

Abschlussbericht zur wissenschaftlichen Begleitung der Fortbildung „Vielfalt fördern“

Projektlaufzeit 2012–2018

Daniela Rzejak & Prof. Dr. Frank Lipowsky

Kassel, November 2019

Inhaltsverzeichnis

Vorwort	1
1 Einleitung	8
2 Lehrerfortbildungen gestalten und evaluieren	10
2.1 Angebots-Nutzungs-Modell der Lehrerfortbildung	10
2.2 Ebenen des Fortbildungserfolges	11
3 Vielfalt fördern und leitende Fragestellungen von LIQUID	14
3.1 Ausgangskonzept der Fortbildung	14
3.1.1 Fortbildungsadressaten	15
3.1.2 Fortbildungsdauer	15
3.1.3 Fortbildungsinhalte	16
3.2 Ziele der Fortbildung	17
4 Das Projekt LIQUID: Die Evaluation von Vielfalt fördern	19
4.1 Methodischer Ansatz	19
4.1.1 Erhebungsplan.....	20
4.1.2 Instrumente	26
4.1.3 Auswertungsverfahren	29
4.2 Stichprobe.....	36
4.2.1 Kohorte 1	37
4.2.2 Kohorte 2	39
5 Zum Vergleich von Untersuchungsgruppe und Kontrollgruppe: Zusammenfassung zentraler Befunde	45
5.1 Kohorte 1: Voraussetzungen und Wahrnehmungen der teilnehmenden Lehrpersonen und ein Vergleich der Untersuchungsgruppe mit der Kontrollgruppe	45
5.2 Kohorte 2: Voraussetzungen und Wahrnehmungen der teilnehmenden Lehrpersonen und ein Vergleich der Untersuchungsgruppe mit der Kontrollgruppe	47
5.2.1 Voraussetzungen der Lehrpersonen	47
5.2.2 Wahrnehmung der Fortbildung Vielfalt fördern	48
5.2.3 Wirkungen der Fortbildung	49
6 Voraussetzungen der teilnehmenden Lehrpersonen	52
6.1 Voraussetzungsprofile der Fortbildungsteilnehmerinnen und -teilnehmer.....	52

6.2	Voraussetzungsprofile als erklärendes Merkmal	58
6.3	Teilnahme aufgrund des schulischen Mehrheitsentscheids	64
6.3.1	Fortbildungsklima zwischen den Lehrpersonen.....	66
6.3.2	Kooperative Erprobungen der Lehrpersonen zwischen den Sitzungen des zweiten und dritten Moduls	67
6.3.3	Zufriedenheit mit der Fortbildung	68
6.3.4	Selbstberichteter Lernertrag	69
7	Veränderung der Zusammenarbeit im Kollegium.....	71
7.1	Facetten der Lehrerverkooperation	71
7.1.1	Deskriptive Betrachtung der Veränderungswerte aller Schulen	73
7.1.2	Deskriptive Betrachtung deutlich hervortretender Schulen.....	74
7.2	Intensität und Struktur der kollegialen Zusammenarbeit	83
7.2.1	Veränderung der Kooperationsintensität	84
7.2.2	Einrichtung institutionalisierter Formen der Teamarbeit	92
8	Subjektive Bewertung und Verarbeitung durch die Lehrpersonen	98
8.1	Entwicklung der Motivation zur Diagnostik über das Modul 2	102
8.2	Entwicklung der Selbstwirksamkeitserwartung im Bereich lernbegleitende Diagnostik im Module 2	104
8.3	Entwicklung der Selbstwirksamkeitserwartung im Bereich Klassenführung im Rahmen der Module 3 und 4.....	106
8.4	Entwicklung der Selbstwirksamkeitserwartung im Bereich adaptive Unterrichtsgestaltung im Rahmen der Module 3 und 4	108
8.5	Transferabsichten am Ende der Fortbildung	110
9	Lerngelegenheiten der Fortbildung	114
9.1	Zeit für die Fortbildungsinhalte und Freiräume für Teilnehmerorientierung	114
9.2	Intensität der Bearbeitung von Fortbildungsthemen	115
9.2.1	Deskriptive Betrachtung der Thematisierungsintensität.....	115
9.2.2	Die Thematisierungsintensität als erklärendes Merkmal	118
10	Enthusiasmus des Moderatorenteams.....	121
10.1	Erklärung der Fortbildungsnutzung und -bewertung durch den Enthusiasmus des Moderatorenteams.....	123

10.1.1 Vorhersage der systematischen Verarbeitung von Fortbildungsinhalten während Modul 2	123
10.1.2 Vorhersage der kooperativen Erprobungen zwischen Fortbildungssitzungen von Modul 2	124
10.1.3 Vorhersage des selbstberichteten Lernertrags der Lehrpersonen am Ende von Modul 2	125
10.2 Beziehung zwischen der Fortbildungsmotivation der Lehrpersonen und dem Enthusiasmus des Moderatorenteams	127
10.2.1 Motivationsfacette: Externale Erwartungsanpassung	128
10.2.2 Motivationsfacette: Schulentwicklungsorientierung	129
10.2.3 Motivationsfacette: Soziale Interaktion	131
11 Einbezug von Erfahrungen der teilnehmenden Lehrpersonen	132
11.1 Motivationsfacette: Externale Erwartungsanpassung	133
11.2 Motivationsfacette: Schulentwicklungsorientierung	134
11.3 Motivationsfacette: Soziale Interaktion	135
12 Moderatorenschulung und Fortbildungsmaterial.....	138
13 Unterrichtliche Anregungen zur Reflexion über das Lernen.....	140
13.1 Fach Mathematik	141
13.2 Fach Deutsch.....	142
14 Befunde der Vertiefungsstudie	144
14.1 Motivation	145
14.1.1 Ausdruck extrinsischer Motivation	145
14.1.2 Ausdruck einer Schulentwicklungsorientierung als Motivationsfacette.....	147
14.1.3 Ausdruck von sozialer Interaktion als Motivationsfacette	148
14.1.4 Motivation im Verlauf des Fortbildungsprozesses	149
14.2 Relevante Themen	153
14.3 Lernförderliche Eigenschaften von Moderatorinnen und Moderatoren.....	159
15 Resümee und Empfehlungen	166
Literaturverzeichnis	177
Abbildungsverzeichnis	182
Tabellenverzeichnis	188

Vorwort

Erkenntnisse aus der Evaluation von Vielfalt fördern

Die Vielfalt in den Klassenzimmern ist in den letzten Jahren deutlich gestiegen. Das stellt Schulen und Lehrkräfte vor große Herausforderungen mit Blick auf das Ziel, Kindern und Jugendlichen in ihrer Unterschiedlichkeit gerecht zu werden und faire Lernchancen zu eröffnen. Am Anfang von *Vielfalt fördern* stand deshalb die Idee, Lehrerinnen und Lehrer bei der Weiterentwicklung ihres Unterrichts so zu unterstützen, dass es ihnen gelingt, ihre Schülerinnen und Schüler individuell zu fördern und zu fordern – und dies mit allen Lehrkräften umzusetzen, die eine Lerngruppe unterrichten. Deshalb haben sich das Schulministerium in Nordrhein-Westfalen und die Bertelsmann Stiftung 2011 zusammengetan, um gemeinsam eine Fortbildung zur individuellen Förderung zu entwickeln und damit einen Beitrag zu mehr Chancengerechtigkeit zu leisten. Als weiterer Partner ist 2015 die Qualitäts- und UnterstützungsAgentur - Landesinstitut für Schule (QUA-LiS NRW) in die Kooperation eingetreten. Nachdem die Fortbildung in zehn Regionen der fünf Regierungsbezirke pilotiert und sukzessive allen nordrhein-westfälischen Schulen der Primarstufe und der Sekundarstufe I zugänglich gemacht wurde, ist sie seit 2017 in die alleinige Verantwortung des Landes NRW übergegangen und wird nun von der QUA-LiS gesteuert.

Die Fortbildung mit dem Langtitel *Kinder und Jugendliche in ihrer Vielfalt fördern – Unterricht fokussiert auf individuelle Förderung weiterentwickeln* ist unter wissenschaftlicher Leitung von Prof. Dr. Christian Fischer von der Westfälischen Wilhelms-Universität Münster und Prof. in Dr. Bettina Amrhein von der Universität Bielefeld entstanden. Die vier Module – Teamentwicklung, Diagnostik und Didaktik 1 & 2 – wurden von einer Entwicklungsgruppe aus erfahrenen Fortbildnerinnen und Fortbildnern und innovativen Lehrkräften erarbeitet. Wissenschaftliche Unterstützung erhielten sie durch Inputs zu spezifischen Themen. Im Schuljahr 2012/13 haben Schulen in den ersten Pilotregionen mit der Fortbildung begonnen. Heute, im Schuljahr 2019/20, wird *Vielfalt fördern* in weiterentwickelter Form im siebten Durchgang angeboten.

Für die vorliegende Evaluation von *Vielfalt fördern* konnten wir Prof. Dr. Frank Lipowsky und Daniela Rzejak von der Universität Kassel gewinnen. Prof. Dr. Lipowsky, der für seine Forschungen zur Wirksamkeit von Lehrerfortbildungen, zu lernwirksamem Unterricht und zur Professionalisierung von Lehrpersonen bekannt ist, und Frau Rzejak, die u. a. zu der Frage promoviert, welchen Einfluss motivationale Voraussetzungen von Lehrkräften auf den Bildungsprozess haben, sind in dieser Evaluation der Frage nachgegangen, wie die Fortbildung in unterschiedlichen Bereichen wirkt. Insbesondere haben sie untersucht, welche Rolle die Motivation der an der Fortbildung teilnehmenden Lehrkräfte spielt – sowohl auf Ebene der

einzelnen Lehrkraft als auch auf Ebene des Kollegiums –, welche Auswirkungen die Innovationsbereitschaft eines Kollegiums hat und inwiefern Moderatorinnen und Moderatoren den Fortbildungserfolg mit beeinflussen können.

Die Erkenntnisse sind für *Vielfalt fördern* in mehrfacher Hinsicht hilfreich und wichtig. Erfreulich ist, dass sie viele Weiterentwicklungen bestätigen, die wir im Laufe der Zeit auf der Basis von Rückmeldungen aus den am Projekt beteiligten Schulen und von Lehrkräften vorgenommen haben. So lassen die Befunde der Evaluation vermuten, dass es für den Prozess der Schulentwicklung förderlicher sein kann, die Fortbildung mit einem Teilkollegium anstatt mit dem Gesamtkollegium zu beginnen. Ebenso stützen die Befunde von Rzejak und Lipowsky, dass die Entscheidung richtig war, die Moderatorinnen und Moderatoren in einem längeren Qualifizierungsprozess auf ihren Einsatz vorzubereiten. In vielen Bereichen haben die Evaluatoren aus wissenschaftlicher Sicht Neuland betreten. Dies gilt unter anderem für die Bedeutung der Motivation der Teilnehmenden, aber auch des Moderatorenteams, für den Fortbildungserfolg. Die Erkenntnisse aus der vorliegenden Evaluation sind damit nicht nur für *Vielfalt fördern* selbst relevant, sondern für alle Fortbildungen, die auf unterrichtliche Veränderungen durch ein systematisches Fortbildungskonzept abzielen.

Die Vergleichsstudie zeigt für die Pilotphase Entwicklungsvorteile bei der Zusammenarbeit der fortgebildeten Lehrkräfte, aber keine Vorteile auf Schülerschulebene

Im statistischen Globalvergleich zwischen einer Stichprobe von Schulen aus der Pilotphase des Projekts und einer Kontrollgruppe von Schulen, die nicht an der Fortbildung teilnahmen, zeigt sich, dass *Vielfalt fördern* in der Pilotphase die Zusammenarbeit unter den Lehrkräften verbessert hat. Einerseits kam es auf der Ebene der Lehrkräfte dazu, dass diese ihre individuelle Unterrichtspraxis stärker mit anderen Kolleginnen und Kollegen reflektierten, andererseits wurden auf Schulebene neue Strukturen für Kooperationen etabliert oder bestehende gestärkt. Globalstatistisch zeigt der Vergleich zur Kontrollgruppe für die Pilotphase keine Entwicklungsvorteile der unterrichteten Schülerinnen und Schüler – weder bei der wahrgenommenen Unterrichtsqualität noch bei den Leistungen in Deutsch und Mathematik. Auffällig sind zudem die großen Unterschiede zwischen den fortgebildeten Schulen und innerhalb dieser zwischen den Lehrkräften. Die Fortbildung entfaltet also ihre Wirkung unterschiedlich – eine erwartbare Erkenntnis, denn neben der Heterogenität in der Schülerschaft gibt es diese auch unter den Lehrkräften.

Die motivationalen Voraussetzungen der Lehrkräfte sowie die Motivation im Kollegium spielen eine wesentliche Rolle für den Fortbildungserfolg

Je stärker die teilnehmenden Lehrkräfte intrinsisch motiviert waren, desto positiver wurden einzelne Aspekte der Fortbildung beurteilt. Dies bezog sich auf

- das individuelle Nutzungsverhalten,
- die kooperative Umsetzung und Erprobung,
- die Umsetzung der Fortbildungsinhalte des Diagnostikmoduls,
- die empfundene zeitliche Belastung (je positiver die Voraussetzungen, desto geringer das Belastungsempfinden),
- die Zufriedenheit mit der Fortbildung am Ende des zweiten Moduls und
- den selbstberichteten Lernertrag.

Ob die Fortbildung im Laufe der Zeit als eher positiv wahrgenommen oder zunehmend kritisch betrachtet wurde, hing unter anderem davon ab, ob ein besonders hoher Anteil der Kolleginnen und Kollegen angab, auch aus anderen Gründen als intrinsisch motiviert teilzunehmen – man spricht dann von „extrinsischer Motivation“. An diesen Schulen erlebten Kolleginnen und Kollegen das Fortbildungsklima und die kooperative Erprobung weniger positiv, waren weniger zufrieden mit der Fortbildung und bewerteten ihren Lernertrag als geringer.

Annehmen lässt sich, dass insbesondere bei einer länger dauernden Fortbildung wie *Vielfalt fördern* die motivationalen Voraussetzungen eine größere Rolle spielen, als dies bei kurzfristigen Maßnahmen der Fall ist. „Je länger eine Fortbildung dauert und je anstrengender und aufwendiger die Fortbildung für die teilnehmenden Lehrpersonen wird, desto größer dürfte die Wahrscheinlichkeit werden, dass Lehrpersonen mit ungünstigen motivationalen Voraussetzungen gedanklich aus der Fortbildung ‚aussteigen‘“ (Rzejak & Lipowsky 2019, S. 163).

Die Ergebnisse der vorliegenden Evaluation zeigen, wie wirkmächtig die motivationalen Voraussetzungen der teilnehmenden Lehrkräfte sind. Für Moderatorinnen und Moderatoren lässt sich daraus die Aufgabe ableiten, die Motivation der teilnehmenden Lehrkräfte in den Blick zu nehmen und zu fördern. Dies kann u. a. dadurch geschehen, dass Teilnehmende verstärkt die Möglichkeit erhalten, Schwerpunkte zu wählen und eigene Erfahrungen einzubringen. Als lernendes Projekt hat *Vielfalt fördern* sich schon früh in diese Richtung weiterentwickelt: Die Moderatorenteams wurden verstärkt darauf vorbereitet, sich mit den schulischen Steuergruppen abzustimmen, welche Inhalte in welcher Intensität behandelt werden sollen und welche Vorerfahrungen einzubeziehen sind. Inzwischen besteht für Schulen generell die Möglichkeit, nur einzelne Module und Bausteine der Fortbildung zu behandeln.

Die Evaluationsbefunde werfen die bereits oben kurz angesprochene Frage auf, ob es sinnvoll ist, *Vielfalt fördern* immer mit dem Gesamtkollegium zu beginnen. Gerade weil sich ein besonders hoher Anteil extrinsisch motivierter Kolleginnen und Kollegen im Fortbildungsprozess negativ auswirkt, kann es sinnvoll sein, die Fortbildung nur mit einem Teilkollegium zu beginnen. Für *Vielfalt fördern* haben wir diese Möglichkeit ab dem vierten Durchgang geschaffen – allerdings wird sie nur selten von Schulen genutzt. In der Praxis scheint über die verschiedenen Ebenen des Schulsystems hinweg die Überzeugung zu bestehen, dass Prozesse der Schulentwicklung nur dann gelingen können, wenn eine Schulgemeinschaft sie gemeinsam beginnt. Dass es sich auf einen Schulentwicklungsprozess negativ auswirken kann, wenn sich sofort alle in ähnlicher Weise beteiligen sollen, dies aber aus unterschiedlicher Motivation heraus geschieht, ist wenig bekannt. Die Evaluationsergebnisse zeigen, dass ein Start mit einem Teilkollegium unter bestimmten Voraussetzungen sinnvoller zu sein scheint. Hierbei sollte die Schulleitung hinter dem Vorhaben stehen und das Teilkollegium muss sinnvoll zusammengestellt sein (beispielsweise als Team einer fünften Jahrgangsstufe). Den Prozess gut in die Schulgemeinschaft zu kommunizieren und weitere Interessenten für die Fortbildung zu gewinnen, ist Aufgabe von Schulleitung und der beauftragten Steuergruppe. Welcher Weg für welche Schule der erfolgversprechendere ist, hängt von der einzelnen Schule ab.

Wer das Fortbildungsangebot positiv wahrnimmt und es intensiv nutzt, profitiert stärker davon

Die Analysen zeigen, dass sich eine intensive Wahrnehmung und Nutzung des Angebots lohnt. Diejenigen, die angeben, das Angebot intensiver zu nutzen, hatten

- am Ende von Modul 2 (Diagnostik) eine höhere Motivation zur Diagnostik,
- am Ende von Modul 2 (Diagnostik) bzw. Modul 4 (Didaktik II) eine höhere Selbstwirksamkeitserwartung im Bereich der lernbegleitenden Diagnostik sowie im Bereich der adaptiven Unterrichtsgestaltung und
- eher die Absicht, die gelernten Inhalte am Ende der Fortbildung in den schulischen und unterrichtlichen Alltag zu transferieren.

Das Evaluationsergebnis hilft, *Vielfalt fördern* im Rahmen eines Angebots-Nutzungs-Modells zu sehen: Wie ein Angebot wirkt, wird durch viele Faktoren bestimmt. Die Analysen verdeutlichen dabei die Bedeutung, die der Nutzung zukommt. Inhalte werden beispielsweise vor allem dann transferiert, wenn die Teilnehmerinnen und Teilnehmer sie systematisch nutzen und die Fortbildung als positive Herausforderung wahrnehmen. Daraus folgt die Frage, wie sich Wahrnehmung und Nutzung einer Fortbildung beeinflussen lassen. Unter anderem gilt auch hier, die Motivation der teilnehmenden Lehrkräfte in den Blick zu nehmen und zu fördern.

Vielfalt fördern ist im Laufe der Zeit mit Blick auf die Bedarfe einer Schule anpassungsfähiger geworden. Dadurch sollte nicht nur die Motivation der Beteiligten befördert werden, sondern auch die intensive Nutzung des Fortbildungsangebots. Jenseits dessen wurden auch einzelne Bausteine in Richtung Teilnehmeraktivierung weiterentwickelt. Kollegiale Unterrichtshospitationen sind auf der Grundlage von Erkenntnissen aus der Unterrichtsforschung so weiterentwickelt worden, dass sie inzwischen als niedrigschwelliger und effektiver gelten können. Bei anderen Bausteinen wurden die fachdidaktischen Inhalte weiter ausgearbeitet: Auch dadurch dürfte eine direkte Umsetzung der Fortbildungsinhalte im Unterricht möglich sein. Diese fachbezogenen Weiterentwicklungen stehen im Einklang mit allgemeinen Erkenntnissen der Wirkungsforschung, wie im vorliegenden Bericht von Rzejak und Lipowsky erläutert wird. Denn wirksame Fortbildungen für Lehrkräfte haben oft einen ausgeprägten Fachbezug.

Faktoren für häufigere und intensivere Kooperation: Eine hohe kollektive Innovationsbereitschaft und ein gutes Fortbildungsklima wirken sich positiv aus

Die Zusammenarbeit der Lehrkräfte an den *Vielfalt-fördern*-Schulen wurde von der kollektiven Innovationsbereitschaft eines Kollegiums und dem Fortbildungsklima wesentlich beeinflusst. Letzteres kann durch die Moderatorinnen und Moderatoren positiv mitgestaltet werden. Dazu gehört ein offener Umgang mit anderen Ansichten und Herangehensweisen in einem Kollegium: Fühlen Lehrkräfte sich angesprochen und ernst genommen, sind sie eher bereit, sich mit Inhalten so auseinanderzusetzen, dass Irritationen ausgelöst und eigene Ansichten überprüft, erweitert und sogar verändert werden können. In diese Richtung wurde die Qualifizierung der Moderatorinnen und Moderatoren im Laufe der Zeit weiterentwickelt.

Moderatorenteams sind besonders wirksam, wenn sie von den Lehrkräften als enthusiastisch wahrgenommen werden

Rzejak und Lipowsky haben auch untersucht, wie sich der Enthusiasmus der Moderatorinnen und Moderatoren auf den Fortbildungsprozess und -erfolg auswirkt. Dazu haben sie sowohl die Moderatorinnen und Moderatoren selbst als auch die teilnehmenden Lehrkräfte befragt. Zusammenhänge mit Merkmalen des Fortbildungsprozesses und -erfolgs zeigen sich nur bei der Einschätzung der Teilnehmerinnen und Teilnehmer.

Erleben die Lehrkräfte ein Moderatorenteam als enthusiastisch, spielt dies eine bedeutsame Rolle

- für die systematische Verarbeitung von Fortbildungsinhalten im Modul 2,
- für die Intensität der Erprobungsphasen zwischen zwei Modulbausteinen und
- für den selbstberichteten Lernertrag am Ende des zweiten Moduls.

Je enthusiastischer ein Moderatorenteam erlebt wurde, umso positiver wurden die drei Bereiche von den Teilnehmenden bewertet.

Moderatorinnen und Moderatoren spielen innerhalb des Angebots-Nutzungs-Modells eine wichtige Rolle. Bislang ist jedoch kaum erforscht, wie wichtig sie für den Fortbildungserfolg sind, was u. a. daran liegt, dass es nur wenige Fortbildungen gibt, die durch viele Moderatorinnen und Moderatoren durchgeführt und die auch evaluiert wurden. Mit den Ergebnissen der vorliegenden Evaluation wird deshalb Neuland betreten.

In *Vielfalt fördern* ist die Qualifizierung der Moderatorinnen und Moderatoren im Laufe der Zeit verändert worden. In den ersten Durchgängen wurden die Moderatorinnen und Moderatoren ein halbes Jahr lang ausgebildet, bevor sie selbst die Fortbildung an Schulen durchführten. Die weitere Ausbildung verlief parallel zur Umsetzung der Fortbildung an Schulen. Der „Vorsprung“ der in der Evaluation untersuchten Moderatorenteam betrug gegenüber den Kollegien dementsprechend nur ein halbes Jahr. Das wurde in der Evaluation sowohl von Moderatorinnen und Moderatoren als auch von den Lehrkräften als zu gering betrachtet. Seit dem vierten Durchgang werden die Moderatorenteam ein ganzes Jahr lang ausgebildet, bevor sie an die Schulen gehen: Sie haben dadurch länger die Möglichkeit, Inhalte zu vertiefen und sie in ihrem eigenen Unterricht umzusetzen und zu erproben.

Die Fortbildung *Vielfalt fördern* wurde im Lichte der Erfahrungen der Pilotphase kontinuierlich weiterentwickelt und in vielen Bereichen so gestaltet, dass verstärkt an konkreten Fallbeispielen aus den Schulen und mit fachdidaktischem Material gearbeitet wird. Dies dürfte auch den Moderatorenteam helfen, an spezifischen unterrichtlichen Verhaltensweisen der fortzubildenden Lehrkräfte anzusetzen. Darüber hinaus dient ein Teil der Qualifizierung der Moderatorinnen und Moderatoren der Reflexion: Sind sie zur Fortbildung der Kollegien an den Schulen, dann wird in der Gruppe reflektiert, was an den einzelnen Schulen gut funktioniert und wo es gehakt hat. Durch den kollegialen Austausch können die Moderatorinnen und Moderatoren ihr Handlungsrepertoire kontinuierlich erweitern.

In der Evaluation von *Vielfalt fördern* wird zum einen hervorgehoben, dass der Erfolg einer Maßnahme nicht nur davon abhängt, wie durchdacht ein Angebot ist. Es braucht auch Moderatorinnen und Moderatoren, die authentisch und kompetent, diagnostisch versiert, empathisch und in der Lage sind, auf heterogene Kollegien und ihre Bedarfe flexibel eingehen zu können.

Die Erkenntnisse aus der Evaluation haben uns als Projektpartner in der Einschätzung bestärkt, dass die Weiterentwicklungen der letzten Jahre in die richtige Richtung gegangen sind. Mit der Veröffentlichung des Evaluationsberichts möchten wir aber auch einen Beitrag zur Beantwortung der Frage leisten, welche Erkenntnisse sich für die Lehrerfortbildung allgemein ableiten lassen. Rzejak und Lipowsky haben weitgehend Neuland betreten, indem sie auf die

Bedeutung der Teilnehmermotivation eingegangen sind und untersucht haben, wie der Enthusiasmus von Moderatorinnen und Moderatoren sich auf den Fortbildungserfolg auswirkt.

Wir bedanken uns herzlich bei Prof. Dr. Lipowsky und Frau Rzejak für ihre sorgfältige Arbeit und die differenzierten Erkenntnisse, die sie den Projektleiterinnen schon während der langjährigen Forschung immer wieder zeitnah zur Verfügung gestellt haben. Mit den Ergebnissen haben Tamara Sturm-Schubert (QUA-LiS NRW) und Angela Müncher (Bertelsmann Stiftung) gemeinsam mit den für die Konzeption der Fortbildung verantwortlichen Wissenschaftlerinnen und Wissenschaftlern sowie Praktikerinnen und Praktikern immer wieder eingeschlagene Pfade überprüft und neue Ideen aufgenommen: Ihnen sei ebenfalls ein herzlicher Dank ausgesprochen, denn ohne diese Bereitschaft zum Lernen wäre *Vielfalt fördern* als lernende Fortbildung nicht zu dem Erfolg geworden, der aus Sicht der beteiligten Schulen und Lehrkräfte in vielen Fällen erzielt wurde. Wir hoffen, mit der Veröffentlichung des Abschlussberichts zur wissenschaftlichen Begleitung der Fortbildung *Vielfalt fördern* den Fachdiskurs zu bereichern. Denn wir sind überzeugt davon, dass wirksame Lehrerfortbildung ein zentraler Schlüssel ist, um Unterricht in heterogenen Klassen und mit Blick auf das Lernen der Schülerinnen und Schüler weiter zu entwickeln.

