde u.a. mithilfe konfirmatorischer Faktorenanalysen untersucht und überarbeitet, sodass ein reliables und valides Testinstrument entwickelt werden konnte.

Beide Fragebögen richteten sich jedoch an Personen aus jeweils einer Berufsgruppe (Lehrpersonen). Im vorliegenden Fall sollten jedoch drei Berufsgruppen (Regellehrpersonen, Schulische Heilpädagog:innen, Logopäd:innen) parallel befragt werden. Außerdem war es notwendig, den Befragungsgegenstand in einigen Items zu präzisieren, da es hier nicht um Kooperation im allgemeinen Schulalltag, sondern um die disziplinenübergreifende Zusammenarbeit im Kontext der unterrichtsintegrierten sprachunterstützenden Maßnahmen geht.

Die erste umfassende Anpassung bezüglich der Itemformulierungen war, die Begriffe «Unterricht» bzw. «Sprachtherapie» berufsgruppenspezifisch zu ersetzen, da klar ist, dass Regellehrpersonen keine Sprachtherapie machen und Logopäd:innen keinen Unterricht. So wurden basierend auf dem Modell «Maßnahmen der primären, sekundären und tertiären Prävention in der inklusiven sprachlichen Bildung» von Sallat, Hofbauer und Jurleta (2017) die Begriffe «sprachsensibler Unterricht», «Sprachförderung» und «Sprachtherapie» verwendet und berufsgruppenspezifisch eingesetzt:

#### **Beispiel Skala «unterrichtsintegrierter sprachunterstützender Maßnahmen in verschiedenen Fachbereichen»**

- Version Regellehrpersonen: «Welche Fachbereiche nutzen Sie, um sprachsensiblen Unterricht im Klassenzimmer durchzuführen?»
- Version Schulische Heilpädagog:innen: «Welche Fachbereiche nutzen Sie, um Sprachförderung im Klassenzimmer durchzuführen?»
- Version Logopäd:innen: «Welche Fachbereiche nutzen Sie, um Sprachtherapie im Klassenzimmer durchzuführen?»

Auf diese Weise wurden Parallelversionen der Fragebogenitems erstellt, die bis auf die beschriebene Differenzierung keine weiteren Unterschiede in der Formulierung aufweisen.

Dennoch war es nötig, im zweiten Teil des Fragebogens zur Häufigkeit und Intensität der multiprofessionellen Zusammenarbeit zwischen den Berufsgruppen zu unterscheiden. So lautet bspw.

ein Item im Original: «Mit meinen Kollegen erstelle ich gemeinsam Arbeitsblätter.». In den kognitiven Pretests wurde jedoch schnell deutlich, dass hierauf keine präzise Antwort gegeben werden konnte, da es sehr stark darauf ankam, wer der oder die Kolleg:in ist – eine Regellehrperson, eine Schulische Heilpädagog:in oder eine Logopäd:in. So musste auch hier berufsgruppenspezifisch gefragt werden:

#### **Beispiel Skala «Ko-Konstruktion»**

- Version Regellehrpersonen: «Ich erstelle Arbeitsblätter gemeinsam...»
  - «mit den Schulischen Heilpädagog:innen»
  - «mit den Logopäd:innen»
- Version Schulische Heilpädagog:innen: «Ich erstelle Arbeitsblätter gemeinsam...»
  - «mit den Regellehrpersonen»
  - «mit den Logopäd:innen»
- Version Logopäd:innen: «Ich erstelle Arbeitsblätter gemeinsam...»
  - «mit den Regellehrpersonen»
  - «mit den Schulischen Heilpädagog:innen»

Diese Itemformulierung wurde ebenfalls in den kognitiven Interviews vorgetestet. Es zeigte sich, dass der Informationsgehalt dieser Frage deutlich anstieg, da bedeutende Unterschiede in den Häufigkeitsangaben bestehen, je nachdem mit wem zusammengearbeitet wird.

Im dritten Teil des Fragebogens war es jedoch nicht nötig, zwischen den Berufsgruppen zu unterscheiden, da hier eher Haltungen und Überzeugungen erfragt werden und nicht konkrete Häufigkeiten. Dennoch musste eine Formulierung gefunden werden, die deutlich macht, dass es nicht um die Kooperation mit den Personen aus der eigenen Berufsgruppe, sondern eben mit jenen aus den anderen Berufsgruppen geht.

#### **Beispiel Skala «Interesse»**

- Alle Versionen: «Mir ist die Kooperation mit den anderen Fachkräften wichtig.»

Lediglich im einleitenden Satz zur jeweiligen Subskala wurde z.B. der Verweis «Wie bewerten Sie die Kooperation mit den anderen Fachkräften (... , ...)?» mit der Ergänzung «(Schulische Heilpädagog:innen, Logopäd:innen)» bei den Regellehrpersonen, «(Regellehrpersonen, Logopäd:innen)» bei den Schulischen Heilpädagog:innen und «(Regellehrpersonen, Schulische Heilpädagog:innen)» bei den Logopäd:innen ergänzt.

Die konsequente Umsetzung der beschriebenen Anpassungen führt zu berufsgruppenspezifischen Fragebogenversionen, die sich in berufsgruppenspezifischen Termini unterscheiden, aber ansonsten die gleichen Sachverhalte erfragen. Die Originalitems wurden weitestgehend beibehalten, dem aktuellen Anliegen jedoch in angemessener Weise angepasst, sodass davon ausgegangen wird, dass die drei Teile des Fragebogens das jeweilige Konstrukt abzubilden vermögen. Wie die obigen Reliabilitätsanalysen der Subskalen zeigen, scheinen diese – bis auf wenige Ausnahmen – für alle drei Berufsgruppen gut geeignet zu sein.

## **7.2 Kriteriumsvalidität**

Mit der Kriteriumsvalidität wird der Zusammenhang der Fragebogendaten mit Außenkriterien untersucht (Bühner, 2011, S. 63). Ein offensichtliches Kriterium, das zu unterschiedlichen Fragebogenergebnissen führen sollte, ist die Berufsgruppenzugehörigkeit. Aufgrund vorliegender

Forschungsergebnisse (vgl. Blechschmidt et al. 2013; Kolbe & Reh 2008) wird postuliert, dass Regellehrpersonen sprachunterstützende Maßnahmen (sM) häufiger unterrichtsintegriert umsetzen als Schulische Heilpädagog:innen, und diese wiederum häufiger als Logopäd:innen:

H1: Regellehrpersonen > Schulische Heilpädagog:innen > Logopäd:innen

Auch bezüglich der Häufigkeit und Intensität der multiprofessionellen Kooperation werden berufsgruppenspezifische Unterschiede angenommen, und zwar derart, dass Regellehrpersonen und Logopäd:innen häufiger weniger intensive Zusammenarbeitsformen (z.B. Austausch) wählen, Schulische Heilpädagog:innen jedoch die Synchronisation und Ko-Konstruktion bevorzugen (vgl. Albers, 2014; Blechschmidt et al., 2013). Beim Team-Teaching als intensivster Form der Kooperation (vgl. Kolbe & Reh, 2008) sollten diese Unterschiede besonders deutlich werden. Schulische Heilpädagog:innen führen Team-Teaching bedeutend häufiger durch als Personen aus den beiden anderen Berufsgruppen:

H2: Regellehrpersonen < Schulische Heilpädagog:innen > Logopäd:innen

Aufgrund der genannten Zusammenhänge wird ebenso vermutet, dass es berufsgruppenspezifische Unterschiede in den Einstellungen zur multiprofessionellen Kooperation gibt, da schon die Ausbildung, die Arbeitsbelastung und die Erfahrungen mit Kooperation unterschiedlich gestaltet sind (vgl. Drossel & Willems, 2014). Es kann jedoch nicht vorhergesagt werden, wie die Unterschiede ausgeprägt sein werden:

H3: Regellehrpersonen  $\neq$  Schulische Heilpädagog:innen  $\neq$  Logopäd:innen

Die Gruppenunterschiede werden mithilfe von einfaktoriellen Varianzanalysen (ANOVA) berechnet, und zwar lediglich auf Grundlage der Subskalen und Items, die sich im Rahmen der Reliabilitätsanalyse als brauchbar erwiesen haben. Da die Stichproben der verschiedenen Berufsgruppen unterschiedlich groß sind, müssen die Voraussetzungen (Normalverteilung, Homogenität der Varianzen) besonders gut geprüft werden (vgl. Bortz & Schuster, 2010, S. 214; Field, 2013, S. 444). Bei Verletzung der Voraussetzungen wird auf den non-parametrischen Kruskal-Wallis-Test zurückgegriffen.

Für den ersten Teil des Fragebogens wird ein Summenscore kreiert und gemittelt. Der zweite Teil des Fragebogens wird lediglich durch die Subskala «Team-Teaching» repräsentiert, da die übrigen Kooperationsformen berufsgruppenspezifisch beantwortet wurden und sich nicht direkt miteinander vergleichen lassen. Da Team-Teaching jedoch als intensivste Form der Zusammenarbeit gilt (Kolbe & Reh, 2008), lässt sich dies auch theoretisch rechtfertigen. Der dritte Teil wird wiederum durch einen gemittelten Summenscore repräsentiert.

Bezüglich des ersten Teils des Fragebogens (unterrichtsintegrierte sprachunterstützende Maßnahmen) sind die Daten eindeutig nicht normalverteilt (Shapiro-Wilk:  $W(233) = .93$ ,  $p < .001$ ). Zwar liegt gemäß dem Levene-Test Varianzhomogenität vor ( $F(2, 230) = 2.60$ ,  $p = .077$ ), aber Field (2013, S. 445) warnt bei ungleichen Gruppengrößen eindeutig davor, sich auf diesen Test zu verlassen. Da der Levene-Test außerdem knapp ausfällt, werden entsprechende Gruppenunterschiede mit dem Kruskal-Wallis-Test überprüft. Dieser ermittelt hoch signifikante Gruppenunterschiede ( $H(2) = 136.50$ ,  $p < .001$ ). Post-hoc gerechnete (und Bonferroni-korrigierte) Gruppenvergleiche offenbaren, dass sich v.a. Logopäd:innen von den Regellehrpersonen ( $p < .001$ ,  $r = .75$ ,) und den Schulischen Heilpädagog:innen ( $p < .001$ ,  $r = .60$ ) unterscheiden. Beide Effektstärken wurden mit dem Dunn-

Bonferroni-Test ermittelt und sind als groß zu interpretieren. Regellehrpersonen und Schulische Heilpädagog:innen unterscheiden sich hinsichtlich der Häufigkeit unterrichtsintegrierter sprachunterstützender Maßnahmen nicht voneinander ( $p = .428$ ). Logopäd:innen führen unterrichtsintegrierte sprachunterstützende Maßnahmen also bedeutend seltener durch als die beiden anderen Berufsgruppen, die sich diesbezüglich nicht unterscheiden. Die H1 kann demnach nur teilweise angenommen werden.

Auch der Team-Teaching-Score, der den zweiten Teil des Fragebogens repräsentiert, ist nicht normalverteilt (Shapiro-Wilk:  $W(262) = .88$ ,  $p < .001$ ). Der Test auf Varianzhomogenität wird ebenfalls signifikant ( $F(2, 259) = 20.08$ ,  $p < .001$ ), weswegen auch hier der Kruskal-Wallis-Test zum Einsatz kommt. Das hochsignifikante Testergebnis ( $H(2) = 131.65$ ,  $p < .001$ ) sowie die Bonferroni-korrigierten Post-hoc-Vergleiche führen zur vorläufigen Annahme der H2:

- Regellehrpersonen vs. Schulische Heilpädagog:innen:  $p < .01$ ,  $r = .30$
- Regellehrpersonen vs. Logopäd:innen:  $p < .001$ ,  $r = .46$
- Schulische Heilpädagog:innen vs. Logopäd:innen:  $p < .001$ ,  $r = .75$

Schulische Heilpädagog:innen arbeiten wie angenommen von allen drei Berufsgruppen am häufigsten im Team-Teaching, aber auch Regellehrpersonen tun dies häufiger als Logopäd:innen. Dabei haben die Unterschiede zwischen den Regellehrpersonen und den Schulischen Heilpädagog:innen eine mittlere Effektstärke, die zwischen den Logopäd:innen und den anderen beiden Berufsgruppen (schon) starke Effekte.

Zur Überprüfung der H3 wird wiederum ein Summenscore herangezogen. Dieser ist nicht normalverteilt (Shapiro-Wilk:  $W(246) = .98$ ,  $p < .01$ ). Auch hier scheint Varianzhomogenität gegeben ( $F(2, 243) = 2.33$ ,  $p = .099$ ), allerdings wird der Empfehlung von Field (s.o.) noch einmal entsprochen und der Kruskal-Wallis-Test eingesetzt:  $H(2) = 16.31$ ,  $p < .001$ . Die Bonferroni-korrigierten Post-hoc-Vergleiche ergeben keinen Unterschied zwischen Regellehrpersonen und Logopäd:innen ( $p = 1.00$ ), Regellehrpersonen unterscheiden sich aber wiederum von den Schulischen Heilpädagog:innen ( $p < .01$ ,  $r = .31$ ) und letztere auch von den Logopäd:innen ( $p < .001$ ,  $r = .27$ ). Beide Effekte sind als «mittel» einzustufen. Die Einstellungen der Schulischen Heilpädagog:innen zur Zusammenarbeit sind im Durchschnitt positiver ausgeprägt als bei den Regellehrpersonen und Logopäd:innen (welche sich nicht voneinander unterscheiden). Die H3 muss also in Teilen abgelehnt werden.

Die gefundenen Unterschiede zwischen den Berufsgruppen entsprechen teils den formulierten Erwartungen, teils widersprechen sie ihnen. Da die Forschungsgrundlage für die getroffenen Annahmen ohnehin jedoch recht «dünn» war, wird dies eher als Beleg für die Notwendigkeit gesehen, diese Forschungslücke zu schließen. Der Fragebogen an sich vermag es jedoch, berufsgruppenspezifische Unterschiede aufzudecken.

### 7.3 Konstruktvalidität

Die Überprüfung der Konstruktvalidität über die konvergente oder die diskriminante Validität ist in diesem Fall nicht sinnvoll, da die verwendeten Fragebögen (vgl. Reber, 2012; Reber & Blechschmidt, 2014; Drossel, 2015) inhaltlich verändert wurden. Somit sind Korrelationen mit den vorhandenen Untersuchungsergebnissen nicht interpretierbar.

Stattdessen wird die faktorielle Validität überprüft, um zu testen, ob die ursprünglichen Messmodelle sich auch in dieser Untersuchung in gleicher oder ähnlicher Weise bestätigen lassen (Bühner, 2011, S. 64). Im Rahmen der konfirmatorischen Faktoranalysen (CFA) werden die Variablen als kontinuierlich behandelt, wenn die verwendete Likertskala mind. fünf Skalenstufen hat. Wenn sie hingegen weniger

Stufen aufweisen, werden sie als kategoriale Daten betrachtet (s. Byrne, 2010, S. 128; Reinecke, 2014, S. 198).

Wie auch bei klassischen inferenzstatistischen Verfahren wird für die Durchführung von CFAs eine Normalverteilung vorausgesetzt. Da CFAs auf Tests von Varianzen und Kovarianzen basieren und diese weniger von der Schiefe der Verteilung, sondern von deren Kurtosis beeinflusst werden, sind v.a. letztere Werte entscheidend (Byrne, 2010, S. 98). Dabei sind Werte  $> \pm 2$  als kritisch zu bewerten (ebd., S. 99). Liegen ungünstige Kurtosis-Werte vor, sollten bei kontinuierlichen Variablen statt der Standardschätzer alternative Schätzer eingesetzt werden, wie z.B. der MLM-Schätzer (maximum likelihood mean adjusted; vgl. Byrne, 2010, S. 100) oder der MLR-Schätzer (robust maximum likelihood estimation; vgl. Wang & Wang, 2012, S. 60). Da der MLM-Schätzer aber nicht mit fehlenden Werten umgehen kann und alle zu untersuchenden Variablen Missings aufweisen, wird der MLR-Schätzer bevorzugt (ebd.). Für die kategorialen Variablen wird entsprechend der WLSMV-Schätzer (weighted least square mean and variance adjusted; ebd., S. 68f.) eingesetzt. Wenn nicht anders erwähnt, werden die fehlenden Werte mit dem FIML-Verfahren geschätzt.

Da v.a. die verschiedenen Faktorladungen der Items bzw. Indikatoren von Interesse sind, wird entgegengesetzt zur Standardeinstellung von Mplus nicht die erste Faktorladung auf 1 gesetzt, sondern die Varianz des latenten Faktors.

Folgende Kriterien werden für die Einschätzung des Modellgüte herangezogen (vgl. Geiser, 2011, S. 59ff.):

- Chi-Quadrat ( $\chi^2$ ): der Wert sollte eher klein und der Test nicht signifikant sein ( $p > .05$ )
- CFI (Comparative Fit Index): Werte  $< .95$  sind schlecht, Werte zwischen  $.95$  und  $.97$  sind gut und Werte  $> .97$  sind sehr gut
- RMSEA (Root-Mean-Square-Error-of-Approximation): Werte sollten  $< .05$  sein; zusätzlich wird das zugehörige Konfidenzintervall angegeben
- SRMR (Standardized-Root-Mean-Square-Residual): Werte sollten  $< .05$  sein

Alle Auswertungen werden mit der Software *Mplus* Version 8.6 (Muthén & Muthén 2011) vorgenommen.

### 7.3.1 CFA der Skala «sprachunterstützende Maßnahmen im Allgemeinen»

Wie die Reliabilitätsanalyse dieser Skala schon vermuten lässt, ist dieses Messmodell mit großen Problemen behaftet. Die Modellfitmaße fallen gesamthaft schlecht aus:  $\chi^2$ : 81.881 (df = 9,  $p < .001$ ), CFI = .717, RMSEA = .171 [.138, .206] und SRMR = .105. In den Modellmodifikationen werden Korrelationen zwischen den Items *sm\_beo*, *sem\_ber* und *sm\_ein* ausgewiesen. Doch sind andere Modelleigenschaften relevanter (s. Abb. 2, S. 57). So gibt es zwei Faktorladungen (von *sm\_beo* und *sm\_ber*), die unter dem empfohlenen Cut-off-Wert von  $.30$  (vgl. Wang & Wang, 2012, S. 39) liegen, eine hohe Residualvarianz haben und auf die folglich verzichtet werden sollte. Außerdem sind zwei Faktorladungen (*sm\_ein* und *sm\_ber*) negativ, was nur Sinn macht, wenn diese Items ursprünglich negativ gepolt waren und für die Analyse nicht umkodiert worden sind. Das ist hier aber nicht der Fall. D.h. die Items mit den negativen Faktorladungen fügen sich nicht in das Konstrukt ein und müssen umformuliert, wenn nicht sogar aus der Skala entfernt werden. Da die Formulierung der Items parallel zu den übrigen Items der Skala erfolgt ist, ist eine Umformulierung nicht zielführend.



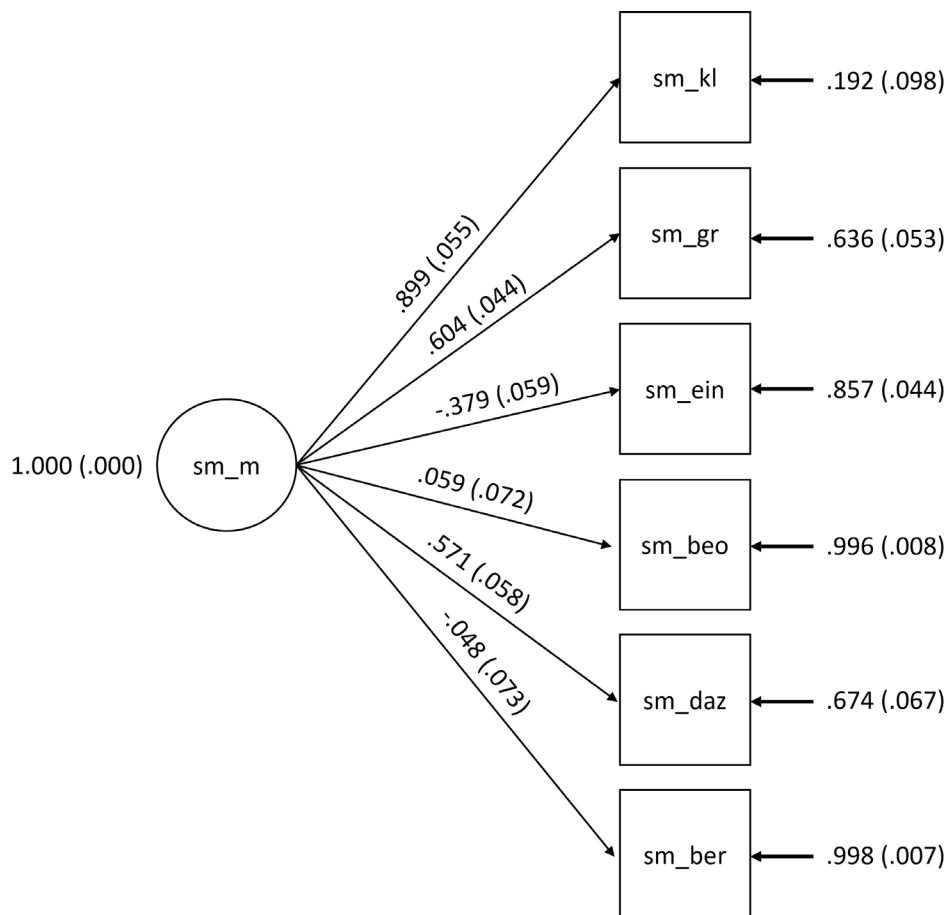


Abbildung 2: Messmodell von "sm\_all" mit standardisierten Werteangaben

Ein Modell ohne das Item sm\_beo weist bessere, aber immer noch ungenügende Fitmaße auf:  $\chi^2$ : 46.837 (df = 5,  $p < .001$ ), CFI = .816, RMSEA = .174 [.131, .221] und SRMR = .081. Die Faktorladung von sm\_ber ist nach wie vor gering und negativ (-.052). Entfernt man wiederum dieses Item, erhält man ein Messmodell mit sehr guten Modellfitmaßen:  $\chi^2$ : 0.156 (df = 2,  $p = .925$ ), CFI = 1.000, RMSEA = .000 [.000, .039] und SRMR = .004. Sinnvoll ist dieses Modell jedoch immer noch nicht, da das Item sm\_ein nach wie vor negativ auf den Faktor lädt (-.380), was inhaltlich nicht gerechtfertigt werden kann. Würde man das Item auch noch entfernen, würde dies in einem gerade identifizierten Modell resultieren, dass mit den Items sm\_kl («sprachunterstützende Maßnahmen in der Klasse»), sm\_gr («sprachunterstützende Maßnahmen in der Gruppe») und sm-daz («DaZ-Unterricht») ein unvollständiges, künstlich reduziertes Abbild der sprachunterstützenden Maßnahmen insgesamt darstellen würde und daher dem Untersuchungszweck nicht dienlich ist. Es scheint, wie schon aus der Reliabilitätsanalyse ersichtlich, dass dieses Messmodell ungeeignet ist, um ein sinnvolles Konstrukt zu erheben. Infolgedessen wird auf diese Skala künftig verzichtet.

