

# **Evaluation der Implementierung des Deutschfördermodells**

im Auftrag des

 **Bundesministerium**  
Bildung, Wissenschaft  
und Forschung

**Abschlussbericht**

**Oktober 2022**

**Christiane Spiel**

**Vera Popper**

**Julia Holzer**

## 1. Abstract

Das Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF) gab Ende 2020 eine wissenschaftliche Evaluationsstudie in Auftrag, die diejenigen Implementierungsfaktoren des Deutschfördermodells identifizieren sollte, die die Verweildauer im ao.-Status beeinflusst haben könnten. Die Deutschkenntnisse zur Überführung in den o.-Status werden mit dem MIKA-D (Messinstrument zur Kompetenzanalyse – Deutsch) geprüft. Die Evaluierung wurde mittels eines Extremgruppendesigns durchgeführt, d.h. durch den Vergleich zweier Gruppen von Schulen, die sich hinsichtlich des Erfolgskriteriums (= möglichst rasche Überführung der ao.-Schüler\*innen in den o.-Status) unterscheiden, jedoch hinsichtlich der Rahmenbedingungen möglichst ähnlich sind. Damit sollte untersucht werden, ob und welche Implementierungsfaktoren zu einer erfolgreichen Deutschförderung beitragen. 110 Schulen wurden zur Teilnahme an der Evaluierung eingeladen. Befragt wurden Schulleitungen, Lehrpersonen in der Deutschförderung (DF) und Lehrpersonen, die in der Primarstufe (Volksschule) unterrichten – damit 3 „Rollen“. Insgesamt füllten 693 Personen aus 93 Schulen das Erhebungsinstrument aus, das geschlossene und offene Fragen enthielt. Im Rücklauf gab es keine Unterschiede zwischen den beiden Gruppen.

Insgesamt zeigen die Ergebnisse der Evaluation über beide Gruppen und Rollen hinweg eine sehr hohe Übereinstimmung. Damit konnte die Annahme, dass sich Schulen mit kürzerer Verweildauer und Schulen mit längerer Verweildauer der Schüler\*innen im ao.-Status systematisch in relevanten Implementierungsfaktoren unterscheiden, nicht bestätigt werden. Alle drei Rollen stimmten in beiden Gruppen tendenziell eher einer längeren Verweildauer im ao.-Status zu als einer kürzeren. Das Deutschfördermodell insgesamt fand eine eher geringe Zustimmung, wobei Deutschförderklassen negativer bewertet wurden als Deutschförderkurse und diese wiederum negativer als integrative Förderung. Auch hinsichtlich der organisatorischen Umsetzung unterschieden sich die Gruppen nicht in ihren Antworten. Ein Großteil der Lehrpersonen DF hatte sich freiwillig für den Einsatz in der Deutschförderung gemeldet und fühlte sich durch das Unterrichten in der Deutschförderung eher nicht überfordert. Auch ihr Selbstwirksamkeitserleben schätzten sie im Mittel eher hoch ein. Jedoch gaben Lehrpersonen DF, die erst knapp vor ihrem Einsatz in der Deutschförderung davon erfahren hatten, durchgängig negativere Antworten. Der MIKA-D, auf Basis dessen über „Erfolg“ versus „Nicht-Erfolg“ in der Deutschförderung entschieden wird, wurde im Mittel als teilweise nützlich für die Feststellung des Erfolgs in der Deutschförderung gesehen. Die Ergebnisse des MIKA-D werden häufig für die Förderplanung und für Evaluation genutzt; häufig wurde jedoch auch „keine Nutzung, nur Aufbewahrung“ angegeben. Sprachbezogene Ziele der Deutschförderung (die in einer Vorstudie identifiziert worden waren) wurden laut Lehrpersonen DF nur zwischen 45.00% und 79.22% erreicht. Ein hoher Prozentsatz der Befragten sah einen Bedarf die Deutschförderung weiterzuentwickeln/zu optimieren. Die meisten Ansätze zur Verbesserung wurden mit Abstand auf der Steuerungsebene gesehen (u.a. kleinere Gruppen, mehr Autonomie und Flexibilität, mehr Ressourcen, integrativ statt segregiert, flexiblerer Aufstieg).

Basierend auf den Befragungsergebnissen werden folgende Empfehlungen für die Weiterentwicklung/Optimierung der Deutschförderung abgeleitet: Weiterentwicklung des MIKA-D zur differenzierten Prüfung der Ziele der Deutschförderung; mehr Autonomie und Flexibilität für Schulen in der Umsetzung des Deutschfördermodells; weitere Förderung von Schüler\*innen, die nach zwei Jahren noch keine ausreichenden Deutschkenntnisse aufweisen; Unterricht in der Deutschförderung nur durch Lehrpersonen mit DaZ-Ausbildung, jedoch Sensibilisierung aller Lehrpersonen für dieses Thema. Die Implementierung des weiterentwickelten Deutschfördermodells sollte unter Einbindung der Stakeholder-Gruppen geplant und umgesetzt sowie durch eine Evaluation begleitet werden. Zusätzlich wird empfohlen ein systematisches und längsschnittliches Monitoring einzurichten und mit Registerdaten zu vernetzen.

## 2. Einleitung

In Österreich werden Schüler\*innen, die dem Unterricht auf Grund unzureichender Sprachkenntnisse nicht folgen können, seit dem Schuljahr 2018/19 in eigenen Deutschförderklassen bzw. in unterrichtsparallelen Deutschförderkursen unterrichtet. Zur Feststellung ihrer Deutschkenntnisse wird der MIKA-D (Messinstrument zur Kompetenzanalyse – Deutsch) eingesetzt. Jene Schüler\*innen, die auf Basis ihres MIKA-D Testergebnisses ungenügende Deutschkenntnisse aufweisen, werden einer Deutschförderklasse zugeteilt. Jene mit mangelhaften Deutschkenntnissen werden einem Deutschförderkurs zugeteilt. Ist die Anzahl der Schüler\*innen mit mangelhaften oder ungenügenden Deutschkenntnissen geringer als 8, ist die Ausnahme vorgesehen, dass die Schüler\*innen integrativ in der Regelklasse nach den jeweiligen Lehrplan-Zusätzen unterrichtet werden (diese drei Maßnahmen – Deutschförderklassen, Deutschförderkurse und integrative Deutschförderung – werden als „Deutschfördermodell“ bezeichnet). Alle Schüler\*innen, die bei Schuleintritt laut MIKA-D Testergebnis ungenügende oder mangelhafte Deutschkenntnisse aufweisen, werden als Schüler\*innen im außerordentlichen Status (kurz ao.-Status) geführt. Als Erfolgsindikator der Deutschförderung gilt der möglichst rasche Umstieg in den ordentlichen Status (kurz o.-Status). Ob ein Umstieg möglich ist und die Deutschkenntnisse als „ausreichend“ beurteilt werden können, wird jeweils zu Semesterende durch erneute Testung mit dem MIKA-D festgestellt. Nach vier Semestern wechseln die Schüler\*innen automatisch in den o.-Status, auch wenn sie laut MIKA-D ungenügende oder mangelhafte Deutschkenntnisse aufweisen. Eine weiterführende Förderung ist gesetzlich nicht vorgesehen. Die MIKA-D Testung erfolgt dezentral in den Schulen, wobei das Ergebnis auf Schulebene eingemeldet wird und dem BMBWF damit aggregiert vorliegt (siehe BMBWF, 2022).

Bei der Einführung neuer Maßnahmen in komplexen Interaktionsfeldern wie dem Schulsystem, kommt der Implementierung der Maßnahmen eine zentrale Rolle ein. Damit ist die tatsächliche Umsetzung der Maßnahme in der schulischen Praxis gemeint. Relevante Implementierungsfaktoren sind eine differenzierte Information an alle beteiligten Gruppen sowie eine entsprechende Kommunikation, die sorgfältige Durchführung der geplanten Prozessschritte von der Pilotierung bis zum Rollout der Maßnahme, die Berücksichtigung von kontextspezifischen Rahmenbedingungen sowie das Vorwissen und die Einstellungen jener Personen, die die Maßnahme umsetzen sollen. Die Erreichung der Maßnahmenziele hängt wesentlich davon ab, ob diese Implementierungsfaktoren berücksichtigt werden (siehe z.B. Fixsen & Blase, 2009; Schober et al., 2019). Aufgrund der Komplexität von Implementierung und damit der Notwendigkeit einer wissenschaftlichen Fundierung hat sich – ausgehend von der Evaluationsforschung – die Implementierungsforschung etabliert, die sich zum Ziel setzt, generalisierbare Erfolgsfaktoren zu identifizieren bzw. Methoden vorzuschlagen, wie die Implementierung von Maßnahmen erfolgreich gestaltet werden kann.

Da die Ziele des Deutschfördermodells in dem dafür vorgesehenen Zeitraum von zwei Jahren nur eingeschränkt erreicht wurden, hat das BMBWF eine externe wissenschaftliche Studie in Auftrag gegeben. In dieser sollten relevante Faktoren für die Implementierung der Deutschförderung analysiert werden. Daraus sollten Schlussfolgerungen gezogen werden, welche Implementierungsfaktoren die Verweildauer im ao.-Status (d.h. die Ergebnisse der MIKA-D Testungen) beeinflusst haben könnten. Aus den Ergebnissen sollten - soweit möglich - Schlussfolgerungen gezogen werden, wie das derzeitige gesetzliche Deutschfördermodell weiterentwickelt/ optimiert werden kann.

Die vorliegende Evaluation bezieht Ergebnisse einer Vorstudie mit ein. Diese beinhaltete die Explikation der Zielsetzungen der Deutschförderung und ihrer derzeitigen Implementierung in den Schulen sowie Überlegungen zum Forschungsdesign der Evaluierung. In der Vorstudie wurden zwei Workshops mit Expert\*innen aus Schulen, Schulaufsicht, Bildungsministerium und Hochschulen

durchgeführt (Spiel, Popper & Holzer, 2021; der vollständige Workshopbericht<sup>1</sup> kann [hier](#) aufgerufen werden). Auf Basis der mit den Expert\*innen in den beiden Workshops erarbeiteten Inhalte wurden sprachbezogene und flankierende Ziele der Deutschförderung, sowie Gelingensfaktoren für die Implementierung und relevante Kontextbedingungen für eine erfolgreiche Deutschförderung abgeleitet. Auf Basis dieser Ergebnisse der Workshops wurde das weiter unten vorgestellte Evaluationsinstrumentarium entwickelt.

Zudem wurde im Vorfeld der Evaluation gemeinsam mit Expert\*innen des BMBWF die für die Evaluation der Implementierung der Deutschförderung vorhandene Datengrundlage gesichtet und beurteilt. Es zeigte sich, dass die derzeit vorliegenden Daten (z.B. die MIKA-D Ergebnisse) nicht ausreichend differenziert genug waren, um Fragestellungen auf Schüler\*innen- oder Klassenebene zu untersuchen. Daher wurde festgelegt, die Evaluation auf Fragestellungen zu beschränken, die sinnvoll auf Schulebene untersucht werden können. Dafür noch notwendige Daten wurden durch das BMBWF zusätzlich erhoben und für die Studie aufbereitet.

Das Evaluationsteam schlug vor, die Evaluierung mittels eines Extremgruppendesigns durchzuführen. Damit ist ein Vergleich von zwei Gruppen gemeint, die sich hinsichtlich des Erfolgskriteriums deutlich voneinander unterscheiden, hinsichtlich ihrer Rahmenbedingungen einander jedoch möglichst ähnlich sind. Damit kann untersucht werden, ob und welche Implementierungsfaktoren zur erfolgreichen Deutschförderung beitragen.

Anhand der Bildungsverlaufsstatistik der Statistik Austria sowie des Bildungsdokumentationssystems und des Stellenplans des BMBWF wurden in einem nächsten Schritt Möglichkeiten exploriert, Stichproben zur Beantwortung der leitenden Fragestellungen zu ziehen. Das Vorgehen zur Ziehung der Stichproben aus der Gesamtheit aller Schulen, die für die Evaluierung der Deutschförderung eingesetzten Erhebungsinstrumente sowie die Methodik der Datenanalyse werden im nächsten Kapitel im Detail beschrieben.

### **3. Methodisches Vorgehen**

#### **Evaluationsdesign**

Die Evaluation der Implementierung der Deutschförderung erfolgte mittels des Vergleichs von zwei Gruppen von Schulstandorten, die sich hinsichtlich ihres Ergebnisses in Bezug auf die Deutschförderung unterscheiden, sich hinsichtlich ihrer Rahmenbedingungen jedoch möglichst ähnlich sind.

Die Festlegung der zwei Extremgruppen erfolgte entsprechend des Ziels der Deutschförderung (möglichst kurze Verweildauer der Schüler\*innen im ao.-Status): Gruppe 1 „kürzere Verweildauer im ao.-Status“ und Gruppe 2 „längere Verweildauer im ao.-Status“. Diese beiden Gruppen wurden hinsichtlich Implementierungsfaktoren sowie Kontextvariablen verglichen. Daraus sollten Schlussfolgerungen abgeleitet werden, welche Implementierungsfaktoren die Verweildauer im ao.-Status (festgestellt mittels der MIKA-D Testungen) beeinflusst haben könnten. Dazu wurden Schulleitungen, Lehrpersonen in der Deutschförderung (Lehrpersonen DF) und Lehrpersonen mit Volksschul- bzw. Primarstufenlehramt (Lehrpersonen VS) – im Folgenden als Rollen bezeichnet – im

---

<sup>1</sup> Spiel, C., Popper, V., & Holzer, J. (2021). *Bericht der Workshopergebnisse zu den Themen Zieleexplikation und Implementierung. Vorprojekt zur Evaluation der Deutschförderung.*  
[https://www.bmbwf.gv.at/dam/jcr:dc3a144a-7cef-4a48-ac8c-4c09b6ca14ad/dfk\\_eval.pdf](https://www.bmbwf.gv.at/dam/jcr:dc3a144a-7cef-4a48-ac8c-4c09b6ca14ad/dfk_eval.pdf)

Sommersemester 2022 mittels Online-Fragebögen zur Implementierung des Deutschfördermodells an ihren Schulstandorten befragt.