Ulrich Janzen

Arbeitsbereichsleiter

Professionalisierung I - Zentrale Entwicklungsarbeiten und Personalentwicklung

in der Lehrerfortbildung,

Qualitätsrahmen Fortbildung,

Fortbildungsberichterstattung

Qualitäts- und Unterstützungs-Agentur - Landesinstitut für Schule Nordrhein-Westfalen

Ulrich Kober

Programmdirektor

Integration und Bildung

Bertelsmann Stiftung

Bernd Müller

Referatsleiter

Lehrerfortbildung und Qualifizierung im Schulwesen

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen

1 Einleitung

„Jeder junge Mensch hat ohne Rücksicht auf seine wirtschaftliche Lage und Herkunft und sein Geschlecht ein Recht auf schulische Bildung, Erziehung und individuelle Förderung“ lautet es unmittelbar in §1 Absatz 1 des nordrhein-westfälischen Schulgesetzes (Schulministerium Nordrhein-Westfalen, 2016). Blickt man weiter in den ersten Abschnitt des Gesetzestextes, in dem der Auftrag von Schule formuliert ist, heißt es, dass die individuellen Voraussetzungen der Schülerinnen und Schüler von der Schule zu berücksichtigen sind (§2 Absatz 4) und Schülerinnen und Schüler insbesondere lernen sollen, „selbstständig und eigenverantwortlich zu handeln“ und „für sich und gemeinsam mit anderen zu lernen und Leistungen zu erbringen“ (§2 Absatz 6). Ferner heißt es in §2 Absatz 9 auch, dass Unterricht die Lernfreude erhalten und weiter fördern soll.

Individuelle Voraussetzungen von Lernenden im Unterricht berücksichtigen zu können, setzt voraus, dass Lehrpersonen in der Lage sind, diese wahrzunehmen, kategorial einzuordnen und entsprechend der Diagnose den Unterricht zu gestalten. Dabei wird (gesetzlich) ein Unterricht erwartet, der motivierend gestaltet ist, kooperative Lernformen umfasst und das selbstständige, eigenverantwortliche Lernen von Schülerinnen und Schülern fördert. Unter anderem, um Schulen und Lehrpersonen bei der Erfüllung dieser Erwartungen zu unterstützen und sie dafür zu qualifizieren, initiierten das Ministerium für Schule und Weiterbildung (jetzt: Ministerium für Schule und Bildung) von Nordrhein-Westfalen und die Bertelsmann Stiftung das Fortbildungsprojekt „Kinder und Jugendliche in ihrer Vielfalt fördern – Unterricht fokussiert auf individuelle Förderung weiterentwickeln“ mit dem Kurztitel *Vielfalt fördern*. An der Entwicklung des Fortbildungskonzeptes arbeitete eine Gruppe von Personen aus der Fortbildungs- und Unterrichtspraxis unter wissenschaftlicher Leitung von Prof. Dr. Christian Fischer von der Universität Münster. Die Universität Kassel, unter Leitung von Prof. Dr. Frank Lipowsky, erhielt den Auftrag, eine wissenschaftliche Evaluation von *Vielfalt fördern* durchzuführen.

Wesentliche Ergebnisse aus dem Projekt *LIQUID*, der *Evaluation* von *Vielfalt fördern* werden in dem vorliegenden Abschlussbericht dargestellt. Zunächst werden dafür die Fortbildung *Vielfalt fördern* (Abschnitt 3) und das Evaluationsprojekt *LIQUID* (Abschnitt 4) beschrieben. In Abschnitt 4.1 wird die methodische Vorgehensweise – Design, Instrumente, Auswertungsverfahren – erläutert. Eine Darstellung der Stichproben, d. h. eine Beschreibung der in die Evaluation einbezogenen Personengruppen, erfolgt in Abschnitt 4.2. Hieran anschließend werden in Abschnitt 5 zentrale Befunde aus der Evaluation von Kohorte 1 und Kohorte 2 zusammengefasst, die aus dem Vergleich der Untersuchungsgruppe (Fortbildungsgruppe) mit der Kontrollgruppe resultieren. Da sich hier keine Hinweise auf unterschiedliche Entwicklungen einer der beiden Gruppen ergaben, wird in den nachfolgenden

Abschnitten 6–13 die Gruppe der fortgebildeten Lehrpersonen aus Kohorte 2 und deren Schülerinnen und Schüler detaillierter in den Blick genommen. Die entsprechenden Analysen werden geleitet von der Frage, ob Hinweise darauf bestehen, dass es innerhalb der Gesamtgruppe der fortgebildeten Lehrpersonen und ihrer Klassen Subgruppen gibt, die von der Fortbildung mehr profitiert haben als andere. Hierzu werden zunächst die Voraussetzungen der teilnehmenden Lehrpersonen und Kollegien zu Beginn des Fortbildungsprozesses betrachtet (Abschnitt 6). In Abschnitt 7 werden Veränderungen in der Zusammenarbeit der Kollegien analysiert, während der Fokus in Abschnitt 8 stärker auf den subjektiven Wahrnehmungen und auf der Nutzung der Lerngelegenheiten von *Vielfalt fördern* liegt. Im Mittelpunkt steht dabei die Frage, ob unterschiedliche Entwicklungen der teilnehmenden Lehrpersonen während des Verlaufs von *Vielfalt fördern* durch die individuellen Wahrnehmungen, Bewertungen und durch Unterschiede in der Nutzung des Fortbildungsangebots erklärt werden können. Die Lerngelegenheiten von *Vielfalt fördern* sind auch Gegenstand des Abschnitts 9, bevor anschließend im 10. Abschnitt des Berichts die Bedeutung des Enthusiasmus der Moderatorinnen und Moderatoren untersucht wird. Wie sich der Einbezug von Erfahrungen der Teilnehmerinnen und Teilnehmer als ein Merkmal der Fortbildungsgestaltung auf die Motivation der teilnehmenden Lehrpersonen auswirkt, aktiv in der Fortbildung mitzuarbeiten, wird in Abschnitt 11 untersucht. In Abschnitt 12 wird dargestellt, wie die Moderatorinnen und Moderatoren von Kohorte 2 ihre eigene Qualifizierung und das Fortbildungsangebot bewerten. Eine neue Perspektive wird in Abschnitt 13 eingenommen: Hier wird betrachtet, ob sich unterschiedliche Entwicklungen im Strategiewissen der Schülerinnen und Schüler in den Fächern Mathematik und Deutsch durch die unterrichtliche Anregung zur Reflexion über das Lernen erklären lassen. Befunde aus der vertiefenden Interviewstudie, die bei Kohorte 2 durchgeführt wurde, stellt Abschnitt 14 dar, bevor in Abschnitt 15 ein zusammenfassendes Resümee formuliert wird und Empfehlungen für zukünftige Fortbildungen gegeben werden.

2 Lehrerfortbildungen gestalten und evaluieren

Bevor das konkrete Lehrerfortbildungsprojekt *Vielfalt fördern* und das daran angeschlossene Evaluationsprojekt *LIQUID* erläutert werden, werden allgemeine Hinweise zur Fortbildung von Lehrpersonen und zur Evaluation dieser Angebote gegeben.

2.1 Angebots-Nutzungs-Modell der Lehrerfortbildung

Lipowsky (2014) hat ein Rahmenmodell skizziert (s. Abbildung 1), das Faktoren, die sich auf das Lernen und den Fortbildungserfolg von Lehrpersonen auswirken dürften, sowie deren Zusammenhänge systematisch und vereinfacht abbildet.

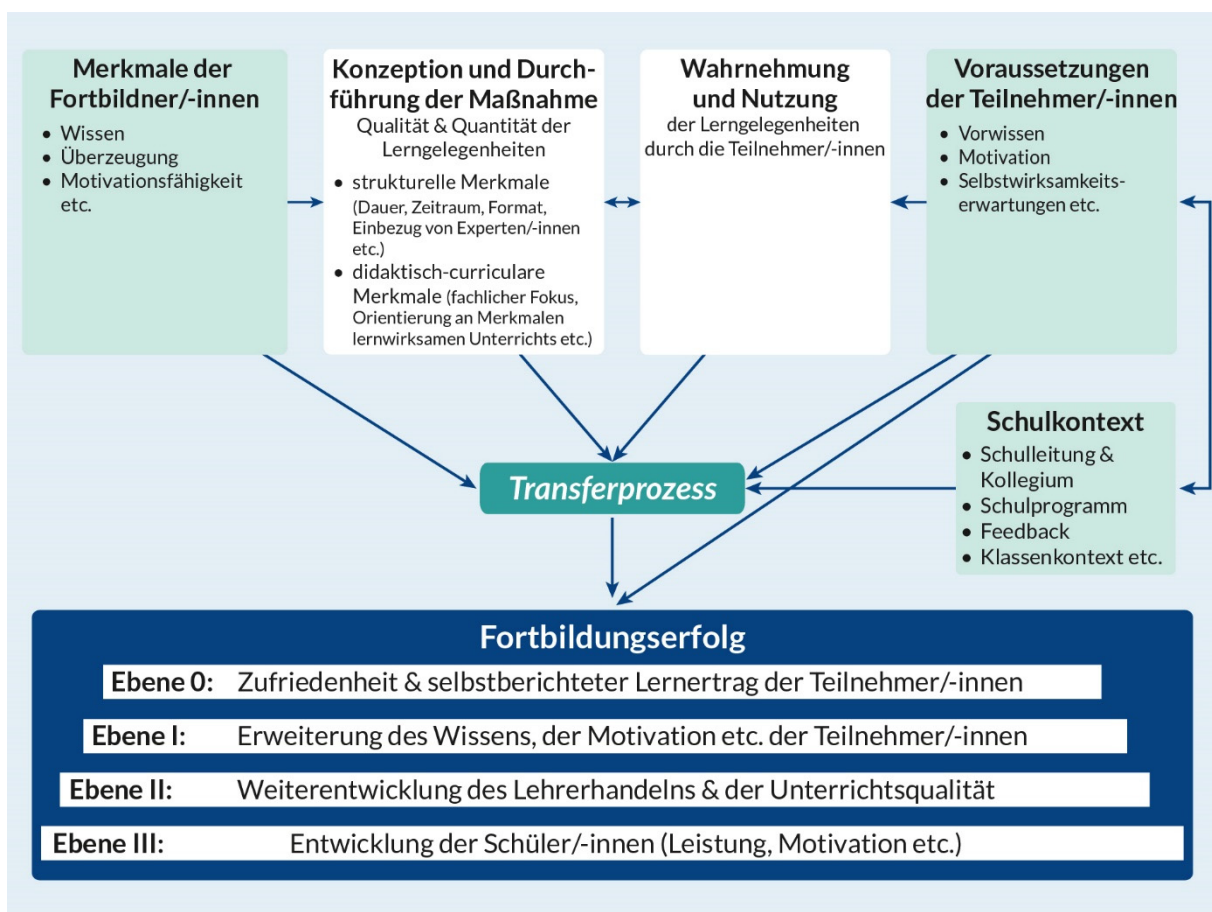


Abbildung 1. Vereinfachtes Angebots-Nutzungs-Modell, angelehnt an Lipowsky (2014)

Es handelt sich dabei um ein Angebots-Nutzungs-Modell, d. h., es wird zwischen dem Fortbildungsangebot auf der einen Seite und der Wahrnehmung und Nutzung dieses Angebots auf der anderen Seite unterschieden. Im Mittelpunkt steht die Annahme, dass die Wahrnehmung und Nutzung der angebotenen Lerngelegenheiten im Rahmen einer Fortbildung den Erfolg und Transfer wesentlich bestimmen. Auf die Wahrnehmung und Nutzung der Fortbildung durch die teilnehmenden Lehrpersonen wirken sich auf der einen Seite deren individuelle Voraussetzungen und die schulischen Rahmenbedingungen aus. Auf der anderen Seite steht

das Angebot, das wesentlich von der konzeptionellen Anlage der Fortbildung und den Merkmalen der Moderatorinnen und Moderatoren, z. B. deren Wissen, deren Überzeugungen und deren Motivierungsfähigkeit, bestimmt wird. In den letzten 15 Jahren konnten einige empirische Befunde dazu vorgelegt werden, durch welche Merkmale sich Fortbildungsangebote für Lehrpersonen auszeichnen, die mit positiven Effekten auf der Ebene der Lehrpersonen und/oder deren Schülerinnen und Schülern einhergehen (zsf. z. B. Darling-Hammond, Hyler & Gardner, 2017; Lipowsky & Rzejak, 2015, 2017). Zu den lernförderlichen Fortbildungsmerkmalen zählt eine längere Fortbildungsdauer, durch die ein Wechsel von unterschiedlichen Bildungsphasen (Input/Erarbeitung, Erprobung, Reflexion) ermöglicht wird. Eine intensive Zusammenarbeit von Lehrpersonen im Sinne des Konzepts der professionellen Lerngemeinschaft und ein genauer Blick auf die Lern- und Verstehensprozesse von Schülerinnen und Schülern gelten als weitere Merkmale, die mit positiven Fortbildungserträgen einhergehen. Darüber hinaus verdeutlichen wirksame Fortbildungen den teilnehmenden Lehrpersonen häufig die Wechselwirkungen zwischen dem eigenen unterrichtlichen Handeln und dem Lernen der Schülerinnen und Schüler und es findet eine inhaltliche Orientierung des Fortbildungscurriculums an Befunden aus der Unterrichtsforschung statt.

2.2 Ebenen des Fortbildungserfolges

In der Evaluationsforschung existieren unterschiedliche Modelle, um potenzielle Ziele und Wirkungen einer Maßnahme hinsichtlich ihrer Reichweite zu systematisieren. Ein prominentes Modell ist das Vier-Ebenen-Modell von Kirkpatrick (1994), in dem die Ebenen *Reaction*, *Learning*, *Behavior* und *Results* unterschieden werden, um die Wirksamkeit einer Fortbildung zu evaluieren (s. **Fehler! Verweisquelle konnte nicht gefunden werden.**).

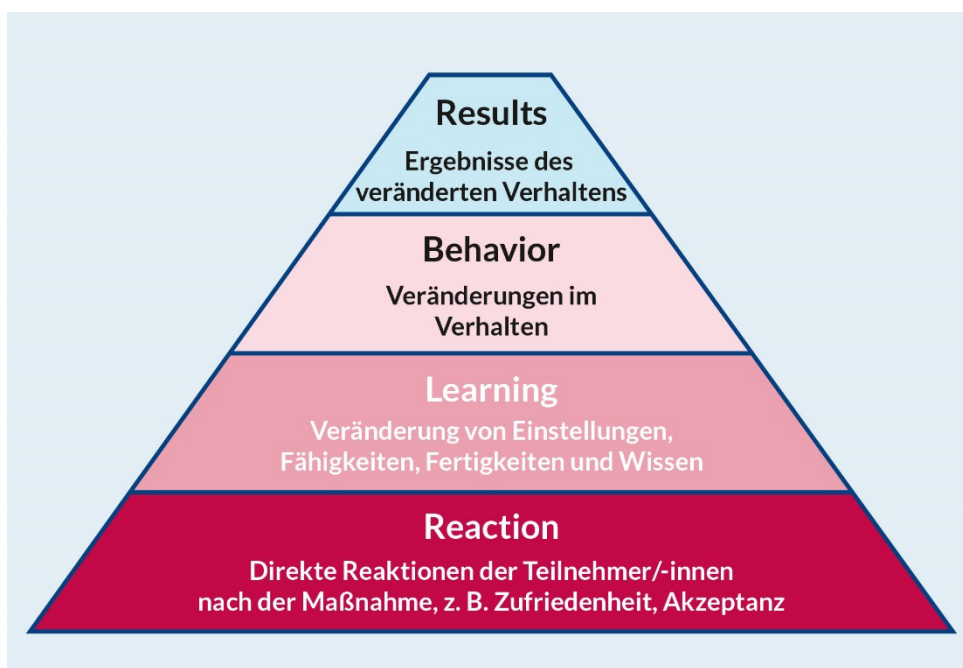


Abbildung 2. Vier-Ebenen-Evaluationsmodell nach Kirkpatrick (1994)

Bei der *ersten Ebene* handelt es sich um Wirkungen, die in unmittelbarem Anschluss an eine Fortbildung gemessen werden, d. h., es geht um direkte subjektive Reaktionen der Teilnehmerinnen und Teilnehmer, wie beispielsweise deren Zufriedenheit, die Akzeptanz und die Relevanz der Fortbildung sowie den selbstberichteten Lernertrag. Entsprechende Bewertungen, die das subjektive Erleben von fortgebildeten Lehrpersonen repräsentieren, lassen jedoch nur bedingt Rückschlüsse auf Entwicklungen und Veränderungen auf den höheren Ebenen zu und sollten daher um Wirksamkeitskriterien, die auf den Ebenen zwei bis vier anzusiedeln sind, ergänzt werden.

Auf der *zweiten Ebene* stehen Wirkungen, die sich im Wissen, in den Überzeugungen und/oder in affektiv-motivationalen Merkmalen der Teilnehmerinnen und Teilnehmer zeigen.

Als Wirksamkeitskriterien auf der *dritten Ebene* werden Veränderungen im eigentlichen Handeln der Fortbildungsteilnehmerinnen und -teilnehmer und damit auch in der Qualität von Unterricht angesehen. Diese Ebene wird wegen ihrer Komplexität und wegen der hohen methodischen Anforderungen an ein entsprechendes Untersuchungsdesign, z. B. Videografie von Unterricht, Unterrichtsbeobachtungen durch externe Beobachter/-innen oder eine Befragung von Schülerinnen und Schülern zur Wahrnehmung des Unterrichts, selten untersucht.

Bezogen auf Lehrerfortbildungen lassen sich Effekte auf die eigentlichen Adressaten des Unterrichts, die Schülerinnen und Schüler, als *vierte Ebene* der Wirksamkeit bezeichnen. Abhängig von der Konzeption und den Inhalten der Fortbildung sind sowohl leistungsbezogene Kriterien als auch ein Einfluss auf affektiv-motivationale Merkmale der Schülerinnen und Schüler vorstellbar (z. B. Lipowsky & Rzejak, 2015, 2017).

Die Wirksamkeitskriterien der zweiten, dritten und vierten Ebene finden sich auch in dem linearen Prozessmodell von Guskey (1985) wieder (s. Abbildung 3), wobei sich die Abfolge von dem Modell von Kirkpatrick unterscheidet.

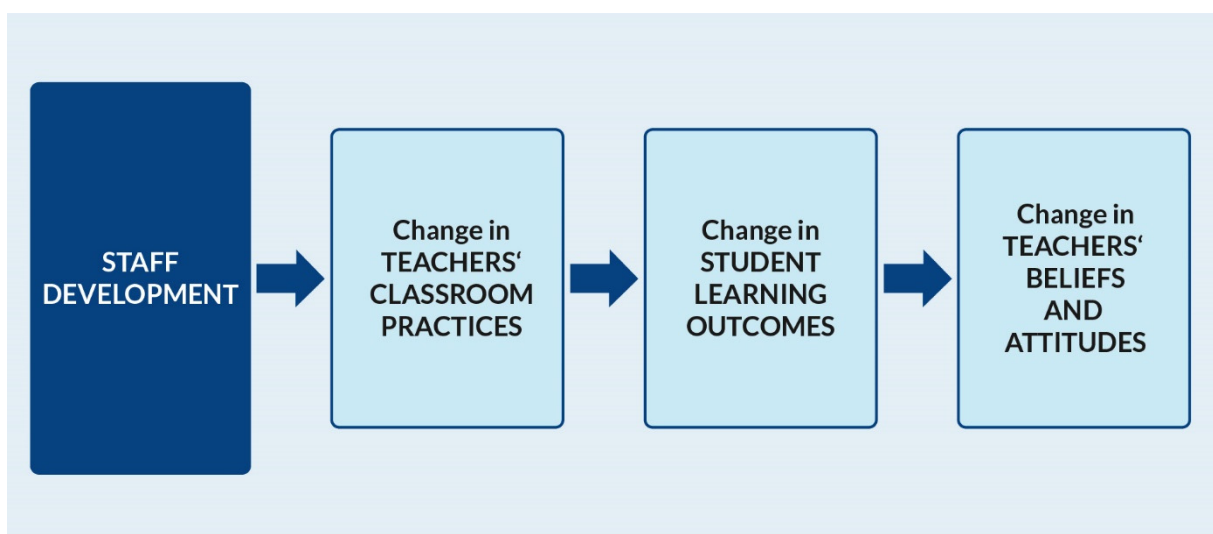


Abbildung 3. Modell zum Fortbildungsprozess von Lehrpersonen nach Guskey (1985)

Dem Modell folgend führt eine Fortbildung zunächst zu Veränderungen im unterrichtlichen Handeln von Lehrpersonen, woraufhin sich das Lernen der Schülerinnen und Schüler verändert. Erst im letzten Schritt zeigt sich ein Wandel in den Einstellungen und Überzeugungen von Lehrpersonen. Empirisch wurden die Wirkmechanismen zwischen den Merkmalen der unterschiedlichen Ebenen und die genauen Wirkprozesse bislang nicht vollständig untersucht. Jedoch ist anzunehmen, dass die Professionalisierung von Lehrpersonen nicht ausschließlich in der von Guskey (1985) angenommenen linearen Weise erfolgt, sondern beispielsweise die Veränderung von Überzeugungen und Einstellungen auch vor einer Weiterentwicklung der unterrichtlichen Praxis eintreten kann.

3 Vielfalt fördern und leitende Fragestellungen von LIQUID

Zum Umgang mit Heterogenität ist in den letzten zehn Jahren eine Vielzahl theoretischer und praxisbezogener Beiträge entstanden (Wischer, 2007). Der Bedarf an einer systematischen Qualifizierung zum Umgang mit Vielfalt in Schulen besteht bis heute (Gebhardt, Schwab, Nusser & Hessels, 2015). Der IQB-Ländervergleich aus dem Jahr 2011 (Richter, Kuhl, Reimers & Pant, 2012) zeigt beispielsweise, dass über 60 % der befragten Deutsch- und Mathematiklehrkräfte aus der Primarstufe einen subjektiven Fortbildungsbedarf im Bereich der Binnendifferenzierung und individuellen Förderung wahrnehmen. In einer Studie von Behr (2017), in der auch Lehrpersonen an Schulen mit Sekundarstufe I beteiligt waren, gaben 78 % der Lehrpersonen einen mittleren bis hohen Bedarf an dem Fortbildungsthema „*Unterricht in heterogenen Klassen*“ an und auch zum Thema „*Methoden der Schülerbeurteilung, Diagnostik*“ bekundeten 66 % der Lehrpersonen dieser Studie einen mittleren bis hohen Fortbildungsbedarf. Ausgehend von diesen Befunden lässt sich annehmen, dass *Vielfalt fördern* mit seinen Inhalten ein wichtiges Angebot für Lehrpersonen darstellen und bestehende Fortbildungsbedarfe bedienen dürfte.

Für das Projekt *Vielfalt fördern* haben die Bertelsmann Stiftung und das Ministerium für Schule und Weiterbildung sowie im weiteren Verlauf auch die Qualitäts- und UnterstützungsAgentur - Landesinstitut für Schule (QUA-LiS) in Nordrhein-Westfalen kooperiert, um Schulen mit Sekundarstufe I ab dem Schuljahr 2012/13 eine Fortbildung zum Umgang mit der wachsenden Heterogenität ihrer Schülerinnen und Schüler anbieten zu können.

Vielfalt fördern begann zunächst in den fünf Pilotregionen Bonn, Borken, Düsseldorf, Gütersloh und Hamm. Im Schuljahr 2013/2014 fand eine Ausweitung auf die fünf weiteren Pilotregionen Bielefeld, Gelsenkirchen, Kleve, den Hochsauerlandkreis und den Oberbergischen Kreis statt. Die Erweiterung auf neue Regionen in Nordrhein-Westfalen wurde zu jedem Schuljahr fortgesetzt. Seit dem Ablauf des Schuljahres 2016/17 liegt die Verantwortung für die Durchführung von *Vielfalt fördern* allein beim Land Nordrhein-Westfalen. Die Zusammenarbeit zwischen dem Land und der Bertelsmann Stiftung setzt sich aber in dem seit März 2018 zugänglichen MOOC (Massive Open Online Course) von *Vielfalt fördern* fort. *MOOC Vielfalt fördern* stellt einen digitalen Selbstlernkurs dar, der Einblicke in die Inhalte der Fortbildungsmodule bietet.

3.1 Ausgangskonzept der Fortbildung

Das Projekt *Vielfalt fördern* versteht sich nicht nur als ein wachsendes, sondern auch als ein lernendes und sich kontinuierlich weiterentwickelndes Projekt, wie auch die Ausweitung von einer „analogen“ Form zu einer digitalen Variante deutlich macht. Diese Offenheit für Verän-

derungen und Verbesserungen hat zur Folge, dass einzelne Merkmale der Fortbildungskonzeption, die den Ausgangspunkt der Evaluation bildeten, zum gegenwärtigen Zeitpunkt bereits angepasst und verändert wurden.

3.1.1 Fortbildungsadressaten

Das schulintern gestaltete Fortbildungsangebot von *Vielfalt fördern* richtet sich vorrangig an ganze Kollegien von allgemeinbildenden Schulen mit Sekundarstufe I. Großen Kollegien wird jedoch die Möglichkeit eingeräumt, dass die Fortbildung schrittweise begonnen wird, indem Lehrpersonen einzelner Jahrgangsstufen nacheinander die Maßnahme beginnen. Schulinterne Fortbildungen gelten als ein wichtiges Instrument der Schulentwicklung. Es wird angenommen, dass der Transfer von der Fortbildung in die schulische Praxis besser gelingt (Behr, 2017), allerdings dürfte eine schulinterne Fortbildung nicht per se wirksamer sein als eine Fortbildung, an der Lehrpersonen verschiedener Schulen teilnehmen. Um an *Vielfalt fördern* teilnehmen zu können, müssen sich Kollegien mit einer eindeutigen Mehrheit für die Teilnahme entscheiden und zudem eine Arbeits- oder Steuergruppe an der Schule einrichten, welche die Organisation und Implementierung der Fortbildung koordiniert. Ein Orientierungsworkshop im Vorfeld der eigentlichen Fortbildung dient den Schulen als Entscheidungsgrundlage für die Lehrer- und Schulkonferenz.

3.1.2 Fortbildungsdauer

Nach der Entscheidung für eine Teilnahme an *Vielfalt fördern*, aber noch vor dem eigentlichen Beginn der Fortbildung nehmen die Schulen an einem Pädagogischen Tag teil, der einer ersten Klärung des Grundverständnisses zur individuellen Förderung und kollegialen Zusammenarbeit im Kollegium dient.

Der anschließende Fortbildungsprozess umfasst vier Module, die jeweils für ein Schulhalbjahr konzipiert sind. Jedes Fortbildungsmodul besteht dabei aus mehreren Präsenzphasen mit einer Gesamtdauer von ca. 2.5 Tagen. Insgesamt umfasst die Fortbildung *Vielfalt fördern* 80 Stunden Präsenzzeit, wobei das Programm vorsieht, dass die teilnehmenden Lehrpersonen auch zwischen den Fortbildungssitzungen an den Inhalten arbeiten. Bei einem idealtypischen Verlauf findet demnach ein kontinuierlicher Fortbildungsprozess über einen Zeitraum von insgesamt zwei Jahren statt.