### 7.3.2 CFA der Skala «sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten»

Das Messmodell von «sm\_all» (s. Abb. 3, S. 58) hat folgende Werte:  $\chi^2 = 29.188$  (df = 9,  $p < .001$ ), CFI = .913, RMSEA = .090 [.055, .128] und SRMR = .054. Die Werte sind insgesamt eher ungenügend. Zwar werden keine Modifikationsindices ausgegeben, dennoch finden sich Hinweise, wie das Modell angepasst werden könnte. So unterscheiden sich zwar alle Faktorladungen signifikant von 0, allerdings fällt der Signifikanztest für das Item rfl\_all denkbar knapp aus. Auch die Faktorladung an sich liegt mit .140 unter dem Cut-off-Wert von .30. Entsprechend hoch ist die Residualvarianz des Items (.980), was

bedeutet, dass das Item lediglich 2% zur Varianzaufklärung dieser Skala beiträgt. In ähnlicher Weise gilt dies auch für das Item kom\_all, das eine Residualvarianz von .903 hat. Da Ziel dieser Überprüfung ist, den Fragebogen zu optimieren und das Item rfl\_all schon in der Reliabilitätsanalyse problematisch war, wird das Modell noch einmal ohne dieses berechnet.

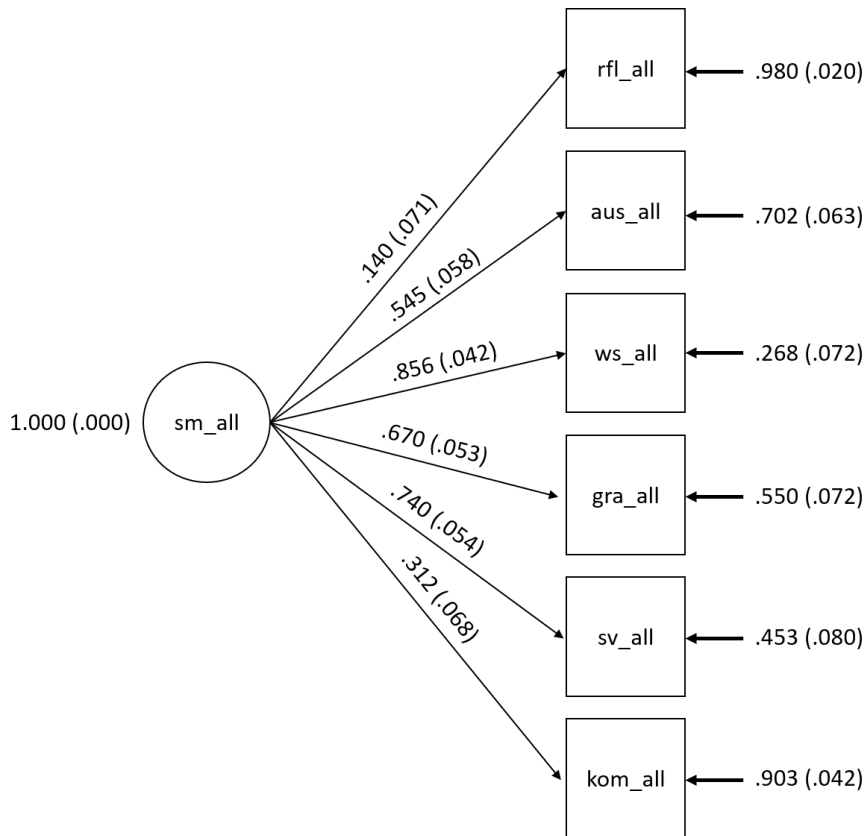


Abbildung 3: Messmodell von "sm\_all" mit standardisierten Werteangaben

Mit dieser Modifikation verbessern sich die Gütekriterien deutlich:  $\chi^2$ -Wert: 11.463 (df = 5, p < .05), CFI = .965, RMSEA = .069 [.011, .122] und SRMR = .033. Der CFI und der SRMR weisen nun gute bis sehr gute Werte auf, die übrigen Kriterien sind noch nicht zufriedenstellend. Vergleicht man die AIC (Akaike Information Criterion)-Werte der beiden Modelle (vgl. Brown, 2015, S. 156), wird das Modell ohne rfl\_all (AIC = 2726.139) begünstigt, da der Wert deutlich niedriger ausfällt, als im Ursprungsmodell (AIC = 3664.561).

Auch das Item kom\_all liegt beinahe unter dem Cut-off-Wert von .30. Da es jedoch in der folgenden Skala, die parallel zu dieser Skala aufgebaut ist, eine bedeutende Rolle spielt, wird auf das Item kom\_all vorerst nicht verzichtet. Letzten Endes wird also das Item rfl\_all aufgegeben und alle übrigen Items beibehalten, sodass das in Abbildung 4 (S. 59) dargestellte Messmodell entsteht.

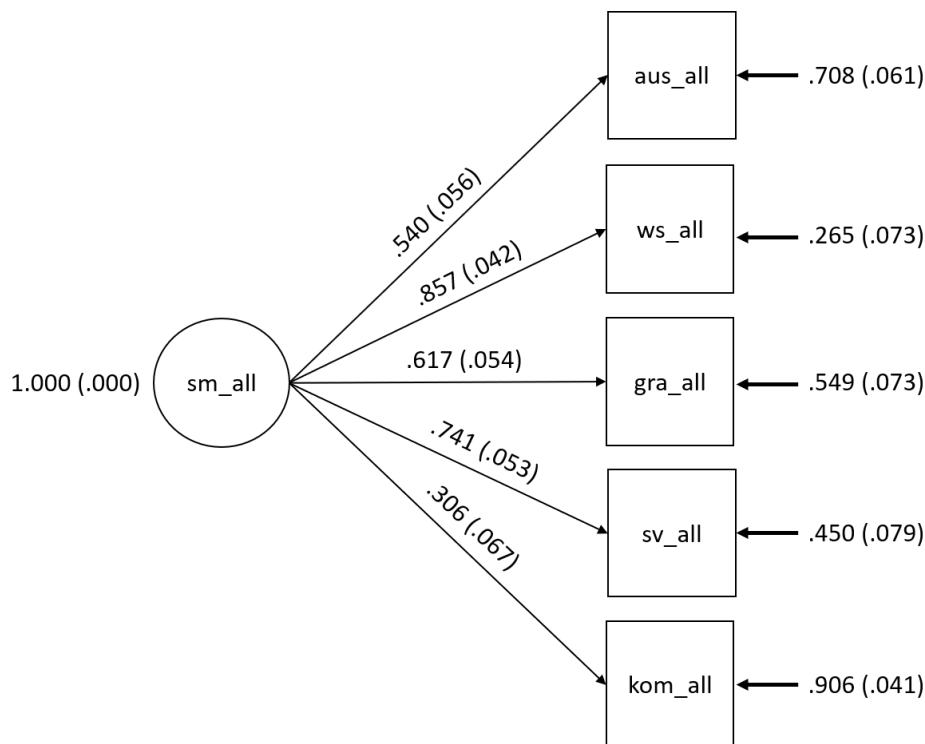


Abbildung 4: Messmodell von "sm\_all" ohne "rfl\_all" mit standardisierten Werteangaben

### 7.3.3 CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten»

Das Messmodell in Abbildung 5 (S. 60) hat mit einem  $\chi^2 = 23.652$ ,  $df = 9$ ,  $p < .01$ , einem CFI = .983, RMSEA = .077 [.040, .116] und SRMR = .021 z.T. sehr gute Werte. Alle Faktorladungen sind signifikant und fallen auch hoch genug aus, sodass kein Indikator aufgegeben werden muss. In den Modifikationsindices wird jedoch auf eine bedeutende Korrelation (.381) der Items rfl\_kl und aus\_kl hingewiesen. Der  $\chi^2$ -Wert ließe sich um 15.522 Punkte verbessern, würde man diese Korrelation frei schätzen lassen. Allerdings lässt sich dies inhaltlich nicht rechtfertigen. Auffälligkeiten des Redeflusses lassen sich eindeutig von Auffälligkeiten der Aussprache unterscheiden und müssen auch unterschiedlich behandelt werden. Aus fachlicher Sicht haben diese beiden Aspekte nur sehr wenig gemeinsam. Betrachtet man die berufsgruppenspezifischen Mittelwerte, fällt auf, dass alle Berufsgruppen angeben, wöchentlich mit Auffälligkeiten der Aussprache zu tun zu haben. Auffälligkeiten des Redeflusses begegnen den Regellehrpersonen und Schulischen Heilpädagog:innen scheinbar noch monatlich, während Logopäd:innen angeben, lediglich mehrmals im Jahr damit zu tun zu haben. Diese Beobachtung bestätigt den Verdacht, dass Logopäd:innen als Fachpersonen für Auffälligkeiten des Redeflusses hier sehr viel differenziertere Antworten geben können als die beiden anderen Gruppen. Für Regellehrpersonen und Schulische Heilpädagog:innen scheint hingegen eine große Nähe von Auffälligkeiten des Redeflusses und der Aussprache gegeben zu sein.

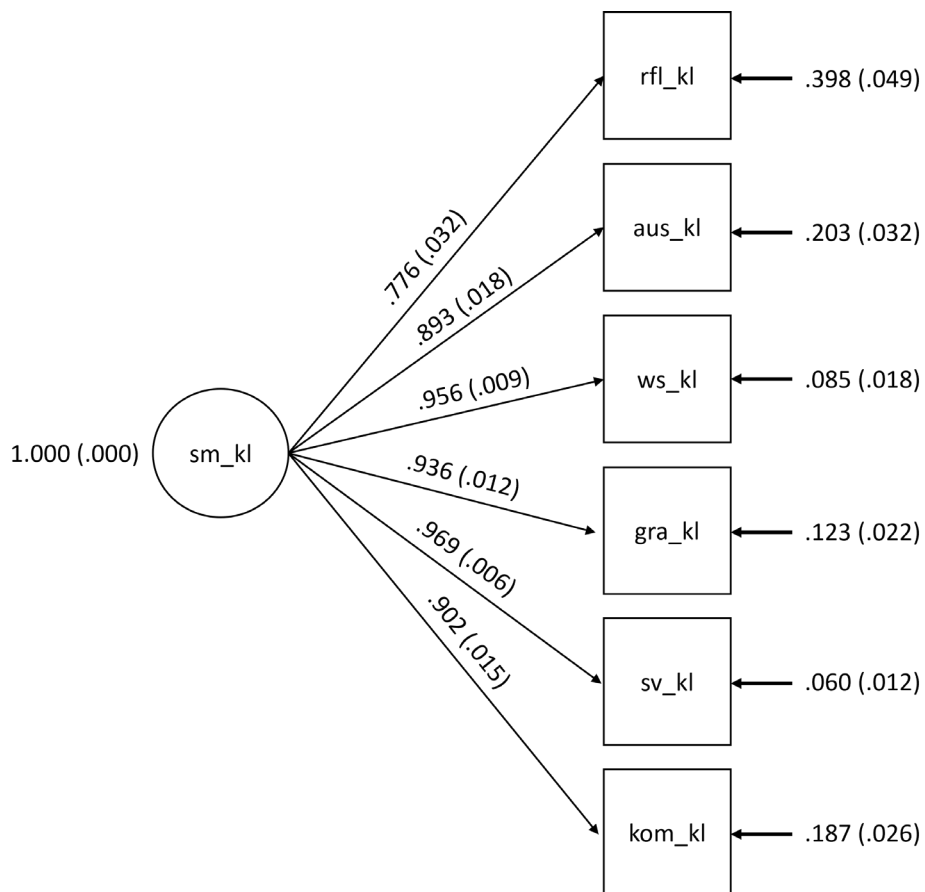


Abbildung 5: Messmodell von "sm\_kl" mit standardisierten Werteangaben

Da in der vorhergehenden Skala ohnehin festgestellt werden konnte, dass das Item rfl\_all eliminiert werden sollte, um das Messmodell zu verbessern, wird auch hier das Item rfl\_kl eliminiert, da die Skala kohärent zur vorhergehenden aufgebaut ist. Die Gütekriterien des neuen Messmodells verbessern sich deutlich:  $\chi^2 = 1.196$ ,  $df = 5$ ,  $p = .987$ , einem CFI = 1.000, RMSEA = .000 [.000, .013] und SRMR = .003. Der entsprechende AIC beträgt 3325.694 (Ursprungsmodell: AIC = 4066.500) und bestätigt wiederum, dass das Messmodell ohne das Item rfl\_kl dem ursprünglichen Modell überlegen ist. So wird aus Überlegungen zur Skalenkonstruktion sowie mit Blick auf die verbesserten Modellfits auch hier das entsprechende Item eliminiert (siehe Abb. 6, S. 61).

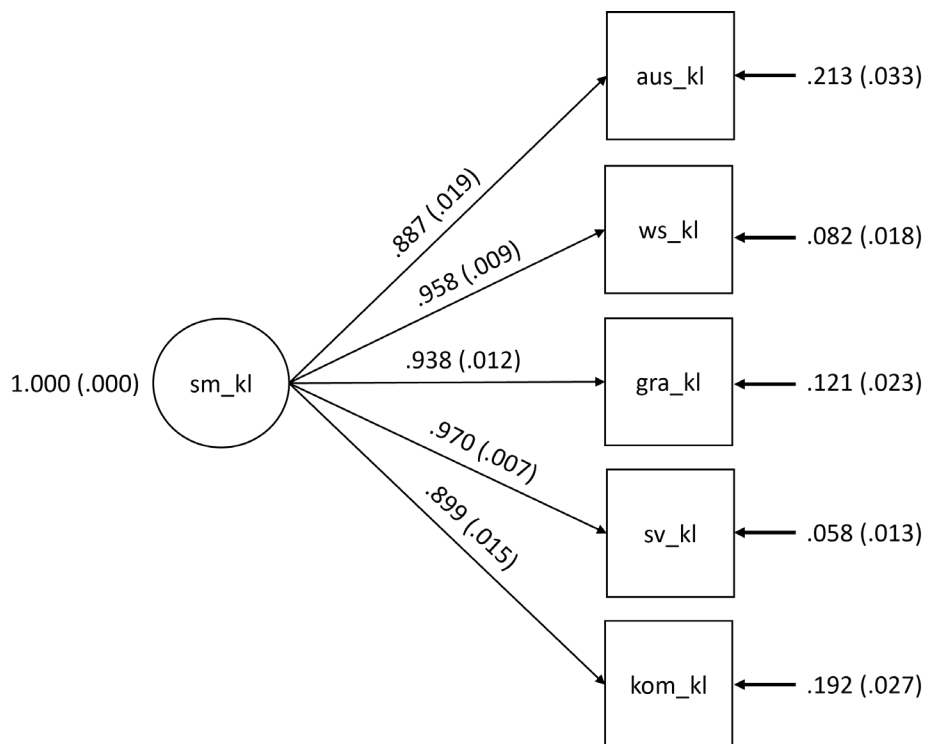


Abbildung 6: Messmodell von "sm\_kl" ohne "rfl\_kl" mit standardisierten Werteangaben

#### 7.3.4 CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Unterrichtsformen»

Das Messmodell zum Faktor sm\_uf (s. Abb. 7, S. 62) weist gemischte Modellgütekriterien auf:  $\chi^2 = 11.218$  (df = 2,  $p < .01$ ), CFI = .973, RMSEA = .130 [.063, .208] und SRMR = .024. Während die Werte für CFI und SRMR beide hervorragend sind, fällt v.a. der RMSEA-Wert auf, der für eine schlechte Modellpassung spricht. Die Faktorladungen werden jedoch sämtlich signifikant und fallen zudem hoch aus. Als Modellmodifikationen werden zwei Korrelationen (uf\_ein mit uf\_gru und uf\_fron mit uf\_off) vorgeschlagen. Die zusätzliche Schätzung der beiden Korrelationen würde den  $\chi^2$ -Wert um 10.355 bzw. 10.341 Punkte verbessern (also in etwa gleich viel). Diese Korrelationen sind aber inhaltlich nicht zu rechtfertigen. So schließen sich Maßnahmen während Einzelarbeitsphasen (uf\_ein) und in Gruppenarbeitsphasen (uf\_gru) gegenseitig aus. Gleiches gilt für Frontalunterricht (uf\_fron) und offenen Unterricht (uf\_off). Wiederum liegt der Verdacht nahe, dass die Items durch die teilnehmenden Personen nicht klar zugeordnet werden konnten. Im Rahmen der Missinganalyse sowie der Reliabilitätsanalyse ist die Gruppe der Schulischen Heilpädagog:innen aufgefallen, die bezüglich der Beantwortung der Items Schwierigkeiten gehabt zu haben scheint. Eine Verbesserung des Messmodells kann also eher durch eine Umformulierung der Items bzw. durch zusätzliche Hilfestellungen beim Beantworten der Fragen erreicht werden.

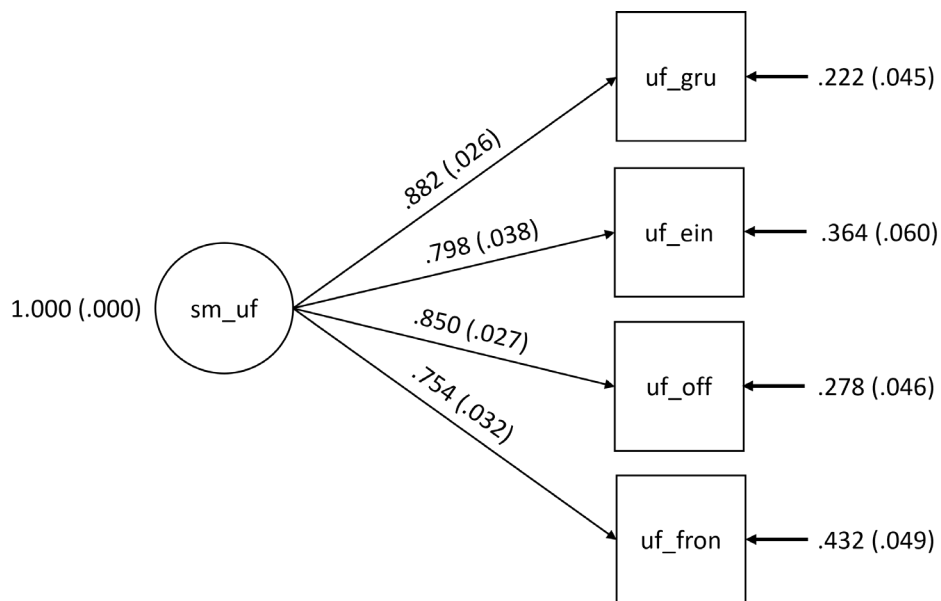


Abbildung 7: Messmodell von "sm\_uf" mit standardisierten Werteangaben

An dieser Stelle sollte erwähnt werden, dass sowohl der  $\chi^2$ -Wert als auch der RMSEA-Wert negativ durch kleine Stichprobengrößen beeinflusst werden, was hier der Fall sein könnte. Von daher werden hier die CFI- sowie SRMR-Werte stärker gewichtet, die nicht von der Stichprobengröße abhängig sind (vgl. Wang & Wang, 2012, S. 18ff.). Von daher wird das Messmodell nicht modifiziert. Die Werte sind insgesamt betrachtet zufriedenstellend. Dennoch sollten Änderungen der Itemformulierungen vorgenommen werden, um die Gütekriterien verbessern zu können.

### 7.3.5 CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Fachbereichen»

Bezüglich des Messmodells von sm\_fb (s. Abb. 8, S. 63) muss vorab erwähnt werden, dass der Datensatz an dieser Stelle viele fehlende Werte aufweist – allerdings sind dies nicht einfach nur Fälle, in denen Items nicht beantwortet wurden, sondern auch solche, in denen die «trifft nicht zu»-Option gewählt wurde, weil das entsprechende Schulfach nicht am Arbeitsort oder nicht auf der jeweiligen Stufe angeboten wird. So gibt es ganze 63 Fälle, von denen keine Angaben vorliegen – die Stichprobe wird entsprechend kleiner ( $n = 215$ ). Dies erklärt wahrscheinlich schon den ungenügenden Modellfit:  $\chi^2 = 109.418$  ( $df = 27$ ,  $p < .001$ ), CFI = .863, RMSEA = .119 [.096, .143] und SRMR = .068. Die Faktorladungen sind alle signifikant und liegen über dem kritischen Wert von .30. Bei den Modifikationsindices finden sich jedoch Hinweise auf verschiedene Korrelationen zwischen den Indikatoren. Dabei würde die Schätzung der Korrelation zwischen fb\_mus und fb\_spo (.960) zu einer Verbesserung des  $\chi^2$ -Werts von 33.525 Punkten führen. Der Wert lässt erkennen, dass diese beiden Items von den Teilnehmenden nahezu gleichbehandelt wurden. Wahrscheinlich sind die Fachbereiche Musik einerseits und Bewegung und Sport andererseits prototypisch für die am wenigsten sprachbezogenen Unterrichtsfächer. Dies wird auch durch einen geschätzten Zusammenhang von fb\_mus mit fb\_deu (Deutsch als eindeutig sprachbezogenes Fach) deutlich, die negativ miteinander korrelieren (-.519).

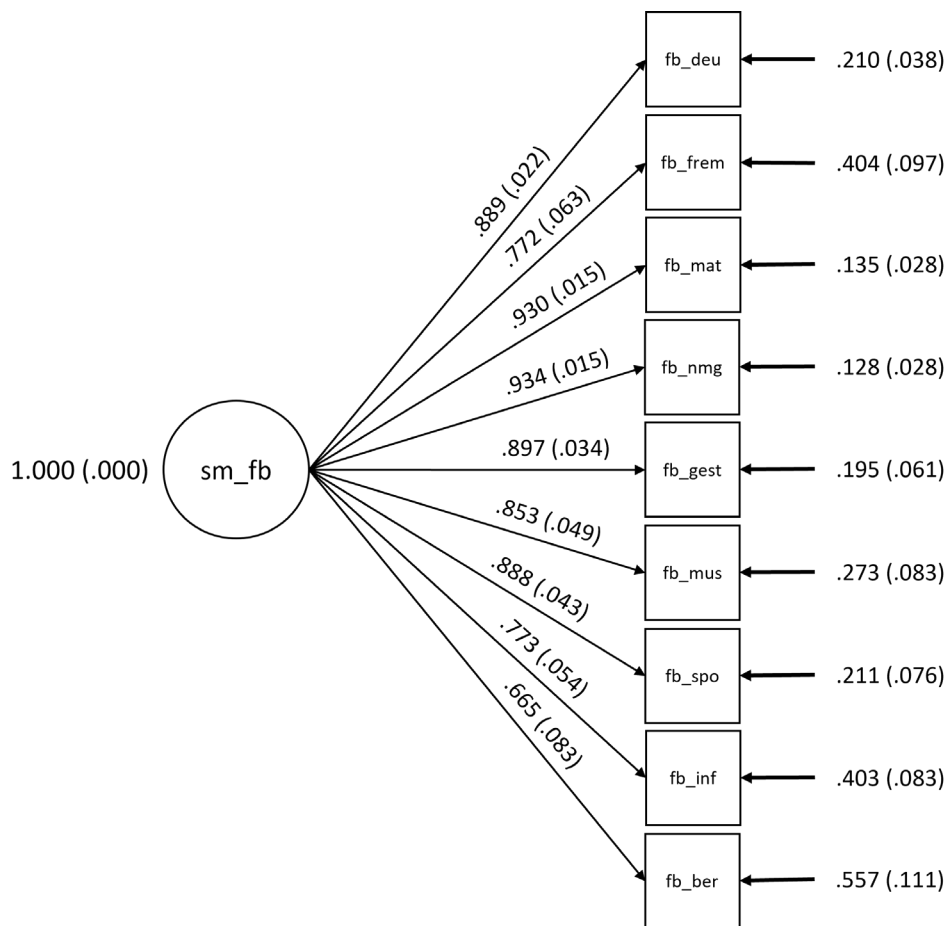


Abbildung 8: Messmodell von "sm\_fb" mit standardisierten Werteangaben

Berücksichtigt man die Korrelation von fb\_mus und fb\_spo im Modell, verbessert sich der Modellfit deutlich:  $\chi^2 = 67.746$  (df = 26,  $p < .001$ ), CFI = .931, RMSEA = .086 [.062, .112] und SRMR = .065. Das Ergebnis des  $\chi^2$ -Differenztests (TRd = 26.223 und  $\Delta df = 1$ , unter Berücksichtigung des Satorra-Bentler-skalierten SB- $\chi^2$ , wie im Folgenden auch berechnet mit [www.thestatisticalmind.com/calculators/SBChiSquareDifferenceTest.html](http://www.thestatisticalmind.com/calculators/SBChiSquareDifferenceTest.html)) ist hochsignifikant ( $p < .001$ ), was für das alternative, weniger restriktive Modell spricht (s. Abb. 9, S. 64).