### **Kriterien für die Stichprobenziehung**

Für die Aufbereitung der Datengrundlage und die Festlegung der Stichproben war ein Projektteam des BMBWF verantwortlich. Anhand der Bildungsverlaufsstatistik der Statistik Austria sowie des Bildungsdokumentationssystems und des Stellenplans des BMBWF schlugen die Expert\*innen des BMBWF Möglichkeiten zur Ziehung von Extremgruppen vor. Die Vorschläge wurden mit dem Evaluationsteam abgestimmt. Ergebnis dieses Prozesses war die im Folgenden beschriebene Herangehensweise.

Zunächst wurde als Auswahlkriterium ein Grenzwert von mindestens 15 ao.-Schüler\*innen im Schuljahr 2019/20 an den Schulen gesetzt, um nur Schulstandorte für die Stichprobe zu selektieren, die kontinuierliche Erfahrung in der Umsetzung der Deutschförderung aufweisen. Hintergrund für die Festlegung des Grenzwertes von 15 ist, dass Schulstandorte unter dem Grenzwert in der Vergangenheit starke Schwankungen in der jährlichen Anzahl von ao.-Schüler\*innen gezeigt hatten und damit nur über eingeschränkte Erfahrungen mit der Deutschförderung verfügen. Das Auswahlkriterium von mindestens 15 ao.-Schüler\*innen im Schuljahr 2019/20 erfüllten nur Volksschulen. Da die Implementierung des gesamten Deutschfördermodells untersucht werden sollte, musste auch sichergestellt werden, dass im Zeitraum der Erhebung eine Mindestanzahl an Schüler\*innen, die nach dem Deutschförderklassenlehrplan beschult werden, am Standort vorhanden war. Als weiteres Kriterium wurde daher festgelegt, dass mindestens 35% der ao.-Schüler\*innen, die sich im ersten Jahr des ao.-Status befanden, nach dem Deutschförderklassen-Lehrplan unterrichtet wurden.

### **Kriterien für die Bildung der Extremgruppen**

Auf Basis der Gesamtstichprobe (d.h. jenen Schulen, die den obigen Auswahlkriterien entsprachen) wurden Kriterien zur Festlegung der zu vergleichenden Gruppen ermittelt. Diese Kriterien orientierten sich an der Verweildauer der ao.-Schüler\*innen im ao.-Status sowie an der verfügbaren Datenlage. Diesbezüglich ist anzumerken, dass zum Zeitpunkt der Stichprobenziehung im Herbst 2021 die Ergebnisse des MIKA-D noch nicht für alle Kohorten seit Einführung des Deutschfördermodells 2018/19 vollständig vorhanden waren. So fehlten beispielsweise noch Informationen, wie viele erstmalige ao.-Schüler\*innen des Schuljahres 2019/20 bis zum 4. Semester, d.h. vor automatischem Ende des ao.-Status in den o.-Status wechseln konnten. Ausgehend von den ao.-Schüler\*innen, die am Schuljahresanfang 2018/19 bzw. 2019/20 erstmalig im ao.-Status waren, wurde eine Bildungsverlaufstatistik erstellt und berechnet, wie hoch der Anteil der Schüler\*innen ist, die bis zum vierten Semester (erstmalige ao.-Schüler\*innen im Schuljahr 2018/19) bzw. bis zum dritten Semester (erstmalige ao.-Schüler\*innen im Schuljahr 2019/20) in den o.-Status wechseln konnten.

Um der Anforderung eines Extremgruppenvergleiches gerecht zu werden, wurden Kriterien hinsichtlich des Umstieges gewählt, die eine Einteilung in Standorte mit „kürzerer“ und „längerer“ Verweildauer im ao.-Status zuließen. Dabei wurde von dem durchschnittlichen Anteil der Schüler\*innen, die zu einem Beobachtungszeitpunkt bereits im ordentlichen Status waren, ausgegangen. Für die Schüler\*innen des Schuljahres 2018/19 im vierten Semester beträgt dieser Mittelwert aller Standorte in der Gesamtstichprobe 56% und für die Schüler\*innen des Schuljahres 2019/20 im dritten Semester 29%.

Die Festlegung von Grenzwerten für die beiden Gruppen folgte einerseits dem Bestreben eine größtmögliche Differenz der Schulstandorte zwischen den Gruppen zu erzeugen und andererseits der Notwendigkeit eine sinnvoll analysierbare Mindestanzahl an Schulstandorten für die Stichprobe zu

erhalten. Dabei wurden in einem iterativen Vorgehen verschiedene Grenzwerte analysiert und letztlich folgende Kriterien festgelegt:

#### *Gruppe 1 - kürzere Verweildauer im ao.-Status*

- mehr als 60% der erstmaligen ao.-Schüler\*innen des Schuljahres 2018/19 wechseln bis zum Schuljahresende 2019/20 in den o.-Status.
- mehr als 35% der erstmaligen ao.- Schüler\*innen des Schuljahres 2019/20 wechseln bis zum Schuljahresanfang 2020/21 in den o.-Status.

Damit umfasst Gruppe 1 Standorte, an denen überdurchschnittlich viele Schüler\*innen zum Beobachtungszeitpunkt bereits im ordentlichen Status waren.

#### *Gruppe 2 - längere Verweildauer im ao.-Status:*

- weniger als 40% der erstmaligen ao.- Schüler\*innen des Schuljahres 2018/19 wechseln bis zum Schuljahresende 2019/20 in den o.-Status.
- weniger als 20% der erstmaligen ao.-Schüler\*innen des Schuljahres 2019/20 wechseln bis zum Schuljahresanfang 2020/21 in den o.-Status

Damit umfasst Gruppe 2 Standorte, an denen unterdurchschnittlich viele Schüler\*innen zum Beobachtungszeitpunkt bereits im ordentlichen Status waren.

67 Schulstandorte erfüllten alle Kriterien für Gruppe 1; 49 Schulstandorte alle Kriterien für Gruppe 2. Im nächsten Schritt sollte sichergestellt werden, dass die Gruppen hinsichtlich standortspezifischer Variablen ähnlich sind, um Verzerrungen der Ergebnisse auf Grund von Rahmenbedingungen (Bundesland, Anzahl der Klassen, Anzahl der Schüler\*innen gesamt, Anzahl der ao.-Schüler\*innen, Verhältnis der ao.- zu den o.-Schüler\*innen, Urbanitätsgrad, Sozialindex, integrative / nicht-integrative Beschulung) zu vermeiden. Fünf Schulen hätten durch ihre Standortbedingungen die Vergleichbarkeit der Gruppen verzerrt und wurden daher ausgeschlossen.

Ein weiterer Standort wurde ausgeschlossen, da dort zu Beginn der Datenerhebung im Sommersemester 2022 keine ao.-Schüler\*innen mehr geführt wurden. Schließlich wurden 110 Schulen zur Teilnahme eingeladen und Links zu den Online-Fragebögen an alle drei Rollen (Schulleitungen, Lehrpersonen DF, Lehrpersonen VS) verschickt.

#### **Stichprobenbeschreibung**

Die Teilnahme an der Studie war völlig freiwillig. Den Befragten wurde absolute Vertraulichkeit zugesichert. Das bedeutete, dass die Rohdaten aller Erhebungen beim Evaluationsteam verbleiben und die Auswertungen immer in aggregierter Form<sup>2</sup> erfolgen. An 93 der 110 Schulen nahm zumindest eine Person teil. Dies entspricht 84,55% der Schulen. 17 Schulen nahmen trotz zweier Erinnerungsmails nicht an der Evaluation teil. Die Gesamtstichprobe umfasst schließlich 693 Personen (siehe Tabelle 1). An 80 der 110 Schulen (72,73%) beteiligte sich die Schulleitung. An 72 der 110 Schulen (65,45%) beteiligte sich zumindest eine Lehrperson DF. An 70 der 110 Schulen (71,82%) beteiligte sich zumindest eine Lehrperson VS. Auf Schulebene nahmen durchschnittlich 43,79% der Lehrpersonen DF und 21,60% der Lehrpersonen VS teil. Der Rücklauf unterschied sich in Bezug auf die beiden gebildeten

---

<sup>2</sup> Aggregierte Daten sind zusammengeführte Daten von mehreren Personen, die nicht mehr auf eine Person zurückführbar sind.

Gruppen (Gruppe 1, „kürzere Verweildauer“ bzw. Gruppe 2 „längere Verweildauer“) nicht signifikant (d.h. statistisch bedeutsam).

**Tabelle 1 – Stichprobenbeschreibung**

	N			Geschlecht	Alter	Dienstjahr
	Gesamt	Gruppe 1	Gruppe 2			
Schulleitungen	80	51	29	88.8% weiblich 11.3% männlich 0% divers	M = 53.29 SD = 6.18 Range = 35 - 64	M = 8.91 SD = 6.69 Range = 1 - 33
Lehrpersonen DF	167	106	61	95.2% weiblich 4.8% männlich 0% divers	M = 43.82 SD = 12.05 Range = 21 - 63	M = 17.01 SD = 13.26 Range = 1 - 43
Lehrpersonen VS	446	288	158	95.1% weiblich 4.5% männlich 0% divers	M = 40.35 SD = 10.77 Range = 22 - 64	M = 15.47 SD = 11.23, Range = 1 - 53

**Anmerkung zur Interpretation der Stichprobenbeschreibung sowie der Ergebnisse.** Der Mittelwert (M) beschreibt den im Durchschnitt angegebenen Wert. Die Standardabweichung (SD) ist ein Kennwert für die durchschnittliche Abweichung vom Mittelwert. Sie gibt an, wie übereinstimmend (kleinere Standardabweichung) oder verschieden (größere Standardabweichung) die Angaben der Befragten sind. Die Range beschreibt die Spannweite aller gemachten Angaben (d.h. kleinster bis größter Wert).

Lehrpersonen DF in der Stichprobe verfügten durchschnittlich über 2.54 Schuljahre Erfahrung mit dem derzeitigen gesetzlichen Deutschfördermodell (SD = 1.18; Range = 1 – 4). Die Erfahrung der Lehrpersonen DF in der Deutschförderung vor Einführung des derzeitigen gesetzlichen Deutschfördermodells betrug durchschnittlich 9.85 Jahre (SD = 6.85, Range = 1 – 34). In beiden Fällen gab es keine signifikanten Unterschiede zwischen den Gruppen.

Abschließend ist hinsichtlich der Datenerhebung zu erwähnen, dass diese im Sommersemester 2022 nach wie vor im Kontext der COVID-19 Pandemie stattfand. So waren zum Zeitpunkt der Erhebung teilweise Einschränkungen des Unterrichts an den Schulen gegeben, sowie Hygienevorschriften in Kraft. Auch von den Lockdowns, Schulschließungen und einer Reihe weiterer Einschränkungen ab März 2020 waren die Schulen hinsichtlich des Unterrichts sowie auch der Implementierung des Deutschfördermodells betroffen. Dies ist inhaltlich bei der Interpretation der Ergebnisse zu berücksichtigen. Nichtsdestotrotz – zumal alle Schulen gleichermaßen von dieser Situation betroffen waren – ist nicht anzunehmen, dass es auf Grund der Kontextbedingungen der Pandemie systematische Unterschiede zwischen den Gruppen gab, die den Extremgruppenvergleich verzerren würden.

### **Erhebungsinstrumente**

Die Erhebungsinstrumente für alle drei Rollen (Schulleitungen, Lehrpersonen DF, Lehrpersonen VS) wurden auf Basis der Ergebnisse der beiden im Vorprojekt durchgeführten Workshops (siehe [Workshopbericht](#)) sowie auf Basis einschlägiger wissenschaftlicher Forschung<sup>3</sup> erstellt. Sie beinhalten

<sup>3</sup> Siehe Kapitel „Weiterführende Literatur“ für einen Überblick über einige relevante Theorien und einschlägige wissenschaftliche Befunde in Zusammenhang mit dem vorliegenden Evaluationsprojekt.

überwiegend geschlossene Fragen (d.h. mit vorgegebenen Antwortmöglichkeiten) zu Implementierungsfaktoren, Kontextvariablen sowie Ansätzen zur Optimierung des Deutschfördermodells. Zusätzlich wurden einige offene Fragen gestellt, die es ermöglichten, eigene Meinungen und Einschätzungen zu formulieren. Um die Passung der Erhebungsinstrumente zu prüfen, wurden sie in zwei Pilotschulen (je einer Schule aus den beiden Gruppen) vorgegeben. Die Befragung der Schulleitungen erfolgte mittels Interviews via zoom, wobei die Fragen des Online-Fragebogens vorgegeben wurden. Diese Form der Pilotierung hat den Vorteil, dass Rückfragen gestellt werden können. Rückmeldungen und Kommentare wurden schriftlich festgehalten. Auch die geplante Organisation der Erhebung wurde im Rahmen der Interviews durchbesprochen, um die Meinung der Schulleitungen zum Vorgehen, insbesondere in Bezug auf die reibungslose Umsetzbarkeit der Datenerhebung, einzuholen. Auf Basis der Interviews mit den Schulleitungen wurden geringfügige sprachliche Adaptionen und kleine Ergänzungen im Schulleitungsfragebogen vorgenommen (sowie bei parallelen Items in den beiden Lehrpersonenfragebögen). Beide Schulleitungen erachteten die geplante Herangehensweise als reibungslos durchführbar. Die Lehrpersonen DF und VS der Pilotschulen wurden online befragt. Eine Kommentarfunktion im Online-Fragebogen ermöglichte es, dass die Lehrpersonen Unklarheiten und Feedback äußern konnten. Zusätzlich wurde dadurch auch die Organisation der Erhebung an den Schulen getestet. Die Lehrpersonen machten keine Kommentare zu den Erhebungsinstrumenten.

### **Vorgehen bei der Datenanalyse**

Sämtliche Analysen wurden entsprechend wissenschaftlicher Standards durchgeführt. Daher werden statistische Kennwerte berichtet, mit welchen vermutlich nicht alle Leser\*innen vertraut sind. Diese werden im Folgenden kurz erläutert. Der Bericht ist jedoch so verfasst, dass die Ergebnisse auch ohne statistische Kenntnisse verständlich sind.