Im Projektverlauf zeigte sich allerdings, dass der realisierte Fortbildungsprozess in beiden Kohorten stark zwischen den einzelnen Schulen variierte und sich teilweise über einen deutlich längeren Zeitraum erstreckte, als dies bei einem idealtypischen Fortbildungsverlauf der Fall gewesen wäre.

3.1.3 Fortbildungsinhalte

Die vier Module von *Vielfalt fördern* haben jeweils einen unterschiedlichen inhaltlichen Schwerpunkt (s. Abbildung 4). Das erste Fortbildungsmodul stellt das kooperative Lernen der teilnehmenden Lehrpersonen selbst in den Mittelpunkt. In Anlehnung an das Konzept der professionellen Lerngemeinschaften sollen Kooperationsprozesse im Kollegium gefördert und intensiviert werden. Ein Schwerpunkt liegt hierbei auf der Deprivatisierung von Unterricht durch die Durchführung kollegialer Unterrichtshospitationen.

Im zweiten Modul wird das Thema pädagogische Diagnostik behandelt, wobei der Fokus auf dem Prozess der Diagnostik liegt. Das zweite Modul wurde bereits nach dem ersten Fortbildungsdurchlauf, der im Jahr 2012 begonnen hatte, so angepasst, dass fachspezifischen Fragen nun ein größeres Gewicht eingeräumt wird. Eine weitere Überarbeitung, die ab der dritten Kohorte gilt, besteht darin, dass theoretische Anteile, z. B. das Thema der Hypothesenformulierung, zugunsten einer stärkeren Anbindung an den Unterricht reduziert wurden.

Die Module drei und vier sehen mit dem Themenfeld Didaktik einen stärkeren Bezug zum Unterrichten und zur Fachspezifik von Lern- und Lehrprozessen vor. In den Modulen werden z. B. die Gestaltung kognitiv anregender Aufgabenformate oder Möglichkeiten des kooperativen Lernens von Schülerinnen und Schülern im Fachunterricht behandelt. Das Thema *Classroom Management*, das ursprünglich innerhalb der Didaktikmodule angesiedelt war, ist seit der zweiten Fortbildungskohorte ein Rahmenthema, das alle vier Module durchzieht und als roter Faden fungiert.

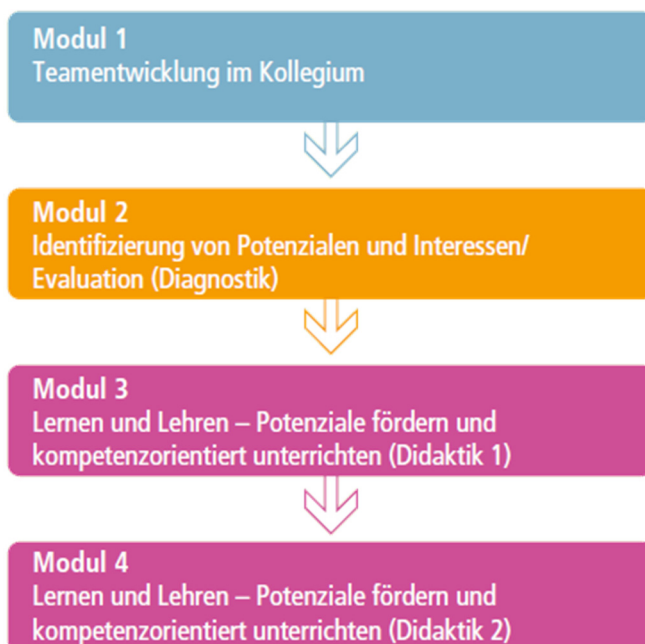


Abbildung 4. Inhalte der Module 1–4 (Müncher & Sturm-Schubert, 2015)

Durch standardisierte Präsentationen und Materialien soll sichergestellt werden, dass die Inhalte und die Qualität der Fortbildung zwischen den Schulen vergleichbar sind. Darüber hinaus

erhalten die Moderatorinnen und Moderatoren von *Vielfalt fördern*, die in der Regel in Teams eingesetzt werden, vor der Fortbildungsdurchführung eine eigene Schulung über die jeweiligen Inhalte der Module.

Das Angebot für Schulen, mit *Vielfalt fördern* an einer langfristig angelegten Fortbildung teilzunehmen, und die systematische Qualifizierung der eingesetzten Moderatoren stellen zwei wesentliche Besonderheiten dieser Fortbildung im Vergleich mit gängigen Angeboten für Lehrkräfte in der deutschen Fortbildungslandschaft dar.

3.2 Ziele der Fortbildung

Die übergeordnete Zielsetzung, die mit dem Fortbildungsprogramm *Vielfalt fördern* verfolgt wird, ist es, Schulen und Lehrpersonen darin zu unterstützen, besser mit der Heterogenität im schulischen Kontext umzugehen und Vielfalt als Chance zu begreifen. Auf diese Weise soll ein *„Beitrag zur Verbesserung von Chancengleichheit, sozialer Gerechtigkeit und Leistungsfähigkeit des Schulsystems in Nordrhein-Westfalen geleistet werden“* (Ministerium für Schule und Unterricht & Bertelsmann Stiftung, 2014).

Ergänzend zu dieser übergeordneten Leitidee sind für die einzelnen Module konkretere Ziele definiert worden, die den unterschiedlichen Ebenen des Fortbildungserfolges (s. Abschnitt 2.2) zugeordnet werden können. Eine ausführlichere Darstellung der Fortbildungsziele findet sich in dem Leitfaden von Müncher und Sturm-Schubert (2015) sowie auf der Projekthomepage (www.vielfalt-fördern.de). Im Folgenden werden nur exemplarisch einzelne Ziele der Module genannt. So soll beispielsweise durch das erste Modul zur Teambildung erreicht werden, dass die Lehrpersonen der teilnehmenden Kollegien das Verfahren der kollegialen Hospitation häufiger anwenden und intensiver zusammenarbeiten, um gemeinsam den Unterricht weiterzuentwickeln und das Lernen der Schülerinnen und Schüler bestmöglich zu unterstützen. Durch das Modul 2 soll sich der diagnostische Blick der teilnehmenden Lehrpersonen erweitern, d. h., sie sollen unterschiedliche Diagnosebereiche und -instrumente kennenlernen. Am Ende des zweiten Moduls sollen die teilnehmenden Lehrpersonen unter anderem dazu in der Lage sein, eine systematische pädagogische Diagnostik in ihrem schulischen Alltag anzuwenden. Hierzu gehört beispielsweise, dass eine Lehrperson am Anfang eines Diagnoseprozesses überprüfbare Hypothesen formuliert sowie anschließend erforderliche Informationen und Daten einholt, auswertet und interpretiert, um daraus Folgerungen und Maßnahmen abzuleiten. Da es im dritten und vierten Modul um Fragen des Lernens und Lehrens geht, beziehen sich die hier formulierten Ziele konkreter auf die Gestaltung von Unterricht und das Lernen von Schülerinnen und Schülern.

Durch das Modul sollen die Lehrpersonen lernen, aktuelle didaktische und fachdidaktische Standards bei der Konstruktion differenzierender Aufgaben zu berücksichtigen und jeder

Schülerin sowie jedem Schüler ein passendes Lernangebot zu bieten. Die Unterrichtsgestaltung der fortgebildeten Lehrpersonen soll am Ende des Moduls in stärkerem Maße individuelle und gleichzeitig auch kooperative Lernformen beinhalten und wichtige Merkmale eines effektiven Classroom Management sollen konsequenter und routinierter umgesetzt werden als zu Beginn der Fortbildung. Die Teilnehmerinnen und Teilnehmer von *Vielfalt fördern* sollen durch das dritte und vierte Modul darüber hinaus darin unterstützt werden, in ihrem Unterricht ein lernförderliches Klima zu schaffen; sie sollen Formen und Strategien selbstregulierten Lernens kennenlernen und diese auch in ihrem eigenen Unterricht einsetzen.

Vermittelt über die Professionalisierung der Lehrpersonen und deren Unterricht soll *Vielfalt fördern* sich auch auf das Lernen der Schülerinnen und Schüler günstig auswirken. Auf der Ebene der Schülerinnen und Schüler formuliert *Vielfalt fördern* unter anderem Ziele, die sich auf die Nutzung von Lehrerfeedback durch die Lernenden und deren Fähigkeiten beziehen, den eigenen Lernprozess selbstständig zu steuern. So sollen Schülerinnen und Schüler lernen, aus Rückmeldungen ihrer Lehrkraft Konsequenzen für das Lernen zu ziehen und ihren Lernprozess aus einer metakognitiven Perspektive zu betrachten. Neben dem positiven Einfluss auf das Lernverhalten derjenigen Schülerinnen und Schüler, deren Lehrpersonen fortgebildet werden, intendiert *Vielfalt fördern* auch positive Wirkungen auf deren fachliche Leistungsentwicklung.

4 Das Projekt LIQUID: Die Evaluation von *Vielfalt fördern*

Die Evaluation *LIQUID* wurde von 2012 bis 2018 in zwei Fortbildungskohorten von Prof. Dr. Frank Lipowsky und Daniela Rzejak von der Universität Kassel durchgeführt.

Die Evaluation der ersten Kohorte umfasst den Zeitraum 2012–2014 und das erste Modul von *Vielfalt fördern* zur Teambildung. Von 2013–2018 und mit Fokus auf die Module 2–4 wurde die zweite Fortbildungskohorte untersucht. Mit dem Projekt war das Ziel verbunden, bereits in der Pilotphase die Akzeptanz und Nutzung von *Vielfalt fördern* durch die teilnehmenden Lehrpersonen sowie die Wirkungen auf unterschiedliche schulische Bereiche zu untersuchen. *LIQUID* konnte als quasi-experimentelles, längsschnittlich angelegtes Evaluationsprojekt realisiert werden, in das unterschiedliche Personengruppen (Schulleiterinnen und Schulleiter, Lehrkräfte, Schülerinnen und Schüler, Moderatorinnen und Moderatoren) einbezogen wurden (s. Abschnitt 4.1).

Einen Ausgangspunkt für die Konzeption der Evaluation bilden die Ziele, die von der Bertelsmann Stiftung und dem Ministerium für Schule und Weiterbildung in Nordrhein-Westfalen für *Vielfalt fördern* definiert wurden (s. Abschnitt 3.2). Im Mittelpunkt steht die Frage, ob sich in der Gruppe, die an *Vielfalt fördern* teilnimmt, andere Entwicklungen zeigen als bei einer nicht an *Vielfalt fördern* teilnehmenden Kontrollgruppe. Wesentliches Ziel der Evaluation ist auch die Untersuchung einer weiterführenden Frage: Welche Einflussgrößen wirken sich auf den Fortbildungsprozess aus und können den Fortbildungserfolg befördern oder hemmen? Hier bildet das Angebots-Nutzungs-Modell aus der Fortbildungsforschung, das auf Faktoren verweist, die das Lernen und den Ertrag einer Fortbildung beeinflussen könnten (s. Abschnitt 2.1), einen wichtigen Bezugsrahmen. Diese weiterführende Fragestellung, die unter Einbezug von Daten der zweiten Fortbildungskohorte untersucht wird, bildet den Schwerpunkt des vorliegenden Berichts. Bevor Ergebnisse aus der Evaluation präsentiert werden, werden in den folgenden Abschnitten zunächst methodische Hintergrundinformationen zum Projekt *LIQUID* gegeben. Zunächst wird in Abschnitt 4.1.1 der Erhebungsplan erläutert, bevor die Instrumente (s. Abschnitt 4.1.2) und die Auswertungsverfahren (s. Abschnitt 4.1.3) dargestellt werden.

4.1 Methodischer Ansatz

Die Evaluation wurde als quantitative Untersuchung mit einem quasi-experimentellen Prä-Post-Design mit Kontrollgruppe (KG) angelegt.

Ein *Prä-Post-Design mit Kontrollgruppe* war für die Evaluation erforderlich, um Veränderungen und Entwicklungen bei der Untersuchungsgruppe (UG) über den Zeitraum der Fortbildung hinweg untersuchen und mit größerer Wahrscheinlichkeit auf die Teilnahme an *Vielfalt fördern* zurückführen zu können. Einzelne Merkmale wurden hierfür im Rahmen der Evaluation zu

mindestens zwei Zeitpunkten (prä, post) mit dem gleichen Erhebungsinstrument (Abschnitt 4.1.2) erfasst. Bei der KG handelt es sich um Schulleiterinnen und Schulleiter, Lehrpersonen sowie Schülerinnen und Schüler, die während des Zeitraums der Evaluation nicht an der Fortbildung *Vielfalt fördern* teilgenommen hatten. Vorausgesetzt, dass die UG und KG in ihrer Ausgangslage vor dem Fortbildungsbeginn vergleichbar sind, können Unterschiede zwischen UG und KG im Evaluationskriterium am Ende der Fortbildung wahrscheinlicher auf die Teilnahme an der Fortbildung zurückgeführt und andere Faktoren, die eine Entwicklung bewirkt haben könnten (z. B. zunehmende Erfahrungen im Laufe der Zeit, informelles Lernen am Arbeitsplatz, Veränderungen im Schuljahresrhythmus etc.), hingegen eher ausgeschlossen werden. Die *Prozesserhebungen*, die zwischen den Prä- und Posterhebungen der einzelnen Module lagen, wurden nur bei der UG durchgeführt, um Rückmeldungen zur Wahrnehmung und Nutzung des Fortbildungsangebots zu erhalten.

Als *quasi-experimentell* ist die Evaluation deshalb zu bezeichnen, da eine zufällige Zuordnung von Personen zu der UG und der KG nicht vorgenommen werden konnte, sondern „natürliche Gruppen“ gegeben sind. Im vorliegenden Projekt stellen die Lehrerkollegien und die Schulklassen solche „natürlichen Gruppen“ dar. Bei Untersuchungen mit quasi-experimentellem Design ist es erforderlich, dass zusätzlich zu den Merkmalen, die als Wirksamkeitskriterien untersucht werden, weitere Merkmale (Störvariablen) erfasst werden, die abgesehen von der Gruppenzugehörigkeit (UG vs. KG) die Zielvariablen beeinflussen können.

Im Verlauf der Evaluation bei der zweiten Fortbildungskohorte wurde entschieden, das quantitativ angelegte Design für die UG um eine qualitativ ausgerichtete *Vertiefungsstudie* zu erweitern. Die Erweiterung war mit der Intention verbunden, die quantitativen Befunde zu stützen und zu ergänzen. Hierzu wurden im Zeitraum von Juni 2016 bis Februar 2017 ausschließlich mit einzelnen UG-Schulen der zweiten Kohorte leitfadengestützte Gruppeninterviews durchgeführt. Die Interviews fanden acht bis 18 Monate nach Abschluss der Fortbildung an den Schulen statt.

4.1.1 Erhebungsplan

Die Entwicklung des Evaluationsdesigns orientierte sich an dem Konzept von *Vielfalt fördern* und betrachtet jedes Modul als einen separaten inhaltlichen Block. Für jedes inhaltliche Modul sah das Evaluationsdesign eine *Prä-, Prozess- und Posterhebung*, d. h. drei Erhebungskomponenten vor (s. Abbildung 5), die an den Fortbildungsverlauf der jeweiligen Schule gebunden waren.

	<i>Auftakt des Moduls</i>	<i>z. B. vorletzte Sitzung</i>	<i>Abschluss des Moduls</i>
	10/2012 <i>K1 t₁</i>	<i>K1 t₂</i>	10/2013 <i>K1 t₃</i>
LPS UG	Prä-FB	Prozess-FB	Post-FB
LPS KG	Prä-FB	-	Post-FB

Abbildung 5. Prä-Prozess-Posterhebungsstruktur am Beispiel von Kohorte 1. Anmerkung: LPS = Lehrpersonen, UG = Untersuchungsgruppe, KG = Kontrollgruppe, K1 = Kohorte 1, t₁, t₂ oder t₃ = Erhebungszeitpunkt 1, 2 oder 3, FB = Fragebogen

Das dritte und vierte Modul wurde allerdings aufgrund der inhaltlichen Nähe – beide Module thematisieren Fragen des Lernens und Lehrens – und zugunsten einer geringeren Belastung für die Befragungsteilnehmerinnen und -teilnehmer zusammengefasst und wie ein Modul behandelt. Es wurde daher nur eine Präerhebung am Anfang von Modul 3, eine Prozesserhebung zu Beginn von Modul 4 und eine Postbefragung am Ende von Modul 4 durchgeführt (s. Abbildung 6). Die Durchführung der Evaluation in zwei Kohorten resultierte hingegen nicht aus forschungsmethodischen Überlegungen, sondern stellt eine Konsequenz administrativer Entscheidungsprozesse dar, die im Vorfeld der Evaluation stattfanden.

In Abbildung 6 wird das Erhebungsdesign der Fortbildungsstudie vollständig dargestellt. Das Grunddesign – bestehend aus Prä-, Prozess- und Posterhebung – wurde bei beiden Kohorten identisch gehalten. Variationen ergaben sich vor allem im Hinblick auf die fokussierten Module und damit auf die Modulhalte: In Kohorte 1 wurde ausschließlich das Modul 1 zur Teambildung untersucht, in Kohorte 2 wurden die Module 2 (Diagnostik) und die Module 3 und 4 (Lernen und Lehren) evaluiert.

Das Ziel der Präerhebungen am Anfang eines jeweiligen Moduls war es, differenzierte Informationen über die Ausgangssituationen zu erhalten, mit denen die Lehrpersonen das Fortbildungsprojekt und das jeweilige Modul begannen. Es wurde z. B. erfasst, wie intensiv und mit welchen Ausrichtungen die Lehrpersonen vor dem Beginn des Teambildungsmoduls (K1 t₁) bereits zusammenarbeiteten und über welche prozessbezogenen Diagnosekompetenzen die Lehrpersonen bereits am Anfang des zweiten Moduls zu Fragen der Diagnostik (K2 t₁) verfügten. Die Posterhebungen am Ende eines Moduls ermöglichen es, unter Berücksichtigung der in den Präerhebungen ermittelten Ausgangslagen, Aussagen über stattgefundenen Veränderungen und Entwicklungen aufgrund der Fortbildungsteilnahme zu treffen. Dafür war es erforderlich, dass die gleichen Instrumente wie bei der Präbefragung eingesetzt wurden. Bei der UG verfolgten die Posterhebungen zusätzlich das Ziel, eine zusammenfassende Bewertung des jeweiligen Moduls von *Vielfalt fördern* zu erfassen. Die Fragebögen der Posterhebungen enthielten für die UG daher noch zusätzliche Fragen, beispielsweise zur Kosten-Nutzen-Einschätzung und dem selbstberichteten Lernertrag. Eine besondere Funktion hat zudem die jeweilige erste Präerhebung (t₁) in den beiden Kohorten. Die erste Befragung diente auch dazu,

demografische Angaben zu erfassen, um die Stichprobe beider Gruppen (UG und KG) beschreiben zu können. Bei den Lehrpersonen der UG wurden darüber hinaus auch motivationale Merkmale, die sich auf die Teilnahme an *Vielfalt fördern* beziehen, zum ersten Messzeitpunkt erhoben. Aufgrund der Intention, die motivationalen Voraussetzungen der teilnehmenden Lehrpersonen zu erfassen, wurde auch bei der zweiten Kohorte eine t_1 -Präerhebung am Anfang von Modul 1 durchgeführt, obwohl dieses Modul keinen Gegenstand der Evaluation darstellt. Bei den Lehrpersonen der KG der zweiten Kohorte wurde auf diesen Erhebungszeitpunkt verzichtet und die entsprechenden Hintergrundvariablen wurden stattdessen etwa ein halbes Jahr nach der Befragung der UG zusammen mit den Merkmalen der Präerhebung zum zweiten Modul (t_2) erfasst (s. Abbildung 6).

Die Prozesserhebungen, die im Zeitraum zwischen der Prä- und Posterhebung eines jeweiligen Moduls stattfanden (K1 t_3 , K2 t_3 , K2 t_6), wurden mit dem Ziel durchgeführt, Informationen über die Wahrnehmung und Nutzung des Fortbildungsangebots zu erhalten, z. B. über die wahrgenommene Strukturiertheit der Sitzungen, über die Relevanz der Inhalte und über deren Verarbeitung. Aus der Funktion der Prozesserhebungen für die Evaluation ergibt sich, dass es diese nur für die Lehrpersonen der UG sowie für die Moderatorinnen und Moderatoren gab.

Die Moderatorinnen und Moderatoren, die in Teams die Fortbildung in den Kollegien der UG durchführten, wurden in der Regel parallel zu den Lehrpersonen im Rahmen einer Fortbildungssitzung befragt.¹ Die Erhebungsunterlagen für beide Personengruppen – Lehrpersonen und Moderatorinnen/Moderatoren – wurden zu den jeweiligen Erhebungszeitpunkten direkt an die Schulen geschickt. Durch die Integration der Befragungen in die Fortbildungssitzungen sollte bei den Lehrpersonen und den Moderatorinnen und Moderatoren einem zeitlichen Mehraufwand durch die Evaluation entgegengewirkt werden. Die Eingangserhebungen (prä) wurden hierfür jeweils am Anfang des ersten Bausteins eines Moduls durchgeführt. Sowohl den Fragebogen der Prozesserhebungen als auch den Fragebogen der Abschlusserhebungen (post) bearbeiteten die Lehrpersonen und Moderatorinnen/Moderatoren am Ende eines Fortbildungsbausteins. Die zeitliche Anbindung der Erhebungen an die schulischen Fortbildungstermine führte dazu, dass sich zusammen mit den Verzögerungen im Fortbildungsverlauf auch die Erhebungszeiträume der Fortbildungsstudie beziehungsweise der Lehrerbefragungen bei der UG verlängerten.

¹ Eine Ausnahme stellte eine computerbasierte Erhebung zum fünften Messzeitpunkt (t_5) bei den Moderatorinnen und Moderatoren der zweiten Kohorte dar, an welcher diese in einem vorgegebenen Zeitraum teilnehmen sollten.

Auf die Erhebungen bei den Lehrpersonen der KG wirkten sich die Verzögerungen im Fortbildungsverlauf nicht aus. Hier fanden die Befragungen – zeitlich orientiert an einen idealtypischen Fortbildungsverlauf – in der Regel angeleitet durch eine Mitarbeiterin oder einen Mitarbeiter der Universität Kassel im Rahmen einer Lehrerkonferenz statt.²

Während die grundlegende Struktur und der Ablauf der Erhebungen bei der Evaluation der ersten und der zweiten Kohorte identisch waren, bestehen zwischen den beiden Evaluationen auch kleinere Unterschiede, die im Folgenden kurz erläutert werden.

Die Schulleiterinnen und Schulleiter der ersten Kohorte wurden einmalig zu Beginn der Fortbildung (K1 t₁) befragt. Diese Befragung diente der ökonomischen Erfassung von wichtigen Schulmerkmalen, die für alle Lehrpersonen einer Schule in vergleichbarer Weise gelten sollten. Da sich *Vielfalt fördern* aber nicht ausdrücklich an die Gruppe der Schulleitung wendet und keine expliziten Ziele für die Schulleitung formuliert sind, stand die Untersuchung von Entwicklungen bei dieser Personengruppe nicht im Zentrum der Evaluation. Über einen längeren Zeitraum betrachtet – wie bei der zweiten Kohorte – sind jedoch auch indirekte Wirkungen von *Vielfalt fördern* auf schulische Rahmenbedingungen, beispielsweise aufgrund von neuen Erfahrungen und Formen der Zusammenarbeit unter den Lehrpersonen, denkbar. Bei der zweiten Kohorte fanden aus diesem Grund zwei Schulleitererhebungen statt, eine Erhebung am Anfang des Fortbildungsprozesses (K2 t₁) und eine Erhebung am Ende von *Vielfalt fördern* (K2 t₇). Die entsprechende Post-Erhebung erfolgte vornehmlich, um die Perspektive der Schulleitung auf die Fortbildung *Vielfalt fördern* zu erfassen und um Hinweise auf neu eingerichtete schulische Strukturen zu erhalten.

Ein weiterer Unterschied zwischen den beiden Evaluationskohorten liegt in der Unterteilung in die sogenannte *Fortbildungsstudie* und die sogenannte *Unterrichtsstudie*, die nur bei der zweiten Kohorte besteht. In der Fortbildungsstudie stehen die Entwicklungen der Lehrpersonen in der UG sowie deren Wahrnehmung und Nutzung des Fortbildungsangebots im Mittelpunkt des Interesses. Bei der Unterrichtsstudie werden Schülerinnen und Schüler in die Evaluation einbezogen, um zu prüfen, inwieweit *Vielfalt fördern* zu einer Veränderung des Unterrichts und des Lernens der Schülerinnen und Schüler beigetragen hat. Untersucht wurden sowohl Merkmale der Unterrichtsqualität als auch leistungsbezogene Merkmale der Schülerinnen und Schüler. Da Wirkungen auf den Unterricht und das Lernen der Schülerinnen und Schüler insbesondere aufgrund der Fortbildungsmodule 2–4 und nicht aufgrund des Moduls 1 erwartet wurden, wurde auf diesen aufwendigen Erhebungsteil bei der ersten Evaluationskohorte verzichtet.

² In Ausnahmefällen wurden die Erhebungsmaterialien an die Schulen gesandt, durch die Schulleitung verteilt und ausgefüllt wieder zurückgeschickt.

Der Fokus wurde in der Unterrichtsstudie auf den Unterricht im Fach Mathematik und im Fach Deutsch gelegt, wobei die Lernenden einer Klasse für die Unterrichtsstudie zufällig auf beide Fächer aufgeteilt und entweder der Erhebung im Fach Mathematik oder der Erhebung im Fach Deutsch zugewiesen wurden. Die Zuteilung zu diesem Fach blieb für die Schülerinnen und Schüler über beide Erhebungszeitpunkte im Regelfall konstant, auch wenn sich die Klassenzusammensetzung beispielsweise durch Abgänge oder Zugänge zwischen beiden Messzeitpunkten veränderte. Eine Ausnahme bildete der Fall, dass der Rücklauf einer Fachgruppe pro Klasse zum ersten Erhebungszeitpunkt zu gering war. Unter diesen Umständen wurde in der betroffenen Klasse zum zweiten Erhebungszeitpunkt nur noch ein Fach erhoben, wobei dann alle Lernenden der Klasse an dieser Erhebung teilnahmen.