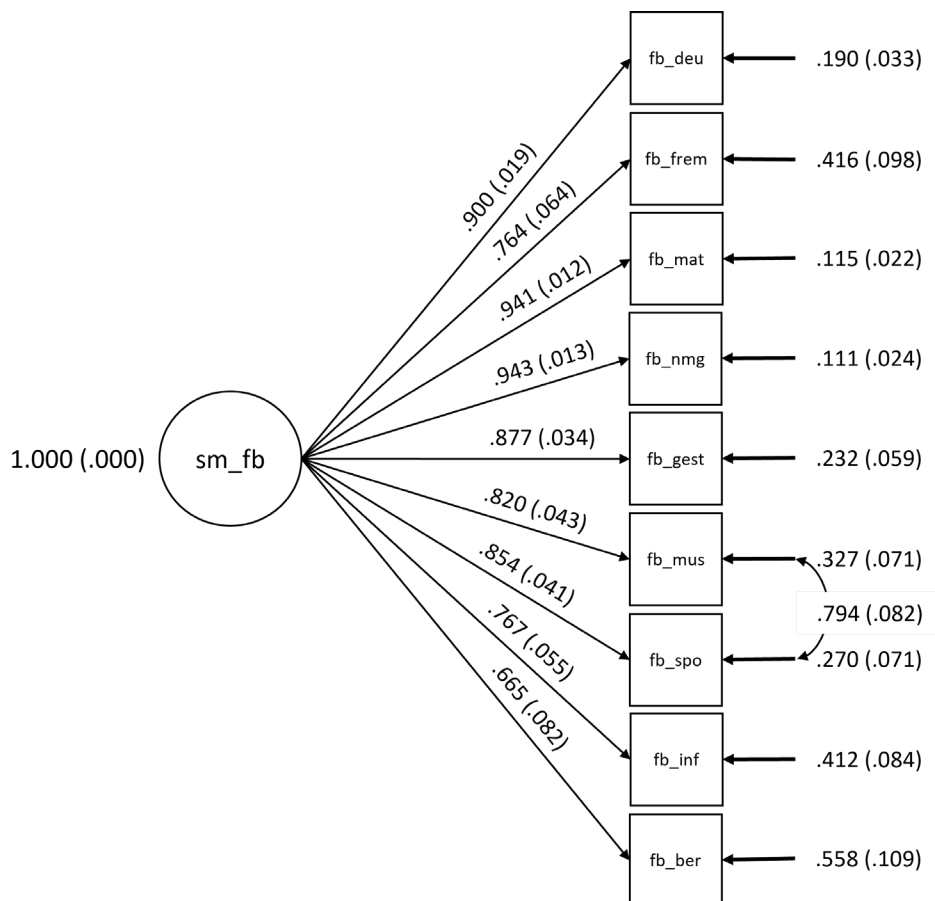


Abbildung 9: Messmodell von "sm\_fb" mit standardisierten Werteangaben unter Einbezug der Korrelation von fb\_mus und fb\_spo

Die Gütekriterien des alternativen Messmodells fallen aber nach wie vor ungenügend aus und weitere Modellmodifikationen werden nicht vorgeschlagen. Der Einsatz dieser Skala ist fraglich. Allerdings könnten die ungenügenden Werte auch auf die stark reduzierte und stark verzerrte Stichprobe zurückzuführen sein. Von daher wird diese Skala vorläufig beibehalten.

### 7.3.6 CFA der Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Erwerbsphasen»

Das Messmodell ist mit lediglich drei Indikatoren (pha\_ein, pha\_ueb und pha\_anw) gerade identifiziert, sodass keine Modellfit-Maße ausgegeben werden. Erst muss eine zusätzliche Restriktion im Modell eingeführt werden, um eine Überidentifizierung des Modells zu erreichen und die Berechnung von Modellgütekriterien zu ermöglichen. Bei der Sichtung der Eigenschaften der Indikatoren fällt auf, dass die Items pha\_ein und pha\_anw sehr ähnliche Residualvarianzen aufzeigen (.508 bzw .471), weswegen diese nun auf den gleichen Wert fixiert werden (.492). Das resultierende Modell (s. Abb. 10, S. 65) hat hervorragende Werte:  $\chi^2 = 0.136$  (df = 1, p = .713), CFI = 1.000, RMSEA = .000 [.000, .116] und SRMR = .005. Die Faktorladungen sind alle signifikant und sind hoch. Es werden keine Modellmodifikationen vorgeschlagen.



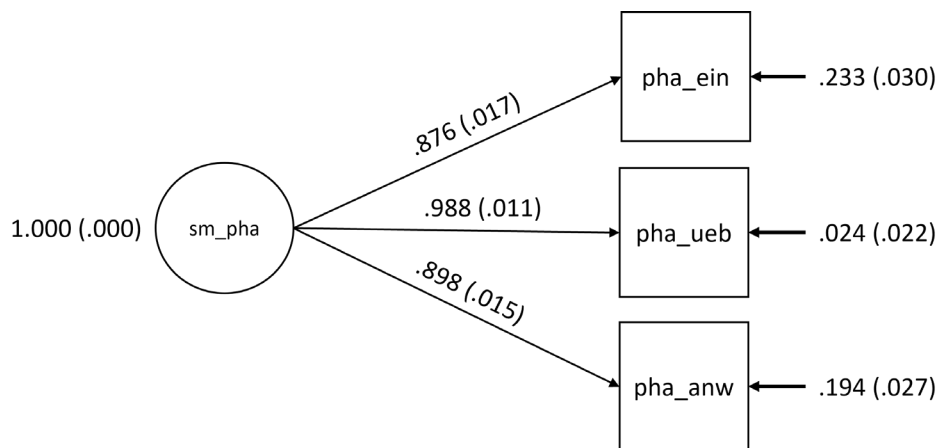


Abbildung 10: Messmodell von "sm\_pha" mit standardisierten Werteangaben

### 7.3.7 CFA der Skala «sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers»

Auf den ersten Blick scheint es sich hier um ein Messmodell mit hervorragenden Gütekriterien zu handeln ( $\chi^2 = 1.103$  (df = 2, p = .576), CFI = 1.000, RMSEA = .000 [.000, .100] und SRMR = .015), doch täuscht dieser Eindruck. Das Konstrukt krankt nämlich an grundlegenden Missspezifikationen des Modells, sog. «Heywood Cases» (vgl. Geiser, 2011, S. 73; Wang & Wang, 2012, S. 22), weswegen auch auf eine Darstellung des Messmodells verzichtet wird: negativen Varianzen und negativen Residualvarianzen. Gemäß Geiser (2011, S. 73) sind solche Missspezifikationen nicht durch einfache Modellanpassungen zu beheben, sondern müssen durch eine grundsätzliche Überarbeitung der Skala beseitigt werden. Da sich dieses Ergebnis sehr gut mit der Reliabilitätsanalyse derselben Skala deckt, unterstreicht dies noch einmal die Entscheidung zur Verwerfung dieser Skala.

### 7.3.8 CFA der Skala «Austausch»

Die Austausch-Skala basiert auf drei unterschiedlich zusammengesetzten Stichproben, je nachdem, wer das Ziel des Austauschs ist (Regellehrpersonen, Schulische Heilpädagog:innen oder Logopäd:innen). Dadurch reduziert sich die Stichprobengröße z.T. deutlich, was sich v.a. auf die Modellfitmaße auswirkt, die diesbezüglich empfindlich sind ( $\chi^2$  und RMSEA). Wird dies in Betracht gezogen, kann man mit den Modellfitmaßen des Messmodells «Austausch mit Regellehrpersonen» zufrieden sein:  $\chi^2 = 8.442$  (df = 2, p < .05), CFI = .971, RMSEA = .122 [.046, .212] und SRMR = .036. Der CFI und der SRMR liefern sehr gute Werte. Die Faktorladungen liegen sämtlich über dem Cut-off-Wert von .30, Modellmodifikationen werden keine vorgeschlagen. Es fällt zwar auf, dass das Item aus1\_lp eine relativ hohe Residualvarianz aufweist, jedoch ist diese noch akzeptabel.

Das Messmodell «Austausch mit Schulischen Heilpädagog:innen» weist bessere Werte auf:  $\chi^2 = 6.029$  (df = 2, p < .05), CFI = .988, RMSEA = .098 [.005, .192] und SRMR = .021. Wiederum sind v.a. die Werte sehr gut ausgefallen, die nicht von der Stichprobengröße beeinflusst werden (CFI und SRMR). Aber auch die anderen Werte sind in diesem Messmodell besser ausgefallen. Ebenso sind die Faktorladungen insgesamt höher und die Residualvarianzen kleiner. Modellmodifikationen werden wiederum keine vorgeschlagen.

Beim letzten Messmodell («Austausch mit Logopäd:innen») reduziert sich die Stichprobengröße beträchtlich (n = 121), da der größte Teil der Stichprobe ausgeschlossen werden muss. Dies sollte sich wiederum nachteilig auf die erwähnten Modellfitmaße auswirken. Tatsächlich fallen jedoch sämtliche Fitmaße sehr gut aus:  $\chi^2 = .303$  (df = 2, p = .859), CFI = 1.000, RMSEA = .000 [.000, .096] und SRMR =

.003. Auch die Faktorladungen sind allesamt hoch bei kleinen Residualvarianzen. Wiederum werden keine Modellmodifikationen empfohlen.

Die Skala scheint somit zwar unterschiedlich gut, aber doch in eindeutiger Weise ein Konstrukt zu erfassen, das als «Austausch» bezeichnet werden kann. Die Ergebnisse der drei Messmodelle sind in Abbildung 11 zusammengefasst.

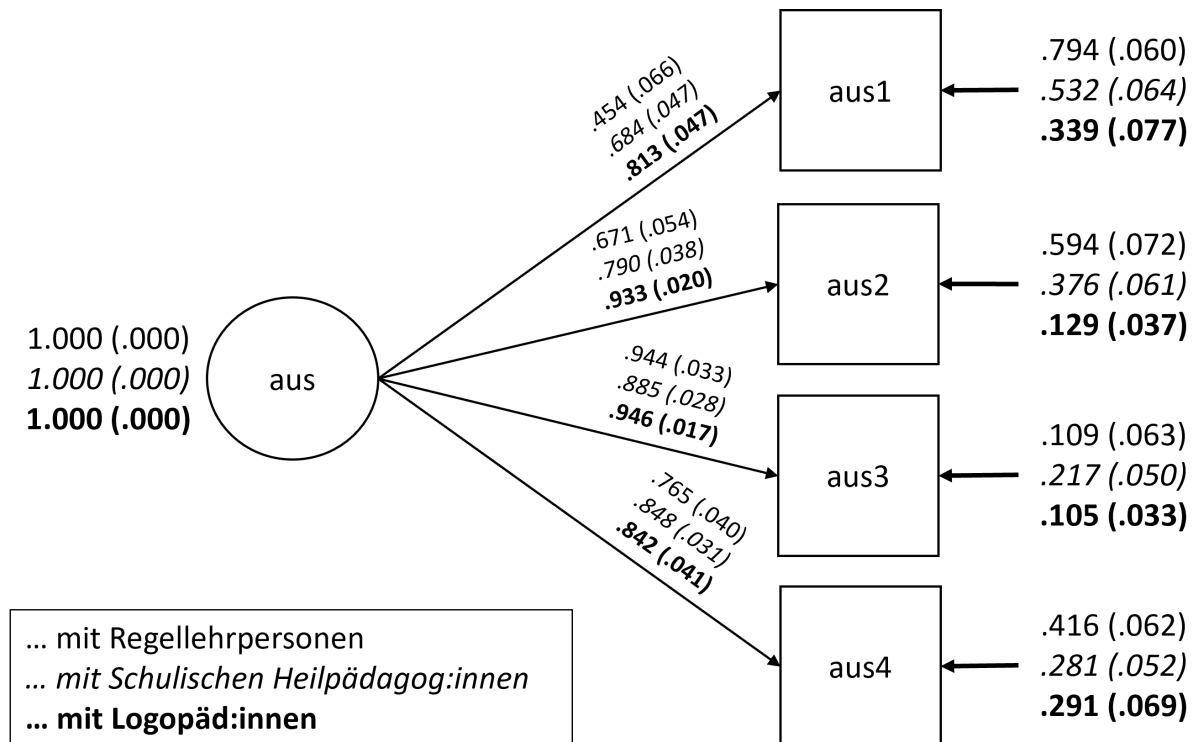


Abbildung 11: Messmodell von "aus" mit standardisierten Werteangaben für die jeweiligen Zielgruppen

Ob die unterschiedlichen Ausprägungen der Messmodelle bezüglich der Zielgruppen des Austauschs auch zu Beeinträchtigungen der Messinvarianz führen, lässt sich mit der vorliegenden Stichprobe nicht überprüfen, da diese zu klein ist. Die entsprechenden Untersuchungen werden voraussichtlich erst mit der Stichprobe der Hauptuntersuchung vorgenommen werden können.

### 7.3.9 CFA der Skala «Synchronisation»

Wie auch bei der Skala «Austausch» müssen die Daten hier bezüglich der unterschiedlichen Zielgruppen der Synchronisation (Regellehrpersonen, Schulische Heilpädagog:innen und Logopäd:innen) analysiert werden. Dies hat wiederum Einfluss auf die zugrundeliegende Stichprobengröße mit den oben geschilderten Einschränkungen. Betrachtet man also zunächst das Messmodell «Synchronisation mit den Regellehrpersonen» fallen wiederum bezüglich der von der Stichprobe wenig beeinflussten Modellfitmaße sehr gute Werte auf:  $\chi^2 = 12.633$  (df = 2, p < .01), CFI = .972, RMSEA = .156 [.082, .243] und SRMR = .025. Sämtliche Faktorladungen sind signifikant und fallen hoch aus. Unter den Modellmodifikationen werden zwei Korrelationen ausgewiesen (syn3\_lp mit syn1\_lp und syn4\_lp mit syn2\_lp), die den  $\chi^2$ -Wert jeweils um 10.148 Punkte verbessern würden. Dabei handelt es sich um die Absprache von Inhalten/Themen (syn1\_lp) und Terminen (syn3\_lp) einerseits sowie Absprachen zu Anforderungen von Aufträgen (syn2\_lp) und Absprachen von Zielen und Schwerpunkten (syn4\_lp) andererseits. Inhaltlich gesehen lässt sich die Schätzung einer entsprechenden Korrelation eher schwer rechtfertigen. Ob diese Modellmodifikation dennoch in

Betracht gezogen wird, sollte erst nach Sichtung der Ergebnisse bezüglich der anderen Zielgruppen entschieden werden.

Für das Messmodell «Synchronisation mit Schulischen Heilpädagog:innen» lauten die Modellgütekriterien wie folgt:  $\chi^2 = 25.408$  (df = 2, p < .001), CFI = .927, RMSEA = .236 [.160, .322] und SRMR = .038. Einzig der SRMR liefert einen zufriedenstellenden Wert, die anderen Werte sind auffällig. Die Faktorladungen werden alle signifikant und sind auch relativ hoch, aber aus den angegebenen Modellmodifikationen wird deutlich, dass das Messmodell Schwächen hat. Es werden vier Korrelationen angegeben, die eine Verbesserung des  $\chi^2$ -Werts von bis zu 20.177 Punkten ermöglichen. Die Korrelationen sind jedoch z.T. positiv (syn3\_sh mit syn2\_sh sowie syn4\_sh mit syn1\_sh) und z.T. negativ (syn3\_sh mit syn1\_sh sowie syn4\_sh mit syn2\_sh), was innerhalb der Skala nur schwer zu rechtfertigen ist, da die Skala ausschließlich positiv gepolt ist und positive und negative Korrelationen nebeneinander nicht sinnvoll existieren können. Außerdem stehen diese Ergebnisse denjenigen von «Synchronisation mit Regellehrpersonen» entgegen. Bevor wiederum Modellanpassungen erwogen werden, wird das Modell erst noch in Bezug auf die dritte Zielgruppe analysiert.

Trotz der deutlich eingeschränkten Stichprobengröße (n = 121) liefert dieses Messmodell («Synchronisation mit den Logopäd:innen») bessere Modellfitmaße als das vorhergehende:  $\chi^2 = 8.944$  (df = 2, p < .05), CFI = .966, RMSEA = .169 [.068, .289] und SRMR = .028. Der SRMR ist sehr gut und auch der CFI bewegt sich in einem akzeptablen Bereich. Die Faktorladungen werden alle signifikant und sind hoch. Modellmodifikationen werden keine angegeben.

Das Messmodell «Synchronisation» weist unterschiedlich gute Modellfitmaße auf, je nachdem wer die Zielgruppe dieser Form von Kooperation ist. V.a. bezüglich der Zielgruppe Schulische Heilpädagog:innen hat das Messmodell Schwächen. Dass es im Prinzip jedoch gut geeignet ist, ein Konstrukt «Synchronisation» zu erheben, zeigt sich bei der Zielgruppe Logopäd:innen, weswegen auch vorerst darauf verzichtet wird, das Messmodell zu modifizieren. Dies wird unter Umständen erst vorgenommen, wenn auch die Messinvarianz zwischen den Gruppen überprüft werden kann, was eine größere Stichprobe voraussetzt.

Das Messmodell ist in Abbildung 12 (S. 68) zusammengefasst.

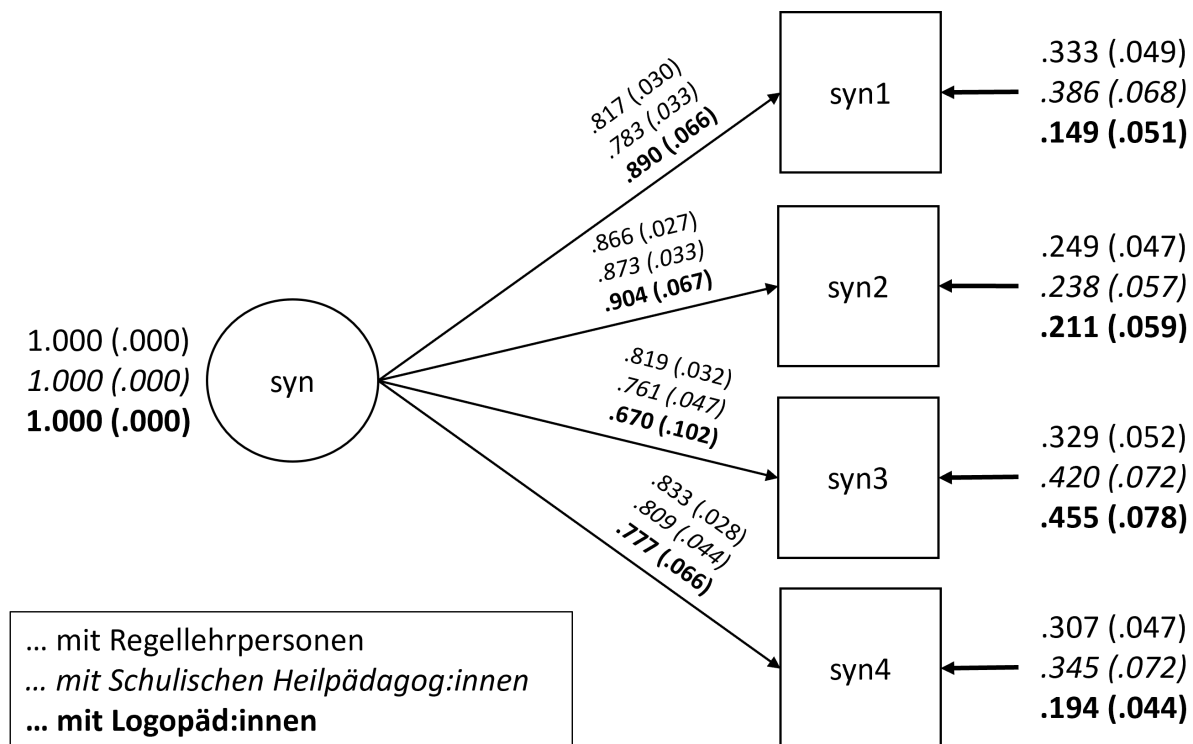


Abbildung 12: Messmodell von "syn" mit standardisierten Werteangaben für die jeweiligen Zielgruppen

### 7.3.10 CFA der Skala «Ko-Konstruktion»

Noch einmal müssen die CFAs unter Berücksichtigung der unterschiedlichen Zielgruppen von «Ko-Konstruktion» vorgenommen werden, was eine Reduktion der Stichprobengröße bedeutet und die geschilderten Konsequenzen hat. Das Messmodell «Ko-Konstruktion mit Regellehrpersonen» weist jedoch alles in allem zufriedenstellende Modellfitmaße auf:  $\chi^2 = 27.832$  (df = 14,  $p < .05$ ), CFI = .952, RMSEA = .067 [.029, .104] und SRMR = .051 (AIC = 3255.439). Die Werte sind insgesamt als gut, aber eben nicht als sehr gut einzuordnen. Auch wenn sich alle Faktorladungen signifikant von Null unterscheiden, fallen zwei Items auf, die ungenügende Faktorladungen haben: kok2\_lp (.161) und kok3\_lp (.268). Sollte sich dies auch in den Parallelversionen dieses Messmodells so zeigen, kann man erwägen, diese beiden Items zu eliminieren. Auch kok1\_lp hat eine eher geringe Faktorladung bzw. eine hohe Residualvarianz, bewegt sich aber noch in einem akzeptablen Bereich. Außerdem wird eine Korrelation der Items kok6\_lp («gemeinsames Erstellen von Arbeitsblättern») und kok1\_lp («Durchführung gegenseitiger Hospitationen») in den Modellmodifikationen als bedeutend eingestuft. Inhaltlich scheinen diese Items jedoch wenig miteinander gemein zu haben, weswegen auch diese Korrelation vorerst nicht geschätzt wird.