Die meisten Befragungsthemen wurden mittels mehrerer geschlossenen Fragen (Items) erfasst. Diese wurden, wie in der psychologischen Forschung üblich, zu Mittelwertskalen zusammengefasst. Um dies statistisch abzusichern, wurde das in der Psychologie übliche Maß der internen Konsistenz (Koeffizient  $\alpha$ ) angewandt. Dieser Koeffizient kann Werte zwischen 0 und 1 annehmen. Entsprechend üblicher Cutoff-Kriterien wurden Werte über  $\alpha = .60$  als akzeptabel eingestuft. Im nächsten Schritt wurden sämtliche Ergebnisse separat für die Gruppen deskriptivstatistisch dargestellt und anschließend durch Signifikanztests verglichen. Je nach Beschaffenheit der Daten wurden t-Tests oder Chi-Quadrat-Tests angewandt. Als Signifikanzniveau für die statistische Bedeutsamkeit wurde der übliche Kennwert ( $p = .05$ ) festgelegt<sup>4</sup>. Zudem stützt sich die Interpretation der Daten auf die ermittelten Effektstärken (Cohen's  $d$  bzw. Cramer's  $V$ ), wobei Werte um  $d = 0.2$ ,  $d = 0.5$  und  $d = 0.8$ ; sowie  $V = 0.1$ ,  $V = 0.3$  und  $V = 0.5$  kleine, mittlere und große Effekte widerspiegeln und entsprechend auf die praktische Bedeutsamkeit der Ergebnisse verweisen.

Die Auswertung der Antworten auf die offenen Fragen erfolgte durch zusammenfassende Inhaltsanalysen. Hierzu wurden die Antworten inhaltlich gruppiert und auf knappe, nur auf den Inhalt beschränkte Kurztexte (Paraphrasen) reduziert. Die Paraphrasen sowie die Beteiligung an der Beantwortung der offenen Fragen wurden separat für die beiden Gruppen ausgezählt und hinsichtlich Gemeinsamkeiten und Unterschieden verglichen.

---

<sup>4</sup> Im Sinne der Offenlegung von Limitationen verweisen wir hinsichtlich der quantitativen Analysen auf eine mögliche Alpha-Fehler-Kumulierung: Werden mehrere Hypothesen an derselben Stichprobe gleichzeitig getestet, so steigt die Wahrscheinlichkeit, dass Hypothesen (inkorrekt) als zutreffend angenommen werden. Aufgrund der für die durchgeführten Analysen eher geringen Stichprobengrößen haben wir jedoch auf die sehr konservative Vorgehensweise der Bonferroni-Korrektur verzichtet.

## 4. Ergebnisse

Im Folgenden werden die Ergebnisse der Evaluation der Implementierung der Deutschförderung entlang der Themen der Befragung für beide Gruppen und alle drei Rollen dargestellt. Zu jedem der Themen wurden geschlossene Fragen gestellt sowie teilweise auch offene (hier waren Mehrfachantworten möglich). Zur Beantwortung der geschlossenen Fragen wurden fünf Antwortmöglichkeiten vorgegeben, die von 1 (= keine Zustimmung) bis 5 (= Zustimmung) reichen. Die Antworten auf die offenen Fragen wurden nur dann in den Bericht aufgenommen, wenn sich von jeder Gruppe mindestens 5 Personen beteiligt hatten (d.h. mindestens 10 Antworten pro Frage). Die Antworten wurden im ersten Schritt paraphrasiert und strukturiert. Danach wurden sie hinsichtlich der Häufigkeiten der Angaben gereiht. Im Bericht werden häufig vorkommende Angaben, die von mehreren Befragten inhaltlich übereinstimmend getätigt wurden, präsentiert. Handelt es sich um einzelne Äußerungen, wird im Speziellen darauf verwiesen.

Da es eine Vielzahl an Detailergebnissen gibt, werden die Ergebnisse im Folgenden, soweit sie inhaltlich übereinstimmen, zusammengefasst und vorwiegend narrativ berichtet. Die entsprechenden statistischen Kennwerte sind in den jeweiligen Tabellen zu finden.

Der Fragebogen umfasste folgende 10 Themen:

1. Einstellung zum Deutschfördermodell und Image des Deutschfördermodells
2. Organisatorische Umsetzung an den jeweiligen Schulstandorten
3. Heterogenität der Schüler\*innen in der Deutschförderung
4. Kooperation und Maßnahmen zur Qualitätssicherung
5. Lehrpersonenmerkmale mit Fokus auf Deutschförderung
6. MIKA-D (Messinstrument zur Kompetenzanalyse – Deutsch)
7. USB DaZ (Unterricht begleitende Sprachstandsbeobachtung Deutsch als Zweitsprache)
8. Zielerreichung der Deutschförderung
9. Vorschläge zur Verbesserung und Weiterentwicklung der Deutschförderung
10. Ergänzende Aspekte aus Sicht der Befragten

Insgesamt zeigen die Ergebnisse sowohl über die beiden Extremgruppen hinweg, als auch für die drei Rollen eine sehr hohe Übereinstimmung. Unterschiede wurden nur selten festgestellt.

### 1. Einstellung zum Deutschfördermodell und Image des Deutschfördermodells

Die Schulleitungen und die Lehrpersonen DF wurden sowohl nach ihrer Einstellung zum Deutschfördermodell gefragt, als auch nach ihrer Einstellung zu den einzelnen Maßnahmen – Deutschförderklassen, Deutschförderkurse, integrative Umsetzung der Deutschförderung. Die Ergebnisse zeigten über Gruppen und Rollen hinweg ähnliche Einschätzungen: Deutschförderklassen werden negativer bewertet als Deutschförderkurse und diese wiederum negativer als die integrative Deutschförderung.

Die Mittelwerte in der Bewertung der integrativen Deutschförderung drückten eher Zustimmung zu dieser Maßnahme aus; während das Deutschfördermodell insgesamt bei beiden Rollen eher geringe Zustimmung fand. Signifikante Unterschiede zwischen beiden Gruppen gab es bei den Lehrpersonen DF bzgl. der Einstellung zu den Deutschförderklassen und bei der Einstellung zum Deutschförderungsmodell. In beiden Fällen berichteten Lehrpersonen DF in Gruppe 2 (= längere Verweildauer im ao.-Status) positivere Einstellungen als Lehrpersonen DF in Gruppe 1 (= kürzere Verweildauer im ao.-Status).

**Tabelle 2 – Einstellung zum Deutschfördermodell**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Einstellung zu Deutschförderklassen</b>			
Schulleitungen	M = 2.33, SD = 1.21	M = 2.55, SD = 1.15	$t(78) = -.80, p = .43; d = -.18$
Lehrpersonen DF	M = 2.61, SD = 1.13	M = 3.03, SD = 1.32	$t(165) = -2.17, p = .03; d = -.35$
<b>Einstellung zu Deutschförderkursen</b>			
Schulleitungen	M = 3.61, SD = 1.10	M = 3.55, SD = 1.21	$t(78) = .21, p = .84; d = .05$
Lehrpersonen DF	M = 3.71, SD = 1.02	M = 3.93, SD = .96	$t(165) = -1.41, p = .16; d = -.23$
<b>Einstellung zur integrativen Deutschförderung</b>			
Schulleitungen	M = 3.94, SD = 1.10	M = 4.38, SD = .86	$t(78) = -1.86, p = .07; d = -.43$
Lehrpersonen DF	M = 3.79, SD = 1.03	M = 3.84, SD = 1.04	$t(165) = -.26, p = .79; d = -.04$
<b>Einstellung zum Deutschfördermodell</b>			
Schulleitungen	M = 2.14, SD = .94	M = 2.41, SD = 1.09	$t(78) = -1.20, p = .24; d = -.28$
Lehrpersonen DF	M = 2.21, SD = .87	M = 2.72, SD = 1.11	Welch: $t(102.39) = -3.10, p < .01; d = -.53$

Lehrpersonen VS wurden danach gefragt, ob sie selbst gerne in der derzeitigen Deutschförderung tätig wären. In beiden Gruppen fand diese Frage eine eher geringe Zustimmung ( $M_{\text{Gruppe1}} = 1.90, SD_{\text{Gruppe1}} = 0.95; M_{\text{Gruppe2}} = 1.96, SD_{\text{Gruppe2}} = 0.92$ ).

Sowohl Schulleitungen als auch Lehrpersonen DF und VS wurden Fragen nach dem Image der Deutschförderung gestellt. Konkret wurden sie gefragt, ob sie der Aussage zustimmen, dass Lehrpersonen, die in der Deutschförderung unterrichten, die Anerkennung der Schulpartner haben (erfasst mittels einer Skala). Alle Befragten stimmten tendenziell eher zu. Es gab keine signifikanten Unterschiede zwischen den Gruppen.

**Tabelle 3 – Image der Lehrpersonen in der Deutschförderung**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
Schulleitungen	M = 3.88, SD = 1.15	M = 3.73, SD = .90	$t(71) = .57, p = .57; d = .14$
Lehrpersonen DF	M = 3.51, SD = 1.15	M = 3.49, SD = 1.28	$t(135) = .10, p = .92; d = .02$
Lehrpersonen VS	M = 3.66, SD = 1.25	M = 3.63, SD = 1.20	$t(348) = .16, p = .87; d = .02$

Die Schulleitungen wurden auch gefragt, wie man das Image der Deutschförderung verbessern könnte. In Gruppe 1 beantworteten 40 Schulleitungen diese Frage; in Gruppe 2 waren es 24. Es gab keine Gruppenunterschiede in der relativen Beteiligung. Die Angaben, wie man das Image der Deutschförderung verbessern könnte, stimmten in beiden Gruppen hoch überein. Häufig wurden „mehr Ressourcen“, „integrativ statt segregiert“ und „mehr Autonomie“ genannt. Die Antwort „kleinere Gruppen“ wurde ebenfalls von Schulleitungen in beiden Gruppen gegeben, häufiger in Gruppe 2.

Da die Vorstudie ergeben hatte, dass sich sowohl innerhalb der Lehrpersonen als auch der Schulleitungen die Einstellungen zum Verbleib der Schüler\*innen in der Deutschförderung (ao.-Status)

unterscheiden, wurden alle Rollen dazu befragt. Alle drei Rollen stimmten tendenziell eher einer möglichst langen Verweildauer zu, als einer kurzen. Signifikante Unterschiede zwischen den beiden Gruppen gab es hinsichtlich der Zustimmung zur möglichst langen Verweildauer bei den Lehrpersonen DF. Lehrpersonen DF in Gruppe 2 stimmen der möglichst langen Verweildauer im ao.-Status tendenziell höher zu als Lehrpersonen DF in Gruppe 1.

**Tabelle 4 – Zustimmung zur möglichst langen Verweildauer der Schüler\*innen im ao.-Status**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
Schulleitungen	M = 3.55, SD = 1.26	M = 3.89, SD = 1.40	$t(74) = -1.08, p = .29; d = -.26$
Lehrpersonen DF	M = 3.60, SD = 1.32	M = 4.05, SD = 1.10	<i>Welch</i> : $t(136.86) = -2.33, p = .02; d = -.37$
Lehrpersonen VS	M = 3.69, SD = 1.27	M = 3.76, SD = 1.21	$t(404) = -.58, p = .56; d = -.06$

Zusätzlich wurden Befragte, die einer möglichst langen Verweildauer zugestimmt hatten („5“) oder eher zugestimmt hatten („4“) in einer offenen Frage nach einer Begründung gefragt. In den Beteiligungen (= Antworten zu der offenen Frage) gab es für alle drei Rollen keine signifikanten Unterschiede zwischen den Gruppen. In beiden Gruppen wurden auch sehr ähnliche Begründungen angegeben, die sich im Wesentlichen nur durch die sprachlichen Formulierungen unterschieden.

In Gruppe 1 stimmten 27 Schulleitungen der langen Verweildauer zu („5“) oder eher zu („4“). 7 davon gaben eine Begründung an. In Gruppe 2 stimmten 18 zu („5“) oder eher zu („4“). Von diesen begründeten 5 Schulleitungen ihre Antwort. Häufige Begründungen waren „Kinder brauchen Zeit, um sich mit der neuen Sprache auseinanderzusetzen“; „zusätzliche Stunden zur Deutschförderung gibt es nur mit ao-Status“; „Kinder erwecken oft den Eindruck die Sprache zu beherrschen – jedoch ist sinnerfassendes Lesen schwierig“.

In Gruppe 1 stimmten 57 Lehrpersonen DF einer langen Verweildauer zu („5“) oder eher zu („4“). 11 davon begründeten ihre Antwort. In Gruppe 2 stimmten 43 zu („5“) oder eher zu („4“). 11 begründeten ihre Antwort. Häufige Begründungen waren „Fördermaßnahmen können nicht nach 2 Jahren einfach beendet werden, manche brauchen längere Förderung, Schulautonomie wäre hilfreich“; „Deutschförderung muss schleichend weniger werden und nicht von 100% auf 0%“; „die deutsche Sprache ist sehr komplex; bessere, dem Wissensstand entsprechende Förderung ist in Kleingruppen möglich“.

156 Lehrpersonen VS in Gruppe 1 stimmten einer längeren Verweildauer in der Deutschförderung zu („5“) oder eher zu („4“). 23 davon begründeten ihre Antwort. In Gruppe 2 waren es 89, wovon 9 ihre Antwort begründeten. Häufige Begründungen lauteten „durch ao.-Status werden Schüler\*innen nicht überfordert, was gut für die Lernmotivation ist“; „die MIKA-D Tests sind nicht sehr zuverlässig und die Zeit könnte besser genutzt werden“; „es sollte beim Spracherwerb berücksichtigt werden, dass Kinder unterschiedlich schnell lernen“; „bei vielen Schüler\*innen ist Deutsch noch zu schlecht, um in allen Fächern richtig beurteilt werden zu können; bei der Beurteilung sind die Noten schlecht, was sich negativ auf die Motivation auswirkt“. Zusätzlich wurde noch häufig angegeben „es ist vorteilhaft, wenn es in der Klasse mehrere Kinder mit der Unterrichtssprache als Muttersprache gibt, weil so Sprachvorbilder vorhanden sind“.