Die Eingangserhebung der Unterrichtsstudie (K2 ust₁) wurde zum Ende der 5. Jahrgangsstufe im Schuljahr 2013/2014 in einem sechswöchigen Zeitfenster durchgeführt. Die Abschlusserhebung erfolgte am Anfang des Schuljahres 2015/2016 in den Monaten August/September. Die Schülerinnen und Schüler befanden sich demnach zum zweiten Messzeitpunkt (K2 ust₂) am Beginn des 7. Schuljahres. Der Zeitraum, der durch die Unterrichtsstudie betrachtet wurde, ist demnach das 6. Schuljahr. Bei einem idealtypischen Fortbildungsprozess, d. h. einer Dauer von einem Schulhalbjahr pro Fortbildungsmodul, wären die beiden Erhebungszeitpunkte der Unterrichtsstudie zeitlich so an die Fortbildung angebunden gewesen, dass die Lehrpersonen der UG zum Zeitpunkt der ersten Schülererhebung das dritte Modul noch nicht begonnen und zum Zeitpunkt der zweiten Schülererhebung das vierte Modul abgeschlossen hätten. Die Abweichungen vom idealtypischen Fortbilverlauf hatten jedoch zur Folge, dass die Lehrpersonen der UG in einigen Fällen die Module 3 und 4 zum Lernen und Lehren zum Zeitpunkt der zweiten Schülererhebung noch nicht vollständig durchlaufen hatten. Dies ist vor dem Hintergrund problematisch, dass insbesondere diese Module positive Effekte auf den Unterricht und auf das Lehren und Lernen beabsichtigen. Dennoch war es erforderlich, die Erhebungszeiträume bei der Unterrichtsstudie konzeptgetreu durchzuführen, um den zeitlichen Rahmen, über den hinweg Entwicklungen betrachtet wurden (6. Schuljahr), zwischen den Schülerinnen und Schülern vergleichbar zu belassen. Um eine möglichst hohe Vergleichbarkeit der Erhebungssituationen zu erzielen, erfolgte die Durchführung in den Klassen zu beiden Erhebungszeitpunkten durch geschulte Mitarbeiterinnen und Mitarbeiter der Universität Kassel.

Auch die Vertiefungsstudie, in deren Rahmen Gruppeninterviews mit Schulen der UG geführt wurden, stellt eine Besonderheit der zweiten Kohorte dar. Mit ausgewählten Schulen der UG wurde im Zeitraum von Juni 2016 bis Februar 2017 ein leitfadengestütztes Gruppeninterview durchgeführt. Das Interview fand in der Regel acht bis 18 Monate nach Abschluss der Fortbildung an den Schulen statt und diente dazu, die Daten der quantitativen Erhebungen durch weitere Rückmeldungen aus den Schulen zu ergänzen und Hinweise auf langfristige Wirkungen von *Vielfalt fördern* zu erhalten.

		Auftakt Modul 1	Vorletzte Sitzung	Abschluss Modul 1	Auftakt Modul 2	Vorletzte Sitzung	Abschluss Modul 2	Auftakt Modul 3	Auftakt Modul 4	Abschluss Modul 4	
Fortbildungsstudie Kohorte 1		10/2012 K1 t₁		10/2013 K1 t₂							
	LPS UG	Prä-FB	Prozess-FB	Post-FB							
	LPS KG	Prä-FB		Post-FB							
	SL (UG, KG)	Prä-FB									
	Mod	Prä-FB	Prozess-FB	Post-FB							
Fortbildungsstudie Kohorte 2		10/2013 K2 t₁			K2 t₂		K2 t₃	K2 t₄	K2 t₅	K2 t₆	11/2016 K2 t₇
	LPS UG	Prä-FB			Prä-FB	Prozess-FB	Post-FB	Prä-FB	Prozess-FB	Post-FB	
	LPS KG	Prä-FB					Post-FB	Prä-FB		Post-FB	
	SL (UG, KG)	Prä-FB								Post-FB	
	Mod	Prä-FB			Prä-FB	Prozess-FB	Post-FB	Prä-FB	Prozess-FB	Post-FB	
Unterrichtsstudie Kohorte 2		2014 Ende 5. Jg. K2 ust₁					2015 Anfang 7. Jg. K2 ust₂				
	SuS-M (UG, KG)				Prä-FB und Leistungstest					Post-FB und Leistungstest	
	SuS-D (UG, KG)				Prä-FB und Leistungstest					Post-FB und Leistungstest	
	F-LPS (M, D)				Prä-FB					Post-FB	

Abbildung 6. Erhebungsdesign bei Kohorte 1 (K1) und Kohorte 2 (K2). Anmerkung: SuS-M = Schülerinnen und Schüler, untersucht im Fach Mathematik, SuS-D = Schülerinnen und Schüler, untersucht im Fach Deutsch, F-LPS (M, D) = Fachlehrpersonen Mathematik und Deutsch, FB = Fragebogen

4.1.2 Instrumente

Ein umfassender technischer Bericht zu den eingesetzten Instrumenten (z. B. inklusive aller Quellen, Itemformulierungen und -trennschärfen, Skalenreliabilität und -mittelwerte) wird an dieser Stelle nicht gegeben, da dies den Rahmen des Berichts sprengen und dessen Intention verfehlen würde. Die Instrumente werden daher im Folgenden nur konzise beschrieben. Merkmale und Konstrukte, die für den vorliegenden Bericht von besonderer Bedeutung sind, werden bei der ersten Bezugnahme im Ergebnisteil separat vorgestellt.

4.1.2.1 Fragebögen zur Erfassung von Einstellungen, Motivationen und Wahrnehmungen

Die Erhebungen des Projekts *LIQUID* erfolgten für alle Personengruppen mit Ausnahme einer computerbasierten Befragung der Moderatorinnen und Moderatoren und der qualitativ angelegten Vertiefungsstudie bei der Kohorte 2 (s. Abschnitt 4.1.1) schriftlich und papierbasiert. Um die Spezifika von *Vielfalt fördern* berücksichtigen und abbilden zu können, wurden die eingesetzten Instrumente für die Evaluation neu entwickelt. Dabei wurden sowohl erprobte Skalen aus früheren Studien übernommen und adaptiert als auch neuentwickelte Skalen eingesetzt. In erster Linie beinhalteten die Fragebögen Skalen mit mehrfach abgestuften Antwortkategorien, die jeweils durch eine verbale Verankerung gekennzeichnet waren (s. Abbildung 7).

Bei der Bearbeitung des Erhebungsinstruments war jede Person (Lehrerin/Lehrer, Schülerin/Schüler, Schulleiterin/Schulleiter, Moderatorin/Moderator) dazu aufgefordert, einzuschätzen (engl.: rate), in welchem Maß die jeweilige Aussage (Item) auf sie zutrifft. Die verbale Verankerung differierte hierbei zwischen unterschiedlichen Merkmalen, war bei den Lehrer-, Schulleiter- und Moderatorenfragebögen aber überwiegend sechsstufig (z. B. [1] = *trifft gar nicht zu* bis [6] = *trifft voll und ganz zu*; [1] = *absolut unsicher* bis [6] = *absolut sicher*).

09. Wie zufrieden sind Sie mit Ihrem Beruf?	trifft gar nicht zu	trifft größtenteils nicht zu	trifft eher nicht zu	trifft eher zu	trifft größtenteils zu	trifft voll und ganz zu
Für mich überwiegen in unserem Beruf eindeutig die Vorteile.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Wenn ich noch einmal wählen könnte, würde ich sofort wieder Lehrer werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich habe mir schon einmal überlegt, ob es für mich nicht besser gewesen wäre, einen anderen Beruf zu ergreifen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Für mich gibt es keinen besseren Beruf.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich bereue manchmal, Lehrer geworden zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 7. Anordnung von Items und Antwortkategorien am Beispiel der Berufszufriedenheit von Lehrpersonen

Da die kognitiven Anforderungen bei der Beantwortung eines Items mit einer zunehmenden Anzahl an Antwortkategorien eher wachsen, wurde den Schülerinnen und Schülern nur ein vierstufiges Antwortformat vorgegeben. Die Polung wurde in allen Befragungsinstrumenten identisch gehalten, d. h., ein Wert von 1.00 ist jeweils als geringe Ausprägung zu interpretieren und ein Wert von 6.00 (4.00 bei den Schülerinnen und Schülern) als sehr hohe Ausprägung. Abhängig von der Anzahl an Antwortkategorien liegt das theoretische Skalenmittel bei $M = 3.50$ (sechsstufiges Antwortformat) oder bei $M = 2.50$ (vierstufiges Antwortformat).

Um die Genauigkeit (Reliabilität) zu erhöhen, mit der ein Merkmal gemessen wird, wurden diese überwiegend mit mehreren Aussagen oder Fragen erfasst, über die zum Teil auch unterschiedliche Merkmalsfacetten abgebildet wurden. Durch die Erfassung eines Merkmals über mehrere Items können Messfehler reduziert werden. Hierfür wurden im Projekt *LIQUID*, wie üblich, die einzelnen Items zu einer Skala zusammengefasst, indem in der Regel ein Mittelwert über die Werte der zu einer Skala gehörenden Items gebildet wurde. Der Fragebogenausschnitt in Abbildung 7 zeigt beispielsweise fünf einzelne Items, die nach vorangehender statistischer Prüfung zu einer Mittelwertskala, der Berufszufriedenheit, zusammengefasst wurden.

Neben den Ratingskalen beinhalteten die Instrumente zum Teil auch Einzelitems mit einem itemspezifischen Antwortformat, z. B. bei der Erfassung von demografischen oder schulbezogenen Merkmalen sowie bei Wissensfragen.

Anders als bei Fragebögen, durch die beispielsweise Einstellungen, Motivationen und/oder Wahrnehmungen erfasst werden, gibt es bei Wissens- und Leistungstests in der Regel eindeutig richtige und falsche Lösungen. Wissens- und Leistungstests wurden im Projekt *LIQUID* bei den Lehrpersonen und bei den Schülerinnen und Schülern der zweiten Kohorte eingesetzt.

4.1.2.2 Tests zur Erfassung des Wissens der Lehrpersonen über Diagnostik und lernwirksame Unterrichtsmerkmale

Bei den Lehrpersonen wurde am Anfang und am Ende von Modul 2 (K2 t₂, K2 t₄) z. B. das *diagnostische Wissen* über einen Test aus elf Fragen im Multiple-Choice-Format gemessen, wobei zum Teil Mehrfachnennungen möglich waren. Insgesamt konnten die Lehrpersonen 47 Punkte in dem Test erreichen. Am Anfang von Modul 3 und am Ende von Modul 4 (K2 t₅, K2 t₇) wurde unter anderem das Wissen der Lehrpersonen über Merkmale lernwirksamen Unterrichts erfasst. Die Lehrpersonen wurden mit Merkmalen von Unterricht konfrontiert und aufgefordert, jeweils zu beurteilen, in welchem Ausmaß es sich um ein Merkmal lernwirksamen Unterrichts handelt. Unter den präsentierten Merkmalen wurden auch einige genannt, die auf der Basis des Forschungsstands nicht als Merkmale lernwirksamen Unterrichts gelten. Bewertet wurde, ob Merkmale lernwirksamen Unterrichts und Merkmale, die nicht als lernwirksam

gelten, korrekt identifiziert wurden. Insgesamt konnte hierbei eine Gesamtpunktzahl von 14 Punkten erreicht werden.

4.1.2.3 Tests zur Erfassung der Leistungen der Schülerinnen und Schüler in Mathematik und Deutsch

Die Schülerinnen und Schüler wurden zu den beiden Messzeitpunkten der Unterrichtsstudie entweder in ihren fachlichen Leistungen in Mathematik *oder* in ihren Leseleistungen im Fach Deutsch getestet. Für die Testungen wurden standardisierte Verfahren eingesetzt, die von einer geschulten Mitarbeiterin oder einem geschulten Mitarbeiter der Universität Kassel angeleitet wurden, um den Erhebungsablauf zwischen den Klassen vergleichbar zu gestalten. Während der Erhebung war jedoch auch eine Lehrperson der Schülerinnen und Schüler anwesend.

Da *Vielfalt fördern* für das Fach Mathematik keine fachlichen und kompetenzspezifischen Schwerpunkte setzte³ und die eingesetzten Leistungstests nach Möglichkeit curricular valide sein sollten, wurde über die in den Tests eingesetzten Aufgaben ein möglichst breites Spektrum an Leitideen und Kompetenzen abgedeckt. Um die Leistungsentwicklung erfassen zu können, wurde in beiden Tests (prä, post) ein Teil der Aufgaben, sogenannte Ankeraufgaben, konstant gehalten. Der überwiegende Teil der Aufgaben wurde den Testinstrumenten DEMAT 5+ (Götz, Lingel & Schneider, 2013a) und DEMAT 6+ (Götz, Lingel & Schneider, 2013b) entnommen. Darüber hinaus stammten einige weitere Aufgaben aus den national einheitlichen Vergleichsarbeiten (www.iqb.hu-berlin.de/vera) sowie aus den Leistungstests des Forschungsprojekts PALMA (Projekt zur Analyse der Leistungsentwicklung in Mathematik; http://www.uni-regensburg.de/Fakultaeten/nat_Fak_I/BIQUA/newsf.html).

Das Textverständnis der Schülerinnen und Schüler wurde zu beiden Erhebungszeitpunkten mit dem standardisierten Lesetest *LESEN 6-7* (Bäuerlein, Lenhard & Schneider, 2012) gemessen. Hierfür wurden beide in dem Test vorgesehenen Texte – ein expositorischer und ein narrativer Text – eingesetzt. Die Lernenden wurden nach der Textlektüre jeweils aufgefordert, Multiple-Choice-Verständnisfragen mit je einer richtigen Antwortoption und vier Distraktoren zu beantworten. Die Antworten wurden hinsichtlich ihrer Korrektheit (richtig vs. falsch) ausgewertet.

4.1.2.4 Tests zur Erfassung des Strategiewissens der Schülerinnen und Schüler in Mathematik und Deutsch

Zusätzlich zu den Leistungstests in den Fächern Mathematik und Deutsch (s. Abschnitt 4.1.2.3) wurde zu beiden Erhebungszeitpunkten – am Ende der 5. Klasse und am Anfang der

³ Dies markiert einen Unterschied zum Fach Deutsch. Hier legte *Vielfalt fördern* explizit einen Fokus auf die Leseförderung.

7. Klasse – auch das metakognitive Strategiewissen der Schülerinnen und Schüler erfasst. Hierfür wurden in beiden Fächern erprobte Instrumente eingesetzt. Im Fach Mathematik kam der mathematische Strategiewisstest für 5. und 6. Klassen (MAESTRA 5-6+) von Lingel, Götz, Artelt und Schneider (2014) zum Einsatz und im Fach Deutsch beziehungsweise im Bereich Lesen und Arbeiten mit Texten ein Testverfahren von Neuenhaus (2011).

Die beiden Testverfahren basieren auf der identischen Grundidee, um das metakognitive Strategiewissen der Schülerinnen und Schüler zu erfassen (Artelt & Neuenhaus, 2010; Lingel, 2016): Die Lernenden werden in dem Test mit konkreten Lernsituationen sowie jeweils mit einer Auswahl von Strategien konfrontiert. Die verschiedenen Strategien sollen hinsichtlich ihrer Angemessenheit und Nützlichkeit bewertet werden. Zur Bewertung der Strategien vergeben die Schülerinnen und Schüler Schulnoten von 1 = sehr gut bis 6 = unbefriedigend. Diese Strategiebewertungen der Schülerinnen und Schüler werden bei der Testauswertung paarweise betrachtet und dahingehend geprüft, ob für eine angemessenere Strategie tatsächlich auch eine bessere Note vergeben wurde. Den Maßstab für die Bewertung der Rangfolge der verschiedenen Strategien bilden Experteneinschätzungen. Für eine korrekte Beurteilung der Relation zwischen zwei alternativen Strategien, d. h. für die Übereinstimmung mit den Expertenurteilen, wird jeweils ein Punkt vergeben.

In beiden Verfahren werden den Lernenden fünf fachspezifische Lern- und Leistungsszenarien zusammen mit jeweils fünf bis sechs strategischen Handlungsalternativen präsentiert, die zu bewerten sind.

4.1.2.5 Interview

Für die Gruppeninterviews, die im Rahmen der Vertiefungsstudie geführt wurden, wurde ein Interviewleitfaden entwickelt, an dem sich der Ablauf der Interviews orientierte. Der Leitfaden umfasste neben einer einleitenden Warm-up-Phase und einer Abschlussphase insgesamt fünf Themenbereiche: Konzeption, Motivation, Moderatorinnen und Moderatoren, Wirkungen, außerschulische Zusammenarbeit. Als Interviewdauer waren ca. zwei Stunden veranschlagt. Um sicherzustellen, dass in diesem Zeitfenster alle fünf Oberthemen angesprochen werden, wurden innerhalb der fünf Bereiche jeweils Unterthemen definiert, die bei Überschreitung einer bestimmten Interviewdauer übersprungen wurden.

4.1.3 Auswertungsverfahren

Aufgrund der unterschiedlichen Erhebungsinstrumente, Stichproben und Fragestellungen wurden die erfassten Daten mit unterschiedlichen Verfahren ausgewertet. Da die Evaluation aber vornehmlich quantitativ angelegt war, dominierten deskriptive und inferenzstatistische Analysemethoden. Die Verfahren, die zur inhaltlichen Auswertung der Fragestellungen und Hypothesen eingesetzt wurden, werden im Folgenden kurz erläutert.

Auf eine Darstellung der vorbereitenden Analyseverfahren (z. B. Faktoren-, Reliabilitäts- und IRT-Analysen), durch die unter anderem die Struktur und Güte der eingesetzten Instrumente überprüft wurde, wird an dieser Stelle verzichtet.

4.1.3.1 Verfahren zur Auswertung der quantitativen Daten

Um zu untersuchen, ob die Voraussetzungen, d. h. die Ausgangslage der UG und KG vergleichbar sind, wurde abhängig von dem Merkmal und der Stichprobe der Chi-Quadrat-Test, der Mann-Whitney-U-Test oder der t-Test für unabhängige Stichproben eingesetzt.

Mittels *Chi-Quadrat-Test* werden Häufigkeitsverteilungen auf signifikante Unterschiede überprüft. Das Verfahren wurde im Projekt *LIQUID* eingesetzt, um zu untersuchen, ob nominale Merkmale, wie z. B. das Geschlecht oder die Muttersprache, in verschiedenen Gruppen unterschiedlich verteilt sind. Als Prüfgröße wird Chi-Quadrat (χ^2) angegeben.

Der *Mann-Whitney-U-Test* (kurz: *U-Test*) ist ein nonparametrisches Verfahren, um Hypothesen zur Unterschiedlichkeit von zwei unabhängigen Gruppen zu prüfen. Bei nonparametrischen Verfahren handelt es sich um Verfahren, die keine spezifische Voraussetzung an die Verteilungsform der Daten stellen und insbesondere bei kleinen Stichproben eingesetzt werden. Unter Anwendung des U-Tests wurde im Projekt *LIQUID* beispielsweise untersucht, ob die Schulen der UG und der KG in Merkmalen des schulischen Kontexts vergleichbar sind. Als Prüfgröße dient U, für das die zweiseitige Signifikanz angegeben wird. Um die praktische Bedeutsamkeit eines signifikanten Ergebnisses einschätzen zu können, wird die *Effektstärke Cohens d* bestimmt.⁴ Ein Wert ab $d = 0.20$ gilt als ein Indiz für einen schwachen Effekt. Ein *Cohens d* ab $d = 0.50$ kann als mittlerer Effekt interpretiert werden und ab einem $d = 0.80$ spricht man von einem starken Effekt (Bortz & Weber, 2005).

Mittels *t-Test für unabhängige Stichproben* kann ebenfalls überprüft werden, ob sich die Mittelwerte eines Merkmals zwischen zwei Gruppen unterscheiden. Im Rahmen der Auswertungen für den vorliegenden Bericht wird das Verfahren unter anderem genutzt, um zu überprüfen, ob sich das Durchschnittsalter der Lehrpersonen in der UG von dem Durchschnittsalter in der KG bedeutsam unterscheidet.

Wenn Unterschiede zwischen mehr als zwei Gruppen (z. B. UG vs. KG, Frauen vs. Männer) untersucht werden sollen, kann der t-Test nicht mehr angewandt werden. Die Überprüfung von Unterschieden erfolgt in solchen Fällen mit Hilfe von *einfaktoriellen Varianzanalysen*. Im vorliegenden Bericht kam das Verfahren beispielsweise zum Einsatz, um zu untersuchen, ob sich Lehrpersonen, die hinsichtlich ihrer motivationalen Voraussetzungen gruppiert wurden, in der

⁴ Die Berechnung der Effektstärke erfolgt unter Zuhilfenahme der Internetseite <http://www.psychometrica.de/effektstaerke.html>.

Wahrnehmung, Nutzung und in den Fortbildungserträgen unterscheiden. Um bei einem signifikanten Unterschied zu prüfen, welche Gruppen sich voneinander unterscheiden, wurde ein sogenannter A-posteriori-Einzelvergleich (Post-hoc-Test) durchgeführt. Da die Gruppen, die miteinander zu vergleichen sind, unterschiedlich groß sind, wurde der Scheffé-Test genutzt, obwohl dieser häufig konservativer ist, d. h. größere Mittelwertdifferenzen erforderlich sind, damit ein Unterschied signifikant ausfällt. Um die praktische Bedeutsamkeit eines signifikanten Unterschieds beurteilen zu können, wird Eta-Quadrat (η^2) angegeben. Ein $\eta^2 > .01$ gilt als schwacher Effekt, ein $\eta^2 > .06$ als mittlerer Effekt und ab einem $\eta^2 > .14$ spricht man von einem starken Effekt.

Darüber hinaus wurden die Daten mit *regressionsanalytischen Verfahren* ausgewertet. Mit einer linearen Regression untersucht man, ob die Ausprägung eines abhängigen Merkmals (Kriterium) durch ein anderes unabhängiges Merkmal (Prädiktor) erklärt werden kann. Eine Regressionsanalyse ist nur sinnvoll, wenn die Prädiktorvariable und die Kriteriumsvariable miteinander zusammenhängen (auch: korrelieren). Wenn nur ein Prädiktor als erklärende Variable berücksichtigt wird, spricht man von einer einfachen linearen Regression (s. Abbildung 8).

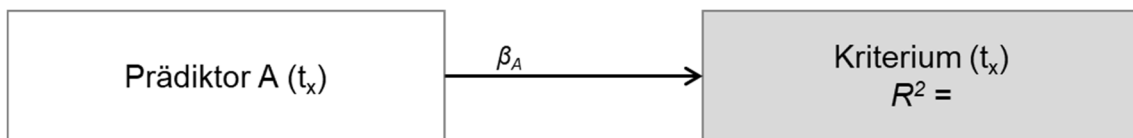


Abbildung 8. Beispielabbildung für eine einfache lineare Regression

Werden hingegen mehrere Prädiktoren in das Modell aufgenommen, um die Ausprägung im Kriterium besser erklären zu können, bezeichnet man das Vorgehen als multiple lineare Regression. In Abbildung 9 ist ein Modell mit drei erklärenden Prädiktoren (A, B, C) dargestellt.

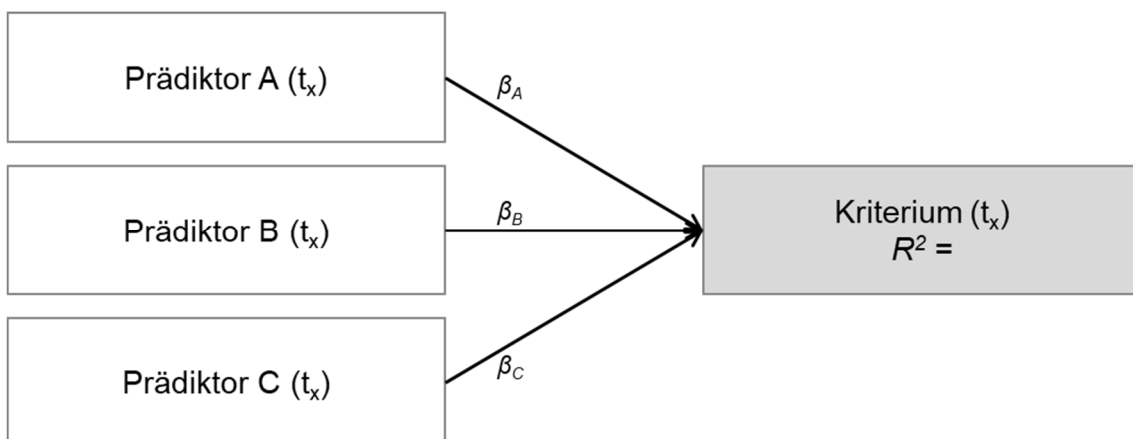


Abbildung 9. Beispielabbildung für eine multiple lineare Regression

Wichtige Kennwerte sind bei Regressionsanalysen der *standardisierte Regressionskoeffizient Beta* (β) und das *Bestimmtheitsmaß* R^2 : Die standardisierten β -Werte sind skalunabhängig und können daher miteinander verglichen werden. Aus dem Regressionskoeffizienten geht die

Richtung (positiv vs. negativ) hervor sowie die Genauigkeit, mit der die Prädiktorvariable das Kriterium vorhersagt. R^2 gibt Auskunft darüber, wie groß der Anteil der Varianz in der Kriteriumsvariable ist, der durch die Prädiktoren des Modells erklärt werden kann. Es ermöglicht damit eine Einschätzung, wie gut das Modell geeignet ist, das jeweilige Kriterium zu erklären. Das Bestimmtheitsmaß R^2 kann Werte zwischen $R^2 = 0$ und $R^2 = 1$ annehmen, wobei ein $R^2 = 0$ bedeuten würde, dass das Modell 0 % der Kriteriumsvarianz erklären kann, und ein $R^2 = 1$, dass es 100 % der Varianz erklären kann.

Eine Variante der Regressionsanalyse stellen *Mediatormodelle* dar (s. Abbildung 10). In Mediatormodellen wird der Einfluss des eigentlichen Prädiktors auf die abhängige Variable über eine dritte Variable vermittelt (auch: mediiert). Wenn aufgrund der Hinzunahme einer Mediatorvariable der Zusammenhang zwischen dem eigentlichen Prädiktor und dem Kriterium nicht mehr signifikant wird, spricht man von einer vollständigen Mediation. Bei einer partiellen Mediation wird der direkte Einfluss vom Prädiktor auf das Kriterium zwar geringer, bleibt aber statistisch bedeutsam.

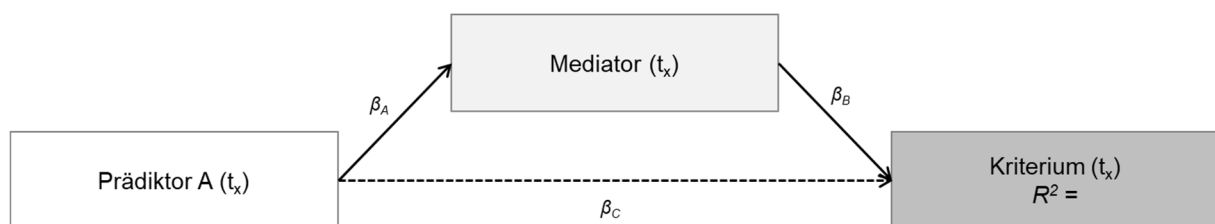


Abbildung 10. Beispielabbildung für ein Mediatormodell

Eine weitere besondere Form der regressionsbasierten Verfahren sind sogenannte *Mehrebenenmodelle*, bei denen die hierarchische Struktur von gegebenen Daten (s. Abbildung 11, Abbildung 12) statistisch berücksichtigt wird. In Mehrebenenmodellen werden die Varianzanteile einer Kriteriumsvariable, die auf unterschiedlichen Analyseebenen liegen, simultan geschätzt (Christ & Schlüter, 2012).