Das Messmodell «Ko-Konstruktion mit Schulischen Heilpädagog:innen» weist weniger gute Werte auf:  $\chi^2 = 39.584$  (df = 14,  $p < .001$ ), CFI = .924, RMSEA = .093 [.060, .128] und SRMR = .064 (AIC = 2859.995). Diese Werte sind insgesamt nicht annehmbar. In diesem Fall haben die Items kok1\_sh (.367), kok2\_sh (.320) und kok3\_sh (.375) zwar tiefe, aber noch akzeptable Werte. Unter den Modellmodifikationen wird dieses Mal eine bedeutende Korrelation der Items kok3\_sh («gemeinsamer Besuch von Weiterbildungen») und kok2\_sh («gemeinsame Durchführung von Elterngesprächen») angegeben, deren zusätzliche Schätzung den  $\chi^2$ -Wert um 19.576 Punkte verbessern würde. Wiederum lässt sich diese Korrelation inhaltlich jedoch nicht rechtfertigen. So wird das Modell vorerst nicht modifiziert. Das Messmodell «Ko-Konstruktion mit Logopäd:innen» hat wiederum schlechtere Werte:  $\chi^2 = 58.913$  (df = 14,  $p < .001$ ), CFI = .776, RMSEA = .162 [.121, .206] und SRMR = .081 (AIC = 1197.227). Diese sind

nicht akzeptabel. Bei Sichtung der Faktorladungen fallen wiederum die Items kok1\_lo (.292), kok2\_lo (.396) und kok3\_lo (.360) mit niedrigen, teils nicht annehmbaren Werten auf. Diese spielen auch eine Rolle bei den Modellmodifikationen, wo verschiedene Korrelationen angegeben werden: kok2\_lo («gemeinsame Durchführung von Elterngesprächen») korreliert dabei negativ mit kok5\_lo («gemeinsame Erprobung von sprachunterstützenden Maßnahmen»), was inhaltlich nicht zu rechtfertigen ist. Den höchsten Einfluss auf die Modellgüte nimmt jedoch die Korrelation von kok2\_lo mit kok1\_lo («Durchführung gegenseitiger Hospitationen»), deren Schätzung den  $\chi^2$ -Wert um 19.965 Punkte verbessern würde. Für eine Zusammenfassung der Messmodelle siehe Abbildung 13.

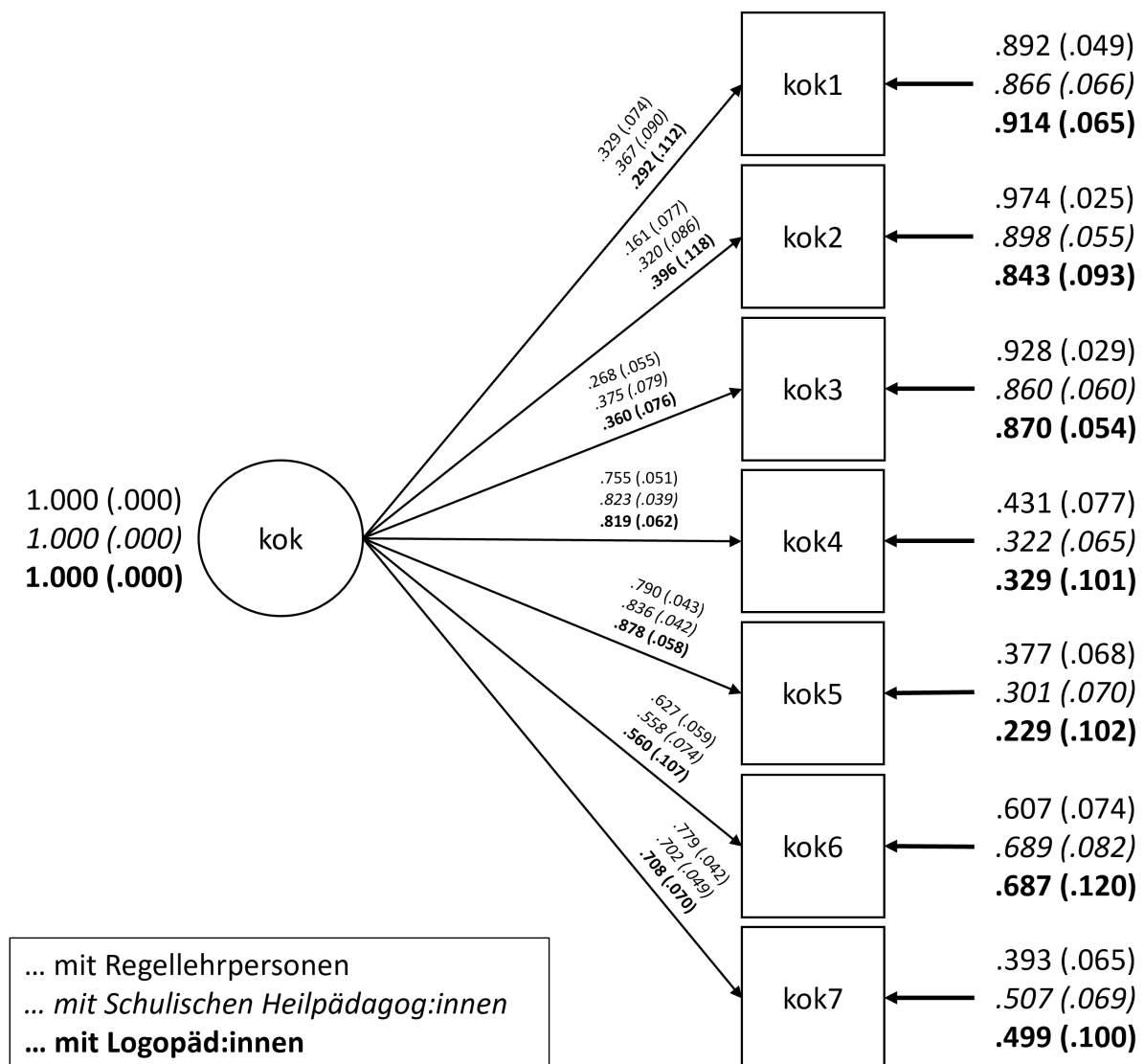


Abbildung 13: Messmodell von "kok" mit standardisierten Werteangaben für die jeweiligen Zielgruppen

In allen drei Sub-Messmodellen ist das Item kok2 mit negativen Einflüssen auf die Modellgüte aufgefallen, weswegen dieses eliminiert wird. Die Werte verbessern sich unmittelbar:

- Ko-Konstruktion mit Regellehrpersonen:  $\chi^2 = 15.379$  (df = 9, p = .081), CFI = .976, RMSEA = .057 [.000, .104] und SRMR = .037; AIC = 2807.188

- Ko-Konstruktion mit Schulischen Heilpädagog:innen:  $\chi^2 = 9.509$  (df = 9, p = .392), CFI = .998, RMSEA = .016 [.000, .080] und SRMR = .034; AIC = 2410.831
- Ko-Konstruktion mit Logopäd:innen:  $\chi^2 = 21.739$  (df = 9, p < .01), CFI = .917, RMSEA = .108 [.050, .166] und SRMR = .047; AIC = 1006.534; AIC = 1006.534

Zwar haben sich die Modellwerte in allen drei Fällen deutlich verbessert, dennoch sind die Werte für die Zielgruppe der Logopäd:innen eher noch ungenügend. Die Items kok1 und kok3 fallen nach wie vor mit niedrigen, teils ungenügend hohen Faktorladungen auf. Das Item kok1 ist zudem Teil von verschiedenen, inhaltlich nicht sinnvollen Korrelationen, weswegen dieses Item ebenfalls entfernt wird:

- Ko-Konstruktion mit Regellehrpersonen:  $\chi^2 = 3.020$  (df = 5, p = .697), CFI = 1.000, RMSEA = .000 [.000, .071] und SRMR = .017; AIC = 2290.018
- Ko-Konstruktion mit Schulischen Heilpädagog:innen:  $\chi^2 = 2.295$  (df = 5, p = .807), CFI = 1.000, RMSEA = .000 [.000, .060] und SRMR = .015; AIC = 2038.413
- Ko-Konstruktion mit Logopäd:innen:  $\chi^2 = 17.675$  (df = 5, p < .01), CFI = .915, RMSEA = .144 [.075, .220] und SRMR = .046; AIC = 839.425

Während sich die Messmodelle für die Zielgruppen der Regellehrpersonen und der Schulischen Heilpädagog:innen noch einmal verbessert haben, ist das Modell für die Zielgruppe der Logopäd:innen beinahe unverändert, und damit ungenügend. Das Item kok3 fällt nach wie vor mit einer tiefen Faktorladung auf und ist außerdem Teil einer Korrelation mit kok7 («gemeinsames Erarbeiten von Konzepten»), was wiederum inhaltlich nicht sinnvoll ist. So wird auch dieses Item entfernt:

- Ko-Konstruktion mit Regellehrpersonen:  $\chi^2 = 1.658$  (df = 2, p = .437), CFI = 1.000, RMSEA = .000 [.000, .127] und SRMR = .013; AIC = 1979.350
- Ko-Konstruktion mit Schulischen Heilpädagog:innen:  $\chi^2 = 1.243$  (df = 2, p = .537), CFI = 1.000, RMSEA = .000 [.000, .120] und SRMR = .010; AIC = 1680.091
- Ko-Konstruktion mit Logopäd:innen:  $\chi^2 = 3.887$  (df = 2, p = .143), CFI = .982, RMSEA = .089 [.000, .221] und SRMR = .026; AIC = 669.332

Mit dieser letzten Reduktion der Skala erreichen auch die Modellfitmaße bezüglich der Zielgruppe der Logopäd:innen sehr gute Werte. Die AIC-Werte haben sich kontinuierlich verbessert. Das neue, reduzierte Messmodell ist in Abbildung 14 (S. 71) dargestellt.

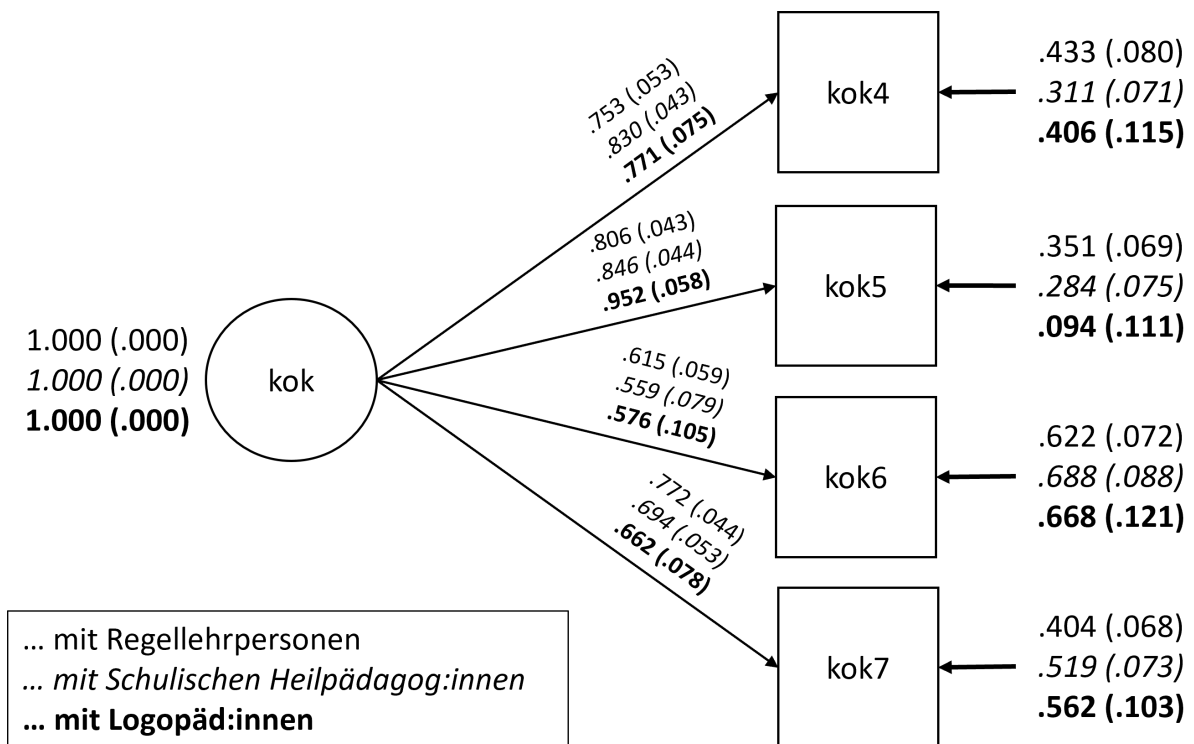


Abbildung 14: Messmodell von "kok" mit standardisierten Werteangaben für die jeweiligen Zielgruppen, ohne die Items kok1, kok2 und kok3

### 7.3.11 CFA der Skala «Team-Teaching»

Beim Messmodell «Team-Teaching» wird nun nicht mehr zwischen den verschiedenen Zielgruppen unterschieden, sodass wieder von der Gesamtstichprobe ausgegangen wird. Die Modellfitmaße bewegen sich im akzeptablen bis sehr guten Bereich:  $\chi^2 = 18.692$  (df = 5,  $p < .01$ ), CFI = .962, RMSEA = .100 [.054, .150] und SRMR = .031. Sämtliche Faktorladungen werden signifikant und bewegen sich im mittleren bis hohen Bereich. Es werden keine Modellmodifikationen ausgegeben. Von daher können keine batendasierten Modellmodifikationen gerechtfertigt werden. Das Messmodell wird, wie in Abbildung 15 (S. 72) dargestellt, beibehalten.

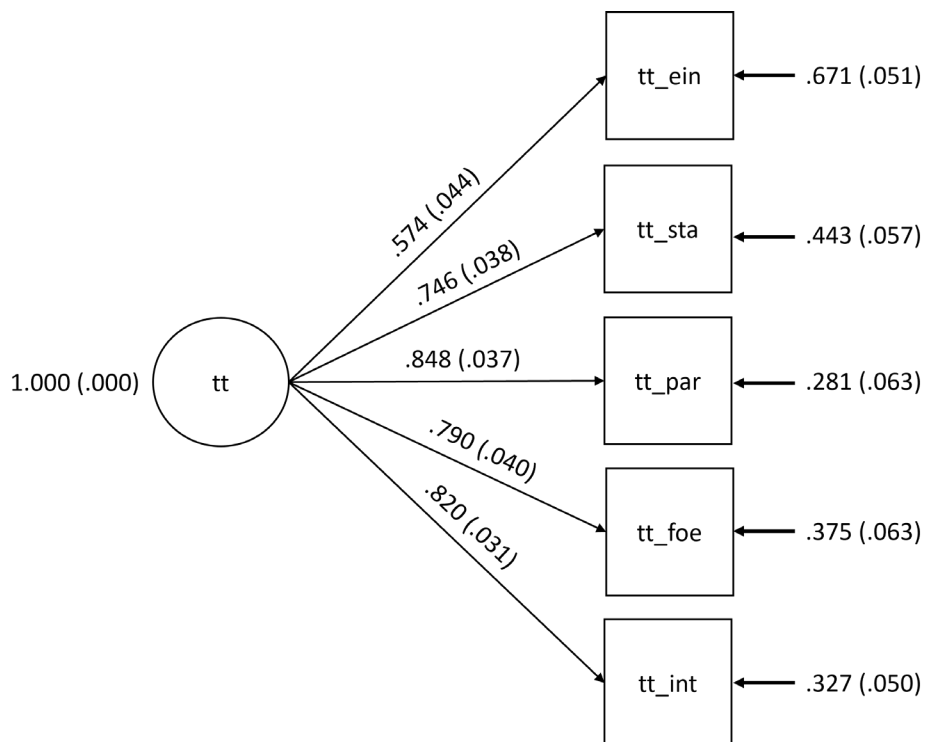


Abbildung 15: Messmodell von "tt" mit standardisierten Werteangaben

### 7.3.12 CFA der Skala «Intention»

Diese und die folgenden Skalen basieren alle auf vierstufigen Likert-Skalen, die gemäß der Fachliteratur als kategoriale Daten behandelt werden müssen (Reinecke, 2014, S. 109). Dementsprechend muss auch ein anderer Schätzer verwendet werden. Byrne (2010, S. 132) empfiehlt den WLSMV-Schätzer, da dieser auch bei nicht-normalverteilten und kleinen Stichproben zuverlässige Ergebnisse liefert.

Das Messmodell von «int» (s. Abb. 16, S. 73) hat keine zufriedenstellenden Modellgütekriterien:  $\chi^2 = 103.321$  (df = 5,  $p < .001$ ), CFI = .931, RMSEA = .267 [.224, .314] und SRMR = .062. Die signifikanten Faktorladungen sind jedoch allesamt hoch. Unter den Modellmodifikationen finden sich mehrere Korrelationen, von denen diejenige zwischen int5 («mehr Team-Teaching») und int1 («mehr unterrichtsintegrierte sprachunterstützende Maßnahmen») eine Verbesserung des  $\chi^2$ -Werts von 71.113 Punkten bedeuten würde. Auch inhaltlich ist dies sinnvoll, da die unterrichtsintegrierte Arbeit zumeist auch die Kooperation (bzw. das Team-Teaching) der beteiligten Fachkräfte erfordert.



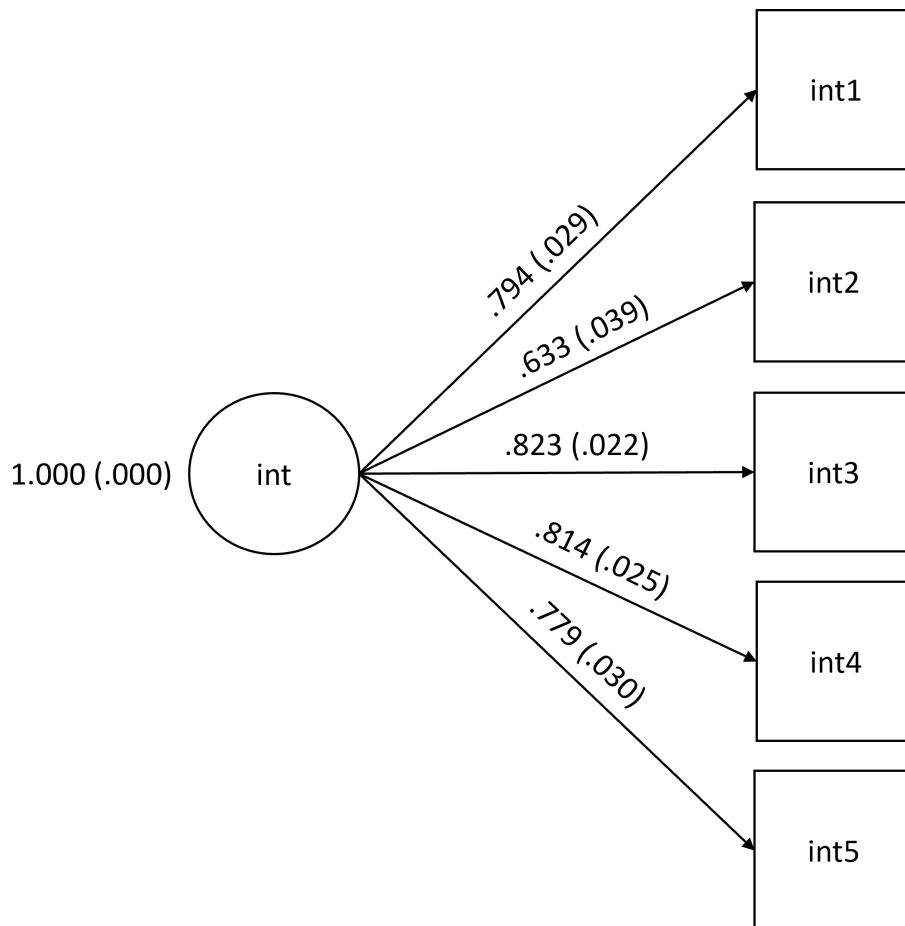


Abbildung 16: Messmodell von "int" mit standardisierten Werteangaben

Das entsprechende Modell hat die Werte:  $\chi^2 = 35.251$  (df = 4,  $p < .001$ ), CFI = .978, RMSEA = .169 [.120, .222] und SRMR = .035. Gemäß dem  $\chi^2$ -Differenztest (gerechnet mit der DIFFTEST-Option in Mplus) ist dieses Modell signifikant besser als das Ausgangsmodell ( $\Delta\chi^2 = 55.046$ ,  $\Delta df = 1$ ,  $p < .001$ ).

Wiederum wird eine Korrelation ermittelt (zwischen int4 [«mehr Ko-Konstruktion»] und int2 [«mehr Austausch»]), die aber inhaltlich nicht sinnvoll ist. Austausch und Ko-Konstruktion haben natürlich etwas gemeinsam, da es sich bei beiden Begriffen um Formen der Kooperation geht. Auf der anderen Seite sind dies jedoch unterschiedlich intensive und auch inhaltlich klar zu differenzierende Zusammenarbeitsformen, weswegen eine Schätzung der Korrelation nicht vorgenommen wird. Die Modellfitmaße des alternativen Modells (s. Abb. 17, S. 74) sind bereits (zumindest teilweise) sehr gut, sodass eine weitere Optimierung ohnehin nicht zwingend erforderlich ist.

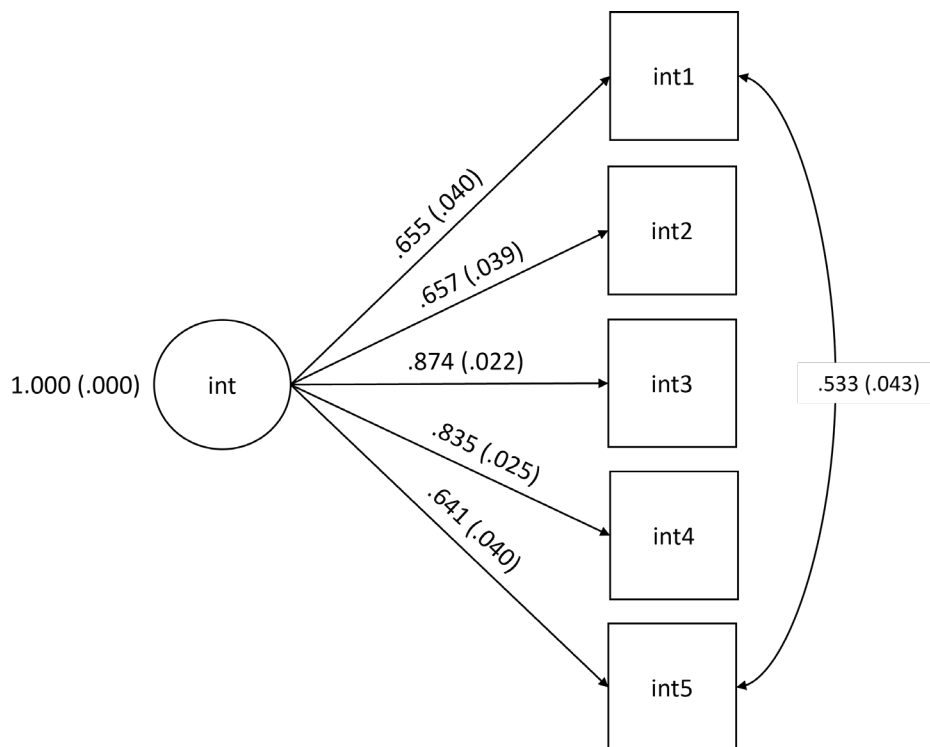


Abbildung 17: Messmodell von "int" mit standardisierten Werteangaben unter Einbezug der Korrelation von int1 und int5

Eine Anpassung der Itemformulierungen sollte jedoch erwogen werden. Wie sich gezeigt hat, sind die Rahmenbedingungen für die einzelnen Berufsgruppen sehr unterschiedlich ausgeprägt. So ist es gerade den Logopäd:innen nicht einfach möglich, die Zusammenarbeit zu intensivieren, da sie an wesentlich mehr Arbeitsorten tätig sind als die anderen Gruppen. Wenn man die Absicht jedoch eher hypothetisch formulieren würde («nach Möglichkeit würde ich gerne XY intensivieren»), werden die Intentionen allenfalls genauer erfasst.