## 2. Organisatorische Umsetzung an den jeweiligen Schulstandorten

Schulleitungen sind für die organisatorische Umsetzung der Deutschförderung an ihrem Standort zuständig; sie wurden daher dazu befragt. Zum Personaleinsatzplan wurde die offene Frage gestellt,

nach welchen Kriterien die Schulleitungen entscheiden, welche Lehrpersonen sie in der Deutschförderung einsetzen. In den Antworten zeigten sich inhaltlich nur geringe Unterschiede zwischen den beiden Gruppen. Am häufigsten wurde „(Zusatz-)ausbildung“ als Kriterium für den Einsatz in der Deutschförderung angegeben. Die Schulleitungen der Gruppe 1 nannten zusätzlich noch „Interesse“, „persönliche Motivation“ und „Erfahrung der Lehrkraft“. In beiden Gruppen wurde auch angegeben, dass aufgrund der Personalknappheit keine Auswahl möglich sei. Dies wurde tendenziell etwas häufiger in Gruppe 2 genannt.

Die Schwierigkeit der Einbettung der Deutschförderung in die Schulorganisation wurde in Gruppe 1 der Schulleitungen höher eingeschätzt als in Gruppe 2 ( $M_{\text{Gruppe1}} = 3.96$ ,  $SD_{\text{Gruppe1}} = 1.07$ ;  $M_{\text{Gruppe2}} = 3.38$ ,  $SD_{\text{Gruppe2}} = 1.21$ ;  $t(77) = 2.22$ ,  $p = .03$ ;  $d = .51$ ).

Von den insgesamt 37 Schulleitungen in Gruppe 1, die Schwierigkeiten in der Einbettung der Deutschförderung in die Schulorganisation angegeben hatten (Antworten „5“ oder „4“) begründeten alle ihre Antwort. In Gruppe 2 gaben 15 Schulleitungen Schwierigkeiten an (Antworten „5“ oder „4“), wovon alle ihre Einschätzung begründeten. Die Antworten zeigten inhaltlich eine hohe Übereinstimmung beider Gruppen. Am häufigsten wurde „Raummangel“ gefolgt von „Personalmangel“ angegeben. Ebenfalls von beiden Gruppen wurden „soziale Probleme“, „Herausreißen aus der Klassengemeinschaft“ und „Schwierigkeiten der Koordination, Organisation“ genannt.

Die Angemessenheit der räumlichen Situation und Ausstattung an den Schulstandorten zur Umsetzung der Deutschförderung wurde mittels einer Skala erhoben. Schulleitungen und Lehrpersonen DF wurden dazu befragt. Die Schulleitungen beider Gruppen beantworteten diese Fragen mit einer leichten Tendenz in Richtung Zustimmung (d.h. Angemessenheit). Bei den Lehrpersonen DF gab es signifikante Gruppenunterschiede. In Gruppe 2 wurde der Angemessenheit höher zugestimmt als in Gruppe 1.

**Tabelle 5 – Angemessenheit der räumlichen Situation und Ausstattung**

	<b>Gruppe 1 - kürzere Verweildauer</b>	<b>Gruppe 2- längere Verweildauer</b>	<b>Signifikanztest</b>
Schulleitungen	M = 3.26, SD = .98	M = 3.54, SD = 1.02	$t(76) = -1.18$ , $p = .24$ ; $d = -.28$
Lehrpersonen DF	M = 2.94, SD = 1.11	M = 3.47, SD = .90	Welch: $t(146.94) = -3.34$ , $p < .01$ ; $d = -.51$

### **3. Heterogenität der Schüler\*innen in der Deutschförderung**

Lehrpersonen DF wurden zum Thema Heterogenität der ao.-Schüler\*innen befragt. Konkret wurden sie gefragt, ob die Deutschförderung durch die Art der Schüler\*innenzusammensetzung erschwert wird: durch zu große Leistungsheterogenität, durch unterschiedliche Begabungen der Schüler\*innen, durch zu viele für diese Schule ungeeignete Schüler\*innen (z.B. Fehlen basaler Fertigkeit der Schulfähigkeit), durch zu viele leistungsschwache Schüler\*innen, oder durch zu hohe Heterogenität in Bezug auf Alter. Gemäß den Antworten der Lehrpersonen DF in beiden Gruppen sind Fördermöglichkeiten für verschiedene Schüler\*innengruppen im Rahmen des Deutschfördermodells tendenziell eher nicht gegeben. Dies gilt für alle vorgegebenen Heterogenitätsdimensionen. D.h. beide Gruppen beantworteten die Frage, ob die Deutschförderung durch die Art der Schüler\*innenzusammensetzung erschwert wird, mit einer Tendenz in Richtung Zustimmung (d.h. Erschwernis). Einen signifikanten Unterschied gab es nur hinsichtlich „zu hohe Heterogenität in Bezug auf Alter“ mit etwas höherer Zustimmung in Gruppe 1.

**Tabelle 6 – Wird die Deutschförderung durch Aspekte der Heterogenität erschwert?**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Lehrpersonen DF</b>			
Zu große Leistungsheterogenität	M = 3.71, SD = 1.12	M = 3.59, SD = 1.30	$t(164) = .65, p = .52; d = .10$
Unterschiedliche Begabung der Schüler*innen	M = 3.45, SD = 1.25	M = 3.36, SD = 1.18	$t(164) = .44, p = .66; d = .07$
Zu viele für diese Schule ungeeignete Schüler*innen (z.B. Fehlen basaler Fertigkeiten der Schulfähigkeit)	M = 3.19, SD = 1.39	M = 3.26, SD = 1.33	$t(164) = -.33, p = .75; d = -.05$
Zu viele leistungsschwache Schüler*innen	M = 3.44, SD = 1.32	M = 3.62, SD = 1.10	Welch: $t(143.99) = -.97, p = .33; d = -.15$
Zu hohe Heterogenität in Bezug auf Alter	M = 3.30, SD = 1.37	M = 2.79, SD = 1.34	$t(164) = 2.32, p = .02; d = .37$

Zusätzlich wurden die Lehrpersonen DF auch nach Verhaltens- und Motivationsproblemen gefragt, die den Unterricht erschweren. Die Fragen wurden zu folgenden Verhaltens- und/ oder Motivationsproblemen gestellt: Schüler\*innen, die den Unterricht stören; verbale Gewalttätigkeit unter den Schüler\*innen; physische Gewalttätigkeit unter den Schüler\*innen; schlechtes Klima unter den Schüler\*innen; großer Anteil an unmotivierten Schüler\*innen; uninteressierte Schüler\*innen. Mit Ausnahme der Frage bzgl. Schüler\*innen, die den Unterricht stören, wurde alle Fragen nach dem Vorkommen von Verhaltens- und Motivationsproblemen tendenziell in Richtung geringe Erschwernis des Unterrichts beantwortet. Es gab keine signifikanten Unterschiede zwischen den beiden Gruppen.

**Tabelle 7 – Wird die Deutschförderung durch Verhaltens- und Motivationsprobleme erschwert?**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Lehrpersonen DF</b>			
Schüler*innen, die den Unterricht stören	M = 3.47, SD = 1.31	M = 3.26, SD = 1.35	$t(165) = .98, p = .33; d = .16$
Verbale Gewalttätigkeit unter den Schüler*innen	M = 2.52, SD = 1.31	M = 2.51, SD = 1.18	$t(165) = .05, p = .96; d = .01$
Physische Gewalttätigkeit unter den Schüler*innen	M = 2.19, SD = 1.24	M = 2.13, SD = 1.04	$t(165) = .31, p = .76; d = .05$
Schlechtes Klima unter den Schüler*innen	M = 2.33, SD = 1.23	M = 2.26, SD = 1.06	$t(165) = .36, p = .72; d = .06$
Großer Anteil an unmotivierten Schüler*innen	M = 2.80, SD = 1.23	M = 2.98, SD = 1.18	$t(165) = -.93, p = .35; d = -.15$

#### 4. Kooperation und Maßnahmen zur Qualitätssicherung

Allen drei Rollen wurden nach Strukturen des Austauschs zwischen Lehrpersonen in der Deutschförderung und dem restlichen Lehrkörper gefragt. Die Schulleitungen berichteten das Vorhandensein solcher Strukturen zu einem sehr hohen Prozentsatz – in Gruppe 1 beantworteten 92.2% die Frage mit „Ja“; in Gruppe 2 waren es 96,6%. Die Zustimmung in den beiden Lehrpersonengruppen war deutlich niedriger (Lehrpersonen DF: Gruppe 1: 50.0%; Gruppe 2: 57.4%; Lehrpersonen VS: Gruppe 1: 62.8%; Gruppe 2: 61.8%). Zwischen den Gruppen gab es keine signifikanten Unterschiede.

Die Frage, wie die Strukturen des Austauschs gestaltet sind, wurde von 73 Schulleitungen (Gruppe 1: 46; Gruppe 2: 27) beantwortet, von 62 Lehrpersonen DF (Gruppe 1: 36; Gruppe 2: 26) und von 148 Lehrpersonen VS (Gruppe 1: 99; Gruppe 2: 49). Es gab keine signifikanten Gruppenunterschiede in der Beteiligung. In beiden Gruppen gaben die Schulleitungen hauptsächlich Teambesprechungen zwischen verschiedenen Lehrpersonen (innerhalb der Deutschförderung und zwischen Lehrpersonen DF und Klassenlehrpersonen) sowie Konferenzen an. Während die Besprechungen bei einigen in regelmäßigen Abständen etabliert sind, finden sie bei anderen nach Bedarf statt. Die meisten Lehrpersonen DF beider Gruppen berichteten von Teambesprechungen. Teilweise werden zusätzlich informelle Kommunikationsformen wie Telefonate und schriftlicher Austausch per Messenger verwendet. Auch bei Konferenzen wird auf die Deutschförderung eingegangen. Laut Lehrpersonen VS findet der meiste Austausch durch Teambesprechungen und/oder Einzelgesprächen mit den Lehrpersonen DF statt. Manche Lehrpersonen berichteten dabei von regelmäßigen Terminen, andere von informellen Treffen und einem Austausch zwischendurch. Viele Lehrpersonen VS berichteten auch von einem Austausch in Konferenzen. In manchen Fällen findet der Austausch auch über SMS/Mails/Telefonate/MS-Teams statt. Viele Lehrpersonen VS berichteten auch von gemeinsamen Absprachen mit den Lehrpersonen DF bezüglich der Planung und Durchführung des Unterrichts und dass die Lehrpersonen DF sie bezüglich der Lernfortschritte der Schüler\*innen auf dem Laufenden halten. Jeweils einzelne Lehrpersonen VS gaben an, dass ihnen die Lehrpersonen DF spezielle Materialien zur Verfügung stellen oder Empfehlungen abgeben und dass Elterngespräche gemeinsam mit den Lehrpersonen DF durchgeführt werden. Die Gruppen unterschieden sich inhaltlich nicht in ihren Aussagen.

Auf Fragen nach dem Ausmaß der Kooperation und Zusammenarbeit zwischen Lehrpersonen DF und Lehrpersonen VS bzgl. ao.-Schüler\*innen gaben beiden Rollen im Mittel an, dass die Kooperation in hohem Ausmaß gegeben ist. Jedoch gaben sowohl Lehrpersonen DF und VS der Gruppe 2 ein noch höheres Ausmaß an als Lehrpersonen DF und VS in Gruppe 1. Das Ausmaß der Kooperation mit den Eltern (erfasst mittels einer Skala) wurde von den Lehrpersonen DF in beiden Gruppen als tendenziell hoch bewertet.

**Tabelle 8 – Kooperation**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Kooperation und fachliche Zusammenarbeit zwischen den Lehrpersonen der ao-Schüler*innen</b>			
Lehrpersonen DF	M = 4.04, SD = .90	M = 4.29, SD = .71	Welch: $t(147.27) = -1.94, p = .05; d = -.29$
Lehrpersonen VS	M = 3.82, SD = 1.18	M = 4.08, SD = .98	Welch: $t(269.82) = -2.08, p = .04; d = -.24$
<b>Kooperation zwischen Lehrpersonen und Eltern der ao.-Schüler*innen</b>			
Lehrpersonen DF	M = 3.19, SD = 1.04	M = 3.50, SD = 1.04	$t(165) = -1.84, p = .07; d = -.30$

Die Schulleitungen wurden (mittels einer Skala) danach gefragt, ob an den Schulstandorten Maßnahmen zur Qualitätssicherung bzgl. Organisations-, Personal- und Unterrichtsentwicklung etabliert sind, um die Qualität der Deutschförderung zu sichern und weiterzuentwickeln. Die Mittelwerte in beiden Gruppen lagen um den Mittelpunkt der Skala („trifft teilweise zu“). Einen signifikanten Unterschied gab es hinsichtlich der Organisationsentwicklung. Schulleitungen in Gruppe 2 stimmten in höherem Ausmaß zu, dass entsprechende Maßnahmen zur Qualitätssicherung etabliert sind.

**Tabelle 9 – Maßnahmen zur Qualitätssicherung**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Schulleitungen</b>			
Organisationsentwicklung	M = 2.65, SD = .82	M = 3.06, SD = 1.00	$t(78) = -1.98, p = .05; d = -.46$
Personalentwicklung	M = 2.83, SD = .84	M = 3.23, SD = .98	$t(78) = -1.93, p = .06; d = -.45$
Unterrichtsentwicklung	M = 3.80, SD = .79	M = 3.78, SD = .77	$t(78) = .09, p = .93; d = .02$

## 5. Lehrpersonenmerkmale mit Fokus auf Deutschförderung

Lehrpersonenmerkmale wurden vorwiegend von Lehrpersonen DF erhoben, da diese für die ao.-Schüler\*innen besondere Relevanz haben; teilweise auch von Lehrpersonen VS.