Als hierarchisch bezeichnet man Stichproben, die eine sogenannte „natürliche Gruppierung“ aufweisen. So bilden z. B. Lehrpersonen eines Kollegiums oder Schülerinnen und Schüler einer Klasse natürliche Gruppen (auch: Klumpen). Lehrpersonen mit denselben schulischen Rahmenbedingungen und Lernende aus einer Klasse, die von derselben (Fach-)Lehrkraft unterrichtet werden, weisen in der Regel eine größere Homogenität auf als Lehrpersonen verschiedener Schulen und Lernende aus verschiedenen Klassen. Mit anderen Worten: Aufgrund „natürlicher“ Gruppierungen sollte sich ein größerer Unterschied zwischen Lehrpersonen beziehungsweise Lernenden verschiedener Schulen beziehungsweise Klassen zeigen als zwischen Lehrpersonen beziehungsweise Lernenden einer Schule beziehungsweise Klasse.

Eine wesentliche Unterscheidung bei hierarchischen Daten sind die verschiedenen Analyseebenen (auch: Analyseeinheiten). Mit Ebene 1 wird die niedrigste Einheit bezeichnet, z. B.

Schülerinnen und Schüler. Ebene 2 bezeichnet die nächsthöhere Aggregatebene, z. B. Klassen, und Ebene 3 wäre dann erneut eine weiter übergeordnete Einheit, z. B. Schulen.

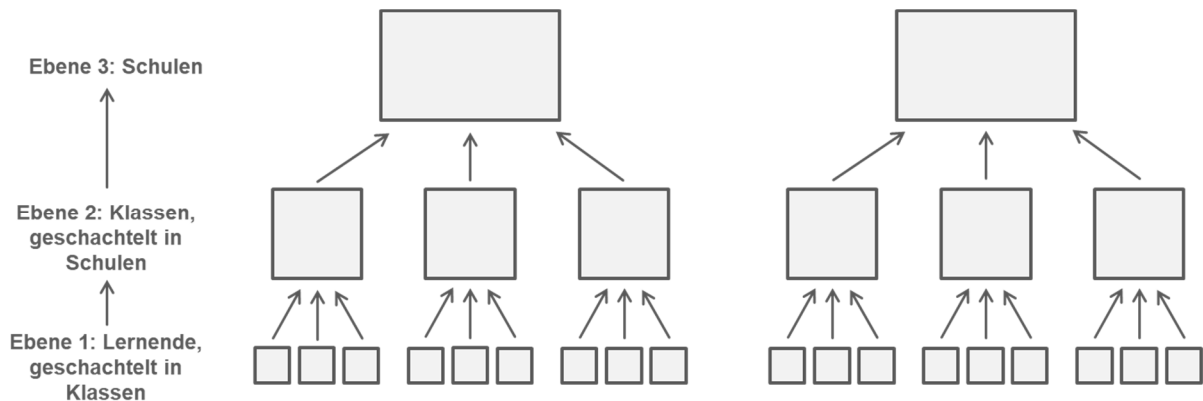


Abbildung 11. Beispiel einer hierarchischen Datenstruktur

Mehrebenenanalysen sind ein komplexes statistisches Verfahren, für dessen Anwendung verschiedene Voraussetzungen erfüllt sein müssen. Eine Bedingung ist, dass eine genügend große Anzahl von Analyseeinheiten auf allen Ebenen vorhanden ist, wobei die genaue Anforderung an die Stichprobengröße von den jeweiligen Fragestellungen und Analysemodellen abhängt (Hox, 2010; Maas & Hox, 2005). Im Rahmen von *LIQUID* wurde bei der Kohorte 2 nach Möglichkeit ein mehrebenenanalytisches Vorgehen mit zwei Analyseebenen (s. Abbildung 12) umgesetzt. Die Zugehörigkeit zur UG oder KG (*Treatmentzugehörigkeit*), was eigentlich ein Merkmal der Schule (Ebene 3) ist, wurde in den Modellen der Unterrichtsstudie als Klassenmerkmal (Ebene 2) behandelt, da für eine verlässliche Modellierung von drei Ebenen eine größere Anzahl an Klassen pro Schule notwendig gewesen wäre.

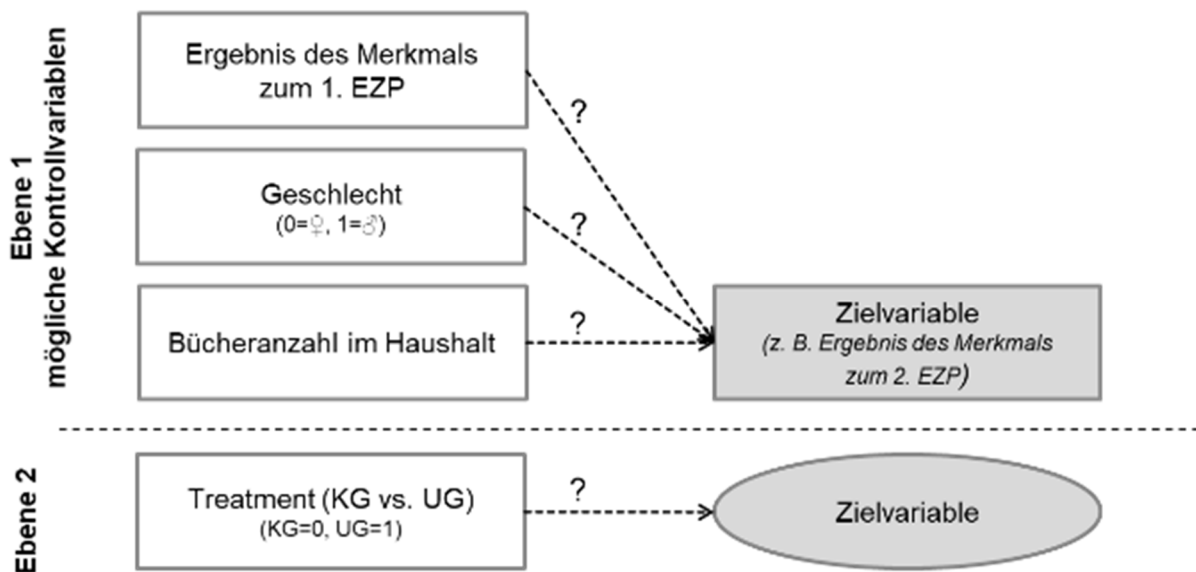


Abbildung 12. Beispiel für ein Mehrebenenmodell mit Prädiktoren auf Ebene 1 und Ebene 2

Bei der Durchführung von Mehrebenenanalysen betrachtet man unterschiedliche Kennwerte: Der Intraklassen-Korrelationskoeffizient 1 (ICC 1) gibt an, wie hoch der Anteil der Gesamtvarianz ist, der durch Gruppenzugehörigkeit erklärt werden kann (Zurbriggen, 2016), also z. B. wie viel Varianz in der Kriteriumsvariable durch die Zugehörigkeit zu einer Klasse erklärt werden kann. Trotz eines vergleichsweise geringen ICC-1-Wertes kann es sinnvoll sein, die Gruppenzugehörigkeit durch die Berechnung von Mehrebenenanalysen zu berücksichtigen, um eine Verzerrung der Ergebnisse zu vermeiden (Geiser, 2010). Der Intraklassen-Korrelationskoeffizient 2 (ICC 2) erlaubt eine Schätzung der Reliabilität auf Gruppenebene. Ein hoher ICC 2 bedeutet, dass es eine kongruente Einschätzung innerhalb einer Klasse gibt und der Klassenmittelwert einen zuverlässigen Wert darstellt (Zurbriggen, 2016). ICC-2-Werte über .70 gelten hierbei konventionell als gut, Werte zwischen .50 und .70 als akzeptabel und Werte unter .50 als marginal und nicht zufriedenstellend (Haehnel, 2011; Lüdtke, Robitzsch, Trautwein & Kunter, 2009). Darüber hinaus wird auch bei den Mehrebenenmodellen der standardisierte Beta-Koeffizient (β) und das Bestimmtheitsmaß R^2 betrachtet.

Ein Verfahren, das eingesetzt wurde, um die Lehrpersonen aus der UG der zweiten Kohorte zu gruppieren, ist die *latente Profilanalyse*. Bei der latenten Profilanalyse werden Personen auf der Basis von mehreren kontinuierlichen Variablen unterschiedlichen latenten Gruppen (auch: Profilen oder Typen) zugeordnet. Dabei wird angenommen, dass die Ausprägung einer Variable allein von der Profiltzugehörigkeit und nicht von den anderen Variablen, die zur Profilbildung herangezogen werden, abhängt. Mit anderen Worten: Die profilbildenden Variablen sollten möglichst nicht beziehungsweise möglichst gering miteinander korrelieren. Jede Gruppe beziehungsweise jedes Profil zeichnet sich durch charakteristische Mittelwerte in den profilbildenden Merkmalen aus. Hierbei sind sich die Mitglieder eines Profils sehr ähnlich, während die Distanzen zu den anderen Profilen im Mittel besonders groß sind. Zum Beispiel konnte unter den Fortbildungsteilnehmerinnen und -teilnehmern von *Vielfalt fördern* eine Gruppe (Profil) identifiziert werden, deren Mitglieder sehr offen für schulische Veränderungen und durch die Chance zur Weiterentwicklung der Schule stark zur Teilnahme an *Vielfalt fördern* motiviert sind. Die Motivationsfacette externe Erwartungsanpassung ist bei den Lehrpersonen dieses Profils hingegen sehr gering ausgeprägt, d. h., sie sind in einem geringen Maß zur Teilnahme an *Vielfalt fördern* motiviert, weil sie dies als eine dienstliche Pflicht empfinden.

Zur Beurteilung der Modellgüte werden bei der Profilanalyse mehrere Kennwerte herangezogen (Geiser, 2010). Die *Entropy* ist ein Globalmaß für die Klassifikationszuverlässigkeit. Werte nahe 1 sprechen für eine hohe Sicherheit. Die *mittlere Zuordnungswahrscheinlichkeit* gibt einen Hinweis darauf, wie sicher die Personen den einzelnen Profilklassen zugeordnet werden können. Der Wert sollte möglichst > 0.80 ausfallen. Die informationstheoretischen Maße *AIC*, *BIC*, *aBIC* beurteilen die Anpassung eines Modells an die Daten sowie die Sparsamkeit des

Modells. Hierbei erfolgt ein deskriptiver Modellvergleich von Modellen mit aufsteigender Klassenanzahl, wobei Modelle mit kleineren Werten zu bevorzugen sind. Anstelle einer Beurteilung des absoluten Modellfits wurde zusätzlich der relative Fit eines Modells für die Entscheidung der adäquatesten Lösung herangezogen. Beim *BRMT*- und *VLMR-Test* wird ein Modell im Vergleich zu einem anderen Modell mit einer Klasse weniger getestet. Ein signifikanter *p*-Wert sagt aus, dass das Modell mit einer weiteren Klasse (Profil) besser auf die Daten passt als das sparsamere Modell mit einer Klasse weniger.

4.1.3.2 Verfahren zur Auswertung der qualitativen Interviewdaten

Im Gegensatz zu den statistischen Verfahren, die zur Auswertung der Daten aus den Befragungs- und Testinstrumenten angewandt wurden, orientiert sich die Analyse der Gruppeninterviews am Vorgehen der qualitativ-strukturierten Inhaltsanalyse. Die einzelnen Schritte sind in Tabelle 1 überblicksartig dargestellt.

Tabelle 1. Ausgewählte Schritte im Rahmen der qualitativ-strukturierten Inhaltsanalyse

Schritt der Inhaltsanalyse	Umsetzung in LIQUID
<i>Erstellen des Kategoriensystems</i>	
Basisstrategie	deduktiv
Aspekte des Materials	Inhalte allgemein
<i>Unterteilung des Materials in Einheiten</i>	
Art der Einheit	Codeeinheiten
Systematik des Vorgehens	simultanes Markieren und Codieren
Anzahl beteiligter Personen	zwei Personen unabhängig voneinander
Größe der Codeeinheiten	Minimalprinzip ⁵
<i>Probekodierung mit anschließender Evaluation und Modifikation des Kategoriensystems</i>	
Anzahl beteiligter Personen	drei Personen unabhängig voneinander (inkl. wissenschaftlichem Projektmitarbeiter)
Art des Codevergleichs	Codebesprechung
<i>Hauptcodierung</i>	
Anzahl beteiligter Personen	zwei Personen unabhängig voneinander (s. o.)
Aufteilung des Materials	beide Personen kodieren gesamtes Material
Umgang mit unterschiedlichen Kategoriezuweisungen	konsensuelle Codierung unter Einbezug einer dritten Person (wiss. Projektmitarbeiter)

⁵ Mit dem Minimalprinzip ist gemeint, dass die Einheiten, für die ein Code vergeben wird, so gewählt werden, dass die Codezuweisung nachvollziehbar, die gewählte Sequenz aber so kurz wie möglich ist.

4.2 Stichprobe

Die Evaluation war multiperspektivisch angelegt. Das bedeutet, dass unterschiedliche Personengruppen befragt wurden, die in den Professionalisierungsprozess eingebunden waren. Die Stichprobe setzt sich in beiden Kohorten aus Lehrerinnen und Lehrern, Moderatorinnen und Moderatoren sowie Schulleiterinnen und Schulleitern zusammen. Bei der zweiten Kohorte wurden im Rahmen der Unterrichtsstudie zudem Schülerinnen und Schüler befragt und getestet.

Die Personen der genannten Gruppen stammen nicht nur von Schulen, die an *Vielfalt fördern* teilgenommen haben, sondern auch von Schulen, die nicht an der Fortbildung teilgenommen haben (s. Abschnitt 4.1.1).

Bei Untersuchungen, die Aussagen über Veränderungen und Entwicklungen treffen wollen und deswegen zu unterschiedlichen Zeitpunkten Erhebungen durchführen, tritt in der Regel der Fall ein, dass sich die Stichprobe über den Projektverlauf hinweg verändert, meist verkleinert. Man spricht im Zusammenhang mit dem Wegfall von Teilnehmerinnen und Teilnehmern von Panelmortalität oder Drop-out, für die/den es unterschiedliche Ursachen geben kann.

Neben der allgemeinen Fluktuation innerhalb von Schulkollegien, z. B. aufgrund von Schwangerschaften, Krankheiten oder Pensionierungen, sind für den Drop-out in der Stichprobe im Verlauf von LIQUID insbesondere zwei Gründe anzuführen:

1. Schulen der UG haben die Fortbildungsteilnahme und in Folge dessen auch die Evaluation vorzeitig beendet. Dies trifft auf vier Schulen in der ersten Kohorte und fünf Schulen in der zweiten Kohorte zu.
2. Kollegien, die *Vielfalt fördern* zunächst als gesamtes Kollegium begonnen hatten, setzten die Fortbildung im weiteren Verlauf nur noch mit einem Teilkollegium fort.
3. Ausschließlich bei der Kohorte 1 trat der Fall ein, dass Befragungen nicht durchgeführt werden konnten, weil Termine von Präsenzveranstaltungen, für die eine Erhebung angedacht war, nicht rechtzeitig an die Projektleitung der Universität Kassel weitergegeben wurden.

Die LIQUID-Stichprobe hat sich aber nicht nur durch den Wegfall ganzer Schulkollegien und einzelner Lehrpersonen verändert, sondern es kamen auch neue Lehrpersonen hinzu, z. B. in Fällen, in denen es sich um Schulen im Aufbau handelte.

Neben diesen Gründen, die überwiegend das gesamte Kollegium oder ein Teilkollegium betrafen, gab es auch Ausfälle auf der Individualebene. Das Fehlen von Daten einer Person kann unterschiedliche Gründe haben:

1. Eine Lehrperson wird im Laufe der Projektlaufzeit pensioniert und scheidet daher aus dem Schuldienst und der Teilnahme an der Evaluation aus.

2. Eine Lehrperson tritt im Verlauf der Fortbildung eine neue Stelle an einer Projektschule an, sodass Daten aus vorangehenden Erhebungszeitpunkten fehlen.
3. Eine Lehrperson ist aufgrund von Krankheit oder einem anderen Termin verhindert, an einem Fortbildungsbaustein, in dem eine Erhebung stattfindet, teilzunehmen.
4. Eine Lehrperson gibt einen fehlerhaften persönlichen Code⁶ an, der dazu dient, die Daten aus unterschiedlichen Erhebungszeitpunkten zusammenzuführen.

4.2.1 Kohorte 1

Für die erste Fortbildungskohorte wurden aus den fünf Pilotregionen Bonn, Borken, Düsseldorf, Gütersloh und Hamm 22 Schulen einbezogen, die im Schuljahr 2012/2013 die Fortbildung begannen, und sechs Schulen, welche als KG-Schulen die Professionalisierungsmaßnahme nicht durchliefen. Die Schulen stammten aus unterschiedlichen Schulformen, wie aus Abbildung 13 hervorgeht.

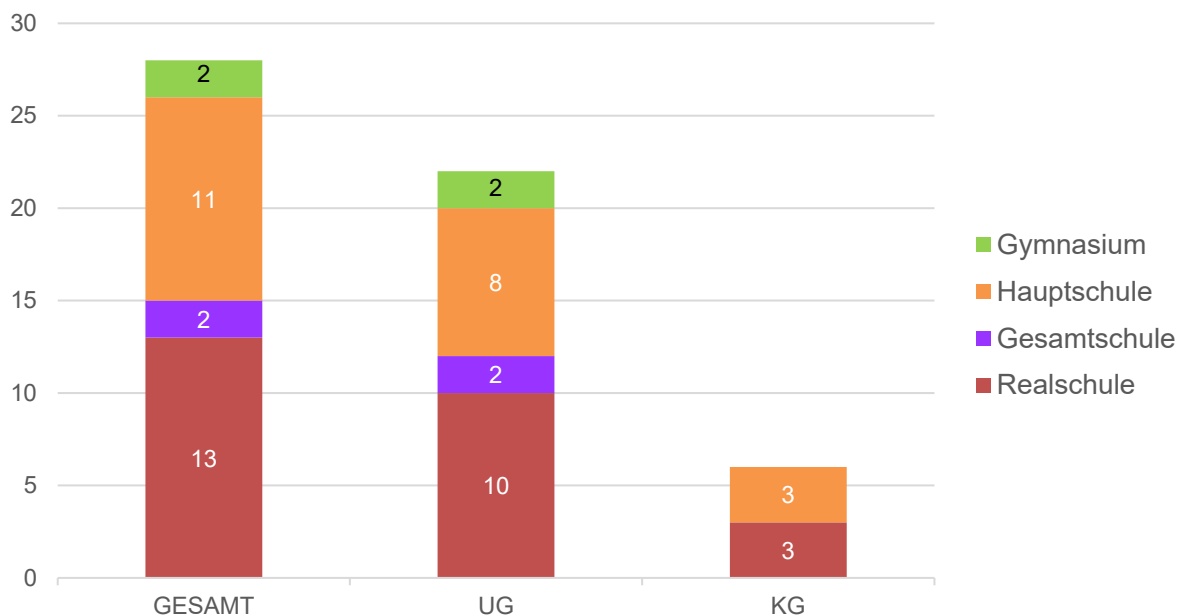


Abbildung 13. Anzahl der Schule nach Schulformen am Anfang der Projektlaufzeit von Kohorte 1

Zum ersten Erhebungszeitpunkt haben 525 Lehrpersonen aus der UG und 200 Lehrpersonen aus der KG an der Befragung teilgenommen. Wie sich der Stichprobenumfang bei den Lehrpersonen über die drei Erhebungszeitpunkte hinweg entwickelt hat, kann Tabelle 2 entnommen werden. Auch wenn bei der Posterhebung (t_3) insgesamt Daten von 521 Lehrpersonen vorliegen, bedeutet dies nicht, dass Entwicklungen für eine entsprechend große Stichprobe

⁶ Die Lehrpersonen wurden am Beginn jedes Fragebogens dazu aufgefordert, einen individuellen Code nach einem vorgegebenen Schema zu erstellen.

untersucht werden konnten, da in die längsschnittlichen Analysen nur die Daten von Lehrpersonen einbezogen werden können, die aufgrund des persönlichen Codes zusammengeführt werden können.

Tabelle 2. Stichprobenumfang bei Kohorte 1

Gruppe	t ₁	t ₂	t ₃
UG	525	192	317
KG	200	-	204

4.2.1.1 Lehrerinnen und Lehrer

In beiden Gruppen sind etwa ein Drittel der Lehrpersonen männlich und zwei Drittel der Lehrpersonen weiblich (UG: 66.5 % weiblich, 33.5 % männlich, KG: 66.8 % weiblich, 33.2 % männlich). Ein signifikanter Gruppenunterschied im Geschlechtsverhältnis besteht demnach nicht ($\chi^2[1,710] = 0.01$, n. s.).

Auch in Bezug auf das Alter der Probanden liegt zwischen beiden Gruppen kein statistisch bedeutsamer Unterschied vor ($t[682] = -1.73$, n. s.). In der UG sind die Teilnehmer durchschnittlich 44.7 Jahre alt ($SD = 11.0$) und in der KG beträgt das Durchschnittsalter 45.8 Jahre ($SD = 10.9$).

Der Abschluss des Vorbereitungsdienstes liegt bei den Lehrpersonen der UG und KG jedoch unterschiedlich weit zurück. Die Lehrpersonen der UG sind mit einer durchschnittlichen Arbeitszeit von 13.4 Jahren ($SD = 11.1$) etwa zwei Jahre kürzer im Beruf tätig als Lehrpersonen der KG ($M = 15.4$, $SD = 11.8$). Es handelt sich um einen kleinen Effekt ($d = 0.18$).

4.2.1.2 Schulleiterinnen und Schulleiter

Von den Schulleiterinnen und Schulleitern haben 20 Personen aus der UG und fünf Personen aus der KG den Eingangsfragebogen bearbeitet. Dies führt dazu, dass zu zwei UG-Schulen und einer KG-Schule keine Hintergrundinformationen zum schulischen Kontext vorliegen.

In beiden Gruppen sind die Geschlechter im gleichen Verhältnis (60.0 % weiblich, 40.0 % männlich) repräsentiert. Auch im Durchschnittsalter unterscheiden sich die Schulleiterinnen/Schulleiter der beiden Gruppen nicht signifikant voneinander (UG: $M = 54.0$, $SD = 6.9$; KG: $M = 55.8$, $SD = 4.6$; $U = -.27$, n. s.). Jedoch fällt auf, dass die Spannweite zwischen der jüngsten Schulleiterin/dem jüngsten Schulleiter und der ältesten Schulleiterin/dem ältesten Schulleiter in der UG mit 22.0 Jahren ($Min. = 41.0$, $Max. = 63.0$) etwa doppelt so groß ist wie die Spannweite in der KG ($Min. = 50.0$, $Max. = 61.0$). Eine Schulleiterfunktion bekleiden die Schulleiterinnen/Schulleiter der UG seit durchschnittlich 7.7 Jahren ($SD = 5.8$) und die Schulleiterinnen/Schulleiter der KG seit 11.4 Jahren ($SD = 8.0$). Auch dieser Unterschied ist statistisch nicht bedeutsam ($U = -1.12$, n. s.). Insgesamt kann bei den demografischen Merkmalen

der Schulleiterinnen und Schulleitern eine Vergleichbarkeit zwischen UG und KG angenommen werden.

4.2.1.3 Moderatorinnen und Moderatoren

Für die Durchführung von *Vielfalt fördern* an den 22 UG-Schulen waren insgesamt 28 Moderatorinnen und Moderatoren verantwortlich, davon waren 67.9 % weiblich und 32.1 % männlich. Alle Moderatorinnen und Moderatoren unterrichteten neben ihrer Tätigkeit in der Fort- und Weiterbildung auch noch als Lehrkräfte an allgemeinbildenden Schulen. Die Mehrheit der Moderatoren (75 %) war bereits vor ihrer Arbeit im Rahmen der Fortbildung *Vielfalt fördern* in der Lehreraus- und/oder -fortbildung beschäftigt und besaß zum Zeitpunkt der Eingangserhebung eine durchschnittliche Moderations-/Fortbildungserfahrung von $M = 7.5$ Jahre ($SD = 6.4$). Im Mittel waren die Moderatorinnen und Moderatoren 45.4 Jahre ($SD = 10.0$) alt, wobei ein Altersunterschied von 34.0 Jahren zwischen der jüngsten Moderatorin/dem jüngsten Moderator (29.0 Jahre) und der ältesten Moderatorin/dem ältesten Moderator (63.0 Jahre) lag.

4.2.2 Kohorte 2

Die zweite Kohorte hat im Schuljahr 2013/2014 mit 21 Schulen in der UG und 16 Schulen in der KG die Evaluation begonnen. Es handelt sich um Realschulen, Gesamtschulen, Sekundarschulen und eine Verbundschule (s. Abbildung 14) aus den Regionen Bielefeld, Gelsenkirchen, Kleve, dem Hochsauerlandkreis und dem Oberbergischen Kreis.

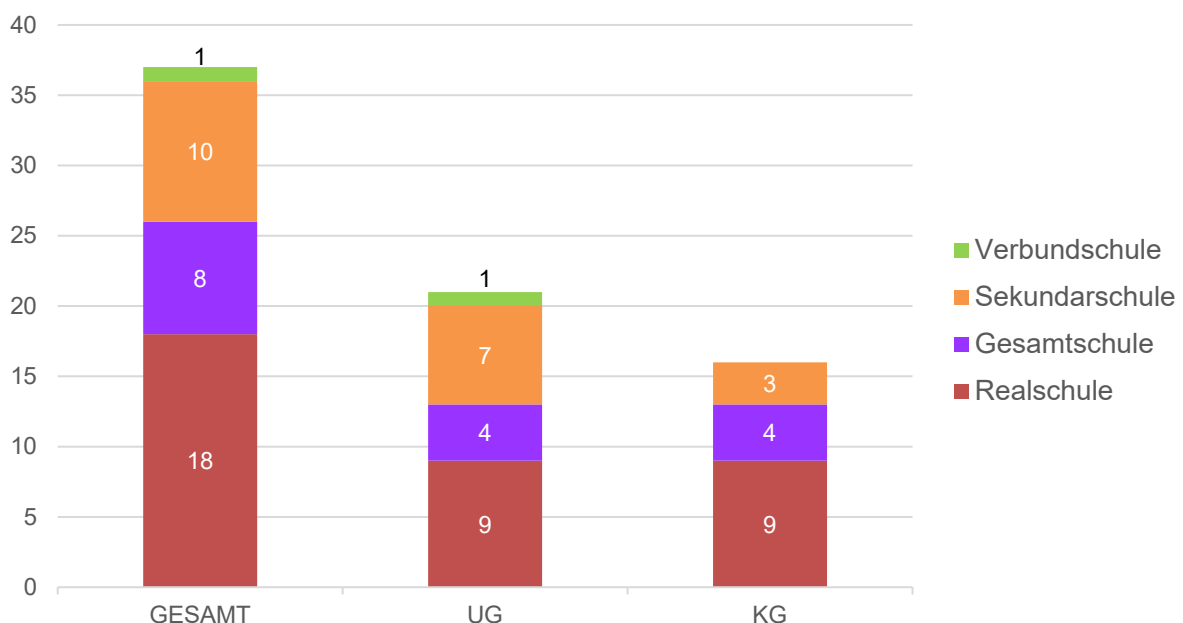


Abbildung 14. Anzahl der Schule nach Schulformen am Anfang der Projektlaufzeit von Kohorte 2

4.2.2.1 Fortbildungsstudie: Lehrerinnen und Lehrer

Obwohl sich die Schulen der UG und die Schulen der KG im Mittel nicht bedeutsam in der Größe des Schulkollegiums unterscheiden ($M_{UG} = 42.9$, $SD_{UG} = 31.5$; $M_{KG} = 53.2$,

$SD_{KG} = 30.0$) und weniger KG-Schulen in der Stichprobe vertreten sind, nehmen mehr Lehrpersonen, die zur KG gehören, an den Erhebungen teil (s. Tabelle 3). Dies könnte unter anderem darauf zurückzuführen sein, dass die Erhebungen in der KG von einer Mitarbeiterin/einem Mitarbeiter der Universität Kassel durchgeführt wurden, was mit einer höheren Verbindlichkeit für die teilnehmenden Lehrpersonen verbunden gewesen sein könnte.