### 7.3.13 CFA der Skala «Interesse»

Das Messmodell von «intr» (s. Abb. 18, S. 75) hat akzeptable bzw. sogar sehr gute Modellfitmaße, wenn man jene fokussiert, die gegenüber der Stichprobengröße weniger anfällig sind (CFI und SRMR):  $\chi^2 = 60.735$  (df = 14,  $p < .001$ ), CFI = .989, RMSEA = .110 [.082, .139] und SRMR = .029. Die Faktorladungen sind sämtlich signifikant und hoch. Unter den Modellmodifikationen wird eine Korrelation der Items intr4 («Ich kooperiere gerne ...») und intr3 («Es macht mir Spass ... zu kooperieren») ausgewiesen, die den  $\chi^2$ -Wert um 42.492 Punkte verbessern könnten. Dies klingt sinnvoll, da diese Items inhaltlich sehr ähnlich sind und nur schwer differenziert werden können.

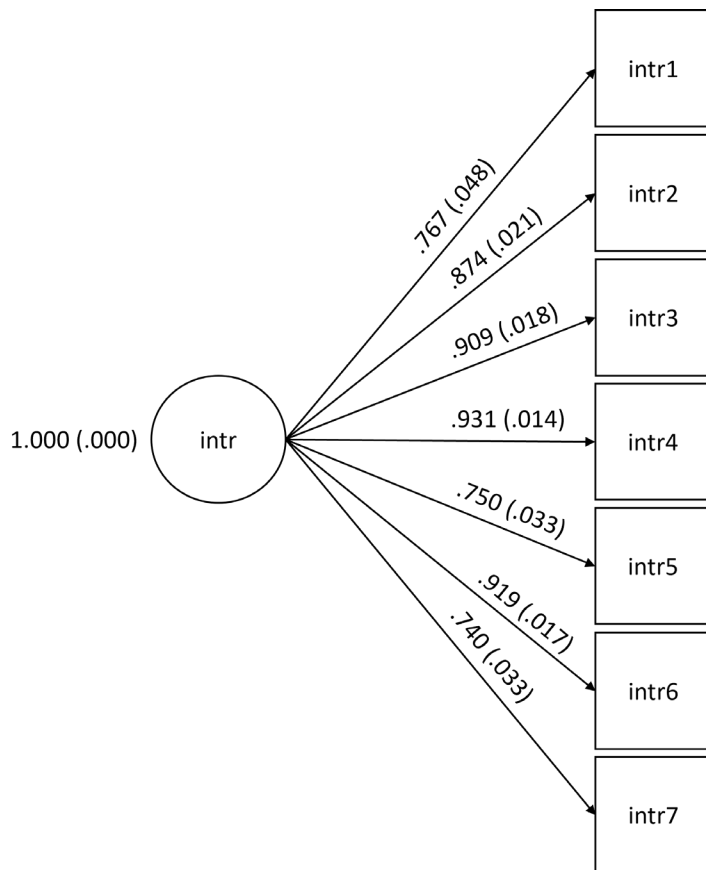


Abbildung 18: Messmodell von "intr" mit standardisierten Werteangaben

Das alternative Modell hat die Werte  $\chi^2 = 20.919$  (df = 13, p = .075), CFI = .998, RMSEA = .047 [.000, .083] und SRMR = .018. Es weist somit durchweg sehr gute Modellfitmaße auf. Der entsprechende  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 19.764$ ,  $\Delta df = 1$ , p < .001) wird signifikant und spricht somit für das alternative, weniger restriktive Modell. Auch werden keine weiteren Modellmodifikationen empfohlen, weswegen vom Modell in Abbildung 19 (S. 76) ausgegangen wird.

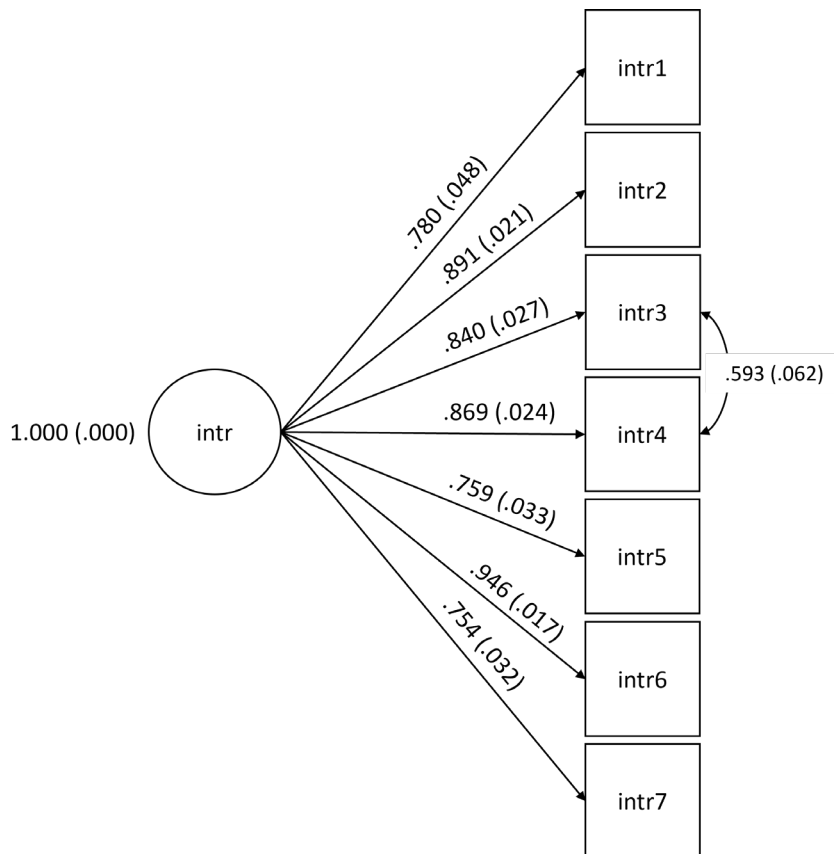


Abbildung 19: Messmodell von "intr" mit standardisierten Werteangaben unter Einbezug der Korrelation von intr3 und intr4

### 7.3.13 CFA der Skala «Kosten»

Die Modellfitmaße dieses Messmodells (s. Abb. 20, S. 77) bewegen sich zwischen gut (CFI), (gerade noch) annehmbar (SRMR) und schlecht ( $\chi^2$  und RMSEA):  $\chi^2 = 131.668$  (df = 20,  $p < .001$ ), CFI = .962, RMSEA = .142 [.120, .166] und SRMR = .063. Die Faktorladungen sind alle signifikant und relativ hoch. Unter den Modellmodifikationen befindet sich eine hohe Korrelation zwischen den Items kost1 («kostet unnötig viel Zeit») und kost2 («kostet unnötig viel Energie»), deren freie Schätzung den  $\chi^2$ -Wert um 41.594 Punkte verbessert. Die inhaltliche Nähe dieser beiden Items ist offensichtlich. So sind diese sowohl durch das Wort «unnötig» verbunden, aber auch durch die Termini «Zeit» und «Energie» zumindest umgangssprachlich eine große Nähe haben («Das kostet mich Zeit und Kraft») und nur selten auseinandergenommen betrachtet werden.

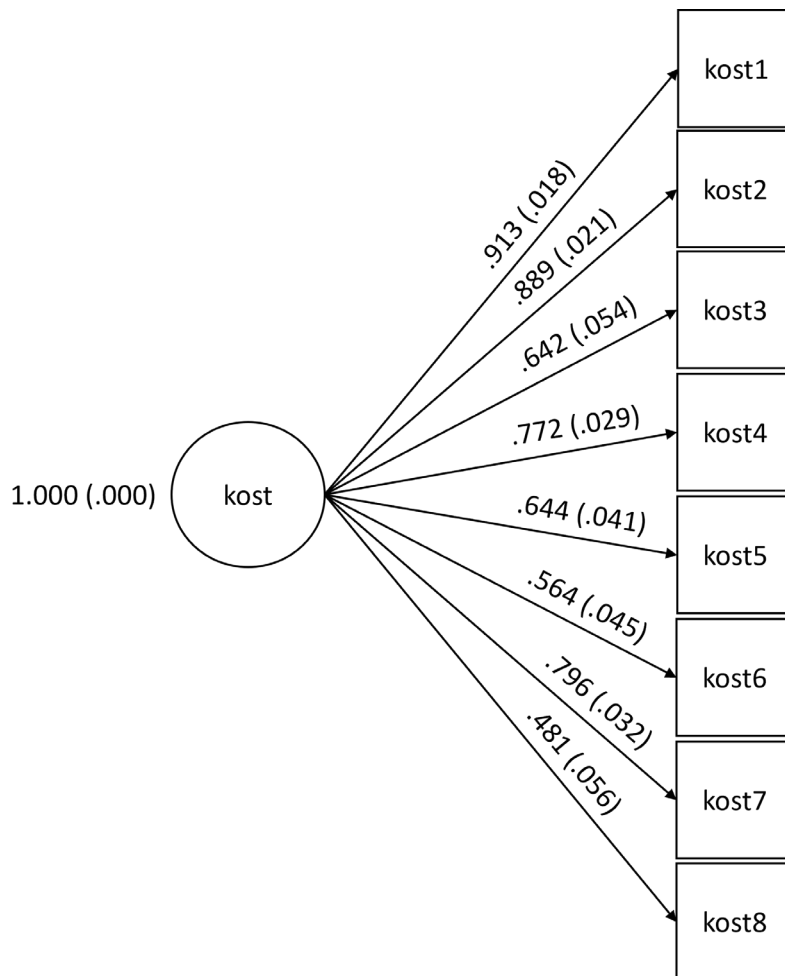


Abbildung 20: Messmodell von "kost" mit standardisierten Werteangaben

Ein Modell, in dem diese Korrelation frei geschätzt wird, hat die Werte  $\chi^2 = 95.376$  (df = 19,  $p < .001$ ), CFI = .974, RMSEA = .121 [.097, .145] und SRMR = .057. Diese sind schon besser, wie auch ein signifikanter  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 27.677$ ,  $\Delta df = 1$ ,  $p < .001$ ) belegt, aber noch nicht überzeugend. Eine weitere Korrelation zwischen kost8 («schlechte Erfahrungen mit Kooperation») und kost3 («Kooperation kostet Überwindung») verspricht eine Verbesserung des  $\chi^2$ -Werts von 28.414 Punkten und ist auch inhaltlich vertretbar: Wenn man schlechte Erfahrungen mit Kooperation gemacht hat, wird es einem künftig schwerer fallen, sich noch einmal darauf einzulassen.

Das entsprechende Messmodell (Modell 3) hat folgende Werte:  $\chi^2 = 70.389$  (df = 18,  $p < .001$ ), CFI = .982, RMSEA = .103 [.078, .129] und SRMR = .046. Der  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 18.963$ ,  $\Delta df = 1$ ,  $p < .001$ ) wird wiederum signifikant und spricht somit für das Beibehalten des Modells 3.

Wiederum findet sich jedoch eine inhaltlich plausible Korrelation in den Modellmodifikationen: kost5 («zeitliche Anforderungen stören private Verhältnisse») und kost4 («Zeit für wichtige Dinge muss eingespart werden»). In beiden Fällen geht es darum, dass Kooperation Zeit beansprucht, die man lieber für andere (wichtige/private) Dinge aufwenden würde. Das neue, vierte Messmodell hat die Werte:  $\chi^2 = 55.402$  (df = 17,  $p < .001$ ), CFI = .987, RMSEA = .090 [.065, .118] und SRMR = .040. Auch dieser  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 10.758$ ,  $\Delta df = 1$ ,  $p < .01$ ) wird signifikant und spricht somit für das Beibehalten des Modells 4. Es werden noch einmal zwei Korrelationen ausgewiesen, die zu einer Verbesserung des  $\chi^2$ -Werts führen würden, aber inhaltlich nicht mehr plausibel sind. Von daher wird aktuell von Messmodell 4 ausgegangen (s. Abb. 21, S. 78).

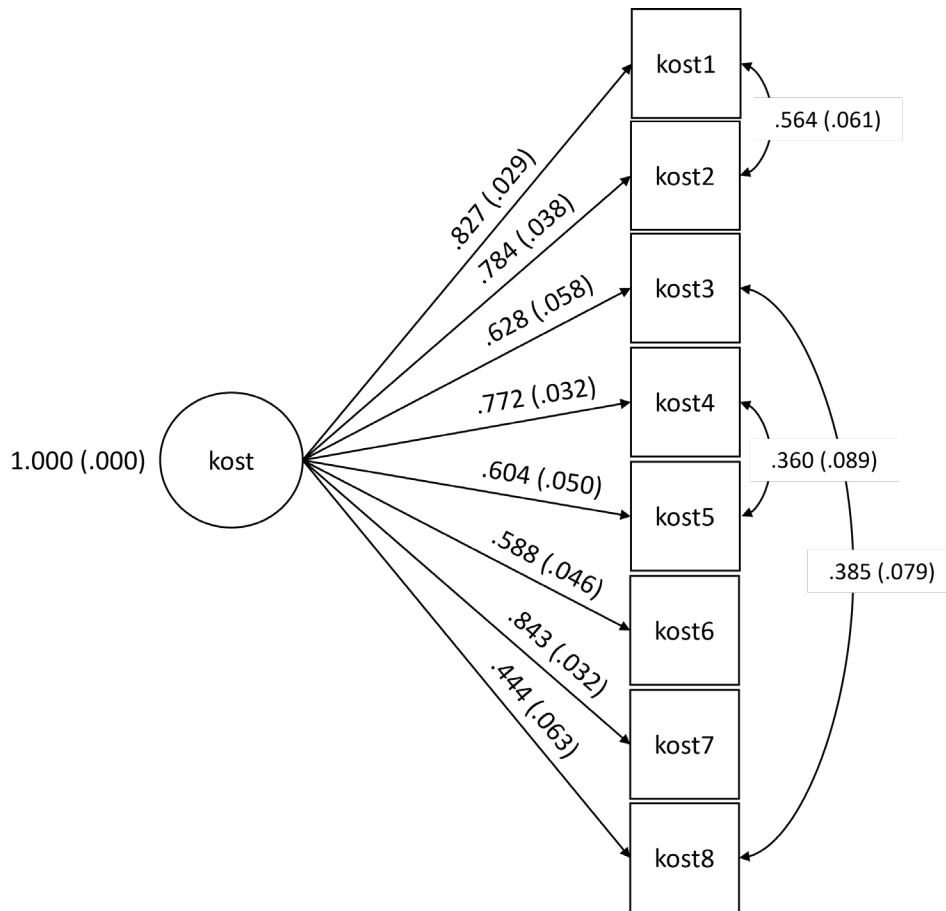


Abbildung 21: Messmodell von "kost" mit standardisierten Werteangaben unter Einbezug der Korrelation von kost1 und kost2, kost3 und kost8 und kost4 und kost5

### 7.3.14 CFA der Skala «Nützlichkeit für die professionelle Entwicklung»

Das Messmodell «nu\_fk» (s. Abb. 22, S. 79) liefert zufriedenstellende Werte:  $\chi^2 = 162.872$  (df = 44,  $p < .001$ ), CFI = .969, RMSEA = .099 [.083, .116] und SRMR = .045. Die Faktorladungen werden alle signifikant und sind hoch. Unter den Modellmodifikationen findet sich v.a. eine Korrelation, die den  $\chi^2$ -Wert um 59.008 Punkte senken würde: nu\_fk10 («Probleme besprechen, die man alleine nicht lösen kann») mit nu\_fk9 («bei Problemen nicht auf sich alleine gestellt sein»). Auch hier zeigt sich wiederum eine große inhaltliche Nähe darin, dass man Probleme nicht alleine bewältigen muss und dass auch so empfindet.

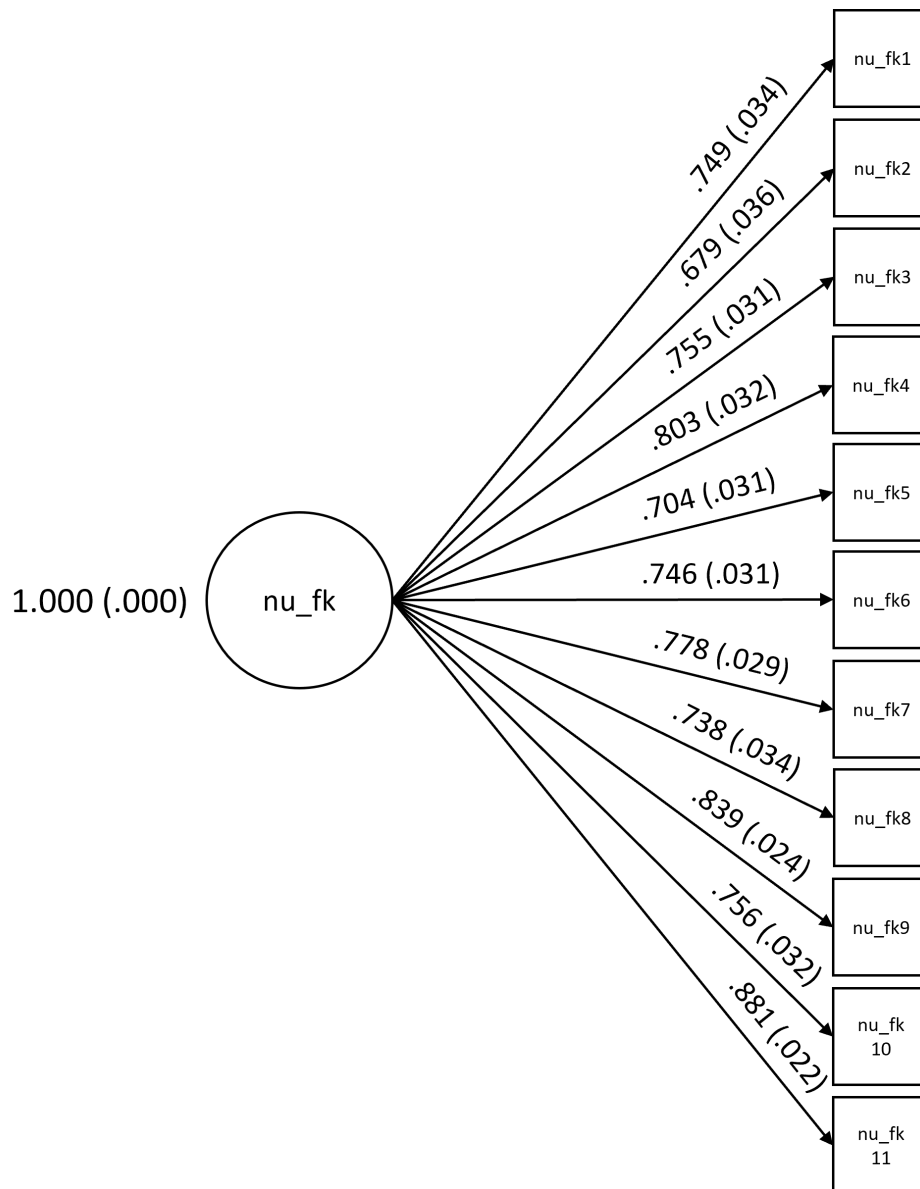


Abbildung 22: Messmodell von "nu\_fk" mit standardisierten Werteangaben

Die Schätzung der entsprechenden Korrelation führt zu folgenden Modellwerten:  $\chi^2 = 106.057$  (df = 43,  $p < .001$ ), CFI = .983, RMSEA = .073 [.056, .091] und SRMR = .037. Die Modellwerte haben sich deutlich verbessert, was durch den signifikanten  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 34.653$ ,  $\Delta df = 1$ ,  $p < .001$ ) bestätigt wird. Weitere Modellmodifikationen führen zu weniger bedeutenden Veränderungen und sind auch inhaltlich weniger gut zu rechtfertigen. So wird das alternative Modell (s. Abb. 23, S. 80) beibehalten

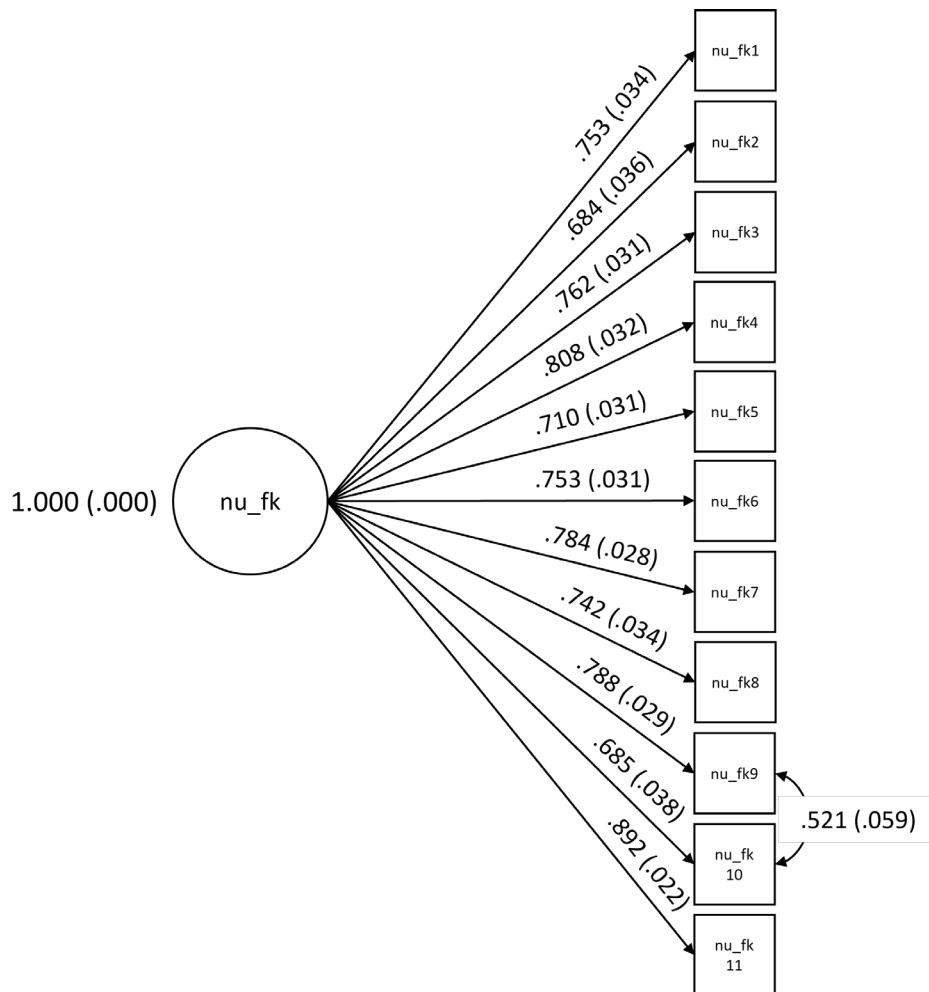


Abbildung 23: Messmodell von "nu\_fk" mit standardisierten Werteangaben unter Einbezug der Korrelation von nu\_fk9 und nu\_fk10

### 7.3.15 CFA der Skala «Nützlichkeit für die Schüler und Schülerinnen sowie die Schule»

Das Messmodell für «nu\_sus» (s. Abb. 24, S. 81) liefert schlechte ( $\chi^2$  und RSMEA), annehmbare (SRMR) bis gute Ergebnisse (CFI):  $\chi^2 = 250.840$  (df = 44,  $p < .001$ ), CFI = .965, RMSEA = .131 [.115, .147] und SRMR = .053. Die signifikanten Faktorladungen sind allesamt hoch einzustufen. Unter den Modellmodifikationen finden sich einige Korrelationen, von denen diejenige zwischen nu\_sus11 («verstärktes Zusammengehörigkeitsgefühl») und nu\_sus3 («besseres Arbeitsklima») eine Veränderung des  $\chi^2$ -Werts von 77.324 Punkten zur Folge haben würde. Im Kontext dieses Fragebogens, der sich fokussiert mit der multiprofessionellen Zusammenarbeit befasst, ist gut vorstellbar, dass ein unmittelbarer Zusammenhang zwischen dem Zusammengehörigkeitsgefühl und dem Arbeitsklima zwischen den verschiedenen Fachkräften besteht.