Die Frage, ob sie sich zu Beginn ihrer Tätigkeit in der Deutschförderung dieser Aufgabe gewachsen gefühlt hatten, beantworteten 36.3% der Lehrpersonen DF in Gruppe 1 mit „eher Ja“ und 33.3% mit „Ja“. In Gruppe 2 waren es 46.6% („eher Ja“), respektive 31.0% („Ja“). Lehrpersonen DF beider Gruppen fühlten sich durch das Unterrichten in der Deutschförderung eher nicht überfordert, mit keinen signifikanten Gruppenunterschieden ( $M_{\text{Gruppe1}} = 2.45, SD_{\text{Gruppe1}} = 0.87; M_{\text{Gruppe2}} = 2.27, SD_{\text{Gruppe2}} = 0.82; t(165) = 1.36, p = .18; d = .22$ ).

Ihr Selbstwirksamkeitserleben hinsichtlich des Unterrichts von Schüler\*innen im ao.-Status (erhoben mittels einer Skala) schätzten die Lehrpersonen DF in beiden Gruppen eher hoch ein. Die Frage nach dem Selbstwirksamkeitserleben wurde auch den Lehrpersonen VS gestellt. Bei diesen gab es signifikante Gruppenunterschiede mit höherer Zustimmung in Gruppe 2. Sowohl Lehrpersonen DF als auch Lehrpersonen VS gaben an, eine sehr positive Beziehung zu den ao.-Schüler\*innen zu haben. Bei den Lehrpersonen DF zeigten sich dabei signifikante Gruppenunterschiede mit höherer Zustimmung in Gruppe 2.

**Tabelle 10 – Selbstwirksamkeit und Lehrer\*innen-Schüler\*innen-Beziehung**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Selbstwirksamkeitserleben hinsichtlich des Unterrichts von ao.-Schüler*innen</b>			
Lehrpersonen DF	M = 3.83 SD = .85	M = 4.07, SD = .73	$t(165) = -1.82, p = .07; d = -.29$
Lehrpersonen VS	M = 3.39, SD = .97	M = 3.59, SD = .80	$Welch: t(273.50) = -1.98, p = .05; d = -.22$
<b>Positive Beziehung zwischen Lehrpersonen und den ao.-Schüler*innen</b>			
Lehrpersonen DF	M = 4.72, SD = .50	M = 4.88, SD = .31	$Welch: t(163.96) = -2.60, p = .01; d = -.37$
Lehrpersonen VS	M = 4.60, SD = .67	M = 4.72, SD = .50	$Welch: t(285.99) = -1.70, p = .09; d = -.19$

Lehrpersonen VS wurden auch Fragen zum sprachbewussten Unterricht gestellt. In beiden Gruppen wurde tendenziell angegeben, mit sprachbewusstem Unterricht vertraut zu sein und entsprechende Prinzipien umzusetzen.

**Tabelle 11 – Sprachbewusster Unterricht**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Lehrpersonen VS</b>			
„Ich bin mit dem Konzept des sprachbewussten / sprachsensiblen Unterrichts vertraut.“	M = 3.48, SD = 1.28	M = 3.58, SD = 1.32	$t(444) = -.83, p = .41; d = -.08$
„Obwohl ich selbst nicht in der Deutschförderung tätig bin, sehe ich es als meine Aufgabe am kontinuierlichen Aufbau der Bildungssprache bei allen Schüler*innen mitzuwirken.“	M = 4.76, SD = .55	M = 4.80, SD = .54	$t(444) = -.81, p = .42; d = -.08$
„Durchgängige Sprachbildung hat in meinem Unterricht in allen Gegenständen einen hohen Stellenwert.“	M = 4.68, SD = .61	M = 4.73, SD = .55	$t(444) = -.75, p = .45; d = -.08$

Die Lehrpersonen DF wurden auch gefragt, ob sie sich freiwillig für die Tätigkeit in der Deutschförderung gemeldet hatten. In Gruppe 1 beantworteten 76.2% die Frage mit „Ja“; in Gruppe 2 waren es 73.8%. Der Unterschied ist statistisch nicht signifikant. Zwischen den Gruppen gab es ebenso keine signifikanten Unterschiede dahingehend, wann sie über ihren Einsatz in der Deutschförderung erfahren hatten. Es zeigten sich jedoch eine Reihe von Unterschieden zwischen den Lehrpersonen DF, die erst wenige Tage/unmittelbar vor Beginn der Tätigkeit erfahren hatten, dass Sie in der Deutschförderung arbeiten werden, und den Lehrpersonen, die es früher erfahren hatten. Die Gruppe, die es spät erfahren hatte, gab geringere Werte in der Selbstwirksamkeit an, geringere fachliche Zusammenarbeit mit den Lehrpersonen, die nicht in der Deutschförderung unterrichten, weniger gut gelingende Kooperation mit den Eltern der ao.-Schüler\*innen, sowie höhere Arbeitsüberforderung.

**Tabelle 12 – Einsatz in der Deutschförderung**

	Gruppe 1 - kürzere Verweildauer		Gruppe 2- längere Verweildauer		Signifikanztest
<b>Lehrpersonen DF</b>					
Haben Sie sich freiwillig für die Tätigkeit in der Deutschförderung gemeldet?	Ja	76.2%	Ja	73.8%	$\chi^2(1) = .12, p = .73, \phi = .03$
	Nein	23.8%	Nein	26.2%	
Wann haben Sie erfahren, dass Sie in der Deutschförderung arbeiten werden?					$\chi^2(3) = 2.19, p = .53, v = .13$
Mehr als 2 Monate vor Beginn der Tätigkeit	65.5%		57.7%		
Mehr als 1 Monat vor Beginn der Tätigkeit	5.7%		3.8%		
2 bis 3 Wochen vor Beginn der Tätigkeit	6.9%		13.5%		
Wenige Tage/unmittelbar vor Beginn der Tätigkeit	21.8%		25.0%		

Lehrpersonen DF	Früher vom Einsatz erfahren > wenige Tage davor	Kurzfristig vom Einsatz erfahren ≤ wenige Tage davor	Signifikanztest
Selbstwirksamkeitserleben hinsichtlich des Unterrichts von ao.-Schüler*innen	M = 4.00, SD = 0.78	M = 3.61, SD = 0.84	$t(138) = -2,43, p = .02; d = .80$
Kooperation und fachliche Zusammenarbeit zwischen den Lehrpersonen der ao-Schüler*innen	M = 4.23, SD = 0.82	M = 3.86, SD = 0.90	$t(137) = -2,20, p = .03; d = .84$
Kooperation zwischen Lehrpersonen und Eltern der ao.-Schüler*innen	M = 3.36, SD = 1.01	M = 2.92, SD = 1.11	$t(138) = -2,08, p = .04; d = 1.0$
Arbeitsüberforderung	M = 2.30, SD = 0.81	M = 2.63, SD = 0.93	$t(138) = 1.98, p = .01; d = .84$

## 6. MIKA-D (Messinstrument zur Kompetenzanalyse – Deutsch)

Im ersten Schritt wurden die Schulleitungen nach der formalen Handhabung des MIKA-D gefragt, d.h. wer die Testungen durchführt und wer die Ergebnisse der MIKA-D Testungen einmeldet. In beiden Gruppen wurde am häufigsten angegeben, dass mehrere Lehrpersonen die MIKA-D Testungen durchführen (Gruppe 1: 70.0%; Gruppe 2: 71.4%). Die zentrale Einmeldung der Ergebnisse wird fast ausschließlich von den Schulleitungen gemacht (Gruppe 1: 90.0%; Gruppe 2: 92.9%).

Der MIKA-D ist fast allen Lehrpersonen bekannt. Von den Lehrpersonen DF kennen ihn in Gruppe 1 98.1%, in Gruppe 2 sind es 93.4%. In Gruppe 1 der Lehrpersonen VS gaben 95.1% an ihn zu kennen, in Gruppe 2 waren es 93.0%. Die durchschnittliche Zustimmung zur Aussage „Die MIKA-D Testergebnisse sind für die Arbeit der Lehrpersonen in der Deutschförderung sehr nützlich.“ lag in beiden Gruppen sowohl für die Lehrpersonen DF als auch für die Lehrpersonen VS um den Mittelpunkt der Skala („stimmt teilweise“). Bei den Lehrpersonen DF gab es einen signifikanten Unterschied zwischen den Gruppen mit höherer Zustimmung in Gruppe 2.

Alle drei Rollen wurden auch gefragt, wie gut der MIKA-D geeignet ist, um den Erfolg der Deutschförderung festzustellen (Lehrpersonen VS wurde diese Frage nur dann gestellt, wenn sie den MIKA-D kennen bzw. durchführen). Wieder lagen die Mittelwerte in beiden Gruppen und für alle Rollen um den Mittelpunkt der Skala („teilweise geeignet“) mit keinen signifikanten Gruppenunterschieden.

Tabelle 13 – MIKA-D

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>„Die MIKA-D Testergebnisse sind für die Arbeit der Lehrpersonen in der Deutschförderung sehr nützlich.“</b>			
Schulleitungen	M = 2.78, SD = 1.19	M = 2.76, SD = 1.22	$t(78) = .09, p = .93; d = .02$
Lehrpersonen DF	M = 2.71, SD = 1.28	M = 3.19, SD = 1.43	$t(151) = -2.11, p = .04; d = -.36$
Lehrpersonen VS	M = 2.86, SD = 1.14	M = 2.73, SD = 1.29	$t(247) = .82, p = .42; d = .11$
<b>„Wie gut ist der MIKA-D aus Ihrer Sicht geeignet, um den Erfolg der Deutschförderung festzustellen?“</b>			
Schulleitungen	M = 3.06, SD = .98	M = 2.79, SD = 1.18	$t(75) = 1.09, p = .28; d = .26$
Lehrpersonen DF	M = 2.84, SD = 1.13	M = 3.02, SD = 1.15	$t(151) = -.93, p = .36; d = -.16$
Lehrpersonen VS	M = 2.96, SD = .99	M = 2.87 SD = 1.08	$t(251) = .62, p = .53; d = .08$

In einer offenen Frage wurden die Schulleitungen und Lehrpersonen DF auch gefragt, was mit den MIKA-D Testergebnissen passiert und wie sie genutzt werden. Die Frage wurde von 60 Schulleitungen (Gruppe 1: 40; Gruppe 2: 20) und von 77 Lehrpersonen DF (Gruppe 1: 50; Gruppe 2: 27) mit keinen signifikanten Gruppenunterschiede in der Beteiligung beantwortet. Die Antworten waren über Rollen und Gruppen hinweg hoch übereinstimmend. Am häufigsten werden die Ergebnisse für die Förderplanung genutzt, gefolgt von der Nutzung für Evaluation und der Antwort „keine Nutzung, nur Aufbewahrung“. Einige Schulleitungen nehmen den MIKA-D auch als unbrauchbar wahr (tendenziell etwas häufiger in Gruppe 2). Die Lehrpersonen DF gaben zusätzlich an, die Testergebnisse auch für Feedback an Eltern und Lehrer\*innenteams zu nutzen (diese Angabe wurde tendenziell etwas häufiger in Gruppe 2 gemacht).

## 7. USB DaZ (Unterricht begleitende Sprachstandsbeobachtung Deutsch als Zweitsprache)

Die Bekanntheit des USB DaZ ist bei den Lehrpersonen DF sehr hoch, ebenso bei den Schulleitungen, wobei der Bekanntheitsgrad in Gruppe 2 signifikant höher ist. Bei Lehrpersonen VS ist der Bekanntheitsgrad des USB DaZ tendenziell niedriger. Etwa die Hälfte der Schulleitungen sowohl in Gruppe 1 und Gruppe 2 gaben an, dass der USB DaZ manchmal eingesetzt wird.

Die Frage, ob der USB DaZ für die Arbeit der Lehrpersonen in der Deutschförderung sehr nützlich ist, wurde sowohl den Schulleitungen als auch den Lehrpersonen DF und VS gestellt. Sämtliche Mittelwerte lagen um den Mittelpunkt der Skala. Auch die Mittelwerte der Lehrpersonen VS auf die Frage nach der Nützlichkeit des USB DaZ für ihre Arbeit im Regelunterricht, lag in beiden Gruppen geringfügig unter dem Mittelpunkt der Skala.

**Tabelle 14 – USB DaZ**

	Gruppe 1 - kürzere Verweildauer		Gruppe 2- längere Verweildauer		Signifikanztest
<b>„Kennen Sie den USB DaZ?“</b>					
Schulleitungen	Ja	82.4%	Ja	100%	$\chi^2(1) = 5.77, p = .02, \phi = -.27$
	Nein	17.6%	Nein	0%	
Lehrpersonen DF	Ja	92.5%	Ja	83.6%	$\chi^2(1) = 3.15, p = .08, \phi = .14$
	Nein	7.5%	Nein	16.4%	
Lehrpersonen VS	Ja	52.1%	Ja	55.1%	$\chi^2(1) = .36, p = .55, \phi = -.03$
	Nein	47.9%	Nein	44.9%	
<b>Schulleitungen - „Wird der USB DaZ an Ihrer Schule eingesetzt?“</b>					$\chi^2(2) = 2.53, p = .28, v = .19$
Ja, regelmäßig	28.6%		41.4%		
Manchmal	47.6%		48.3%		
Nein, gar nicht	23.8%		10.3%		
<b>„Der USB DaZ ist für die Arbeit der Lehrpersonen in der Deutschförderung sehr nützlich.“</b>					
Schulleitungen	M = 3.10, SD = 1.03		M = 3.10, SD = 1.17		$t(48) = 0.00, p = 1; d = 0$
Lehrpersonen DF	M = 2.91, SD = 1.12		M = 3.19, SD = 1.28		$t(90) = -1.12, p = .27; d = -.24$
Lehrpersonen VS	M = 3.31, SD = 1.21		M = 3.10, SD = 1.25		$t(102) = .86, p = .39; d = .17$
<b>„Der USB DaZ ist für meine Arbeit im Regelunterricht sehr nützlich.“</b>					
Lehrpersonen VS	M = 2.75, SD = 1.22		M = 2.86, SD = 1.30		$t(104) = -.46, p = .65; d = -.09$

Die Frage, was mit den USB DaZ Ergebnissen passiert und wie diese genutzt werden, wurde nur den Schulleitungen gestellt. 29 Personen aus Gruppe 1 und 23 Personen aus Gruppe 2 beantworteten diese Frage. Die Antworten wiesen eine hohe Übereinstimmung mit den Antworten bzgl. der Nutzung der Ergebnisse auf. An erster Stelle wurde in beiden Gruppen die Nutzung des USB DaZ für die Förderplanung angegeben; der USB DaZ wurde von Befragten in beiden Gruppen jedoch auch als unbrauchbar wahrgenommen und analog zum MIKA-D wurde auch „keine Nutzung, nur Aufbewahrung“ von den Schulleitungen angegeben. Die Nutzung für Evaluation wurde ebenfalls genannt; etwas häufiger in Gruppe 1 als in Gruppe 2.