Zusätzlich fällt auch bei der Entwicklung des Stichprobenumfangs ein deutlicherer Rückgang in der Personenanzahl bei den Lehrpersonen der UG im Vergleich mit den Lehrpersonen der KG auf (s. Tabelle 3).

Tabelle 3. Stichprobenumfang zu den verschiedenen Erhebungszeitpunkten bei Kohorte 2

Gruppe	t ₁	t ₂	t ₃	t ₄	t ₅	t ₆	t ₇
UG	572	504	438	370	375	335	311
KG	661		-	530	613	-	613

In der UG der zweiten Kohorte sind etwa zwei Drittel der Lehrpersonen (66.1 %) weiblich und ein Drittel der Lehrpersonen (33.9 %) männlich. In der Kontrollgruppe beträgt der Anteil weiblicher Lehrpersonen 60.5 % und 39.5 % der Lehrpersonen sind männlich. Das Ergebnis des Chi-Quadrat-Tests ($\chi^2[1,1214] = 4.17, p < .05$) zeigt einen statistisch bedeutsamen Unterschied in der Geschlechterverteilung. Mit $d = 0.12$ ist der Effekt jedoch klein, d. h., er besitzt kaum eine praktische Bedeutsamkeit und kann vernachlässigt werden.

Auch beim durchschnittlichen Alter ($d = 0.29$) und der Berufserfahrung ($d = 0.34$) zeigen sich signifikante Unterschiede zwischen UG und KG mit schwachem Effekt. Ein t-Test für unabhängige Stichproben zeigt, dass die Lehrpersonen, die nicht an der Fortbildung *Vielfalt fördern* teilnehmen, im Durchschnitt früher geboren und folglich älter sind als die Lehrpersonen, die an der Fortbildung teilnehmen ($t[1199] = 4.99, p < .001$). In der UG sind die Lehrpersonen im Durchschnitt 43.7 Jahre⁷ alt ($SD = 11.4$). In der KG beträgt das Durchschnittsalter der Lehrpersonen 47.0 Jahre ($SD = 11.4$).

Eine Differenz von ca. vier Jahren zeigt sich auch bei der Berufserfahrung der Lehrpersonen. Die Lehrpersonen der UG weisen im Mittel eine um ca. vier Jahre kürzere Berufserfahrung auf ($M = 11.7$ Jahre, $SD = 10.5$) als die Lehrpersonen der KG ($M = 15.5, SD = 11.6$).

Insgesamt zeigen sich bei den Lehrpersonen der UG und KG demnach marginale Unterschiede in demografischen Merkmalen, die aufgrund der geringen Effektstärke jedoch nicht ins Gewicht fallen.

⁷ Zur Ermittlung des Durchschnittsalters wurde das Jahr 2014 als Erhebungsjahr zugrunde gelegt.

4.2.2.2 Fortbildungsstudie: Schulleiterinnen und Schulleiter

Von den Schulleiterinnen und Schulleitern der UG haben 18 Personen an der Eingangserhebung und 14 Personen an der abschließenden Erhebung am Ende der Fortbildung (t_7) teilgenommen. Für die KG liegen zu beiden Erhebungszeitpunkten die Daten von 15 Schulleiterinnen/Schulleitern vor.

In der Untersuchungsgruppe sind 27.8 % der befragten Personen weiblich und 72.2 % männlich. In der Kontrollgruppe sind Schulleiterinnen mit einem Anteil von 42.9 % vertreten und Schulleiter mit 57.1 %. Das Geschlechterverhältnis ist somit etwas ausgeglichener als in der UG. Der Unterschied zwischen der UG und KG erweist sich aus statistischer Sicht als nicht signifikant ($\chi^2[1,32] = 0.79$, n. s.).

Die Schulleiterinnen und Schulleiter der Schulen, die an *Vielfalt fördern* teilnahmen, wurden zwischen 1952 und 1970 geboren und waren demnach bei der Eingangserhebung im Mittel etwa 55.6 Jahre ($SD = 5.1$) alt. Das Geburtsjahr der Schulleiterinnen und Schulleiter aus der KG liegt zwischen 1950 und 1973, womit sich bei der Eingangserhebung 2014 ein Durchschnittsalter von 53.4 Jahren ($SD = 7.1$) ergibt. Der deskriptive Unterschied von ca. zwei Jahren im mittleren Alter ist statistisch nicht bedeutsam, d. h., im Alter unterscheiden sich die Schulleitungen der beiden Gruppen nicht signifikant voneinander ($U = 104.5$, $z = -.82$, n. s.).

Die Tätigkeitsdauer in der Schulleitung ist zwischen der UG und KG vergleichbar ($U = 116.5$, $z = -0.36$, n. s.) und beträgt im Durchschnitt 7.6 Jahre ($M_{UG} = 7.5$, $SD_{UG} = 4.7$, $M_{KG} = 7.8$, $SD_{KG} = 7.2$).

4.2.2.3 Fortbildungsstudie: Moderatorinnen und Moderatoren

Die Moderatorenstichprobe setzt sich zu Beginn des Projekts aus $n = 29$ Personen zusammen. Im zeitlichen Verlauf kam es jedoch auch in den Moderatorenteams zu personellen Veränderungen, sodass die Anzahl der Befragungsteilnehmerinnen und -teilnehmer zu den einzelnen Erhebungszeitpunkten variiert.

Von den 29 Moderatorinnen und Moderatoren, die zu Beginn der Fortbildung befragt wurden, sind zwei Drittel (66.0 %) weiblich und ein Drittel (34.0 %) ist männlich. Im Durchschnitt sind die Moderatorinnen und Moderatoren, die *Vielfalt fördern* an den Schulen der zweiten Kohorte geleitet haben, 48.8 Jahre ($SD = 10.3$) alt. Wie auch bei der ersten Fortbildungskohorte besteht ein hoher Altersunterschied (33.0 Jahre) zwischen der jüngsten Person (29.0 Jahre) und der ältesten Person (62.0 Jahre). Der Abschluss des Vorbereitungsdienstes liegt im Mittel 17.3 Jahre ($SD = 10.6$) zurück. Erwartungsgemäß – aufgrund des genannten Unterschieds im Alter – variieren auch die Berufsjahre von $Min = 2.0$ Jahren bis $Max = 38.0$ Jahren.

4.2.2.4 Unterrichtsstudie: Schülerinnen und Schüler

Als vierte Personengruppe wurden bei der zweiten Kohorte auch Schülerinnen und Schüler einbezogen. Die Lernenden einer Klasse, die beim ersten Erhebungszeitpunkt am Ende des 5. Jahrgangs waren und beim zweiten Erhebungszeitpunkt am Anfang des 7. Jahrgangs, wurden für die Unterrichtsstudie zufällig in zwei Gruppen aufgeteilt. Eine Schülerin/ein Schüler nahm entweder an der Erhebung im Fach Mathematik oder an der Erhebung im Fach Deutsch teil. Für die Analysen konnten nicht die gesamten Daten der Schülerstichprobe genutzt werden, wofür im Wesentlichen die folgenden drei Gründe zu nennen sind:

1. Aus *statistischen Erwägungen* wurden nur Klassenhälften einbezogen, bei denen Daten von mindestens sieben Lernenden im Längsschnitt vorliegen.
2. Aus *inhaltlichen Gründen* ist es nicht plausibel, alle Schulen der UG einzubeziehen, weil sich diese zum Zeitpunkt der abschließenden Schülererhebung in unterschiedlichen Stadien des Fortbildungsprozesses befanden. Da insbesondere von dem dritten und vierten Fortbildungsmodul positive Effekte auf den Unterricht angenommen werden, wurden nur Klassen von Schulen berücksichtigt, deren Lehrende die Fortbildung vollständig oder fast vollständig abgeschlossen hatten. Nicht einbezogen wurden hingegen Klassen der UG, deren Lehrpersonen die Fortbildung vor dem zweiten Messzeitpunkt der Unterrichtsstudie vorzeitig beendet oder das vierte Modul der Fortbildung *Vielfalt fördern* noch nicht begonnen hatten.
3. Ebenfalls aus *inhaltlichen Gründen* wurden Klassen ausgeschlossen, bei denen zwischen der ersten und zweiten Erhebung ein Fachlehrerwechsel stattfand, um diesen als Grund für mögliche Leistungsunterschiede auszuschließen beziehungsweise etwaige Verzerrungen zu vermeiden.

Aus Tabelle 4 geht die Schülerstichprobe für die beiden Fächer Mathematik und Deutsch hervor. Im Fach Mathematik umfasst die Ebene 1 insgesamt 558 Lernende aus 58 Klassen (Ebene 2) und 26 Schulen (Ebene 3). Im Fach Deutsch verteilen sich die 478 Lernenden (Ebene 1) auf 50 Klassen (Ebene 2) aus ebenfalls 26 Schulen (Ebene 3).

Tabelle 4: Stichprobe der Unterrichtsstudie bei Kohorte 2

	Mathematik			Deutsch		
	UG	KG	Gesamt	UG	KG	Gesamt
Schulen	$n = 10$	$n = 16$	$n = 26$	$n = 11$	$n = 15$	$n = 26$
Klassen	$n = 27$	$n = 31$	$n = 58$	$n = 23$	$n = 27$	$n = 50$
Schülerzahl	$n = 263$	$n = 295$	$n = 558$	$n = 226$	$n = 252$	$n = 478$
Alter K_2 ust_1	$M = 11.6$ $(SD = 0.6)$	$M = 11.4$ $(SD = 0.6)$	$M = 11.5$ $(SD = 0.6)$	$M = 11.5$ $(SD = 0.5)$	$M = 11.4$ $(SD = 0.5)$	$M = 11.4$ $(SD = 0.5)$

Für das Fach Mathematik zeigt sich, dass die Schülerinnen und Schüler der UG im Mittel signifikant älter als die Schülerinnen und Schüler der KG sind ($t_1: t[548] = -2.42, p < .05$). Der Effekt ist jedoch nur schwach ($t_1: d = -0.21$). Aufgrund dieses Ergebnisses wird in allen Modellen für das Fach Mathematik zunächst in einem separaten Schritt überprüft, ob das Alter der Schülerinnen und Schüler die Entwicklung der jeweiligen Zielvariable (z. B. Interesse, Leistung) beeinflusst. In dem Fall, dass für das Alter ein signifikanter Effekt ermittelt wurde, wird in den eigentlichen Modellen zur Vorhersage des Fortbildungseffekts das Alter der Schülerinnen und Schüler kontrolliert. Im Fach Deutsch unterscheiden sich die Schülerinnen und Schüler der UG und der KG nicht bedeutsam hinsichtlich des Alters ($t_1: t[463] = -0.33, n. s.$).

In der Gesamtstichprobe für das Fach Mathematik ist der Anteil an Schülerinnen mit $n = 250$ (45.1 %) im Vergleich zum Anteil der Schüler mit $n = 304$ (54.9 %) geringer. Zwischen den beiden Gruppen, d. h. zwischen der UG und der KG, ist das Geschlechterverhältnis allerdings vergleichbar ($\chi^2[1,554] = 0.77, n. s.$). Der Anteil der Mädchen ist mit 47.2 % ($n = 225$) auch in der Stichprobe für das Fach Deutsch etwas geringer als der Jungenanteil (52.8 %, $n = 252$). Die UG und KG unterscheiden sich hier nicht signifikant voneinander ($\chi^2[1,477] = 1.72, n. s.$).

4.2.2.5 Vertiefungsstudie

Die Stichprobe der Vertiefungsstudie, die ausschließlich bei UG-Schulen der zweiten Kohorte durchgeführt wurde, sollte möglichst gut die Stichprobenzusammensetzung der quantitativen Hauptstudie, d. h. der schriftlich befragten Fortbildungsteilnehmerinnen und -teilnehmer, widerspiegeln. Daher wurden Schulen gezielt zur Teilnahme an der Vertiefungsstudie eingeladen, jedoch war es diesen freigestellt, sich für oder gegen eine Teilnahme zu entscheiden, sodass die Stichprobenzusammensetzung nur eingeschränkt beeinflusst werden konnte. Sieben Schulen aus vier Regionen bildeten am Ende die Stichprobe der Vertiefungsstudie, d. h., eine Region der Hauptstudie ist nicht vertreten. Aus Abbildung 15⁸ geht hervor, dass die verschiedenen Schulformen in der Haupt- und in der Vertiefungsstudie mit einem ähnlichen Anteil repräsentiert sind. Die in der Hauptstudie vertretene Verbundschule wurde nicht einbezogen, da eine umfassende Anonymisierung der Daten im Rahmen der Vertiefungsstudie nicht hätte realisiert werden können.

⁸ Die Daten der sieben Schulen aus der Vertiefungsstudie sind in der Abbildung 15 auch in der Angabe für die Gesamtstichprobe der UG enthalten.

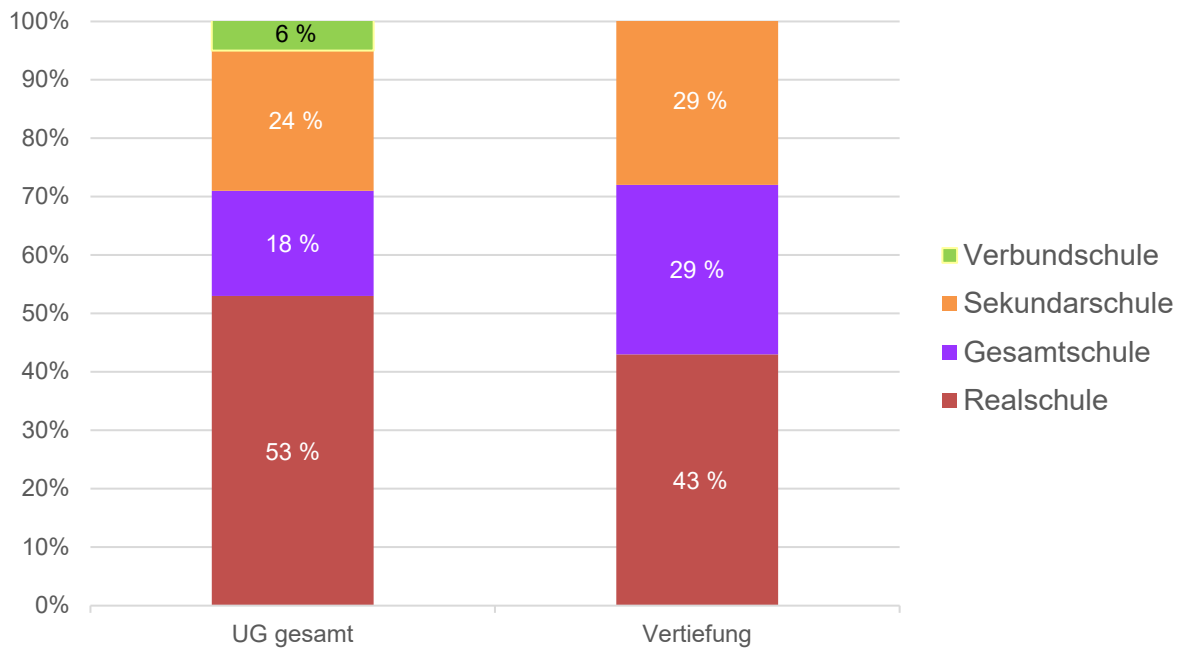


Abbildung 15. Prozentualer Anteil der Schulformen in der Haupt- und Vertiefungsstudie bei Kohorte 2

5 Zum Vergleich von Untersuchungsgruppe und Kontrollgruppe: Zusammenfassung zentraler Befunde

Es werden zunächst die zentralen Befunde aus dem Vergleich der Untersuchungsgruppe mit der Kontrollgruppe von Kohorte 1 und Kohorte 2 zusammengefasst. Da der Globalvergleich kaum bedeutsame Unterschiede auf der Ebene der Lehrpersonen und keine signifikanten Unterschiede auf der Ebene des Unterrichts und bei den Schülerinnen und Schülern zeigt, wird dieser Abschnitt bewusst kurz gehalten.⁹ Motiviert durch die Ergebnisse des Globalvergleichs wurde die Fortbildungsgruppe der zweiten Kohorte von *Vielfalt fördern* detaillierter in den Blick genommen. Die Ergebnisse dieser Analysen werden in diesem Bericht im Anschluss an Abschnitt 5 präsentiert (vgl. Abschnitte 6–11, Abschnitt 13).

5.1 Kohorte 1: Voraussetzungen und Wahrnehmungen der teilnehmenden Lehrpersonen und ein Vergleich der Untersuchungsgruppe mit der Kontrollgruppe

Bei der ersten Kohorte wurde ausschließlich das erste Fortbildungsmodul, d. h. der Aspekt der Teambildung und Zusammenarbeit von Lehrpersonen, untersucht (s. Abschnitt 3.1.3). Die wissenschaftliche Begleitung erfolgte über den Zeitraum von 2012 bis 2014.

Als zentrales Ergebnis kann bei der ersten Fortbildungskohorte konstatiert werden, dass die Lehrpersonen der UG die Fortbildung mit günstigen motivationalen Voraussetzungen begannen. So sind die Lehrpersonen der ersten Kohorte beispielsweise stark zur Teilnahme an *Vielfalt fördern* motiviert, weil sie darin einen Mehrwert für ihre gesamte Schule sehen, weil sie selbst ein tieferes Verständnis für das Thema individuelle Förderung aufbauen und weil sie einen intensiveren Austausch mit Kolleginnen und Kollegen etablieren möchten. Im Vergleich zu diesen drei eher intrinsischen Beweggründen resultiert die Teilnahmemotivation der Lehrpersonen weniger aus einem Pflichtgefühl und der Anpassung an die Entscheidung des Kollegiums.

Diese am Anfang von *Vielfalt fördern* erfasste Teilnahmemotivation konnte in Regressionsmodellen wiederum Unterschiede im selbstberichteten Lernertrag am Ende von Modul 1 erklären. Der selbstberichtete Lernertrag drückt aus, inwiefern eine Lehrperson den Eindruck hat, in dem ersten Modul Neues gelernt zu haben, hilfreiche Impulse für die eigene schulische Praxis erhalten zu haben und sich am Ende des Moduls kompetenter zu fühlen als zu Beginn des

⁹ Genauere Informationen zu Ergebnissen können bei der Projektleitung in Kassel angefragt werden.

Moduls. Es zeigte sich beispielsweise, dass die Motivationsfacetten externale Erwartungsanpassung und soziale Interaktion, welche zu Beginn der Fortbildung erfasst wurden, zusammen 10 % der Unterschiede im selbstberichteten Lernertrag am Ende der Fortbildung erklären. Eine hohe externale Erwartungsanpassung, d. h. eine hohe Motivation, die sich daraus speist, dass eine Lehrperson die Teilnahme als Pflicht empfindet, führt zu einem geringeren selbstberichteten Lernertrag. Ist die Teilnahme an *Vielfalt fördern* hingegen stark dadurch motiviert, einen intensiveren Kontakt zu anderen Kolleginnen und Kollegen aufzubauen (soziale Interaktion), wirkt sich dies positiv auf den selbstberichteten Lernertrag aus.

Insgesamt fiel der selbstberichtete Lernertrag der Lehrpersonen jedoch eher gering aus. Die Relevanz und Nützlichkeit der Inhalte des ersten Moduls wurden hingegen eher hoch bewertet und der zwischen den Fortbildungssitzungen stattfindende Transfer hat den Lehrpersonen durchaus Freude bereitet. Wie weitere einfache Regressionsanalysen zeigen, erklären sowohl die wahrgenommene Relevanz der Fortbildungsinhalte als auch die Bewertung der kooperativen Erprobung der Fortbildungsinhalte zwischen den Fortbildungssitzungen – erfasst in der Prozesserhebung nach der Hälfte des Moduls – Unterschiede im selbstberichteten Lernertrag. Lehrpersonen, die den Fortbildungsinhalten einen höheren Nutzen beimessen, bewerten auch den eigenen Lernertrag am Ende des Moduls höher. Außerdem gilt: Je positiver die Lehrpersonen die Umsetzung der Inhalte von Modul 1 zwischen den einzelnen Bausteinen erlebt haben, desto stärker haben sie am Ende des Moduls 1 den Eindruck, während des Moduls hilfreiche Anregungen und Impulse erhalten zu haben.

Ein überraschendes Ergebnis der ersten Kohorte ist, dass nicht nur zwischen den Schulen beträchtliche Unterschiede in der Wahrnehmung und Bewertung des ersten Moduls von *Vielfalt fördern* festgestellt werden konnten. Bei einer grafischen Betrachtung mittels Boxplots zeigte sich, dass auch innerhalb der Schulen die Einschätzung der Lehrpersonen stark variierte und nahezu das gesamte mögliche Antwortspektrum ausgenutzt wurde. So geben beispielsweise in mehreren Schulen einzelne Lehrpersonen eines Kollegiums am Ende von Modul 1 an, dass sie gar nicht oder größtenteils nicht mit dem Fortbildungsangebot zufrieden sind, was den beiden niedrigsten Antwortoptionen entspricht, während andere Lehrpersonen der gleichen Schule das Modul voll und ganz zufrieden (höchste Antwortkategorie) abschließen.

Darüber hinaus konnte auf einer von insgesamt sechs erfassten Dimensionen der Kooperation (kooperationsfreundliche Rahmenbedingungen, partizipatives Schulklima, geteilte Werte, Deprivatisierung, fachspezifische Zusammenarbeit, Fokus auf das Lernen der Schülerinnen und Schüler) varianzanalytisch ein statistisch bedeutsamer positiver Effekt zugunsten der Fortbildungsgruppe gegenüber der Kontrollgruppe ermittelt werden. Positiv entwickelt hat sich nach dem Selbstbericht der Lehrpersonen, die an *Vielfalt fördern* teilgenommen haben, deren De-

privatisierungspraxis. Ein wesentlicher Aspekt von Deprivatisierung ist die Durchführung kollegialer Unterrichtshospitationen. Im Vergleich mit Lehrpersonen, die zu der Kontrollgruppe gehören, betrachten es die fortgebildeten Lehrpersonen der ersten Kohorte am Ende des ersten Moduls als selbstverständlicher, dass sie von Kolleginnen und/oder Kollegen im eigenen Unterricht besucht werden, und hospitieren nach eigenen Angaben auch häufiger im Unterricht von Kolleginnen und Kollegen. Die Frage, ob es sich dabei um einen nachhaltigen Effekt handelt und ob sich die Kooperation innerhalb von Kollegien durch die gemeinsame Teilnahme an den weiteren Fortbildungsmodulen weiter bedeutsam verändert, kann mit den Daten aus der ersten Evaluationskohorte nicht beantwortet werden. Die Kooperation innerhalb von Kollegien wurde daher auch für die zweite Kohorte untersucht.

5.2 Kohorte 2: Voraussetzungen und Wahrnehmungen der teilnehmenden Lehrpersonen und ein Vergleich der Untersuchungsgruppe mit der Kontrollgruppe

Die Evaluation der zweiten Fortbildungskohorte erfolgte von 2013 bis 2018. Die Evaluation legte einen Schwerpunkt auf das zweite Fortbildungsmodul zum Thema *Diagnostik* sowie das dritte und vierte Fortbildungsmodul zum Thema *Lernen und Lehren*. Darüber hinaus sollte aber auch die Entwicklung der Kooperation innerhalb von Kollegien nach Abschluss des ersten Moduls untersucht werden.

5.2.1 Voraussetzungen der Lehrpersonen

Mehrebenenanalysen ergaben, dass die Lernzielorientierung und die individuelle Innovationsbereitschaft, denen aus theoretischer Sicht eine wichtige Funktion für die Nutzung von Fortbildungsangeboten zugeschrieben werden kann, unter den Lehrpersonen der UG zu Beginn der Fortbildung höher ausgeprägt waren als unter den Lehrpersonen der KG-Schulen. Lehrpersonen, die eine hohe Lernzielorientierung besitzen, ist es wichtig, durch das Unterrichten die Prozesse im Unterricht und komplexe Unterrichtssituationen besser zu verstehen. Eine hohe Innovationsbereitschaft meint, dass eine Lehrperson dazu bereit ist, bei Veränderungen, die die Schule und den Unterricht betreffen, aktiv mitzuarbeiten und diese zu unterstützen.

Bei den Teilnehmerinnen und Teilnehmern der Fortbildung zeigt sich ein bedeutsamer Zusammenhang zwischen der individuellen Innovationsbereitschaft und der Fortbildungsteilnahmemotivation. Die Fortbildungsteilnahmemotivation wurde über mehrere Facetten am Anfang der Fortbildung erfasst. Es wurde beispielsweise untersucht, wie stark die Lehrpersonen zur Teilnahme an *Vielfalt fördern* motiviert sind, weil sie die Schule als Ganzes weiterentwickeln wollen (*Schulentwicklungsorientierung*), weil sie das Bedürfnis nach einem intensiveren Kontakt mit Kolleginnen und Kollegen haben (*soziale Interaktion*) und weil sie dem mehrheitlichen Ent-

schluss des Kollegiums folgen (*externale Erwartungsanpassung*). Während die Schulentwicklungsorientierung eher eine intrinsische Motivationsqualität darstellt, handelt es sich bei der externalen Erwartungsanpassung eher um eine extrinsische Motivationsfacette. Die individuelle Innovationsbereitschaft hängt stark positiv mit der Schulentwicklungsorientierung zusammen, hingegen aber negativ mit der externalen Erwartungsanpassung. Das bedeutet: Je eher Lehrpersonen bereit sind, an innovativen schulischen Neuerungen mitzuarbeiten, desto stärker sind sie auch deshalb zur Teilnahme an *Vielfalt fördern* motiviert, weil die Schule davon profitieren kann, und desto geringer nehmen sie aus Pflichtgefühl an der Fortbildung teil. Da es sich um korrelative Zusammenhänge handelt, gilt aber auch: Je höher die Schulentwicklung ausgeprägt ist, desto höher ist auch die individuelle Innovationsbereitschaft ausgeprägt, beziehungsweise je höher die externale Erwartungsanpassung, desto geringer ist die individuelle Innovationsbereitschaft einer Lehrperson.