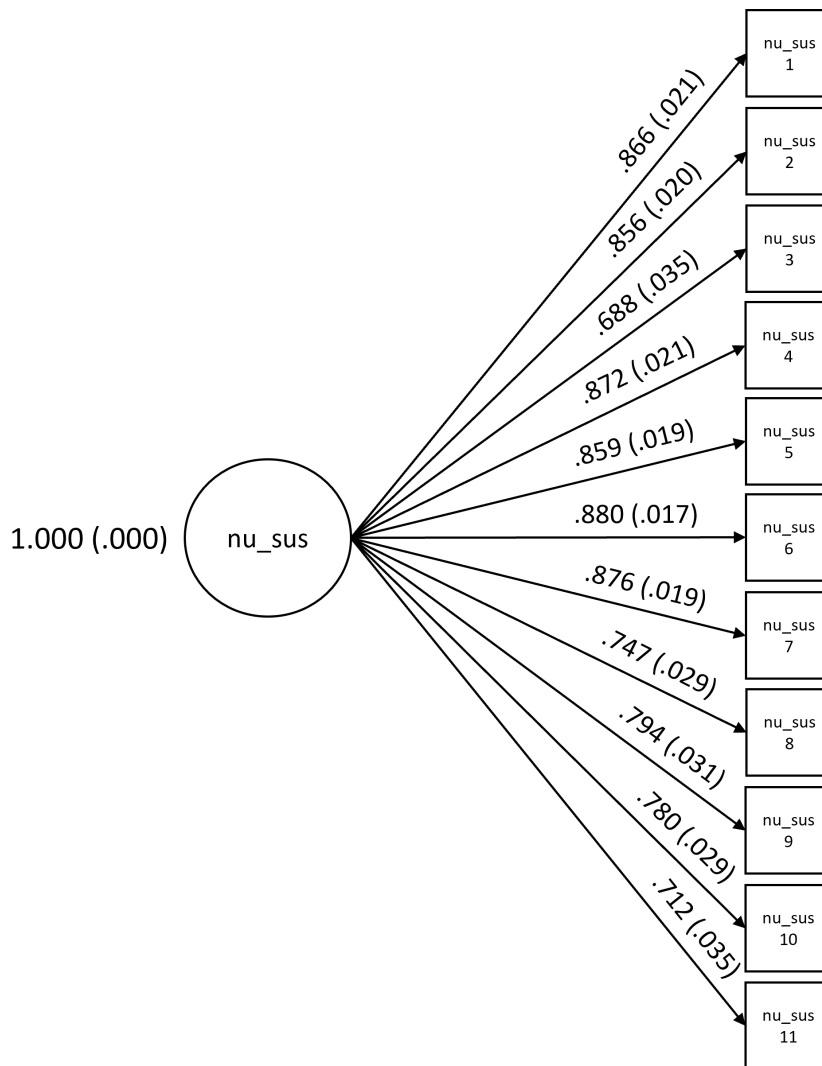


Abbildung 24: Messmodell von "nu\_sus" mit standardisierten Werteangaben

Ein Modell, in dem diese Korrelation frei geschätzt wird, hat folgende Werte:  $\chi^2 = 182.531$  (df = 43,  $p < .001$ ), CFI = .976, RMSEA = .109 [.093, .125] und SRMR = .044. Der CFI und der SRMR bewegen sich nun im sehr guten Bereich. Der  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 38.595$ ,  $\Delta df = 1$ ,  $p < .001$ ) belegt, dass sich die Modelle deutlich unterscheiden. Somit sollte das weniger restriktive Modell, in dem die Korrelation von nu\_sus11 und nu\_sus3 frei geschätzt wird, beibehalten werden.

Auch im alternativen Messmodell werden noch Korrelationen angegeben, die zu einer deutlichen Verbesserung des  $\chi^2$ -Werts von 26.097 Punkten führen können: nu\_sus10 («Weitergabe von Erfahrungen») und nu\_sus9 («fachlicher Austausch»). In beiden Fällen geht es also um einen Austausch/Transfer von Wissen bzw. Erfahrungen. Es scheint also gerechtfertigt, die entsprechende Korrelation zusätzlich schätzen zu lassen. Das entsprechende (dritte) Modell hat die Werte  $\chi^2 = 157.929$  (df = 42,  $p < .001$ ), CFI = .980, RMSEA = .100 [.084, .147] und SRMR = .041. Vergleicht man Modell 2 und 3 im Differenztest, fällt dieser wiederum zugunsten des neuen, dritten Modells aus ( $\Delta\chi^2 = 26.676$ ,  $\Delta df = 1$ ,  $p < .001$ ).

In diesem dritten Modell werden wiederum bedeutende Korrelationen angegeben, deren freie Schätzung zu einer weiteren Steigerung des  $\chi^2$ -Werts von bis zu 32.164 Punkten führen würde. Allerdings sind diese Korrelationen inhaltlich nicht mehr gut begründbar, weswegen auf weitere Modellmodifikationen verzichtet wird. Das endgültige Modell ist in Abbildung 25 (S. 82) dargestellt.

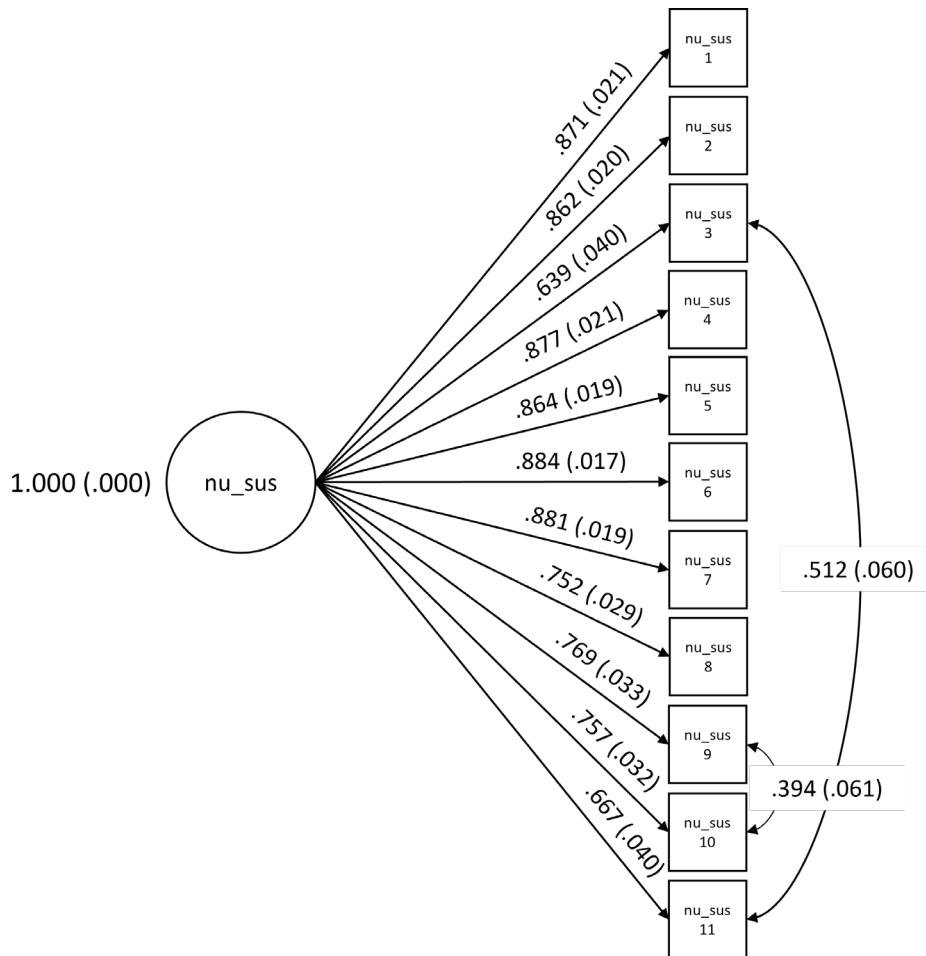


Abbildung 25: Messmodell von "nu\_sus" mit standardisierten Werteangaben unter Einbezug der Korrelation von nu\_sus3 und nu\_sus11 und nu\_sus9 und nu\_sus10

### 7.3.16 CFA der Skala «prozessorientierte Erfolgserwartung»

Das Messmodell «erf\_pro» ist mit drei Indikatoren gerade identifiziert, sodass keine Modellparameter ausgegeben werden. Bei den Faktorladungen fällt jedoch erf\_pro2 auf, das mit einem Wert von .910 fast mit einem Wert von 1 auf den Faktor lädt. Von daher wird diese Faktorladung auf 1 fixiert, um einen zusätzlichen Freiheitsgrad zu gewinnen. Das derart spezifizierte Modell hat folgende Werte:  $\chi^2 = 5.191$  (df = 1, p < .05), CFI = .995, RMSEA = .123 [.037, .236] und SRMR = .023. Das Messmodell in Abbildung 26 (S. 83) hat demnach zufriedenstellende bis sehr gute Werte. Modellmodifikationen werden keine angegeben.

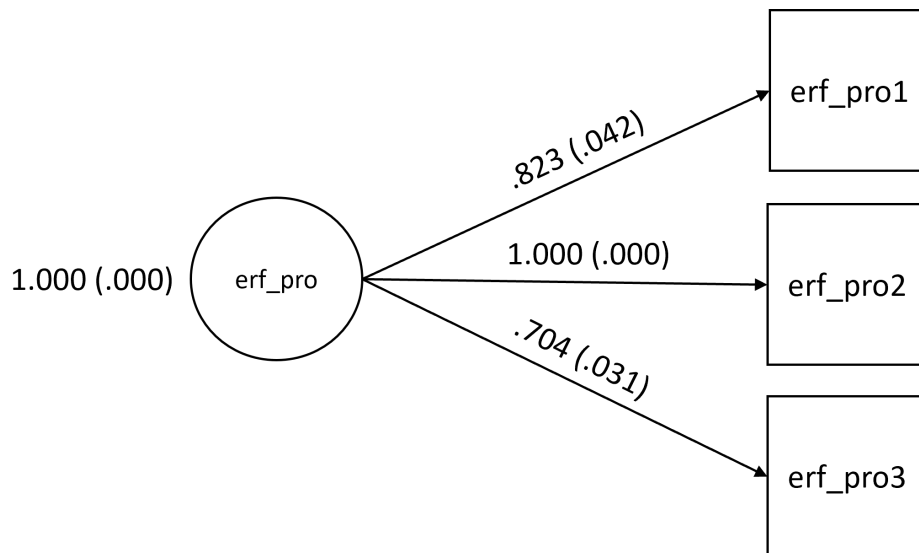


Abbildung 26: Messmodell von "erf\_pro" mit standardisierten Werteangaben

### 7.3.17 CFA der Skala «ergebnisorientierte Erfolgserwartung»

Das Messmodell «erf\_erg» liefert insgesamt zufriedenstellende bis sehr gute Modellfitmaße:  $\chi^2 = 7.560$  (df = 2, p < .05), CFI = .992, RMSEA = .100 [.032, .181] und SRMR = .030. Die Faktorladungen sind signifikant und können als hoch bis sehr hoch eingeordnet werden. Modellmodifikationen werden keine angegeben. Das Modell wird folglich wie in Abbildung 27 sichtbar zum Einsatz kommen.

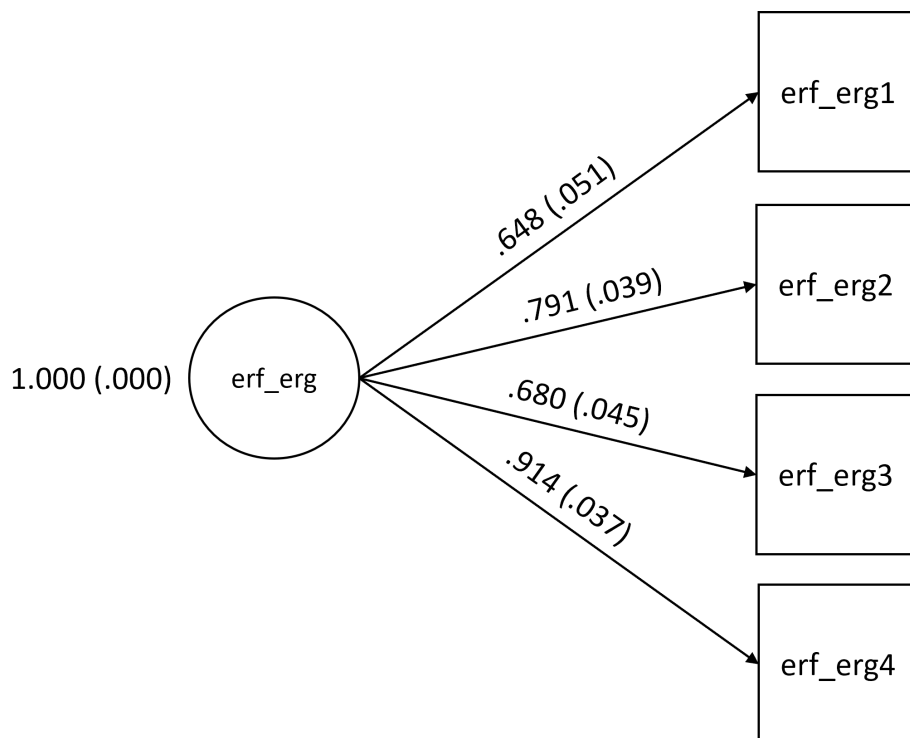


Abbildung 27: Messmodell von "erf\_erg" mit standardisierten Werteangaben

### 7.3.18 CFA der Skala «Zufriedenheit»

Das Messmodell von «zufr» (s. Abb. 28) liefert keine guten Werte:  $\chi^2 = 134.323$  (df = 20,  $p < .001$ ), CFI = .882, RMSEA = .144 [.121, .167] und SRMR = .083 (AIC = 4572.531). V.a. das Item zufr8 fällt mit einer ungenügend großen Faktorladung (.194) auf und wird somit eliminiert.

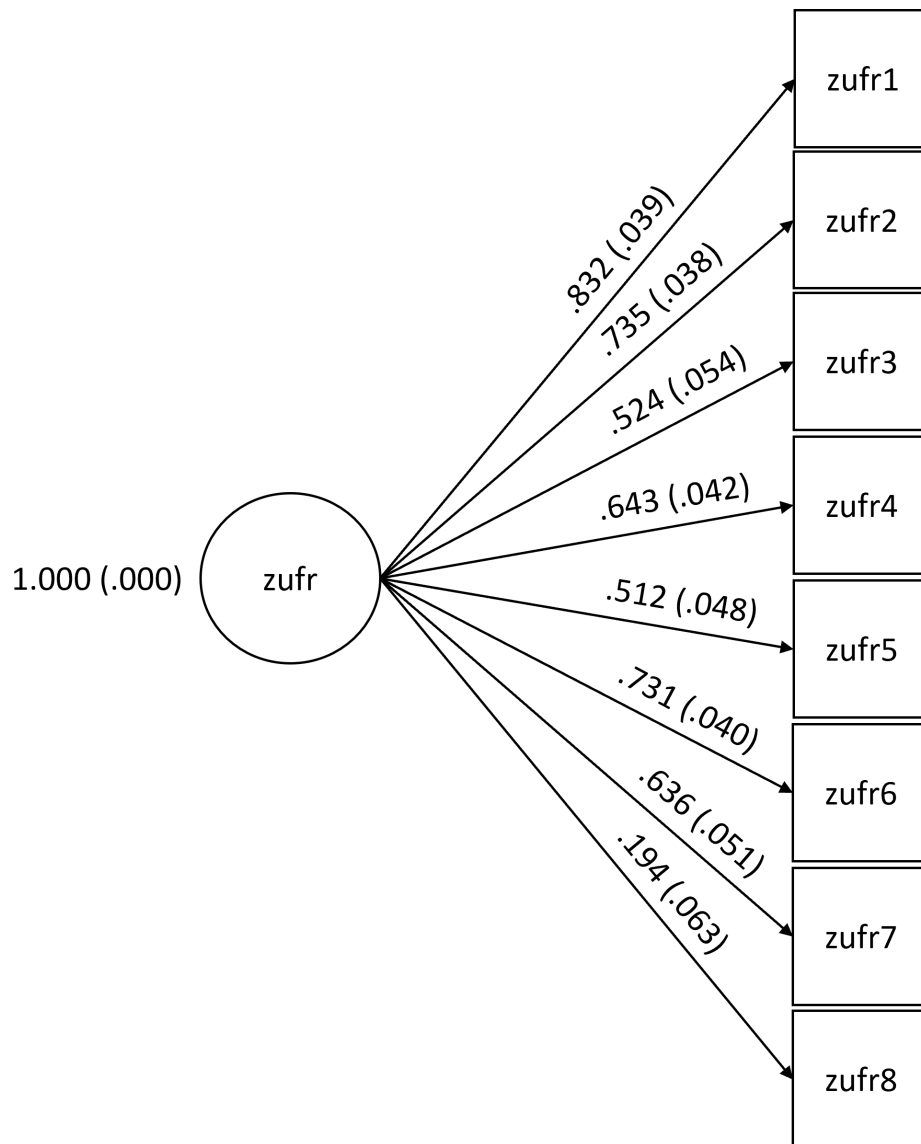


Abbildung 28: Messmodell von "zufr" mit standardisierten Werteangaben

Das Modell ohne zufr8 hat folgende Modellfitmaße:  $\chi^2 = 123.865$  (df = 14,  $p < .001$ ), CFI = .884, RMSEA = .168 [.142, .196] und SRMR = .087 (AIC = 3881.038). Die AIC-Werte, welche unter Verwendung des MLR-Schätzers und im Rahmen einer Monte-Carlo-Simulation errechnet wurden, sprechen für das Modell ohne zufr8. Es sind alle Faktorladungen genügend groß (und signifikant), doch finden sich in den Modellmodifikationen Hinweise auf bedeutende Korrelationen zwischen den Indikatoren. Die Items zufr7 («ausreichend Material zur Verfügung») und zufr6 («ausreichend räumliche Ressourcen») könnten auch als «ausreichende materielle und räumliche Ressourcen» zusammengefasst werden und korrelieren daher stark. Die freie Schätzung dieser Korrelation führt zu einer Verbesserung des  $\chi^2$ -Werts von 54.874 Punkten.

Das entsprechend modifizierte Messmodell hat folgende Parameter:  $\chi^2 = 74.933$  (df = 13,  $p < .001$ ), CFI = .934, RMSEA = .131 [.103, .161] und SRMR = .064. Diese sind zwar noch immer ungenügend, aber schon sichtlich besser als die Werte des Ausgangsmodells, wie auch der signifikante  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 34.356$ ,  $\Delta df = 1$ ,  $p < .001$ ) bestätigt. Wiederum findet sich eine hohe Korrelation zwischen den Items zufr2 («Wertschätzung durch Andere») und zufr1 («Zufriedenheit mit Rolle im Team»), deren freie Schätzung den  $\chi^2$ -Wert um 23.935 Punkte verbessern würde. Beide Items zeichnen sich durch den Bezug auf andere Personen aus. Auch scheint es nachvollziehbar, dass wenn man Wertschätzung durch andere Personen erfährt, man auch eher zufrieden ist mit der eigenen Rolle im Team.

Von daher wird im dritten Modell die Schätzung dieser Korrelation mitberücksichtigt:  $\chi^2 = 55.839$  (df = 12,  $p < .001$ ), CFI = .954, RMSEA = .115 [.085, .146] und SRMR = .053. Diese Werte bewegen sich nun schon in einem akzeptablen Bereich. Der  $\chi^2$ -Differenztest ( $\Delta\chi^2 = 14.47$ ,  $\Delta df = 1$ ,  $p < .001$ ) bestätigt die Überlegenheit des dritten Messmodells. Die Modifikationsindices verweisen noch auf zwei weitere Korrelationen, die aber eher verhältnismäßig wenig zur Verbesserung des  $\chi^2$ -Werts beitragen bzw. auch inhaltlich eher schwer zu rechtfertigen sind. So korreliert zufr4 («Möglichkeiten zur Zusammenarbeit») mit zufr1 («Zufriedenheit mit Rolle im Team») – hier werden zwei unterschiedliche Aspekte vermisch (zeitliche Ressourcen für bzw. Gelegenheiten zur Zusammenarbeit und die Zufriedenheit mit der eigenen Rolle). Von daher wird auf weitere Anpassungen des Modells verzichtet und vom Messmodell in Abbildung 29 ausgegangen.