## 8. Zielerreichung der Deutschförderung

Den Lehrpersonen DF wurden gebeten einzuschätzen, wie hoch der Anteil der Schüler\*innen ist, die die Ziele der Deutschförderung am Ende des außerordentlichen Status erreicht haben. Die Einschätzungen wurden sowohl bzgl. der sprachbezogenen Ziele als auch bzgl. der flankierenden Ziele gemacht. Die Ziele waren im Rahmen der Expert\*innen-Workshops im Vorprojekt expliziert worden.

**Tabelle 15 – Anteil der Schüler\*innen (Prozentsatz), die am Ende des ao.-Status die Ziele der Deutschförderung erreicht haben – aus Sicht der Lehrpersonen DF**

	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
<b>Sprachbezogene Ziele</b>			
Dem Unterricht folgen können	M = 69.87, SD = 19.58	M = 65.38, SD = 19.75	$t(165) = 1.42, p = .16; d = .23$
Sich auf Deutsch verständigen	M = 75.10, SD = 20.74	M = 70.41, SD = 19.04	$t(165) = 1.45, p = .15; d = .23$
Sicherheit in Alltagskommunikation	M = 66.90, SD = 20.91	M = 62.31, SD = 20.79	$t(165) = 1.37, p = .17; d = .22$
Wortschatz für Alltagshandeln	M = 69.94, SD = 18.59	M = 65.41, SD = 22.49	$t(165) = 1.40, p = .16; d = .23$
Bedürfnisse artikulieren können	M = 77.07, SD = 17.72	M = 69.79, SD = 24.42	Welch: $t(96.79) = 2.04, p = .04; d = .36$
Sich trauen, Deutsch zu sprechen	M = 79.22, SD = 18.78	M = 76.28, SD = 21.85	$t(165) = .93, p = .35; d = .15$
Erwerb des Sprachsystems (Lautebene, Grammatik, Alphabetisierung)	M = 62.13, SD = 21.32	M = 58.10, SD = 22.99	$t(165) = 1.14, p = .25; d = .18$
Sprachbezogene Lernstrategien anwenden z.B. Wörterbuchgebrauch	M = 52.31, SD = 26.38	M = 45.00, SD = 27.63	$t(165) = 1.70, p = .09; d = .27$
Mittelwert über alle sprachbezogenen Ziele	M = 69.07, SD = 16.07	M = 64.08, SD = 18.39	$t(165) = 1.83, p = .07; d = .29$
<b>Flankierende Ziele</b>			
Elemente der Schulfähigkeit wie Aufmerksamkeit, Konzentration, etc. erwerben, um dem Unterricht folgen zu können	M = 70.08, SD = 20.37	M = 69.18, SD = 20.40	$t(165) = .27, p = .79; d = .04$
Selbstvertrauen erlangen	M = 72.39, SD = 20.92	M = 71.56, SD = 19.47	$t(165) = .25, p = .80; d = .04$
Repertoire von Erst- und Zweitsprache gemeinsam nutzen können – Mehrsprachigkeits- kompetenz erlangen	M = 56.27, SD = 24.29	M = 56.34, SD = 24.24	$t(165) = -.02, p = .99; d = -.00$
Soziale Kompetenz erlangen	M = 74.70, SD = 19.73	M = 74.28, SD = 19.52	$t(165) = .133, p = .90; d = .02$
Mittelwert über alle flankierenden Ziele	M = 68.36, SD = 16.20	M = 67.84, SD = 17.60	$t(165) = .19, p = .85; d = .03$

In den Einschätzungen der Lehrpersonen DF gab es nur bzgl. des Zieles „Bedürfnisse artikulieren können“ einen signifikanten Unterschied zwischen den beiden Gruppen. Gruppe 1 schätzte den Anteil der Schüler\*innen, die dieses Ziel erreicht haben, höher ein als Gruppe 2.

Insgesamt wurden die Anteile der Schüler\*innen, die am Ende des ao.-Status die sprachbezogenen Ziele erreichen, von den Lehrpersonen DF zwischen 79.22% (sich trauen Deutsch zu sprechen) und 45.00% (sprachbezogene Lernstrategien einsetzen) eingeschätzt. Das bedeutet im Umkehrschluss, dass gemäß der Lehrpersonen DF zwischen 20.78% und 55.00% der Schüler\*innen die sprachbezogenen Ziele am Ende des ao.-Status nicht erreichen. Bei den flankierenden Zielen liegen die Einschätzungen der Zielerreichung zwischen 74.70% (soziale Kompetenz erlangen) und 56.27% (Mehrsprachigkeitskompetenz erlangen). Damit erreichen im Umkehrschluss laut Lehrpersonen DF zwischen 25.30% und 43.73% der Schüler\*innen diese Ziele am Ende des ao.-Status nicht.

## 9. Vorschläge zur Verbesserung und Weiterentwicklung der Deutschförderung

Einleitend in das Themenfeld Verbesserung/ Weiterentwicklung der Deutschförderung wurden Schulleitungen und Lehrpersonen DF gefragt, ob aus ihrer Sicht Bedarf besteht, das derzeitige gesetzliche Deutschfördermodell zu optimieren. Die Antworten zeigen, dass ein hoher Prozentsatz der Befragten einen entsprechenden Bedarf sieht. Bei den Schulleitungen gab es signifikante Gruppenunterschiede: In Gruppe 1 sahen mehr Schulleitungen einen Bedarf zur Optimierung als in Gruppe 2.

**Tabelle 16 – Optimierungsbedarf des Deutschfördermodells**

	Gruppe 1 - kürzere Verweildauer		Gruppe 2- längere Verweildauer		Signifikanztest
<b>„Besteht aus Ihrer Sicht Bedarf, das derzeitige gesetzliche Deutschfördermodell zu optimieren?“</b>					
Schulleitungen	Ja	98%	Ja	79.3%	$\chi^2(1) = 8.12, p < .01, \phi = .32$
	Nein	2%	Nein	20.7%	
Lehrpersonen DF	Ja	90.6%	Ja	85.2%	$\chi^2(1) = 1.09, p = .30, \phi = .08$
	Nein	9.4%	Nein	14.8%	

Um möglichst differenzierte Vorschläge zu erhalten, wurden sowohl Schulleitungen als auch Lehrpersonen DF gebeten anzugeben, auf welchen Wirkebenen (siehe Tabellen 17 und 18) sie Ansätze zur Verbesserung des Deutschfördermodells sehen.

**Tabelle 17 – Ansatzpunkte zur Verbesserung des Deutschfördermodells aus Sicht der Schulleitungen**

Schulleitungen	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
Steuerungsebene / Bund – Bildungsministerium	90.0%	95.7%	$\chi^2(1) = .67, p = .41, \phi = .10$
Steuerungsebene / Länder – Bildungsdirektionen	62.0%	78.3%	$\chi^2(1) = 1.88, p = .17, \phi = .16$
Steuerungsebene / Schulaufsicht, Schulqualitätsmanager*innen	24.0%	47.8%	$\chi^2(1) = 4.14, p = .04, \phi = .24$
Steuerungsebene / Schulerhalter	38.0%	34.8%	$\chi^2(1) = .07, p = .79, \phi = -.03$
Schulebene – Schulleitung	32.0%	26.1%	$\chi^2(1) = .26, p = .61, \phi = -.06$

Schulebene – Kooperation im Kollegium	22.0%	30.4%	$\chi^2(1) = .60, p = .44, \phi = .09$
Schulebene – andere Ansatzpunkte	8.0%	13.0%	$\chi^2(1) = .46, p = .50, \phi = .08$
Lehrpersonen – Ausbildung, Qualifikation	38.0%	34.8%	$\chi^2(1) = .07, p = .79, \phi = -.03$
Lehrpersonen – andere Ansatzpunkte	12.0%	4.3%	$\chi^2(1) = 1.06, p = .30, \phi = -.12$

Die Schulleitungen sahen die meisten Ansätze zur Verbesserung des Deutschfördermodells auf „Steuerungsebene / Bund - Bildungsministerium“ gefolgt von „Steuerungsebene / Länder – Bildungsdirektionen“. Ein deutlich geringerer Prozentsatz der Schulleitungen sah Ansatzmöglichkeiten zur Verbesserung des Deutschfördermodells auf den anderen Ebenen. Hinsichtlich „Steuerungsebene / Schulaufsicht, Schulqualitätsmanager\*innen“ gab es signifikante Unterschiede zwischen den Gruppen. Schulleitungen der Gruppe 1 sahen hier mehr Ansätze zur Verbesserung des Deutschfördermodells als Schulleitungen der Gruppe 2. Bzgl. aller anderen Ebenen gab es keine signifikanten Gruppenunterschiede. Relativ viele Schulleitungen sahen auch auf der Ebene „Lehrpersonen – Ausbildung, Qualifikation“ Ansatzmöglichkeiten zur Verbesserung des Deutschfördermodells. Auf Ebene „Lehrpersonen – andere Ansatzpunkte“ gibt es nach Angaben der Schulleitungen die wenigsten Ansatzmöglichkeiten zur Verbesserung des Deutschfördermodells.

**Tabelle 18 – Ansatzpunkte zur Verbesserung des Deutschfördermodells aus Sicht der Lehrpersonen DF**

Lehrpersonen DF	Gruppe 1 - kürzere Verweildauer	Gruppe 2- längere Verweildauer	Signifikanztest
Steuerungsebene / Bund – Bildungsministerium	79.2%	78.8%	$\chi^2(1) = .00, p = .96, \phi = -.00$
Steuerungsebene / Länder – Bildungsdirektionen	72.9%	55.8%	$\chi^2(1) = 4.48, p = .03, \phi = -.17$
Steuerungsebene / Schulaufsicht, Schulqualitätsmanager*innen	36.5%	21.2%	$\chi^2(1) = 3.69, p = .06, \phi = -.16$
Steuerungsebene / Schulerhalter	21.9%	13.5%	$\chi^2(1) = 1.56, p = .21, \phi = -.10$
Schulebene – Schulleitung	18.8%	17.3%	$\chi^2(1) = .05, p = .83, \phi = -.02$
Schulebene – Kooperation im Kollegium	24.0%	23.1%	$\chi^2(1) = .02, p = .90, \phi = -.01$
Schulebene – andere Ansatzpunkte	10.4%	11.5%	$\chi^2(1) = .04, p = .83, \phi = .02$
Lehrpersonen – Ausbildung, Qualifikation	35.4%	40.4%	$\chi^2(1) = .36, p = .55, \phi = .05$
Lehrpersonen – andere Ansatzpunkte	10.4%	15.4%	$\chi^2(1) = .78, p = .38, \phi = .07$

Der Prozentsatz der Lehrpersonen DF, die Ansatzmöglichkeiten zur Verbesserung des Deutschfördermodells angab, war bzgl. aller Wirkebenen tendenziell niedriger als bei den Schulleitungen. Analog wie bei den Schulleitungen sahen die meisten Lehrpersonen DF Ansatzmöglichkeiten zur Verbesserung des Deutschfördermodells auf „Steuerungsebene / Bund – Bildungsministerium“ gefolgt von „Steuerungsebene / Länder – Bildungsdirektionen“ mit signifikant höheren Werten in Gruppe 1. Hinsichtlich der anderen Ebenen waren, wie bei den Schulleitungen, die Prozentsätze der Lehrpersonen DF, die Fördermöglichkeiten sahen, deutlich niedriger. Die niedrigsten Prozentsätze gab es bei „Schulebene – andere Ansatzpunkte“. Wie auch die Schulleitungen sahen

relativ viele Lehrpersonen DF Ansatzmöglichkeiten zur Verbesserung des Deutschfördermodells auf der Ebene „Lehrpersonen – Ausbildung, Qualifikation“.

Im nächsten Schritt wurden Schulleitungen und Lehrpersonen DF gefragt, welche konkreten Ansätze zur Verbesserung des Deutschfördermodells sie auf den verschiedenen Wirkebenen vorschlagen. Die vielen unterschiedlichen Verbesserungsvorschläge wurden in den zwei Gruppen für beide Rollen (Schulleitungen, Lehrpersonen DF) analysiert. Da es in den genannten Ansätzen weder inhaltliche Unterschiede zwischen Gruppe 1 und Gruppe 2, noch zwischen Schulleitungen und Lehrpersonen DF gab, wurden die Ansätze zur Verbesserung des Deutschfördermodells für die verschiedenen Ebenen über Rollen und Gruppen hinweg zusammengefasst.

Aufgrund der geringen Unterschiede bei den Vorschlägen auf den vier Steuerungsebenen (Bund, Land, Schulaufsicht, Schulerhalterebene) werden diese hier gemeinsam berichtet (für beide Gruppen und beide Rollen: Schulleitungen, Lehrpersonen DF).

Auf den Steuerungsebenen Bund, Land Schulaufsicht und Schulerhalterebene wurden genannt:

- Kleinere Gruppen
- Mehr Autonomie und Flexibilität
- Mehr Ressourcen (Personal, Stundenkontingente, Räumlichkeiten, Hardwareausstattung, Internet, etc.)
- Integrativ statt segregiert
- Flexiblerer Aufstieg (Dauer des ao.-Status)
- Auf wissenschaftlichen Kriterien basierende Förderpolitik
- Vernetzung / Austausch zwischen Schulen (genannt bei Schulaufsichtsebene)

Auch auf der Schulebene (Schulleitung, Kooperation im Kollegium; andere Ansatzpunkte) gab es ebenfalls nur sehr geringe Unterschiede in den Vorschlägen der Schulleitungen und Lehrpersonen DF in beiden Gruppen. Daher werden auch hier die genannten Ansatzpunkte gemeinsam berichtet.