Wie schon bei Kohorte 1 kann die motivationale Ausgangslage der an der Fortbildung teilnehmenden Lehrpersonen der Kohorte 2 im Mittel eher günstig bewertet werden: Die Lehrpersonen erleben eine eher geringe Verpflichtung, an der Fortbildung teilzunehmen. Dagegen ist das Bedürfnis, mit der Fortbildung zur Unterrichtsentwicklung beizutragen und dem eigenen (Lern-)Interesse nachzugehen, höher ausgeprägt.

Regressionsanalysen, bei denen die hierarchische Datenstruktur berücksichtigt wurde (s. Abschnitt 4.1.3.1), ergaben, dass insbesondere die externale Erwartungsanpassung am Anfang der Fortbildung einen bedeutsamen Prädiktor für die Motivation der Lehrpersonen während der Module zum Lernen und Lehren ($K_2 t_6$) darstellt. Lehrpersonen, deren Fortbildungsteilnahme zu Beginn von *Vielfalt fördern* unter anderem stark von dem mehrheitlichen Entscheid des Kollegiums motiviert ist, berichten auch am Anfang von Modul 4, dass sie aufgrund eines Pflichtgefühls in den Bausteinen mitarbeiten. Außerdem zeigt sich unter anderem, dass sich eine hohe externale Erwartungsanpassung zu Beginn von Modul 1 negativ auf die Schulentwicklungsorientierung am Anfang von Modul 4, d. h. ca. 1.5 Jahre später, auswirkt.

Die externale Erwartungsanpassung zu t_1 erwies sich in den Regressionsanalysen zudem als ein negativer Prädiktor für den selbstberichteten Lernertrag der Lehrpersonen am Ende des zweiten Moduls (t_4) und am Ende des vierten Moduls (t_7). Je stärker eine Lehrperson durch ihr Pflichtgefühl zur Teilnahme motiviert ist, desto weniger hat die Lehrperson den Eindruck, durch die Fortbildung etwas gelernt zu haben und kompetenter geworden zu sein.

5.2.2 Wahrnehmung der Fortbildung Vielfalt fördern

Betrachtet man die Fortbildungswahrnehmung, fällt zunächst auf, dass die Lehrpersonen den Nutzen der Fortbildungsinhalte der Module zur Diagnostik sowie zum Lernen und Lehren in der Tendenz eher ungünstig einschätzen, d. h. Urteile vergeben, die im Mittel unterhalb des

theoretischen Mittelwerts liegen. Dieser Befund ist insofern kritisch, als das Erleben eines individuellen Nutzens und/oder Werts als wichtiger Einflussfaktor auf das Lernen gilt. Auffällig ist aber auch hier, dass die wahrgenommene Nützlichkeit der Fortbildungsinhalte von Modul 2 innerhalb von Kollegien stark streut und das Antwortspektrum von 1 „trifft gar nicht zu“ bis 6 „trifft voll und ganz zu“ vollständig ausgenutzt wird. Mittels weiterführender Regressionsanalysen, in denen berücksichtigt wurde, dass die Lehrpersonen in Schulen gruppiert sind, konnte nachgewiesen werden, dass eine höhere Bewertung der Nützlichkeit der Fortbildungsinhalte eines Moduls zu einem höheren selbstberichteten Lernertrag der Teilnehmerinnen und Teilnehmer am Ende dieses Moduls beiträgt.

Positiv zu beurteilen ist, dass die zeitliche Belastung durch die Fortbildungsteilnahme zu beiden Erhebungszeitpunkten, d. h. während Modul 2 (t_3) und am Anfang von Modul 4 (t_6), nahe am theoretischen Mittel liegt. Die Lehrpersonen empfinden den Zeit- und Arbeitsaufwand im Mittel demnach weder als besonders hoch noch als besonders niedrig und es ist ihnen durchaus möglich, sich die Zeit für die Fortbildungsteilnahme einzuräumen. Obwohl *Vielfalt fördern* einen beachtlichen zeitlichen Umfang hat (s. Abschnitt 3.1.2), fühlen sich die Lehrpersonen im Schnitt dadurch nicht besonders zeitlich belastet. Auch hier fällt jedoch auf, dass die zeitliche Belastung aufgrund der Fortbildungsteilnahme nicht nur zwischen Schulen variiert, sondern sich das Erleben zum Teil auch deutlich zwischen den Lehrpersonen eines Schulkollegiums unterscheidet. Während sich in mehreren Schulen zeigt, dass einzelne Lehrpersonen die Arbeitsbelastung aufgrund von *Vielfalt fördern* sehr hoch einschätzen und es schwierig finden, sich die Zeit für die Teilnahme zu reservieren (Antwortkategorie 6, d. h. höchster Wert der Skala), nehmen Lehrpersonen des gleichen Kollegiums, die demnach dieselbe Fortbildung besuchen, diese gar nicht als eine zeitliche Belastung wahr (Antwortkategorie 1, d. h. niedrigster Wert der Skala). Die qualitativen und quantitativen Daten deuten darauf hin, dass eine empfundene zusätzliche Belastung aufgrund der Fortbildungsteilnahme nicht im Widerspruch zum Fortbildungsertrag steht. In Regressionsmodellen, in denen neben der zeitlichen Belastung auch die wahrgenommene Nützlichkeit der Fortbildungsinhalte und die wahrgenommenen Moderatoreigenschaften (z. B. das fachliche Know-how, die Klarheit und Strukturiertheit) einbezogen wurden, trug die erlebte zeitliche Belastung nicht dazu bei, Unterschiede im selbstberichteten Lernertrag der Lehrpersonen zu erklären.

5.2.3 Wirkungen der Fortbildung

Die Analyse der Interviewdaten aus der Vertiefungsstudie weist auf unterschiedliche Aspekte hin, die als Wirkungen der Teilnahme an *Vielfalt fördern* interpretiert werden können. So beschreiben die Interviewteilnehmerinnen und -teilnehmer beispielsweise, dass die Teilnahme an *Vielfalt fördern* mitunter dazu beigetragen habe, dass Lehrpersonen wöchentliche Treffen durchführen, um unter anderem gemeinsame Absprachen für den Unterricht zu treffen und

diesen miteinander abzustimmen. In den Interviews wird zudem von einer intensiveren Zusammenarbeit gesprochen, die im Zusammenhang mit der Teilnahme an *Vielfalt fördern* entstanden sei (s. Abschnitt 14.2). Auch in den quantitativen Befunden zeigt sich deskriptiv ein Vorteil der UG-Schulen gegenüber den KG-Schulen hinsichtlich der Deprivatisierung, gleichwohl konnten aber nach Abschluss des ersten Moduls mehrbenenanalytisch keine unterschiedlichen Entwicklungen zwischen den Schulen der UG und KG nachgewiesen werden. Da die Deprivatisierungspraxis erst am Anfang von Modul 2 erfasst wurde und nicht bereits zu Beginn von Modul 1, kann der Unterschied zugunsten der UG nicht auf die Teilnahme an dem ersten Fortbildungsmodul zurückgeführt werden, da nicht auszuschließen ist, dass die Unterschiede bereits zu Beginn der Fortbildung bestanden.

Ausgehend von deskriptiven Befunden der Schulleiterbefragung konnte festgestellt werden, dass die Teilnahme an *Vielfalt fördern* an allen beteiligten Schulen zur Veränderung von Teamstrukturen geführt hat. Von allen Schulen der UG gab die jeweilige Schulleiterin/der jeweilige Schulleiter die Einrichtung von mindestens einer Teamstruktur seit Beginn von *Vielfalt fördern* an.

Im Bereich der Diagnostik kann für die betrachteten Merkmale (z. B. Diagnosemotivation, prozessdiagnostische Kompetenzen) zwischen den Lehrpersonen, die an *Vielfalt fördern* teilgenommen haben, und den Lehrpersonen, die zur Kontrollgruppe gehören, im Mittel keine unterschiedliche Entwicklung festgestellt werden. Gleiches gilt für die Merkmale, welche im Zusammenhang mit den Modulen zum Lernen und Lehren bei den Lehrpersonen untersucht wurden (z. B. Lehrerselbstwirksamkeitserwartung, Wissen über Merkmale lernwirksamen Unterrichts, selbstberichtete Unterrichtspraxis).

Die mehrbenenanalytischen Auswertungen der Daten aus den Schülererhebungen zeigen, dass sich die Schülerinnen und Schüler der UG und der KG über den betrachteten Zeitraum des 6. Schuljahres in ihren fachlichen Leistungen sowie in ihren Einstellungen zum Lernen in den Fächern Mathematik und Deutsch vergleichbar entwickeln. Auch die Schülerwahrnehmungen beider Gruppen unterscheiden sich nicht: Im Durchschnitt wird der Unterricht im Fach Mathematik und im Fach Deutsch von beiden Schülergruppen vergleichbar beurteilt. Diese Urteile beziehen sich auf die effektive Klassenführung, das unterstützende Unterrichtsklima und auf die Anregung zur Reflexion über den eigenen Lernprozess. Da kooperative Lernbeziehungsweise Arbeitsformen nicht nur unter den Lehrpersonen der UG intensiviert werden sollten, sondern auch ein Thema der Module zum Lernen und Lehren (Modul 3 und 4) darstellen, wurde die Entwicklung der Einstellung gegenüber kooperativen Lernformen bei den Schülerinnen und Schülern untersucht. Auch hierbei ergab sich kein Unterschied darin, wie gerne die Schülerinnen und Schüler der UG und der KG mit Mitschülerinnen und Mitschülern zusammenarbeiten und inwiefern sie der Ansicht sind, durch diese Zusammenarbeit die Unterrichtsinhalte besser zu verstehen und mehr zu lernen. Die Schülerinnen und Schüler von

Lehrpersonen der UG entwickeln im Verlauf des sechsten Schuljahres auch kein größeres Interesse am Fach Mathematik beziehungsweise keine stärkere Lesemotivation als Schülerinnen und Schüler, deren Lehrpersonen nicht an *Vielfalt fördern* teilgenommen hatten.

Es ist somit das Fazit zu ziehen, dass sich Lehrpersonen in der UG im Vergleich zur KG – abgesehen von der Deprivatisierungspraxis – nicht günstiger entwickeln. Auch die Schülerinnen und Schüler der UG profitieren in den untersuchten Dimensionen nicht stärker als Schülerinnen und Schüler der KG. Auffällig ist, dass in beiden Fortbildungskohorten die Wahrnehmungen innerhalb der Kollegien, die an *Vielfalt fördern* teilgenommen haben, erheblich differieren. Dies wurde zum Anlass genommen, die Gruppe der Fortbildungsteilnehmerinnen und -teilnehmer in weiterführenden Analysen genauer zu untersuchen. Die Ergebnisse dieser differenziellen Analysen, die ausschließlich auf Daten der zweiten Kohorte basieren, werden im Folgenden erläutert.

6 Voraussetzungen der teilnehmenden Lehrpersonen

Die Daten aus beiden Evaluationskohorten weisen auf eine bedeutsame Varianz der individuellen Wahrnehmungen und selbstberichteten Nutzung von *Vielfalt fördern* hin: Lehrpersonen eines Kollegiums, die das identische Fortbildungsangebot erhalten, bewerten dieses teilweise sehr unterschiedlich (s. Abschnitt 5.2.2).

Auch Ergebnisse aus anderen Studien liefern Hinweise darauf, dass es wichtig ist, Merkmale der individuellen Lehrkraft, die an der Fortbildung teilnimmt, zu betrachten. So findet beispielsweise Göb (2018) heraus, dass sich eine Fortbildung auf unterschiedliche Weise auf die Lehrpersonen auswirkt und der Wirkungspfad unter anderem von den Voraussetzungen der Teilnehmerinnen und Teilnehmer abhängt. Göb (2018) schlussfolgert, dass die Lernanlässe nicht per se entscheidend sind, sondern diese nur einen indirekten Einfluss auf den Fortbildungsertrag haben, da sie eine elaborierte Auseinandersetzung mit den Fortbildungsinhalten mehr oder weniger begünstigen können. Auch Befunde zur differenziellen Teilnahme an Lehrerfortbildungen, die zeigen, dass Lehrpersonen zu einem unterschiedlichen Fortbildungsverhalten tendieren, wenn es ihnen möglich ist, Fortbildungen selbst auszuwählen (Hoffmann & Richter, 2016; Richter, Engelbert, Weirich & Anand Pant, 2013), können als Hinweis gedeutet werden, dass sich Lehrpersonen abhängig von ihrem individuellen Hintergrund und ihren Voraussetzungen unterschiedlich auf ein Fortbildungsangebot einlassen dürften. Im Fall von *Vielfalt fördern* hatten die Lehrpersonen eines Kollegiums diese Wahlfreiheit nicht, aber es ist plausibel, dass sich individuelle Merkmale nicht nur auf die Wahl zur Teilnahme an einer Fortbildung auswirken, sondern auch auf den eigentlichen Fortbildungsprozess.

Dies führt zu der Frage, ob *Vielfalt fördern* in Abhängigkeit von den Voraussetzungen der Teilnehmerinnen und Teilnehmer unterschiedlich wirkt, d. h., ob individuelle Merkmale einer Lehrperson die Wahrnehmung und Nutzung sowie den Ertrag von *Vielfalt fördern* bedeutsam beeinflussen.

6.1 Voraussetzungsprofile der Fortbildungsteilnehmerinnen und -teilnehmer

Befunde aus der vorliegenden Evaluation und aus anderen Studien (z. B. Gorozidis & Papiannou, 2014; Kao, Wu & Tsai, 2011; Rzejak, Lipowsky & Künsting, 2013) weisen darauf hin, dass die Motivation der teilnehmenden Lehrpersonen ein bedeutsamer Faktor für den Fortbildungsprozess und -erfolg ist. Im Projekt *LIQUID* wurden mit einem Instrument in Anlehnung an Rzejak et al. (2014) unterschiedliche Facetten der Teilnahmemotivation erhoben, beispielsweise die *Erwartungsanpassung*, die *Schulentwicklungsorientierung* und die *soziale Interaktion*. Die Skalen der einzelnen Facetten wurden für die hier beschriebene Profilanalyse aus jeweils drei Items gebildet. Bei der Erwartungsanpassung handelt es sich um eine extrinsische Motivationsqualität (Beispielitem: „An der Fortbildung „Vielfalt fördern“ nehme ich teil, weil es

zu den dienstlichen Verpflichtungen eines Lehrers gehört“, $\alpha = .84$).¹⁰ Eine Lehrperson mit einem hohen Wert auf dieser Skala nimmt eher aus einem Pflichtgefühl heraus an *Vielfalt fördern* teil. In der Gesamtstichprobe liegt der Mittelwert der individuellen Einschätzungen der Lehrpersonen bei $M = 4.05$ ($SD = 1.62$) und damit leicht über dem theoretischen Mittelwert von $M = 3.50$. Das bedeutet, dass die Lehrpersonen auch deshalb an der Fortbildung teilnehmen, weil sie sich verpflichtet fühlen und die Schule sich für die Teilnahme entschieden hat.

Die Schulentwicklungsorientierung misst einen Aspekt der intrinsischen Motivationsqualität, da der Teilnahme an *Vielfalt fördern* ein Wert beigemessen und diese als gewinnbringend betrachtet wird (Beispielitem: „An der Fortbildung „Vielfalt fördern“ nehme ich teil, weil es für die gesamte Schule wichtig ist“, $\alpha = .93$). Es zeigt sich am Mittelwert ($M = 4.77$, $SD = 1.15$), dass die Lehrpersonen dieser Motivationsfacette vergleichsweise stark zustimmen.

Die dritte Motivationsfacette, die zur Profilanalyse herangezogen wird, ist die soziale Interaktion. Diese Motivationsfacette lässt sich eher als eine intrinsische Motivationsfacette betrachten. Lehrpersonen, die hohe Werte auf dieser Motivationsfacette haben, nehmen an *Vielfalt fördern* teil, weil sie gerne zusammen mit anderen Lehrern lernen und weil sie „an den beruflichen Erfahrungen anderer Kollegen teilhaben“ wollen ($\alpha = .78$). Im Mittel ist die soziale Interaktion mit einem $M = 4.09$ ($SD = 1.11$) in der Gesamtstichprobe ähnlich stark ausgeprägt wie die extrinsische Facette externe Erwartungsanpassung. Die Relevanz des Einbezugs der sozialen Interaktion lässt sich gut begründen mit der Selbstbestimmungstheorie von Deci und Ryan (Deci & Ryan, 1993; Ryan & Deci, 2002). Darin wird dem Erleben sozialer Eingebundenheit neben dem Kompetenzerleben und dem Erleben von Autonomie eine wichtige Funktion für die intrinsische Motivation zugeschrieben.

Als ein weiteres motivationales Konstrukt wurde die Lernzielorientierung mit einer von Dickhäuser, Butler und Tönjes (2007) entwickelten Skala aus vier Items (Beispielitem: „In meinem Unterricht geht es mir darum, so viel wie möglich über das Unterrichten zu lernen“, $\alpha = .89$) erfasst. Im Mittel liegt die Lernzielorientierung der Lehrpersonen am Anfang der Fortbildung bei $M = 4.38$ ($SD = 0.97$). Demnach sind die Lehrpersonen eher bestrebt, möglichst viel über das Unterrichten zu lernen und komplizierte Unterrichtssituationen besser zu verstehen. Studien konnten zeigen, dass die Lernzielorientierung von Lehrpersonen das Erleben der beruflichen Situation, die Kompetenzentwicklung und die Unterrichtsgestaltung beeinflusst (Nitsche et al., 2017). So erleben Lehrpersonen mit einer hohen Lernzielorientierung beispielsweise eine geringere berufliche Belastung und berichten von intensiveren Fortbildungsaktivitäten, weshalb eine hohe Lernzielorientierung als eine förderliche Voraussetzung für die Teilnahme an *Vielfalt fördern* betrachtet werden kann.

¹⁰ Cronbachs α ist ein Kennwert zur Beurteilung der internen Konsistenz einer Skala. Im Projekt *LIQUID* wird ein Wert von $\alpha = .60$ als untere Grenze für den Zusammenschluss von Items zu einer Skala festgelegt.

Ferner wurde die allgemeine Innovationsbereitschaft in Anlehnung an Diedrich, Abs und Klieme (2004) mittels vier Items erhoben (Beispielitem: „*Wenn unsere Schule Veränderungen plant, möchte ich gerne aktiv mithelfen*“, $\alpha = .81$). Die Skala ist ein Indikator dafür, wie aufgeschlossen die Lehrpersonen grundsätzlich für Veränderungen und Neuerungen sind, die ihre Schule und ihren Unterricht betreffen, unabhängig von *Vielfalt fördern*. Der Mittelwert von $M = 4.52$ ($SD = 0.89$) liegt über dem theoretischen Skalenmittel von $M = 3.50$ und kann demnach als eine hohe Innovationsbereitschaft der Lehrpersonen interpretiert werden.

Um zu messen, inwiefern die Lehrpersonen bereit sind, sich aktiv in die Fortbildung *Vielfalt fördern* einzubringen und die Lernangebote zu nutzen, auch wenn dies mitunter anstrengend sein könnte, wurde die selbstentwickelte Skala Anstrengungsbereitschaft eingesetzt, die drei Items umfasst (Beispielitem: „*Ich werde bei der Fortbildung engagiert mitarbeiten, auch wenn es anstrengend wird*“, $\alpha = .87$). Aufgrund des Skalenmittelwerts von $M = 4.43$ ($SD = 0.97$) scheinen die Lehrpersonen am Anfang der Fortbildung im Durchschnitt eher dazu bereit zu sein, engagiert in der Fortbildung mitzuarbeiten und das Lernangebot aktiv zu nutzen.

Diese sechs genannten Merkmale (Erwartungsanpassung, Schulentwicklungsorientierung, soziale Interaktion, Lernzielorientierung, individuelle Innovationsbereitschaft, Anstrengungsbereitschaft) sollten für eine latente Profilanalyse genutzt werden, um die Fortbildungsteilnehmerinnen und -teilnehmer hinsichtlich ihrer Voraussetzungen zu gruppieren. Da eine Annahme bei der Profilbildung lautet, dass die Ausprägung eines Merkmals allein von der Profilizugehörigkeit und nicht von den anderen profilbildenden Merkmalen abhängt, sollten die Korrelationen zwischen den Konstrukten nicht zu hoch ausfallen. Die bivariaten Zusammenhänge der sechs Merkmale (s. Tabelle 5) fallen jedoch zum Teil eng aus ($r > .50$). Insbesondere fällt hier die Anstrengungsbereitschaft auf, die stark mit der Innovationsbereitschaft ($r = .56, p < .01$) und den eher intrinsischen Motivationsqualitäten zur Fortbildungsteilnahme ($r = .57 - .66, p < .01$) zusammenhängt. Aus diesem Grund wurde die selbstberichtete Anstrengungsbereitschaft nicht für die Profilbildung genutzt. Da die kollegiale Zusammenarbeit und die gemeinsame Weiterentwicklung von Schule und Unterricht zwei zentrale Momente im Konzept von *Vielfalt fördern* darstellen, wurden diese beiden eher intrinsischen Teilnahmemotivationen (Schulentwicklung, soziale Interaktion) einbezogen, obwohl diese Merkmale stark miteinander korrelieren ($r = .62, p < .01$).

Tabelle 5. Korrelation zwischen den Voraussetzungsmerkmalen für die Profilanalyse

	LERNZO1	ANSB1	IINNOB1	EXTER1	SENTW1	SOZI1
LERNZO1	1	.40**	.32**	-.14**	.41**	.41**
ANSB1	.40**	1	.56**	-.33**	.66**	.57**
IINNOB1	.32**	.56**	1	-.35**	.49**	.40**
EXTER1	-.14**	-.33**	-.35**	1	-.27**	-.27**
SENTW1	.41**	.66**	.49**	-.27**	1	.62**
SOZI1	.41**	.57**	.40**	-.27**	.62**	1

Anmerkung. LERNZO1 = Lernzielorientierung (t_1), ANSB1 = Anstrengungsbereitschaft (t_1), IINNOB1 = individuelle Innovationsbereitschaft (t_1), EXTER1 = externale Erwartungsanpassung (t_1), SENTW1 = Schulentwicklungsorientierung (t_1), SOZI1 = soziale Interaktion (t_1), Fettdruck = starke Korrelationen, ** = $p < .01$, $n = 533-556$.

Aufgrund der statistischen Kennwerte (s. Tabelle 6) und inhaltlicher Interpretierbarkeit wird eine Lösung mit vier disjunkten Profilen gewählt (s. Abbildung 16).

Der Entropywert von .80 (s. Tabelle 6) spricht für eine gute Zuverlässigkeit, mit der die Personen bei diesem Modell in Gruppen klassifiziert werden können. Die Werte von AIC und BIC liegen bei dem Modell mit vier Profilen deutlich unter den AIC- und BIC-Werten des Modells mit drei Profilen. Die Differenz zu den AIC- und BIC-Werten des Modells mit fünf Profilen fällt hingegen eher gering aus. Auch dass der BLRT und LMRT signifikant ausfallen, spricht dafür, dass das Modell mit vier Profilen besser auf die Daten passt als das Modell mit drei Lösungen.

Tabelle 6. Kennwerte der Profillösungen

	Entropy	AIC	BIC	BLRT	LMRT
1 Profil		8453.13	8496.52		
2 Profile	.81	8024.92	8094.34	n. s.	n. s.
3 Profile	.83	7692.89	7788.34	$p < .001$	$p < .001$
4 Profile	.80	7602.16	7723.64	$p < .05$	$p < .05$
5 Profile	.78	7559.19	7706.70	n. s.	n. s.
6 Profile	.80	7528.16	7701.71	n. s.	n. s.

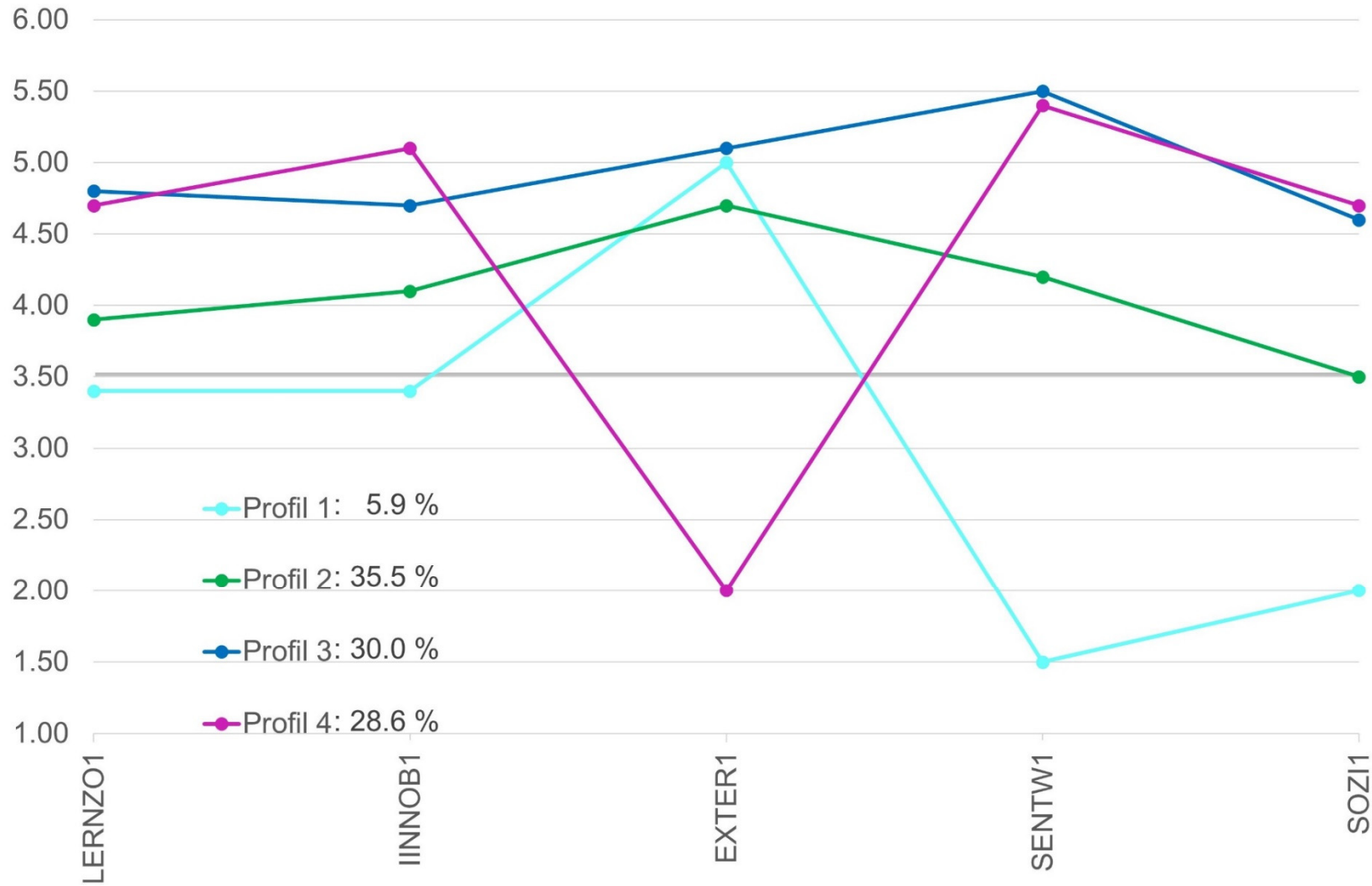


Abbildung 16. Grafische Darstellung der Lösung mit vier Voraussetzungsprofilen. Anmerkung. LERNZO1 = Lernzielorientierung (t_1), IINNOB1 = individuelle Innovationsbereitschaft (t_1), EXTER1 = externe Erwartungsanpassung (t_1), SENTW1 = Schulentwicklungsorientierung (t_1), SOZI1 = soziale Interaktion (t_1)

Wie lassen sich die vier Lehrergruppen (Profile, s. Abbildung 16) näher charakterisieren und beschreiben? Zum ersten Profil (hellblau) gehören nur sehr wenige Lehrpersonen (5.9 %, $n = 33$). Man kann diese Lehrpersonen als extrinsisch motiviert beschreiben: Sie sind Innovationen gegenüber nicht besonders aufgeschlossen ($M = 3.49$) und weisen tendenziell eine eher geringe Lernzielorientierung auf ($M = 3.42$). Diese Lehrpersonen sind demnach eher zurückhaltend, wenn es um die Veränderung von Schule und Unterricht geht, und sie betrachten den eigenen Unterricht weniger als Chance, um ein tieferes Verständnis für das Unterrichten und anspruchsvolle Unterrichtssituationen zu erwerben. An *Vielfalt fördern* nehmen die Lehrpersonen, die dem ersten Cluster zugeordnet werden, insbesondere aufgrund ihres Pflichtgefühls ($M = 5.15$) teil, während sie die Möglichkeit der Schulentwicklung ($M = 1.47$) aufgrund von *Vielfalt fördern* und die Aussicht auf den sozialen Austausch mit Kolleginnen und Kollegen ($M = 1.95$) kaum zur Teilnahme motiviert. Betrachtet man diese Merkmalskomposition, kann man die Ausgangsvoraussetzungen dieser Gruppe als sehr ungünstig bewerten.