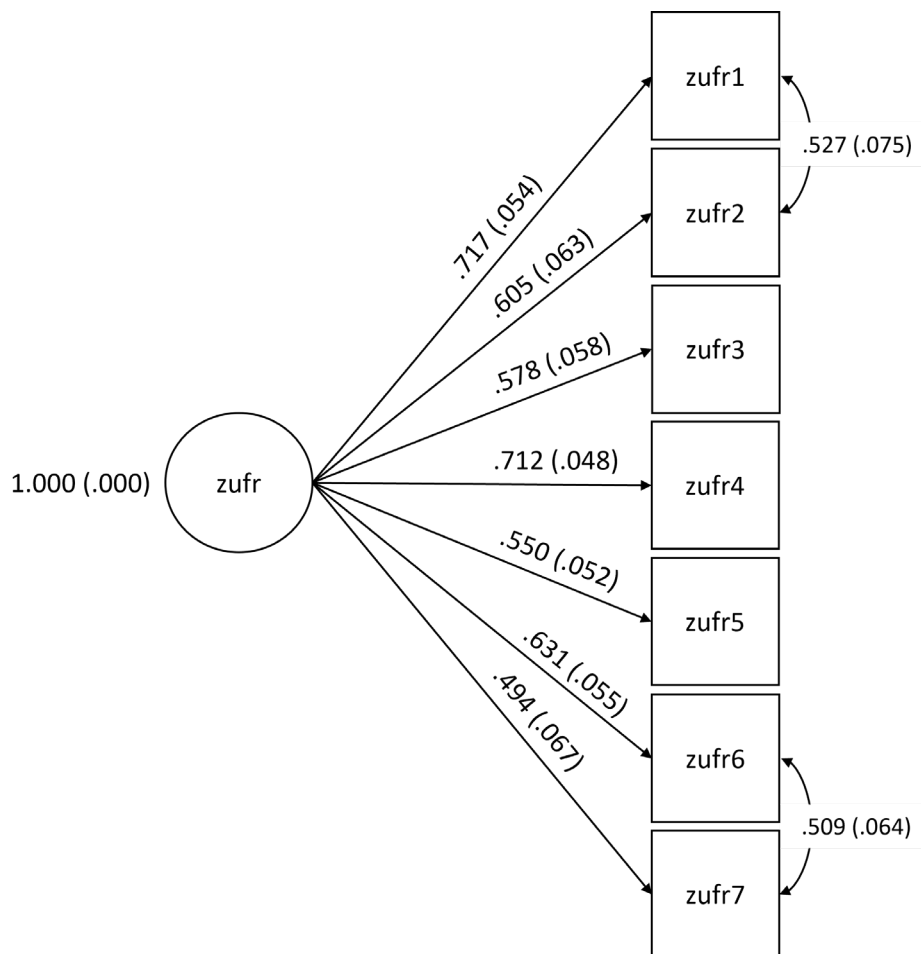


Abbildung 29: Messmodell von "zufr" mit standardisierten Werteangaben unter Einbezug der Korrelation von zufr1 und zufr2 und zufr6 und zufr7

### 7.3.19 Zusammenfassung der CFAs

Die Ergebnisse der konfirmatorischen Faktorenanalysen sind in Tabelle 69 (S. 87f.) noch einmal zusammengefasst. Wie auch in der Reliabilitätsanalyse haben sich auch hier bestimmte Subskalen als ungeeignet zur weiteren Verwendung erwiesen («sprachunterstützende Maßnahmen im Allgemeinen» und «sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers»), da mit den verwendeten Items kein bestimmtes bzw. sinnvolles Konstrukt gemessen werden kann.

Des Weiteren finden sich Subskalen, die abgeändert werden sollten, um das Konstrukt in beabsichtigter Art und Weise messen zu können. Bei den Skalen «(unterrichtsintegrierte) sprachunterstützende Maßnahmen bei Auffälligkeiten» sollte das Item, das den Redefluss erfasst, eliminiert werden, da dies keinen interpretierbaren Informationsgehalt hat und die Subskala ohne diese Items kohärenter ist. Die Skala «unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Unterrichtsformen» liefert v.a. in der Subgruppe der Schulischen Heilpädagog:innen nicht-akzeptable Ergebnisse, weswegen eine Umformulierung der Skala in Erwägung gezogen werden sollte. Die Skalen zur Messung von «Ko-Konstruktion» funktionieren ohne die Items kok1-3 besser, da diese ungenügende Faktorladungen aufwiesen. Die Subskala «Intention» sollte v.a. aus inhaltlichen Gründen überarbeitet werden, da beobachtet werden konnte, dass v.a. Logopäd:innen aufgrund der Rahmenbedingungen gewisse Kooperationsformen kaum wahrnehmen können, auch wenn sie dies wollten. Von daher wird eine Formulierung der Items im Konjunktiv (z.B. «Ich würde mich in Zukunft gerne mehr mit den anderen Fachkräften austauschen») angestrebt. In der Subskala «Zufriedenheit» schließlich ist das Item zufr8 mit einer ungenügenden Faktorladung aufgefallen, welches somit eliminiert wird. Der entsprechende Inhalt («Qualifizierung zur Zusammenarbeit») wird an anderer Stelle erfragt werden.

Die Skalen, die zum Einsatz kommen werden, haben z.T. hervorragende, z.T. aber auch gemischte Werte. Der Fokus wird allerdings auf jene Modellfitmaße gelegt (CFI und SRMR), die auch bei kleinen Stichproben die Güte eines Messmodells zuverlässig belegen können. Dennoch finden sich auch solche Skalen darunter, deren Modellfitmaße noch nicht zufriedenstellend sind, aber für die Erhebung des Gesamtbilds von entscheidender Bedeutung sind. Diese werden ebenfalls beibehalten, wobei zum aktuellen Zeitpunkt noch fraglich ist, ob die Daten, die mit diesen Subskalen erhoben werden, letzten Endes auch in den Analysen berücksichtigt werden. Die konfirmatorischen Faktorenanalysen haben jedoch einen Eindruck davon vermitteln können, welche Konstrukte einwandfrei erhoben werden können, und welche Messungen tendenziell problembehaftet sind. Dieses Wissen wird in künftigen Untersuchungen unter Einsatz der hier analysierten Subskalen einfließen.

Tabelle 69: CFAs der Subskalen des Fragebogens, Angaben zu den Modellfitmaßen und Kommentar zur Verwendung

Subskala	Modellfitmaße	Verwendung
<b>sprachunterstützende Maßnahmen im Allgemeinen</b>	$\chi^2$ : 81.881 (df = 9, p < .001), CFI = .717, RMSEA = .171 [.138, .206] und SRMR = .105	nein, kein sinnvolles Konstrukt messbar
<i>sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten</i>	$\chi^2$ -Wert: 11.463 (df = 5, p < .05), CFI = .965, RMSEA = .069 [.011, .122] und SRMR = .033	ohne rfl_all
<i>unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten</i>	$\chi^2$ = 1.196, df = 5, p = .987, einem CFI = 1.000, RMSEA = .000 [.000, .013] und SRMR = .003	ohne rfl_kl
<i>unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Unterrichtsformen</i>	$\chi^2$ = 11.218 (df = 2, p < .01), CFI = .973, RMSEA = .130 [.063, .208] und SRMR = .024	ja, aber Umformulierung nötig
unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Fachbereichen	$\chi^2$ = 67.746 (df = 26, p < .001), CFI = .931, RMSEA = .086 [.062, .112] und SRMR = .065	ja, unter Vorbehalt
unterrichtsintegrierte sprachunterstützende Maßnahmen in verschiedenen Erwerbsphasen	$\chi^2$ = 0.136 (df = 1, p = .713), CFI = 1.000, RMSEA = .000 [.000, .116] und SRMR = .005	ja
<b>sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers</b>	$\chi^2$ = 1.103 (df = 2, p = .576), CFI = 1.000, RMSEA = .000 [.000, .100] und SRMR = .015	nein, «Heywood Cases»
Austausch mit den Regellehrpersonen	$\chi^2$ = 8.442 (df = 2, p < .05), CFI = .971, RMSEA = .122 [.046, .212] und SRMR = .036	ja
Austausch mit den Schulischen Heilpädagog:innen	$\chi^2$ = 6.029 (df = 2, p < .05), CFI = .988, RMSEA = .098 [.005, .192] und SRMR = .021	ja
Austausch mit den Logopäd:innen	$\chi^2$ = .303 (df = 2, p = .859), CFI = 1.000, RMSEA = .000 [.000, .096] und SRMR = .003	ja
Synchronisation mit den Regellehrpersonen	$\chi^2$ = 12.633 (df = 2, p < .01), CFI = .972, RMSEA = .156 [.082, .243] und SRMR = .025	ja
Synchronisation mit den Schulischen Heilpädagog:innen	$\chi^2$ = 25.408 (df = 2, p < .001), CFI = .927, RMSEA = .236 [.160, .322] und SRMR = .038	ja
Synchronisation mit den Logopäd:innen	$\chi^2$ = 8.944 (df = 2, p < .05), CFI = .966, RMSEA = .169 [.068, .289] und SRMR = .028	ja

Tabelle 69 (forts.): CFAs der Subskalen des Fragebogens, Angaben zu den Modellfitmaßen und Kommentar zur Verwendung

Subskala	Modellfitmaße	Verwendung
<i>Ko-Konstruktion mit den Regellehrpersonen</i>	$\chi^2 = 1.658$ (df = 2, p = .437), CFI = 1.000, RMSEA = .000 [.000, .127] und SRMR = .013	ja, ohne kok1-3
<i>Ko-Konstruktion mit den Schulischen Heilpädagog:innen</i>	$\chi^2 = 1.243$ (df = 2, p = .537), CFI = 1.000, RMSEA = .000 [.000, .120] und SRMR = .010	ja, ohne kok1-3
<i>Ko-Konstruktion mit den Logopäd:innen</i>	$\chi^2 = 3.887$ (df = 2, p = .143), CFI = .982, RMSEA = .089 [.000, .221] und SRMR = .026	ja, ohne kok1-3
Team-Teaching	$\chi^2 = 18.692$ (df = 5, p < .01), CFI = .962, RMSEA = .100 [.054, .150] und SRMR = .031	ja
<i>Intention</i>	$\chi^2 = 35.251$ (df = 4, p < .001), CFI = .978, RMSEA = .169 [.120, .222] und SRMR = .035	ja, aber Umformulierung nötig
Interesse	$\chi^2 = 20.919$ (df = 13, p = .075), CFI = .998, RMSEA = .047 [.000, .083] und SRMR = .018	ja
Kosten	$\chi^2 = 55.402$ (df = 17, p < .001), CFI = .987, RMSEA = .090 [.065, .118] und SRMR = .040	ja
Nützlichkeit für die professionelle Entwicklung	$\chi^2 = 106.057$ (df = 43, p < .001), CFI = .983, RMSEA = .073 [.056, .091] und SRMR = .037	ja
Nützlichkeit für die Schüler und Schülerinnen sowie die Schule	$\chi^2 = 157.929$ (df = 42, p < .001), CFI = .980, RMSEA = .100 [.084, .147] und SRMR = .041	ja
prozessorientierte Erfolgserwartung	$\chi^2 = 5.191$ (df = 1, p < .05), CFI = .995, RMSEA = .123 [.037, .236] und SRMR = .023	ja
ergebnisorientierte Erfolgserwartung	$\chi^2 = 7.560$ (df = 2, p < .05), CFI = .992, RMSEA = .100 [.032, .181] und SRMR = .030	ja
<i>Zufriedenheit</i>	$\chi^2 = 55.839$ (df = 12, p < .001), CFI = .954, RMSEA = .115 [.085, .146] und SRMR = .053	ja, ohne zufr8



## 8 Analyse der Antworttendenzen

Im Folgenden werden verschiedene mögliche Satisficing-Tendenzen der Teilnehmenden analysiert. Gemäß Bogner und Landrock (2015) müssen dabei die folgenden Arten unterschieden werden:

- Soziale Erwünschtheit
- Akquieszenz
- Tendenz zur Mitte und zu Milde/Härte
- Interviewereffekte
- Anwesenheit dritter Personen
- Effekte der Fragereihenfolge
- Reihenfolge der Antwortvorgaben

Interviewereffekte und die Anwesenheit dritter Personen haben in der vorliegenden Untersuchung keine Rolle gespielt bzw. können nicht kontrolliert werden, da die Befragung online und nicht unter Beobachtung durchgeführt wurde. Die übrigen Antworttendenzen sollen jedoch kurz untersucht werden. Dabei stammen alle Hinweise und Empfehlungen aus Bogner und Landrock (2015).

### 8.1 soziale Erwünschtheit

Mit sozialer Erwünschtheit wird das Phänomen beschrieben, dass sich die befragten Personen möglichst positiv darstellen möchten. D.h., dass – orientiert an sozialen Normen – erwünschte Verhaltensweisen übertrieben häufig und unerwünschte Verhaltensweisen übertrieben selten genannt werden. Soziale Erwünschtheit in Befragungen tritt v.a. dann auf, wenn sich die befragte Person und die interviewende Person räumlich nahe sind. Durch die Nennung von vermeintlich erwünschten Antworten schützt sich die befragte Person vor einer möglichen Bloßstellung. In Online-Befragungen herrscht jedoch eine große Distanz vor, was dazu führt, dass das Phänomen der sozialen Erwünschtheit seltener auftritt.

Letzteres ist hier der Fall – durch den Online-Befragungsmodus ist eine große räumliche bzw. soziale Distanz gegeben und die Tendenz zu gefälligen Antworten sollte dadurch minimiert werden. Dennoch besteht die Möglichkeit, mithilfe einer Kurzsкала – bestehend aus sechs Items – die soziale Erwünschtheit zu messen und anschließend statistisch zu kontrollieren (s. KSE-G; Kemper et al., 2014). Dies ist im vorliegenden Fall jedoch nicht zwingend nötig. Erstens ist die soziale Distanz ohnehin hoch und zweitens handelt es sich bei den erfragten Inhalten nicht um sensible Themen, bei denen eine positive Selbstdarstellung unmittelbar zu erwarten wäre.

Das Fazit lautet also dahingehend, dass die soziale Erwünschtheit bei den vorliegenden Antworten eher keinen Einfluss ausgeübt hat.

### 8.2 Akquieszenz

Akquieszenz kann auch als Zustimmungstendenz bezeichnet werden. Das bedeutet, dass Fragen, unabhängig vom Inhalt, mit «trifft zu» oder «sehr wahrscheinlich» beantwortet werden, um Beispiele aus dem vorliegenden Fragebogen aufzugreifen.

Akquieszenten sind v.a. Personen mit gering ausgeprägten kognitiven Fähigkeiten und geringer formaler Bildung (Bogner & Landrock, 2015, S. 4). Da es sich bei der untersuchten Stichprobe aber ausschließlich um Personen mit Hochschul- bzw. äquivalenten Abschlüssen (Regellehrpersonen, Schulische Heilpädagog:innen, Logopäd:innen) handelt, kann dieser Umstand wohl ausgeschlossen werden.

Zustimmungsverhalten kann aber auch durch die interviewende Person provoziert werden, v.a. wenn diese eher schnell vorgeht und wenig Zeit zum Überlegen bietet. Dies ist ebenfalls auszuschließen, da die Befragung, wie bereits erwähnt, online durchgeführt und auch kein Zeitlimit gegeben wurde. Die Personen konnten somit in ihrem eigenen Tempo vorgehen.

Eine weitere Maßnahme, um der Akquieszenz entgegenzuwirken, sind abwechselnd positiv und negativ formulierte Items, die dann entsprechend bestätigend oder ablehnend beantwortet werden müssen, um ein konsistentes Antwortverhalten zu zeigen. Dies wurde im vorliegenden Fragebogen in der Regel nicht umgesetzt. Die Skala «Kosten» jedoch ist generell negativ formuliert. Personen mit einer positiven Einstellung zur Zusammenarbeit müssen die Items auf dieser Skala (und nur auf dieser Skala) also ablehnen. Bei der Analyse der Daten fällt unmittelbar auf, dass dies von den Befragten auch so gemacht wurde: wer in den übrigen Skalen zur Einstellung positiv geantwortet hat, hat bei den «Kosten» negativ geurteilt und umgekehrt.

Die ermittelten Antworten dieser Befragung sind somit wahrscheinlich weniger durch Akquieszenz beeinflusst.

### 8.3 Tendenz zur Mitte und zu Milde/Härte

Die Tendenz zur Mitte beschreibt das Verhalten, dass wenn man sich bei der Beantwortung einer Frage unsicher ist, man zur Wahl der neutralen Mitte der Skala tendiert. Die Tendenz zu Milde/Härte beschreibt das Gegenteil dieses Phänomens – es werden also unabhängig vom Inhalt der Frage die Skalenextreme (unterster/oberster Wert) gewählt.

Die Tendenz zur Wahl der Skalenextreme findet sich eher bei Personen mit geringerer Bildung, was hier – wie gesagt – ausgeschlossen werden kann. Die Tendenz zu Milde/Härte ist (wie auch die soziale Erwünschtheit) vom Interviewmodus abhängig. Um der interviewenden Person zu gefallen, werden v.a. extrem positive bzw. extrem negative Werte gewählt. Auch dies kann hier – wie oben erläutert – ausgeschlossen werden, da es sich um eine Online-Befragung handelt.

Die Tendenz zur Mitte kann in der vorliegenden Befragung aus anderen Gründen kaum vorkommen. Die meisten Skalen bestehen aus vier Stufen, ohne neutrale Mittelposition. Die fünfstufige Skala, die hier zum Einsatz kommt, hat zwar eine Mitte (Stufe 3), diese ist aber nicht neutral, sondern beschreibt eine bestimmte Häufigkeit («1-2 Mal pro Monat»). Selbstverständlich kann nicht ausgeschlossen werden, dass sich die befragten Personen dennoch für die Mittelposition entscheiden. Doch wird auf der anderen Seite auch angenommen, wenn die Skalenstufen eindeutig benannt werden, dass der Versuch unternommen wird, die Häufigkeit möglichst passend anzugeben.

Die Tendenz zur Mitte und zu Milde/Härte hat hier auch wahrscheinlich keinen großen Einfluss genommen.

### 8.4 Effekte der Fragereihenfolge

Die Reihenfolge der Fragen beeinflusst die Antworttendenz insofern, dass Fragen zu ähnlichen Themen auch ähnlich beantwortet werden. Dies hat Vorteile, da bestimmte Gedächtnisinhalte aktiviert werden und dadurch leichter abrufbar sind. Es hat aber auch Nachteile, da Personen dazu neigen, auch bei Folgefragen ähnliche Antworten wie bei der Ausgangsfrage zu geben.

Im vorliegenden Fragebogen wurde bewusst versucht, den positiven Aspekt der Fragereihenfolge zu nutzen. So wurden die Items thematisch gebündelt. Die thematischen Teile des Fragebogens werden auch immer durch einen kurzen Text eingeführt und evtl. durch eine Faktenfrage ergänzt. So soll die Person auf bestimmte Gedächtnisinhalte hingewiesen werden, die für die Beantwortung der Frage hilfreich sind.

Allerdings müssen zwischendurch dieselben Fragen auch differenziert beantwortet werden, nämlich dann, wenn dieselbe Kooperationsform mit jeweils anderen Zielgruppen eingeschätzt werden muss. Hier fällt auf, dass dies den Teilnehmenden gelingt und Fragen nicht immer gleich (gut oder schlecht) beantwortet werden. Auch ist ein deutlicher Unterschied in den Antworttendenzen zwischen den Berufsgruppen sichtbar, weswegen ein allzu starker Einfluss der Fragereihenfolge ausgeschlossen werden kann. Dennoch ist es möglich, dass sich innerhalb einer Berufsgruppe durchaus Fragereihenfolgeeffekte bemerkbar machen.

### 8.5 Reihenfolge der Antwortvorgaben

Die Reihenfolge der Antwortvorgaben zeigt sich darin, dass bei visuell präsentierten Befragungen (wie im vorliegenden Fall) die erste bzw. oberste Antwort gewählt wird. Bei mündlichen Befragungen wird hingegen eher die letzte (gehörte) Antwortalternative gewählt. Dieser Effekt tritt v.a. dann auf, wenn die Antwortalternativen besonders lang und/oder komplex sind. Sobald die Arbeitsgedächtnisleistungen überfordert sind, wird eben die erste bzw. letzte Antwortoption gewählt. Dementsprechend ist darauf zu achten, die Antwortalternativen möglichst einfach zu formulieren. Im vorliegenden Fall werden eher Aussagen verschriftet, die mittels Ratingskalen gesamthaft bestätigt oder abgelehnt werden müssen. Die entsprechenden Skalen sind immer gleich und sehr kurz formuliert («trifft zu» und «trifft nicht zu» bzw. «sehr wahrscheinlich» und «gar nicht wahrscheinlich»). Die Bewertung jeder einzelnen Antwort durch Rating-Skalen führt zwar zu einer Zunahme des Fragebogenumfangs, schützt aber vor dem Effekt der Antwortvorgabenreihenfolge (Bogner & Landrock, 2015).

## 9 Schlussfolgerungen und Implikationen

Bei der Konstruktion und Evaluation des verwendeten Fragebogens mussten zwei Herausforderungen gemeistert werden: 1. Items aus verschiedenen Fragebögen zusammenstellen, die es vermögen, die notwendigen Informationen zur Beantwortung der Forschungsfragen zu erheben, und 2. diese so zu formulieren, dass sie von allen drei Zielgruppen (Regellehrpersonen, Schulischen Heilpädagog:innen und Logopäd:innen) gleichermaßen gut verstanden und auch beantwortet werden können.

Schon die kognitiven Pretests haben gezeigt, dass dies kein leichtes Unterfangen ist. Schnell wurde klar, dass es nicht die perfekten Items gibt, sondern Kompromisse gesucht werden müssen, um das bestmögliche Ergebnis herauszuholen. Dabei wurden Umformulierungen, Erweiterungen und Ergänzungen der Originalitems vorgenommen und erprobt. Auf diese Weise hat man sich einer Fragebogenversion angenähert, die geeignet war, an einer größeren Stichprobe getestet zu werden.

Mit Daten von 278 Testpersonen aus allen drei Berufsgruppen konnten diverse statistische Analysen vorgenommen werden, um u.a. die Itemschwierigkeiten, die Missings, die Reliabilität und die Validität (v.a. die Konstruktvalidität zu überprüfen). Diese Analysen lieferten eindeutige Hinweise darauf, welche Subskalen fehlerhaft konstruiert waren und somit keine brauchbaren Daten lieferten (s. Kap. 6.26 und 7.3.19). Somit werden die Skalen «Sprachunterstützende Massnahmen im Allgemeinen» und «sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers» in Zukunft nicht mehr verwendet. Die Räumlichkeiten abgesehen vom Klassenzimmer, die von den beteiligten Berufsgruppen zusätzlich genutzt werden, werden in einer einfachen Auswahlfrage erhoben, ohne die Häufigkeit der Nutzung einschätzen zu müssen.

Weitere Skalen («sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten», «unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten», «Ko-Konstruktion» und «Zufriedenheit») werden künftig in einer gekürzten Version zum Einsatz kommen, da sich die

Items z.T. als nicht relevant für die Skala erwiesen bzw. das Messmodell ohne diese Items besser funktioniert.

Andere Items müssen umformuliert werden, damit diese besser beantwortet werden können. Die Subskala zur Messung der «Intention» wurde bisher sehr konkret gefasst («Bis zum Ende des Schuljahres nehme ich mir vor...»), was v.a. für die Logopäd:innen unangemessen zu sein scheint. Da aus den Rahmenbedingungen deutlich wird, dass Logopäd:innen an bedeutend mehr Standorten arbeiten und sich dadurch die Zusammenarbeit mit bestimmten Teams schwerer organisieren lässt als bei den anderen zwei Berufsgruppen, ist eine Formulierung im Konjunktiv wahrscheinlich angemessener: «Ich würde in Zukunft gerne häufiger...». So kann dem Wunsch nach einer Intensivierung der Zusammenarbeit mehr Ausdruck verliehen werden, ohne tatsächliche (z.T. hinderliche) Umstände bei der Antwort berücksichtigen zu müssen.

Auch die Skala zu den Unterrichtsformen funktioniert v.a. in der Gruppe der Schulischen Heilpädagog:innen in der jetzigen Form nur schlecht. Es besteht der Verdacht, dass die Unterrichtsformen nicht eindeutig verortet werden konnten, da spezifische Unterrichtsformen («Frontalunterricht», «offener Unterricht») mit Sozialformen («Gruppenarbeit», «Einzelarbeit») vermengt wurden. Dies entspricht allerdings dem Original von Reber und Blechschmidt (2014). Gerade Schulischen Heilpädagog:innen, die didaktisch besonders gut geschult sind, ist diese Vermengung vermutlich aufgefallen und konnte deswegen nicht eindeutig bearbeitet werden. In der überarbeiteten Version des Fragebogens wird deswegen der Fokus auf die Sozialformen gelegt, welche nach Pukas (2008) in «Einzelarbeit», «Partnerarbeit», «Gruppenarbeit», «Frontalunterricht» und «Unterrichtsgespräch» unterteilt und auch durch dessen Illustrationen veranschaulicht werden (s. Abb. 30). So wird hoffentlich für mehr Klarheit gesorgt und die Beantwortung dieses Items für alle Beteiligten vereinfacht.