**Auf Schulebene** wurden folgende Ansätze angegeben (beide Gruppen, beide Rollen):

- Elternarbeit
- Mehr Mitsprache („als Schulleitung“)
- Achtsamer Einsatz kompetenter Kolleg\*innen (Personaleinsatz)
- Mehr Förderstunden
- Weniger Supplieren
- Vermehrter Austausch und stärkere Zusammenarbeit im Team

Eine hohe Übereinstimmung in den Vorschlägen, wo zur Verbesserung des Deutschfördermodells angesetzt werden sollte, gab es auch für die Lehrpersonenebene (Lehrpersonen – Ausbildung, Qualifikation; Lehrpersonen – andere Ansatzpunkte) über beide Gruppen und beide Rollen hinweg.

Auf Lehrpersonenebene wurden folgende Ansätze genannt:

- DaZ-Ausbildung attraktiver machen
- DaZ-Ausbildung für alle Lehrpersonen
- Mehr Praxis / praxisnähere Gestaltung der Ausbildung
- Mehr Aus- und Fortbildungen im Bereich DaZ sowie mehr DaZ-Ausbildungsplätze
- Nur Expert\*innen mit DaZ-Ausbildung einsetzen

Beiden Rollen (Schulleitungen, Lehrpersonen DF) wurden zusätzlich noch zwei offene Fragen gestellt. Die erste Frage lautete: „Gibt es an den Schulen ergänzende Maßnahmen in der Deutschförderung,

die den Umstieg in den Regelunterricht erleichtern?“ Diese Frage wurde von 36 Schulleitungen in Gruppe 1 und von 23 in Gruppe 2 beantwortet. Bei den Lehrpersonen DF waren es 23 (Gruppe 1) und 7 (Gruppe 2). In der Beteiligung gab es weder bei den Schulleitungen noch bei den Lehrpersonen DF signifikante Gruppenunterschiede.

Von Schulleitungen und Lehrpersonen DF wurden in beiden Gruppen „zusätzliche Förderstunden“ respektive „Förderunterricht“ sowie „Teamteaching“ respektive „Teamlehre und Kooperation“ häufig genannt. In beiden Lehrpersonengruppen DF wurde übereinstimmend „integrative Förderung“ am häufigsten angegeben. In Gruppe 1 gaben sowohl Schulleitungen als auch Lehrpersonen DF an, dass es „aufgrund von Ressourcenmangel keine ergänzenden Maßnahmen“ in der Deutschförderung gibt, die den Umstieg in den Regelunterricht erleichtern. Von den Schulleitungen wurden einige weitere Maßnahmen häufiger angegeben. In Gruppe 1: „Ganztagschule“ (am häufigsten) und „Buddy-Systeme“; in Gruppe 2: „Einbindung in den Regelunterricht“ und „Zusammenarbeit mit Schulexternen“. Eine Schulleitung nannte als Erleichterung des Umstiegs auch die Verwendung der Schulbücher für den fachlichen Unterricht.

Die zweite Frage lautete: „Gibt es an den Schulen (weitere) Beispiele guter Praxis, die sich in Bezug auf die Umsetzung der Deutschförderung bewährt haben?“ In Gruppe 1 beantworteten 26 Schulleitungen diese Frage, in Gruppe 2 waren es 16. Bei den Lehrpersonen DF waren es 22 respektive 13. Signifikante Gruppenunterschiede in der Beteiligung gab es keine.

Über beide Gruppen und Rollen hinweg wurden „integrative Förderung“ und „kleine Gruppen“ als Beispiele guter Praxis genannt, die sich in Bezug auf die Umsetzung der Deutschförderung bewährt haben (tendenziell etwas häufiger von den Schulleitungen als von den Lehrpersonen DF). „Individuelle Förderung“ wurde von den Lehrpersonen DF in beiden Gruppen und von den Schulleitungen in Gruppe 1 häufig angegeben. Am häufigsten wurde von den Schulleitungen in beiden Gruppen „mehr Lehrkräfte“ genannt. In Gruppe 1 beantworteten sowohl Schulleitungen als auch Lehrpersonen DF die Frage nach zusätzlichen Beispielen guter Praxis mit „Nein“ und ergänzten, dass die Ressourcen dafür nicht reichen. Zusätzlich noch häufig genannt wurde „Ganztagsbetreuung“ (Schulleitungen, Gruppe 1), „offene Sprachklassen“, „Wertschätzung für die Kinder“, „Online Zusatzunterricht (Schulleitungen, Gruppe 2), sowie Einbeziehung der Eltern (Lehrpersonen DF, Gruppe 1) und „Zusammenarbeit im Team“ (Lehrpersonen DF, Gruppe 2).

## **10. Ergänzende Aspekte aus Sicht der Befragten**

Abschließend wurden alle Rollen gefragt, welche offenen Punkte aus ihrer Sicht mit Blick auf die Implementierung der Deutschförderung relevant sind und im Fragebogen nicht angesprochen wurden.

Sowohl die Schulleitungen als auch die Lehrpersonen DF, die beide bereits um Vorschläge zur Verbesserung/ Weiterentwicklung der Deutschförderung gebeten worden waren, wiederholten zum großen Teil ihre Angaben.

Von den Schulleitungen nannten 23 aus Gruppe 1 und 14 aus Gruppe 2 Punkte, die aus ihrer Sicht bisher noch nicht angesprochen wurden, jedoch mit Blick auf die Implementierung der Deutschförderung relevant sind. Bei den Lehrpersonen DF waren es 45 (Gruppe 1), respektive 17 (Gruppe 2). Für beide Rollen gab es keine signifikanten Gruppenunterschiede in der Beteiligung. Die, trotz der vorigen Fragen, relativ hohe Beteiligung kann als Indikator für die Wichtigkeit des Themas für die Befragten angesehen werden.

Schulleitungen in Gruppe 1 nannten am häufigsten „mehr Flexibilität“, in Gruppe 2 war die häufigste Angabe „kleinere Gruppen“. Ebenfalls häufig wurde in beiden Gruppen „mehr Personal“, „mehr Ressourcen“ und „Wertschätzung aller vorhandenen Sprachen“ angegeben. Schulleitungen in Gruppe

1 nannten zusätzlich noch häufig „mehr Autonomie“ und „angstfreier Umgang mit Mehrsprachigkeit“. In Gruppe 2 wurde noch „nicht segregiert unterrichten“ und „MIKA-D erfüllt Zweck nicht“ häufig angegeben.

Auch Lehrpersonen DF nannten in beiden Gruppen „mehr Personal“ (häufigste Angabe in Gruppe 2) und „kleinere Gruppen“ (häufigste Angabe in Gruppe 1). Ebenfalls in beiden Gruppen wurde „weniger Testungen“ angegeben. Lehrpersonen in Gruppe 1 nannten noch häufig „mehr Stunden“, „integrative Förderung“, „mehr Ressourcen (Geld, Räume)“ und „weniger Administration“. Weitere häufige Angaben von Lehrpersonen DF in Gruppe 2 waren „weniger supplieren“, „flexiblere Aufstiegsmöglichkeiten“ und „mehr speziell ausgebildetes, erfahrenes Personal“.

Auch relativ viele der Lehrpersonen VS beantworteten die gestellte Frage und benannten die aus ihr Sicht offenen Punkte, die mit Blick auf die Implementierung der Deutschförderung relevant sind und im Fragebogen nicht angesprochen wurden. In Gruppe 1 beteiligten sich 88 Lehrpersonen DF; in Gruppe 2 waren es 41. Auch hier gab es keine signifikanten Gruppenunterschiede in der Beteiligung.

Die Antworten der Lehrpersonen VS waren inhaltlich breit gestreut. Sie wiesen jedoch eine relativ hohe Überlappung mit den Verbesserungsvorschlägen auf, die bereits von den Schulleitungen und den Lehrpersonen DF genannt worden waren. In beiden Gruppen häufig genannt wurden „integrative Förderung“ (häufigste Angabe in Gruppe 2), „mehr Personal“ (häufigste Angabe in Gruppe 1), „Sprachvorbilder“, „kleinere Gruppen“ und „mehr Deutschförderstunden“. Zusätzlich wurde in Gruppe 1 noch häufig „Chaos und Ungewissheit im Schulalltag“, „sozialer Ausschluss“, „individuelle Beurteilung/ mehr Autonomie“ und „verpasste Inhalte aus regulärem Unterricht“ angegeben. Häufige Nennungen von Lehrpersonen VS aus Gruppe 2 waren noch „zu viele Nichtmuttersprachliche in Klasse/ Schule“, „MIKA-D überarbeiten“ und „langfristig mehr Förderung“.

## **5. Zusammenfassung und Empfehlungen**

Das Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF) gab Ende 2020 eine wissenschaftliche Evaluationsstudie in Auftrag, die relevante Faktoren für die Implementierung des 2018/19 eingerichteten Deutschfördermodells analysieren sollte. Das Deutschfördermodell besteht aus Deutschförderklassen, Deutschförderkursen und integrativer Förderung, wobei die integrative Förderung nur bei entsprechend geringer Anzahl von ao.- Schüler\*innen am Standort (weniger als 8 Schüler\*innen) Anwendung findet. Ziel des Deutschfördermodells ist es, dass Schüler\*innen, die mangelhafte oder ungenügende Deutschkenntnisse aufweisen und sich daher im außerordentlichen Status (ao.-Status) befinden, möglichst rasch die Unterrichtssprache erlernen und damit in den ordentlichen Status (o.-Status) übergeführt werden können. Die Deutschkenntnisse werden mittels MIKA-D (Messinstrument zur Kompetenzanalyse – Deutsch) festgestellt. Nach vier Semestern wechseln die Schüler\*innen automatisch in den o.-Status, auch wenn sie laut MIKA-D ungenügende oder mangelhafte Deutschkenntnisse aufweisen. Eine weiterführende Förderung ist gesetzlich nicht vorgesehen.

Da das Ziel des Deutschfördermodells – möglichst viele Schüler\*innen vor Ablauf der vier Semester in den o.-Status überzuführen – nicht ausreichend erreicht wurde, sollte die Evaluierung diejenigen Implementierungsfaktoren identifizieren, die die Verweildauer im ao.-Status beeinflusst haben könnten. Aus den Ergebnissen der Evaluierung sollten, soweit möglich, Empfehlungen für eine Weiterentwicklung/ Optimierung des Deutschfördermodells abgeleitet werden.

Die Evaluierung wurde mittels eines Extremgruppendesigns durchgeführt, d.h. durch den Vergleich zweier Gruppen von Schulen, die sich hinsichtlich des Erfolgskriteriums (möglichst rasche Überführung der ao.-Schüler\*innen in den o.-Status) unterscheiden, jedoch hinsichtlich der Rahmenbedingungen

möglichst ähnlich sind. Damit sollte untersucht werden, ob und welche Implementierungsfaktoren zu einer erfolgreichen Deutschförderung beitragen. Die definierten Kriterien wurden von insgesamt 110 Schulen erfüllt (66 Schulen mit kürzerer Verweildauer im ao.-Status = Gruppe 1; 44 Schulen mit längerer Verweildauer im ao.-Status = Gruppe 2). Sie wurden zur Teilnahme an der Evaluierung eingeladen. Befragt wurden jeweils Schulleitungen, Lehrpersonen in der Deutschförderung (DF) und alle weiteren Lehrpersonen mit Lehramt für die Primarstufe/Volksschule (VS), damit drei „Rollen“. In Summe haben 693 Personen aus 93 Schulen an der Befragung teilgenommen. (Details dazu sind dem Abschnitt „Methodisches Vorgehen“ zu entnehmen.)

Insgesamt zeigen die Ergebnisse der Evaluation über beide Gruppen und die drei Rollen (Schulleitungen, Lehrpersonen DF, Lehrpersonen VS) hinweg eine sehr hohe Übereinstimmung. Damit konnte die Annahme, dass sich Schulen mit kürzerer Verweildauer und Schulen mit längerer Verweildauer im ao.-Status in relevanten Implementierungsfaktoren unterscheiden hinsichtlich der allermeisten Faktoren nicht bestätigt werden. Allfällige Gruppenunterschiede wiesen keine durchgängige Systematik auf. Dennoch liefern die Ergebnisse der Evaluation eine Reihe von Ansatzpunkten für die Weiterentwicklung/ Optimierung des Deutschfördermodells. Die zentralen Befunde, die die Basis für die Empfehlungen liefern, werden daher im Folgenden kurz zusammengefasst. (Im Abschnitt „Ergebnisse“ finden sich viele Detailangaben.)

Ausgehend von der Annahme, dass die Einstellung zum Deutschfördermodell einen Einfluss auf die Implementierung haben könnte, wurden dazu eine Reihe von Fragen gestellt. In beiden Gruppen wurden sowohl von den Schulleitungen als auch von den Lehrpersonen DF die Deutschförderklassen negativer bewertet als die Deutschförderkurse und diese wiederum negativer als die integrative Förderung; wobei die Zustimmung zum Deutschfördermodell insgesamt eher gering war (mit höheren Werten in Gruppe 2 der Lehrpersonen DF).

Alle drei Rollen in beiden Gruppen befürworteten auch tendenziell eher eine möglichst lange Verweildauer der Schüler\*innen im ao.-Status als eine kürzere (Lehrpersonen DF in Gruppe 2 mehr als in Gruppe 1). Das ist insofern überraschend, als die möglichst kurze Verweildauer im ao.-Status als Ziel des Deutschfördermodells festgelegt wurde. Die Begründungen für die möglichst lange Verweildauer stimmten über Gruppen und Rollen hinweg überein: „Kinder brauchen Zeit, um sich mit der neuen Sprache auseinanderzusetzen“; „zusätzliche Stunden zur Deutschförderung gibt es nur mit ao.-Status“; „Fördermaßnahmen können nicht nach zwei Jahren einfach beendet werden; Schulautonomie wäre hilfreich“. Sehr ähnliche Aussagen wurden auch bei weiteren offenen Fragen gemacht. (Detailangaben zu dieser und weiterer offener Fragen finden sich unter „Ergebnisse“).