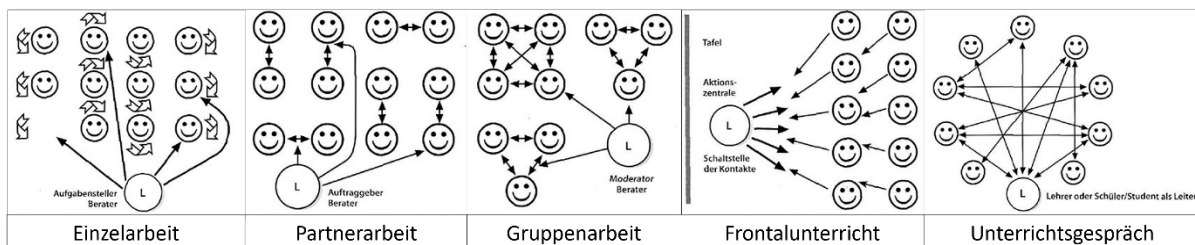


Abbildung 30: Sozialformen im Unterricht nach Pukas (2008)

Außerdem wurde die Subskala «Team-Teaching» durch einen Ausfüllhinweis ergänzt, der deutlich machen soll, dass auch die Team-Teaching-Form «Unterricht und Einzelunterstützung» nicht als spontanes «Aushelfen» in der Klasse zu verstehen ist, sondern als geplante und vorbereitete Maßnahme. Der verwendeten Abbildung geht also die Instruktion «Der Unterricht wird im Vorhinein zusammen geplant und anschliessend in einer der folgenden Formen durchgeführt» voraus. So soll reguliert werden, dass dieses Item im Vergleich zu den anderen Team-Teaching-Formen auffallend häufig gewählt wird.

Da durch diese Anpassungen des Fragebogens insgesamt 22 Items eingespart werden können, wird diese Gelegenheit genutzt, eine bisher nicht eingesetzte Skala zur Bedeutung des Schulleitungshandelns für die multiprofessionelle Kooperation (mit insgesamt acht Items), die ebenfalls aus dem Fragebogen von Drossel (2015) stammt, einzusetzen (s. Tab. 70, S. 93). Auf diese Weise können wichtige inhaltliche Lücken geschlossen werden.

Tabelle 70: Subskala "Strukturelles und personales Schulleitungshandeln" (nach Drossel 2015)

Item	Abkürzung
An unserer Schule gibt es feste Zeiten, um mit den anderen Fachkräften zu kooperieren.	s11
Bei der Erstellung des Stundenplans wird darauf geachtet, dass Kooperationsmöglichkeiten geschaffen werden.	s12
Die Schulleitung versucht Fachkräfte, die eng in Kooperationsformen zusammenarbeiten, in irgendeiner Weise zu entlasten.	s13
Die Schulleitung fördert Kooperation mit den anderen Fachkräften im Rahmen von gemeinsamen Fortbildungen.	s14
Für die Zusammenarbeit stehen Arbeitsräume mit ausreichender Ausstattung zur Verfügung.	s15
Die Schulleitung unterstützt die Kooperation mit den anderen Fachkräften durch Anregungen und Ideen.	s16
Kooperation und Teamarbeit werden an unserer Schule durch die Schulleitung wertgeschätzt.	s17
Die Schulleitung motiviert uns zur Kooperation.	s18

<sup>a</sup> 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft zu

Der Fragebogen vermag es nun, innerhalb von 20-25 Minuten wesentliche Informationen zur Beantwortung der Forschungsfragen zu erheben. Bei der Analyse der Zieldaten ist aber nach wie vor Vorsicht geboten, da noch nicht alle Messmodelle zufriedenstellende Werte lieferten. Außerdem gilt es, mit den Daten aus der Zielstichprobe die Messinvarianz der Skalen zu ermitteln, um zu überprüfen, ob die jeweiligen Konstrukte von den verschiedenen Berufsgruppen überhaupt in gleicher Weise verstanden und beantwortet werden.

Ebenfalls stehen noch Analysen zu den Messmodellen aus: Sind Messmodelle, in denen die einzelnen latenten Konstrukte einer Skala miteinander korrelieren (first order), oder solche, in denen die Konstrukte auf ein latentes Konstrukt höherer Ordnung (second order) laden, besser geeignet, um weiterführende Analysen zu betreiben? Die vorliegende Stichprobe ist jedoch zu klein, um solche komplexen Modelle zu berechnen. Erste Versuche jedoch lassen erahnen, dass manche Skalen (z.B. «unterrichtsintegrierte sprachunterstützende Maßnahmen») besser in einem second-order-Konstrukt abgebildet werden, andere Skalen jedoch eher in einem first-order-Konstrukt (z.B. «Häufigkeit und Intensität der Zusammenarbeit in Bezug auf sprachunterstützende Maßnahmen» und «Einstellungen zur Zusammenarbeit in Bezug auf sprachunterstützende Maßnahmen»). Es handelt sich bei diesen Vermutungen tatsächlich nur um Annäherungen an Umstände, die noch weiter zu ergründen sind – was wiederum eine größere Stichprobe voraussetzt.

# 10 Verzeichnisse

## 10.1 Literaturverzeichnis

- Blanz, M. (2021). *Forschungsmethoden und Statistik für die Soziale Arbeit. Grundlagen und Anwendungen* (2. Auflage). Stuttgart: Verlag W. Kohlhammer.
- Blechsmidt, A., Besset, S., Läber, J., Reber, K. & Allemann, D. (2013). *Logopädie in Zeiten der Inklusion*. Vortrag am Schweizer Heilpädagogik-Kongress, Bern.
- Bogner, K. & Landrock, U. (2015). *Antworttendenzen in standardisierten Umfragen* (SDM Survey Guidelines). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. Zugriff am 04.11.2021. [https://doi.org/10.15465/sdm-sg\\_016](https://doi.org/10.15465/sdm-sg_016)
- Bortz, J. & Döring, N. (2016). *Forschungsmethoden und Evaluation* (5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Berlin, Heidelberg: Springer.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7., vollständig überarbeitete und erweiterte Auflage). Berlin, Heidelberg: Springer.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2<sup>nd</sup> edition). New York, London: The Guilford Press.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2., aktualisierte Aufl.). München, Don Mills: Pearson Studium.
- Byrne, B. M. (2010). *Structural Equation Modeling with Mplus. Basic Concepts, Applications, and Programming*. New York, London: Routledge.
- Dong, Y. & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, (2), 1–17.
- Drossel, K. (2015). *Motivationale Bedingungen von Lehrerverkooperation. Eine empirische Analyse der Zusammenarbeit im Projekt "Ganz In"*. Münster: Waxmann.
- Drossel, K. & Willems, A. S. (2014). Zum Zusammenhang von Formen der Lehrerverkooperation des Schulleitungshandelns und des Kooperationsklimas an Ganztagsgymnasien. In R. Strietholt & W. Bos (Hrsg.), *Empirische Bildungsforschung und evidenzbasierte Reformen im Bildungswesen* (S. 129–154). Münster: Waxmann.
- Elksnin, L. K. & Capilouto, G. J. (1994). Speech-Language Pathologists' Perceptions of Integrated Service Delivery in School Settings. *Language, speech, and hearing services in schools*, 25, 238–267.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics. And sex and drugs and rock 'n' roll* (4th edition). Los Angeles, London, New Delhi: Sage.
- Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (2., durchgesehene Auflage). Wiesbaden: VS Verlag.
- Kemper, C. J., Beierlein, C., Bensch, D., Kovaleva, A., & Rammstedt, B. (2014). *Soziale Erwünschtheit-Gamma (KSE-G)*. Verfügbar unter: <http://www.gesis.org/kurzskalen-psychologischer-merkmale>
- Kolbe, F.-U. & Reh, S. (2008). Kooperation unter Pädagogen. In T. Coelen & H.-U. Otto (Hrsg.), *Grundbegriffe Ganztagsbildung. Das Handbuch* (S. 799–808). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Krosnick, J. A. & Presser, S. (2010). Question and Questionnaire Design. In P. V. Marsden & J. D. Wright (Hrsg.), *Handbook of survey research* (2nd ed, S. 263–312). Bingley, UK: Emerald.
- Leonhart, R. (2013). *Lehrbuch Statistik. Einstieg und Vertiefung* (3., überarbeitete und erweiterte Auflage). Bern, Göttingen: Verlag Hans Huber; Hogrefe.
- Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus User's Guide* (8<sup>th</sup> Edition). Los Angeles, CA: Muthén & Muthén
- Prüfer, P. & Rexroth, M. (2005). *Kognitive Interviews*, Zentrum für Umfragen, Methoden und Analysen -ZUMA-. GESIS-How-to: 15. Zugriff am 16.07.2021. Verfügbar unter: <https://www.ssoar.info/ssoar/handle/document/20147>
- Pukas, D. (2008). *Lernmanagement. Einführung in Lern- und Arbeitstechniken* (3., aktualisierte Aufl.). Rinteln: Merkur-Verl.
- Reber, K. (2012). Sprachheilpädagogik und Sprachtherapie in inklusiven Settings. Perspektiven der Vernetzung zwischen Sprachheilpädagogik, Sprachtherapie und Regelschule. *L.O.G.O.S. Interdisziplinär*, 20(4), 264–275.

- Reber, K. & Blechschmidt, A. (2014). Inklusion und Sprachtherapie/Logopädie/Sprachheilpädagogik in Deutschland. Angebote für Kinder und Jugendliche mit Sprachbehinderungen in Schulen. *Praxis Sprache*, (2), 93–105.
- Reinecke, J. (2014). *Strukturgleichungsmodelle in den Sozialwissenschaften* (2., aktualisierte und erweiterte Auflage). München: DE GRUYTER OLDENBOURG.
- Sallat, S., Hofbauer, C. & Jurleta, R. (2017). *Inklusion an den Schnittstellen von sprachlicher Bildung, Sprachförderung und Sprachtherapie. Eine Expertise der Weiterbildungsinitiative Frühpädagogische Fachkräfte (WIFF)*. München: Deutsches Jugendinstitut e.V. Weiterbildungsinitiative Frühpädagogische Fachkräfte (WIFF).
- Theisel, A. (2014). Ein Faktorenmodell zu Qualitätsmerkmalen des Unterrichts mit sprachbeeinträchtigten Kindern. *Forschung Sprache*, (2), 36–57.
- Wang, J. & Wang, X. (2012). *Structural Equation Modeling. Applications Using Mplus*. Chichester West Sussex, Hoboken N.J.: Wiley.

## 10.2 Abbildungsverzeichnis

Abbildung 1: Anzahl Teilnehmende pro Kanton .....	7
Abbildung 2: Messmodell von "sm_all" mit standardisierten Werteangaben .....	57
Abbildung 3: Messmodell von "sm_all" mit standardisierten Werteangaben .....	58
Abbildung 4: Messmodell von "sm_all" ohne "rfl_all" mit standardisierten Werteangaben .....	59
Abbildung 5: Messmodell von "sm_kl" mit standardisierten Werteangaben .....	60
Abbildung 6: Messmodell von "sm_kl" ohne "rfl_kl" mit standardisierten Werteangaben .....	61
Abbildung 7: Messmodell von "sm_uf" mit standardisierten Werteangaben .....	62
Abbildung 8: Messmodell von "sm_fb" mit standardisierten Werteangaben .....	63
Abbildung 9: Messmodell von "sm_fb" mit standardisierten Werteangaben unter Einbezug der Korrelation von fb_mus und fb_spo .....	64
Abbildung 10: Messmodell von "sm pha" mit standardisierten Werteangaben.....	65
Abbildung 11: Messmodell von "aus" mit standardisierten Werteangaben für die jeweiligen Zielgruppen .....	66
Abbildung 12: Messmodell von "syn" mit standardisierten Werteangaben für die jeweiligen Zielgruppen .....	68
Abbildung 13: Messmodell von "kok" mit standardisierten Werteangaben für die jeweiligen Zielgruppen .....	69
Abbildung 14: Messmodell von "kok" mit standardisierten Werteangaben für die jeweiligen Zielgruppen, ohne die Items kok1, kok2 und kok3 .....	71
Abbildung 15: Messmodell von "tt" mit standardisierten Werteangaben.....	72
Abbildung 16: Messmodell von "int" mit standardisierten Werteangaben .....	73
Abbildung 17: Messmodell von "int" mit standardisierten Werteangaben unter Einbezug der Korrelation von int1 und int5.....	74
Abbildung 18: Messmodell von "intr" mit standardisierten Werteangaben.....	75
Abbildung 19: Messmodell von "intr" mit standardisierten Werteangaben unter Einbezug der Korrelation von intr3 und intr4.....	76
Abbildung 20: Messmodell von "kost" mit standardisierten Werteangaben.....	77
Abbildung 21: Messmodell von "kost" mit standardisierten Werteangaben unter Einbezug der Korrelation von kost1 und kost2, kost3 und kost8 und kost4 und kost5 .....	78
Abbildung 22: Messmodell von "nu_fk" mit standardisierten Werteangaben .....	79
Abbildung 23: Messmodell von "nu_fk" mit standardisierten Werteangaben unter Einbezug der Korrelation von nu_fk9 und nu_fk10.....	80
Abbildung 24: Messmodell von "nu_sus" mit standardisierten Werteangaben .....	81

Abbildung 25: Messmodell von "nu_sus" mit standardisierten Werteangaben unter Einbezug der Korrelation von nu_sus3 und nu_sus11 und nu_sus9 und nu_sus10.....	82
Abbildung 26: Messmodell von "erf_pro" mit standardisierten Werteangaben .....	83
Abbildung 27: Messmodell von "erf_erg" mit standardisierten Werteangaben.....	83
Abbildung 28: Messmodell von "zufr" mit standardisierten Werteangaben .....	84
Abbildung 29: Messmodell von "zufr" mit standardisierten Werteangaben unter Einbezug der Korrelation von zufr1 und zufr2 und zufr6 und zufr7 .....	85
Abbildung 30: Sozialformen im Unterricht nach Pukas (2008).....	92

### 10.3 Tabellenverzeichnis

Tabelle 1: Stichprobenbeschreibung .....	6
Tabelle 2: Häufigkeit sprachunterstützender Maßnahmen (sM) im Allgemeinen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	8
Tabelle 3: Häufigkeit sprachunterstützender Maßnahmen im Allgemeinen bei Auffälligkeiten ..., basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	9
Tabelle 4: Häufigkeit unterrichtsintegrierter Maßnahmen bei Auffälligkeiten ..., basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	9
Tabelle 5: Häufigkeit unterrichtsintegrierter sprachunterstützender Maßnahmen bei verschiedenen Unterrichtsformen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	9
Tabelle 6: Häufigkeit unterrichtsintegrierter sprachunterstützender Maßnahmen in verschiedenen Fachbereichen, basierend auf den Antworten der Gesamtstichprobe unter Ausschluss der «trifft nicht zu»-Antworten .....	10
Tabelle 7: Häufigkeit unterrichtsintegrierter sprachunterstützender Maßnahmen in verschiedenen Erwerbsphasen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	10
Tabelle 8: Häufigkeit sprachunterstützender Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	10
Tabelle 9: Häufigkeit der Kooperationsform "Austausch" mit den Regellehrpersonen, basierend auf den Antworten der Schulischen Heilpädagog:innen und Logopäd:innen (n = 220) nach paarweisem Ausschluss der Missings.....	11
Tabelle 10: Häufigkeit der Kooperationsform "Austausch" mit den Schulischen Heilpädagog:innen, basierend auf den Antworten der Regellehrpersonen und Logopäd:innen (n = 213) nach paarweisem Ausschluss der Missings.....	11
Tabelle 11: Häufigkeit der Kooperationsform "Austausch" mit den Logopäd:innen, basierend auf den Antworten der Regellehrpersonen und Schulischen Heilpädagog:innen (n = 123) nach paarweisem Ausschluss der Missings.....	11
Tabelle 12: Häufigkeit der Kooperationsform "Synchronisation" mit den Regellehrpersonen, basierend auf den Antworten der Schulischen Heilpädagog:innen und Logopäd:innen (n = 220) nach paarweisem Ausschluss der Missings.....	12
Tabelle 13: Häufigkeit der Kooperationsform "Synchronisation" mit den Schulischen Heilpädagog:innen, basierend auf den Antworten der Regellehrpersonen und Logopäd:innen (n = 213) nach paarweisem Ausschluss der Missings.....	12



Tabelle 14: Häufigkeit der Kooperationsform "Synchronisation" mit den Logopäd:innen, basierend auf den Antworten der Regellehrpersonen und Schulischen Heilpädagog:innen (n = 123) nach paarweisem Ausschluss der Missings.....	12
Tabelle 15: Häufigkeit der Kooperationsform "Ko-Konstruktion" mit den Regellehrpersonen, basierend auf den Antworten der Schulischen Heilpädagog:innen und Logopäd:innen (n = 220) nach paarweisem Ausschluss der Missings.....	13
Tabelle 16: Häufigkeit der Kooperationsform "Ko-Konstruktion" mit den Schulischen Heilpädagog:innen, basierend auf den Antworten der Regellehrpersonen und Logopäd:innen (n = 213) nach paarweisem Ausschluss der Missings.....	13
Tabelle 17: Häufigkeit der Kooperationsform "Ko-Konstruktion" mit den Logopäd:innen, basierend auf den Antworten der Regellehrpersonen und Schulischen Heilpädagog:innen (n = 123) nach paarweisem Ausschluss der Missings.....	14
Tabelle 18: Häufigkeit der Kooperationsform "Team-Teaching" im Allgemeinen, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	14
Tabelle 19: Intention, die Häufigkeit der sprachunterstützenden Maßnahmen (SM) und die Zusammenarbeit künftig zu intensivieren, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings .....	14
Tabelle 20: Interesse an multiprofessioneller Kooperation, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	15
Tabelle 21: Empfundene Kosten der multiprofessionellen Kooperation, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings .....	15
Tabelle 22: Wahrgenommene Nützlichkeit der multiprofessionellen Kooperation für die eigene, professionelle Entwicklung, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings .....	16
Tabelle 23: Wahrgenommene Nützlichkeit der multiprofessionellen Kooperation für die Schüler und Schülerinnen sowie die Schule, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings .....	16
Tabelle 24: Einschätzung der prozessorientierten Erfolgserwartung, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	17
Tabelle 25: Einschätzung der ergebnisorientierten Erfolgserwartung, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	17
Tabelle 26: Zufriedenheit mit der Arbeitssituation, basierend auf den Antworten der Gesamtstichprobe (n = 278) nach paarweisem Ausschluss der Missings.....	17
Tabelle 27: Verwendete Likertskalen und deren Beschriftung.....	22
Tabelle 28: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Gesamtstichprobe.....	23
Tabelle 29: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Stichprobe der Regellehrpersonen....	24
Tabelle 30: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Stichprobe der Schulischen Heilpädagog:innen .....	24
Tabelle 31: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen" für die Stichprobe der Logopäd:innen .....	25
Tabelle 32: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Gesamtstichprobe	26
Tabelle 33: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Regellehrpersonen .....	26

Tabelle 34: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Schulischen Heilpädagog:innen .....	27
Tabelle 35: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen im Allgemeinen bei Auffälligkeiten" für die Logopäd:innen .....	27
Tabelle 36: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei Auffälligkeiten" für die Gesamtstichprobe .....	28
Tabelle 37: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Unterrichtsformen" für die Gesamtstichprobe .....	29
Tabelle 38: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Unterrichtsformen" für die Schulischen Heilpädagog:innen .....	29
Tabelle 39: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Fachbereichen" für die Gesamtstichprobe .....	30
Tabelle 40: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "unterrichtsintegrierte sprachunterstützende Maßnahmen bei verschiedenen Erwerbsphasen" für die Gesamtstichprobe .....	31
Tabelle 41: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Gesamtstichprobe .....	32
Tabelle 42: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Regellehrpersonen .....	32
Tabelle 43: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Schulische Heilpädagog:innen .....	33
Tabelle 44: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "sprachunterstützende Maßnahmen in Räumlichkeiten außerhalb des Klassenzimmers" für die Logopäd:innen .....	33
Tabelle 45: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Regellehrpersonen" für die Schulische Heilpädagog:innen und Logopäd:innen .....	34
Tabelle 46: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Regellehrpersonen" für die Schulische Heilpädagog:innen .....	34
Tabelle 47: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Regellehrpersonen" für die Logopäd:innen .....	35
Tabelle 48: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Schulischen Heilpädagog:innen" für die Regellehrpersonen und Logopäd:innen .....	35
Tabelle 49: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Austausch mit den Logopäd:innen" für die Regellehrpersonen und Schulischen Heilpädagog:innen .....	36
Tabelle 50: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Synchronisation mit den Regellehrpersonen" für die Schulische Heilpädagog:innen und Logopäd:innen .....	37
Tabelle 51: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Synchronisation mit den Schulischen Heilpädagog:innen" für die Regellehrpersonen und Logopäd:innen .....	37

Tabelle 52: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Synchronisation mit den Logopäd:innen" für die Regellehrpersonen und Schulischen Heilpädagog:innen .....	38
Tabelle 53: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Konstruktion mit den Regellehrpersonen" für die Schulische Heilpädagog:innen und Logopäd:innen .....	39
Tabelle 54: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Konstruktion mit den Regellehrpersonen" für die Schulische Heilpädagog:innen .....	39
Tabelle 55: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Konstruktion mit den Regellehrpersonen" für die Logopäd:innen.....	40
Tabelle 56: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Konstruktion mit den Schulischen Heilpädagog:innen" für die Regellehrpersonen und Logopäd:innen .....	40
Tabelle 57: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Konstruktion mit den Logopäd:innen" für die Regellehrpersonen und Schulische Heilpädagog:innen .....	41
Tabelle 58: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Team-Teaching" für die Gesamtstichprobe .....	42
Tabelle 59: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Team-Teaching" für die Schulischen Heilpädagog:innen.....	43
Tabelle 60: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Intention" für die Gesamtstichprobe .....	43
Tabelle 61: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Interesse" für die Gesamtstichprobe .....	44
Tabelle 62: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Kosten" für die Gesamtstichprobe.....	45
Tabelle 63: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Nützlichkeit für die professionelle Entwicklung" für die Gesamtstichprobe.....	46
Tabelle 64: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Nützlichkeit für die Schüler und Schülerinnen sowie die Schule" für die Gesamtstichprobe .....	47
Tabelle 65: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "prozessorientierte Erfolgserwartung" für die Gesamtstichprobe.....	48
Tabelle 66: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "ergebnisorientierte Erfolgserwartung" für die Gesamtstichprobe .....	48
Tabelle 67: Schwierigkeiten, Trennschärfen und Cronbachs Alpha der Items der Subskala "Zufriedenheit" für die Gesamtstichprobe .....	49
Tabelle 68: Subskalen des Fragebogens, Angaben zur internen Konsistenz und zu den Itemtrennschärfen .....	51
Tabelle 69: CFAs der Subskalen des Fragebogens, Angaben zu den Modellfitmaßen und Kommentar zur Verwendung.....	87
Tabelle 70: Subskala "Strukturelles und personales Schulleitungshandeln" (nach Drossel 2015).....	93