Weiters wurde auch angenommen, dass die organisatorische Umsetzung der Deutschförderung Hinweise für die Unterschiede in der Verweildauer liefern könnte. Die Antworten der Schulleitungen stimmten jedoch in beiden Gruppen hoch überein. Das häufigste Kriterium für den Einsatz von Lehrpersonen in der Deutschförderung war „(Zusatz)-ausbildung“ (= DaZ-Ausbildung; Deutsch als Zweitsprache). Aufgrund von Personalknappheit sei jedoch häufig keine Auswahl möglich. Als Gründe für die Schwierigkeit der Einbettung der Deutschförderung in die Schulorganisation wurden in beiden Gruppen „Raummangel“ gefolgt von „Personalmangel“ angegeben. Strukturen des Austauschs zwischen Lehrpersonen in der Deutschförderung und dem restlichen Lehrkörper wurden auch durchgängig als etabliert angegeben (am häufigsten Teambesprechungen und Konferenzen). Damit liefern auch die Ergebnisse zur organisatorischen Umsetzung keine Erklärung für die Unterschiede zwischen kürzerer und längerer Verweildauer im ao.-Status.

Den Lehrpersonen DF wurden – da ihnen ein besonderer Stellenwert für die Umsetzung zukommt – eine Reihe tätigkeitsspezifischer Fragen gestellt; auch hier gab es hohe Übereinstimmung zwischen beiden Gruppen. Ein Großteil der Lehrpersonen DF hatte sich bereits zu Beginn der Tätigkeit der

Aufgabe gewachsen gefühlt und fühlte sich durch das Unterrichten in der Deutschförderung auch eher nicht überfordert. Auch ihr Selbstwirksamkeitserleben (hinsichtlich des Unterrichtens in der Deutschförderung) schätzten sie im Mittel eher hoch ein. Ein Großteil hatte sich auch freiwillig für den Einsatz in der Deutschförderung gemeldet. Hinsichtlich dieser Bedingungen dürften demnach gute Voraussetzungen für den Unterricht gegeben sein.

Interessante Unterschiede gab es zwischen den Lehrpersonen, die erst wenige Tage bzw. unmittelbar vor ihrem Einsatz erfahren hatten, dass sie in der Deutschförderung unterrichten werden, und den Lehrpersonen, die es bereits früher erfahren hatten. Die Gruppe, die es spät erfahren hatte, gab geringere Werte in der Selbstwirksamkeit an, geringere fachliche Zusammenarbeit mit den Lehrpersonen, die nicht in der Deutschförderung unterrichten, weniger gut gelingende Kooperation mit den Eltern der ao.-Schüler\*innen, höhere Arbeitsüberforderung sowie geringere Identifikation mit der Tätigkeit. Diese Ergebnisse deuten darauf hin, dass der Zeitpunkt, zu dem Lehrpersonen von ihrem Einsatz erfahren, eine relevante Rolle für die Tätigkeit spielt.

Dem MIKA-D, auf Basis dessen über „Erfolg“ versus „Nicht-Erfolg“ in der Deutschförderung entschieden wird, kommt ohne Zweifel eine zentrale Rolle zu. Fast alle Lehrpersonen kennen ihn (den Evaluatorinnen ist er nicht zugänglich). Der Aussage „Die MIKA-D Testergebnisse sind für die Arbeit der Lehrpersonen in der Deutschförderung sehr nützlich.“ stimmten im Mittel alle Lehrpersonen teilweise zu; auch der Frage, wie gut der MIKA-D zur Feststellung des Erfolgs in der Deutschförderung geeignet ist. Auf die Frage nach der Nutzung der Ergebnisse waren die Antworten heterogen. Einerseits werden die Ergebnisse des MIKA-D häufig für die Förderplanung und für Evaluation genutzt; andererseits wurde auch häufig „keine Nutzung, nur Aufbewahrung“ angegeben. Manche Schulleitungen nehmen ihn auch als unbrauchbar wahr. Mit Blick auf die Bedeutung der Ergebnisse des MIKA-D erscheinen diese Befunde problematisch.

In der Evaluierung wurde zusätzlich auch die Erreichung differenzierter sprachbezogener Ziele der Deutschförderung sowie flankierender Ziele erfragt (diese Ziele waren in der Vorstudie von den Expert\*innen identifiziert worden). Die Anteile der Schüler\*innen, die am Ende des ao.-Status die sprachbezogenen Ziele erreichen, wurden von den Lehrpersonen DF zwischen 79.22% (sich trauen Deutsch zu sprechen) und 45.00% (sprachbezogene Lernstrategien einsetzen) eingeschätzt. Das bedeutet im Umkehrschluss, dass gemäß der Lehrpersonen DF zwischen 20.78% und 55.00% der Schüler\*innen die sprachbezogenen Ziele am Ende des ao.-Status nicht erreichen. Bei den flankierenden Zielen liegen die Einschätzungen der Zielerreichung zwischen 74.70% (soziale Kompetenz erlangen) und 56.27% (Mehrsprachigkeitskompetenz erlangen). Damit erreichen im Umkehrschluss laut Lehrpersonen DF zwischen 25.30% und 43.73% der Schüler\*innen diese Ziele am Ende des ao.-Status nicht. Diese Einschätzungen erscheinen insbesondere mit Blick darauf, dass nach Erreichen des o.-Status (automatisch nach vier Semestern) keine weitere Förderung vorgesehen ist, als problematisch.

Abschließend wurden die Schulleitungen und die Lehrpersonen DF gebeten, Vorschläge zur Verbesserung/ Weiterentwicklung der Deutschförderung zu machen. Ein sehr hoher Prozentsatz insbesondere der Schulleitungen sah einen entsprechenden Bedarf. Die meisten Ansätze zur Verbesserung wurden mit Abstand auf der Steuerungsebene (Bund, Land, Schulerhalter) genannt (u.a. kleinere Gruppen, mehr Autonomie und Flexibilität, mehr Ressourcen, integrativ statt segregiert, flexiblerer Aufstieg), gefolgt von der Schulebene (u.a. Elternarbeit, mehr Mitsprache in der Rolle als Schulleitung, mehr Förderstunden). Auf Lehrpersonenebene wurden u.a. „DaZ-Ausbildung attraktiver machen“, „DaZ-Ausbildung für alle Lehrpersonen“ und „mehr Praxis“ angegeben. Als weitere Maßnahmen in den Schulen, die den Umstieg in den Regelunterricht erleichtern, wurden ähnliche Angaben wie zu früheren Fragen gemacht: u.a. „zusätzliche Förderstunden“, „integrative Förderung“

und „mehr Lehrkräfte“. Eine Schulleitung nannte als Erleichterung des Umstiegs auch die Verwendung der Schulbücher für den fachlichen Unterricht. Häufig wurde auch hier angegeben, dass aufgrund von mangelnden Ressourcen keine zusätzlichen Maßnahmen möglich sind.

Basierend auf den Aussagen der befragten Schulleitungen, Lehrpersonen DF und Lehrpersonen VS werden folgende Maßnahmen zur Weiterentwicklung/ Optimierung der Deutschförderung empfohlen (die Reihenfolge der Empfehlungen entspricht dem Verlauf der Deutschförderung):

- Weiterentwicklung des MIKA-D mit Blick auf das Kriterium, wonach die Schüler\*innen über ausreichende Deutschkenntnisse verfügen, um dem fachlichen Unterricht folgen zu können. Dabei sollten die sprachbezogenen Ziele der Deutschförderung (siehe oben) berücksichtigt werden.
- Ermöglichung von Flexibilität in der Deutschförderung: die Entscheidung über die Art der Förderung (Deutschförderklassen, Deutschförderkurse und integrative Förderung) sollte in der Autonomie der Schulen (Schulleitungen) liegen, da dadurch Rahmenbedingungen (Anzahl der Schüler\*innen, die eine Deutschförderung benötigen, Anzahl von Lehrpersonen, die eine entsprechende Ausbildung haben, Räume, sonstige Ressourcen, etc.) entsprechend berücksichtigt werden können. Insbesondere sollte eine stärkere Integration der ao.-Schüler\*innen unter Berücksichtigung der individuellen Situation der Schüler\*innen ermöglicht werden.
- Der Wechsel in den o.-Status sollte ebenfalls flexibilisiert werden. D.h. zumindest ergänzend zum (weiterentwickelten) MIKA-D sollte die Einschätzung der Lehrperson in der Deutschförderung berücksichtigt werden. In diese Einschätzung sollte auch das Erreichen der flankierenden Ziele (siehe oben) eingehen. Auch hier sollte der individuellen Situation der Schüler\*innen Rechnung getragen werden.
- Das derzeit vorgesehene automatische Ende der Deutschförderung (= Erreichung des o.-Status) nach vier Semestern ist mit Blick auf die Ziele der Deutschförderung nicht nachvollziehbar. In jedem Fall sollte eine weitere Förderung erfolgen, wenn der/die jeweilige Schüler\*in keine ausreichenden Deutschkenntnisse aufweist, um dem fachlichen Unterricht folgen zu können. Auch diese Entscheidung sollte in die Autonomie der Schulen fallen. Für die weitere Förderung sollten die notwendigen Ressourcen zur Verfügung gestellt werden (generell fehlt es häufig an Ressourcen wie Personal, Räumen, etc.).
- In der Deutschförderung sollten nur Personen eingesetzt werden, die die DaZ-Ausbildung (Deutsch als Zweitsprache) haben. Diese sollten möglichst früh über ihren Einsatz in der Deutschförderung verständigt werden. Ergänzend sollten alle Lehrpersonen (auch diejenigen, die in der Sekundarstufe unterrichten) für das Thema sensibilisiert werden. Diagnostische Kompetenzen von Lehrpersonen, z.B. um Einschätzungen zu treffen inwieweit schulische Leistungen auf Fach- oder Sprachkompetenz zurückzuführen sind, wären wichtige Voraussetzungen für die individuelle Förderung und Leistungsbeurteilung im fachlichen Unterricht. Entsprechende Kompetenzen könnten z.B. im Rahmen der Pädagog\*innenbildung vermittelt werden und in den entsprechenden Lehrveranstaltungen Eingang finden. Mit Blick auf die weltpolitische Situation wird es künftig vermutlich immer wichtiger, den Anforderungen im Kontext von Mehrsprachigkeit gerecht zu werden.
- Die Implementierung der Weiterentwicklung/ Optimierung des Deutschfördermodells sollte unter Einbindung der Stakeholder-Gruppen geplant und umgesetzt werden (es wird auch empfohlen, dabei die Kenntnisse aus der Implementationsforschung zu berücksichtigen). Die Implementierung sollte auch durch eine Evaluation begleitet werden, um Aussagen über deren Zielerreichung machen zu können. Zusätzlich wird empfohlen ein systematisches und

längsschnittliches Monitoring einzurichten und mit Registerdaten (z.B. über die weitere Bildungskarriere) zu verknüpfen.

Ergänzend sollten die vielen Angaben der Befragten auf die offenen Fragen (siehe oben) hinsichtlich ihrer Relevanz und Möglichkeit der Umsetzung geprüft werden. Ein Beispiel dafür wäre der Vorschlag einer Schulleitung, auch in Deutschförderklassen die Schulbücher für den fachlichen Unterricht zu verwenden, um den Umstieg in den o.-Status zu erleichtern.

## 6. Referenzen

- BMBWF. (2022). *Deutschförderklassen und Deutschförderkurse*.  
<https://www.bmbwf.gv.at/Themen/schule/schulpraxis/ba/sprabi/dfk.html>
- Fixsen, D. L., & Blase, K. (2009). *Implementation: The missing link between research and practice. National Implementation Research Network – Implementation Brief, 1, 1–2*.  
<https://files.eric.ed.gov/fulltext/ED507422.pdf>
- Schober, B., Schultes, M. T., Kollmayer, M., & Lüftenegger, M. (2019). Implementierung von Reformen im Bildungsbereich. In S. Breit, F. Eder, K. Krainer, C. Schreiner, A. Seel, & C. Spiel (Eds.), *Nationaler Bildungsbericht Österreich 2018 Band 2. Fokussierte Analysen und Zukunftsperspektiven für das Bildungswesen* (pp. 455–483). Leykam.  
[https://doi.org/https://doi.org/10.17888/nbb2018\\_2\\_11](https://doi.org/https://doi.org/10.17888/nbb2018_2_11)
- Spiel, C., Popper, V., & Holzer, J. (2021). *Bericht der Workshopergebnisse zu den Themen Zielexplication und Implementierung. Vorprojekt zur Evaluation der Deutschförderung*.

## 7. Weiterführende Literatur

- DeGEval - Deutsche Gesellschaft für Evaluation. (2021). *Standards für Evaluation*.  
[https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards\\_fuer\\_Evaluation.pdf](https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf)
- Ditton, H. (2009). Evaluation und Qualitätssicherung. In *Handbuch Bildungsforschung* (pp. 607–623). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-91831-0\\_32](https://doi.org/10.1007/978-3-531-91831-0_32)
- Fixsen, D. L., Schultes, M.-T., & Blase, K. A. (2016). Bildung-Psychology and implementation science. *European Journal of Developmental Psychology, 13*(6), 666–680.  
<https://doi.org/10.1080/17405629.2016.1204292>
- Hasselhorn, M., Köller, O., Maaz, K., & Zimmer, K. (2014). Implementation wirksamer Handlungskonzepte im Bildungsbereich als Forschungsaufgabe. *Psychologische Rundschau, 65*(3), 140–149. <https://doi.org/10.1026/0033-3042/a000216>
- Hagenauer, G., & Hascher, T. (2018). Bedingungsfaktoren und Funktionen von Emotionen von Lehrpersonen im Unterricht. *Unterrichtswissenschaft, 46*(2), 141–164.  
<https://doi.org/10.1007/s42010-017-0010-8>
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating Training Programs – The four Levels* (3rd ed.). Berrett-Koehler.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology, 105*(3), 805–820. <https://doi.org/10.1037/a0032583>