Das zweite Profil (grün), dem 35.5 % aller Lehrpersonen ($n = 201$) angehören, zeichnet sich dadurch aus, dass alle Werte über dem theoretischen Skalenmittel liegen, gleichzeitig aber nicht besonders hoch ausfallen. Die Lehrpersonen sind eher lernzielorientiert ($M = 3.97$) und offen gegenüber Innovationen ($M = 4.07$), welche die Schule und den Unterricht betreffen. Von den drei Teilnahmemotivationsfacetten ist die Erwartungsanpassung ($M = 4.68$) am stärksten ausgeprägt. Die Fortbildungsteilnahme erfolgt demnach durchaus extrinsisch motiviert. Hingegen wirkt die Aussicht auf den kollegialen Austausch für die Lehrpersonen des zweiten Profils am wenigsten motivierend ($M = 3.57$) und der Mittelwert für die Schulentwicklungsorientierung liegt dazwischen ($M = 4.22$).

Die Kurve, die das dritte Profil beschreibt (dunkelblau), verläuft weitgehend parallel zur Kurve des zweiten Profils. Alle Werte sind hier allerdings nach oben verschoben. Das bedeutet: Dieses Profil weist sowohl eine hohe intrinsische Motivation als auch eine hohe extrinsische Motivation auf. Besonders deutlich wird der Unterschied zwischen den Lehrpersonen des zweiten und des dritten Profils in der Ausprägung der Motivationsfacette Schulentwicklung. Lehrpersonen, die zum dritten Profil gehören, zeichnen sich dadurch aus, dass sie insbesondere deshalb motiviert sind, an *Vielfalt fördern* teilzunehmen, weil ihnen die Schulentwicklung ($M = 5.47$) wichtig ist. Es trifft auf diese Lehrpersonen demnach im starken Maße zu, dass sie an *Vielfalt fördern* teilnehmen, weil die Schule und deren Entwicklung davon profitiert. Bei diesem Profil zeigt sich, dass sich eine hohe intrinsische Motivation und eine hohe extrinsische Motivation nicht ausschließen müssen, denn auch die Motivationsfacette Erwartungsanpassung ist bei den Lehrpersonen dieses Profils hoch ausgeprägt ($M = 4.98$). Der Mittelwert für die dritte Motivationsfacette, die soziale Interaktion, liegt bei $M = 4.59$. Zudem berichten die 170 Lehrpersonen (30.0 %), die zu dem dritten Profil gehören, auch eine ausgeprägte Lernzielorientierung ($M = 4.74$), d. h. ein hohes Interesse, im eigenen Unterricht ein tieferes Verständnis für das Unterrichten und das komplexe

unterrichtliche Geschehen zu erwerben, sowie eine eher hohe Aufgeschlossenheit gegenüber schulischen und unterrichtlichen Neuerungen und Veränderungen ($M = 4.71$). Im Vergleich zu den Lehrpersonen, die zum ersten und zweiten Profil gehören, sind die Voraussetzungen der Lehrpersonen, die zum dritten Profil gehören, insgesamt günstiger zu bewerten.

Die Lehrpersonen des vierten Profils (28.6 %, $n = 162$, pink) besitzen eine besonders ausgeprägte Offenheit gegenüber innovativen schulischen Veränderungen ($M = 5.10$), die im Vergleich zu den drei anderen Profilen am höchsten ausgeprägt ist. Es überrascht daher nicht, dass auch die Schulentwicklungsorientierung bei den Lehrpersonen des vierten Profils stark ausgeprägt ist ($M = 5.41$). Diese Lehrpersonen sind demnach in einem besonderen Maße zur Teilnahme an *Vielfalt fördern* motiviert, weil die Schule als Ganzes davon profitieren und sich weiterentwickeln kann. Bei der Motivationsfacette Erwartungsanpassung haben die Lehrpersonen des vierten Profils hingegen den niedrigsten Mittelwert ($M = 2.04$). Die Lehrpersonen sind demnach in einem sehr geringen Maß zur Teilnahme an *Vielfalt fördern* motiviert, weil sie dies als eine Pflicht empfinden oder weil der Teilnahme ein Mehrheitsentscheid der Schule zugrunde lag. Die Ausprägungen der Lernzielorientierung ($M = 4.71$), d. h. des Bedürfnisses, durch das Unterrichten ein tiefergehendes Verständnis für Prozesse des Lehrens und Lernens zu erwerben, und der Motivationsfacette soziale Interaktion ($M = 4.67$), also einer Motivation zur Teilnahme an *Vielfalt fördern*, die aus dem Wunsch nach einem intensiveren Kontakt mit Kolleginnen und Kollegen resultiert, sind vergleichbar mit den Werten der Lehrpersonen des dritten Profils. Die Voraussetzungen für die Teilnahme an *Vielfalt fördern* sind für die Lehrpersonen, die dem vierten Profil zugeordnet wurden, am günstigsten.

6.2 Voraussetzungsprofile als erklärendes Merkmal

In dem folgenden Abschnitt wird dargestellt, ob sich die Lehrpersonen der vier identifizierten Profile (s. Abschnitt 6.1) darin unterscheiden, wie sie die Lernangebote von *Vielfalt fördern* wahrgenommen, genutzt und bewertet haben.

Überprüft wurde, ob statistisch signifikante Unterschiede in der Intensität der selbstberichteten Nutzung der Lernangebote (t_3), in der Intensität der kooperativen Umsetzung zwischen den Fortbildungstreffen (t_3), in der Bewertung der Umsetzung der Fortbildungsinhalte (t_3), in der erlebten zeitlichen Belastung (t_3) sowie in der Zufriedenheit mit dem Fortbildungsangebot (t_4) und dem selbstberichteten Lernertrag (t_4) bestehen.

Der Tabelle 7 können jeweils die Reliabilität, die Itemanzahl und ein Beispielitem für die Skalen entnommen werden. Ebenfalls werden die jeweiligen Skalenmittelwerte und Standardabweichungen für die Gesamtstichprobe angegeben.

Tabelle 7. Skalenkennwerte und Beispielitem der abhängigen Variablen

Merkmal	Alpha	Items	Beispielitem	M (SD)
Intensität der Nutzung individueller Lernaktivitäten	.90	12	In und zwischen den Fortbildungssitzungen habe ich die von mir verwendeten Konzepte/Methoden der Diagnostik überdacht.	3.85 (0.87)
Intensität der kooperativen Erprobung	.89	4	Auch zwischen den Fortbildungstreffen habe ich gemeinsam mit Kollegen die Fortbildungsinhalte umgesetzt.	3.37 (1.16)
Bewertung der Umsetzung der Fortbildungsinhalte	.92	4	Die Umsetzung der Fortbildungsinhalte ist erfolgreich.	3.36 (1.14)
Zeitliche Belastung	.78	4	Durch die Teilnahme an der Fortbildung fühle ich mich zeitlich zusätzlich belastet.	3.63 (1.15)
Zufriedenheit	.92	4	Mit der bisherigen Fortbildung bin ich zufrieden.	3.16 (1.28)
Lernertrag	.87	4	Durch die Fortbildung habe ich viele Impulse erhalten, die mir etwas für meinen beruflichen Alltag bringen.	3.13 (1.13)

Es zeigt sich, dass sich die Lehrpersonen der vier Profile signifikant in der Intensität unterscheiden, mit der individuell Lernangebote aus dem zweiten Modul zur Diagnostik genutzt werden ($F [3, 317] = 23,99, p < .001; \eta^2 = .19$). Die Effektstärke von $\eta^2 = .19$ ist als starker Effekt zu interpretieren. Bedeutsame Mittelwertunterschiede liegen aber nicht zwischen allen Profilen vor, sondern nur zwischen dem ersten Profil und den drei anderen Profilen ($M_{\Delta} = -0.81$ bis $-1.43, p < .001$), zwischen dem zweiten Profil und dem dritten Profil ($M_{\Delta} = -0.56, p < .001$) sowie zwischen dem zweiten und dem vierten Profil ($M_{\Delta} = -0.62, p < .001$). Mit anderen Worten: Die Lehrpersonen, die zu Profil 3 und Profil 4 gehören, geben die intensivste Nutzung an, die Lehrpersonen von Profil 1 die geringste (s. Abbildung 17).

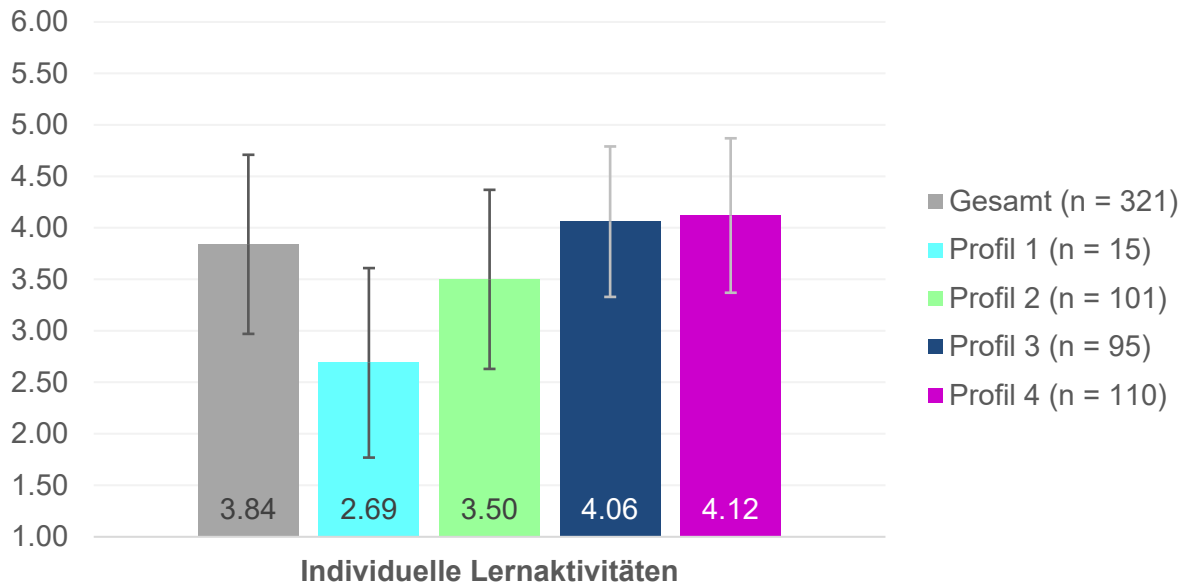


Abbildung 17. Mittelwerte und Standardabweichungen für die Intensität der Nutzung individueller Lernaktivitäten bei den vier Profilen

Da *Vielfalt fördern* ein Fortbildungsangebot ist, bei dem aufgrund der zweijährigen Anlage unterschiedliche Phasen (Input/Erarbeitung, Erprobung, Reflexion) kombiniert werden können und weil der kooperativen Zusammenarbeit ein besonderer Stellenwert beigemessen wird, wurden nicht nur Daten zur individuellen Nutzung erfasst, sondern auch zur Intensität der kooperativen Umsetzung. Die Skala Intensität der kooperativen Erprobung bildet ab, in welchem Maß die Lehrpersonen zusammen mit Kolleginnen und Kollegen auch zwischen den Präsenzveranstaltungen des zweiten Moduls über die Fortbildungsinhalte diskutiert haben und diese an die eigene schulische Situation angepasst und in der Praxis erprobt haben. Die Profile unterscheiden sich signifikant in der kooperativen Umsetzung und Erprobung ($F [3, 324] = 17,82, p < .001; \eta^2 = .14$). Wie schon bei der Intensität der individuellen Nutzung der Lernaktivitäten handelt es sich um einen starken Effekt. Die Intensität der kooperativen Umsetzung ist bei den Profilen 3 und 4, d. h. bei den Lehrpersonen mit günstigeren Ausgangsbedingungen, höher (s. Abbildung 18). Das erste Profil, bei dem die Lehrpersonen sich insbesondere durch eine hohe Ausprägung der Erwartungsanpassung auszeichnen und die wenig intrinsisch zur Teilnahme an *Vielfalt fördern* motiviert sind, weist signifikant niedrigere Werte auf als die Lehrpersonen der drei anderen Profile ($M_{\Delta} = -1.33$ bis $-1.89, p < .001$). Des Weiteren zeigt sich ein bedeutsamer Unterschied zwischen dem Profil 2 und dem Profil 4 ($M_{\Delta} = -0.55, p < .01$).

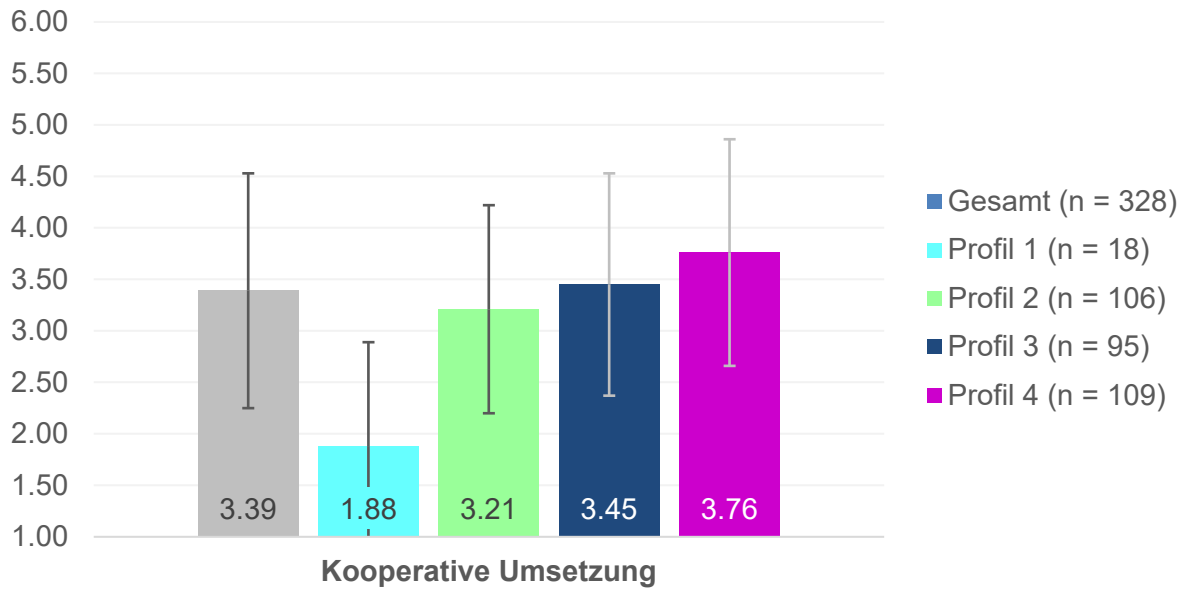


Abbildung 18. Mittelwerte und Standardabweichungen für die Intensität der kooperativen Umsetzung bei den vier Profilen

Auch die Bewertung der Umsetzung der Fortbildungsinhalte des Diagnostikmoduls (t_3) fällt bei den Lehrpersonen der vier Profile nicht gleich aus ($F [3, 306] = 16,67, p < .001; \eta^2 = .14$; s. Abbildung 19). Ähnlich bewerten die Lehrpersonen des ersten und des zweiten Profils sowie die beiden Profile mit den günstigeren Ausgangsbedingungen (Profil 3 und Profil 4) die Umsetzung der Fortbildungsinhalte aus dem zweiten Modul zum Thema Diagnostik. Hingegen zeigen sich zwischen Profil 1 und Profil 3 ($M_\Delta = -1.18, p < .001$) sowie zwischen Profil 1 und Profil 4 ($M_\Delta = -1.45, p < .001$) signifikante Differenzen. Im Mittel unterscheiden sich außerdem die Lehrpersonen von Profil 2 signifikant von den Lehrpersonen, die zu Profil 3 ($M_\Delta = -0.53, p < .01$) oder Profil 4 ($M_\Delta = -0.80, p < .001$) gehören.

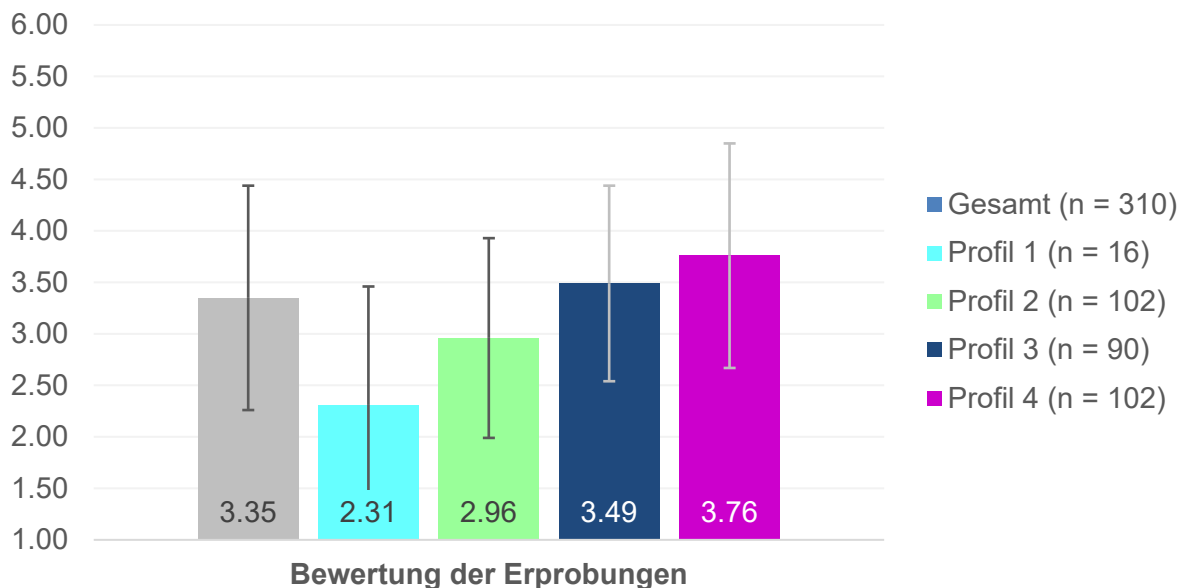


Abbildung 19. Mittelwerte und Standardabweichungen für die Umsetzung der Fortbildungsinhalte bei den vier Profilen

Es wäre plausibel, anzunehmen, dass die Lehrpersonen des dritten und vierten Profils, die eine intensivere individuelle wie auch kollektive Auseinandersetzung mit den Fortbildungsinhalten angeben, die zeitliche Belastung aufgrund der Teilnahme an *Vielfalt fördern* höher empfinden. Die Ergebnisse der Analyse sind jedoch so zu interpretieren, dass mit der Zugehörigkeit zu Profil 1 oder zu Profil 2 im Durchschnitt eine stärker erlebte zeitliche Belastung während des zweiten Moduls (t_3) verbunden ist (s. Abbildung 20).

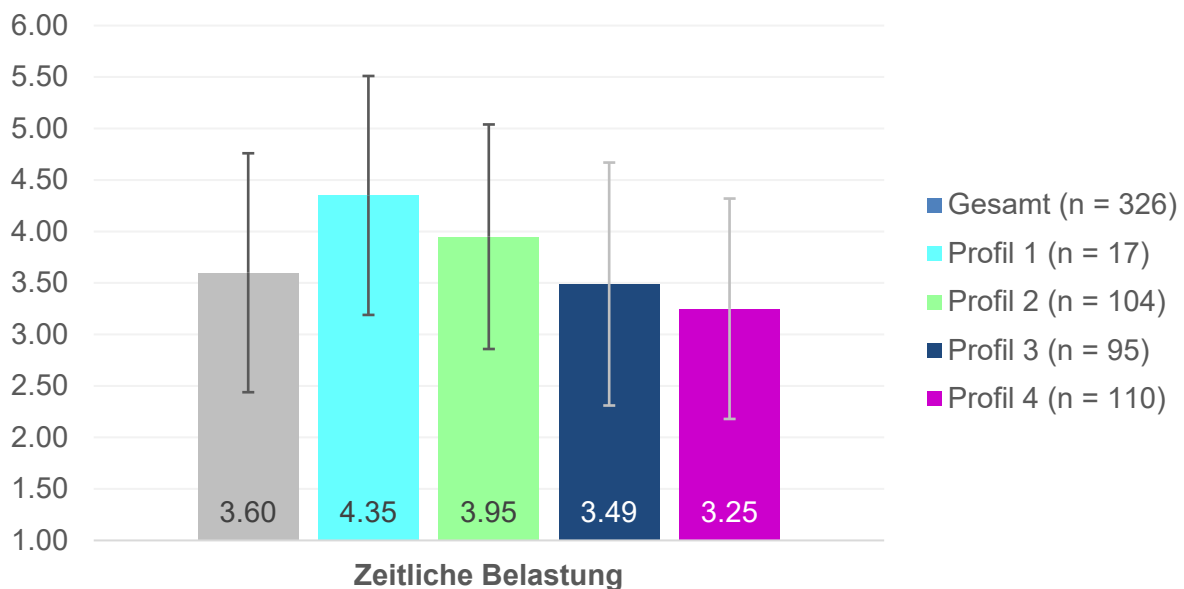


Abbildung 20. Mittelwerte und Standardabweichungen für die zeitliche Belastung bei den vier Profilen

Der Unterschied in der erlebten zeitlichen Belastung im Verlauf des zweiten Moduls ist statistisch bedeutsam ($F [3, 322] = 9,81, p < .001$). Es handelt sich um einen mittleren Effekt ($\eta^2 = .08$). Differenzen zeigen sich erneut zwischen Profil 1 und Profil 3 ($M_\Delta = 0.86, p < .05$) sowie zwischen Profil 1 und Profil 4 ($M_\Delta = 1.10, p < .01$). Außerdem unterscheiden sich das Profil 2 und das Profil 3 ($M_\Delta = 0.45, p < .05$) sowie das Profil 2 und Profil 4 ($M_\Delta = 0.69, p < .001$) signifikant.

Auch in der Zufriedenheit mit dem Fortbildungsangebot weichen die Profile am Ende des zweiten Moduls (t_4) signifikant voneinander ab ($F [3, 258] = 8,16, p < .001, \eta^2 = .09$; s. Abbildung 21). Es zeigt sich erneut, dass eine bedeutsame Differenz zwischen einzelnen Profilen vorliegt. Die Lehrpersonen, die *Vielfalt fördern* mit ungünstigeren Voraussetzungen begonnen haben (Profil 1 und Profil 2), sind am Ende des Diagnostikmoduls (Modul 2) weniger zufrieden mit dem Fortbildungsangebot als Lehrpersonen des Profils 4. Zwischen den Lehrpersonen von Profil 1 und den Lehrpersonen von Profil 4 beträgt die mittlere Differenz $-1.56 (p < .01)$. Die mittlere Differenz fällt zwischen dem zweiten Profil und dem vierten Profil geringer aus ($M_\Delta = -0.66, p < .01$) als zwischen Profil 1 und Profil 4.

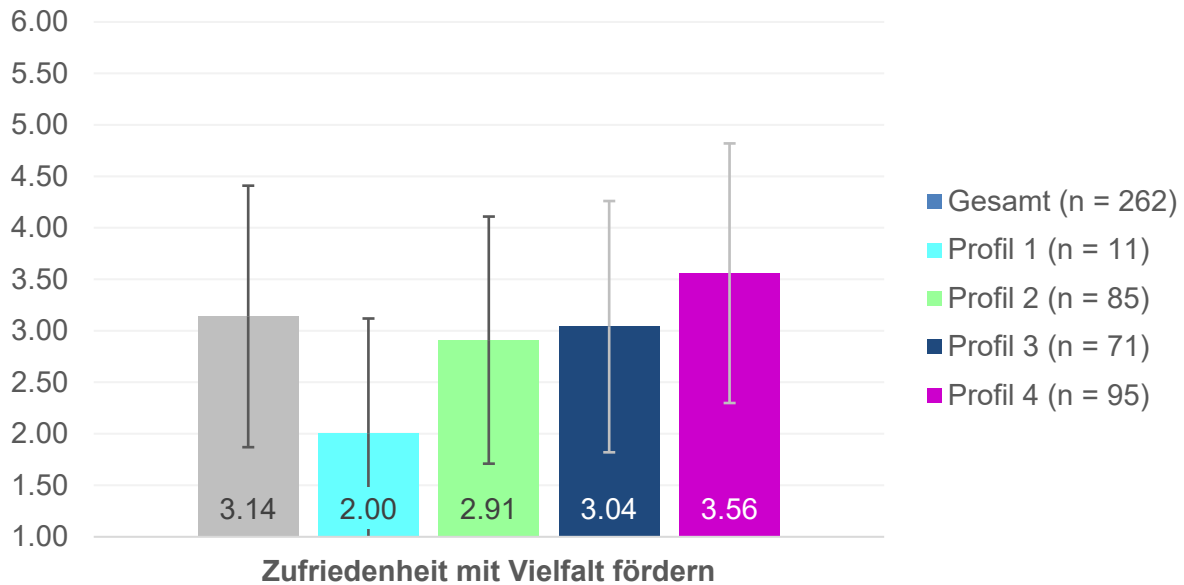


Abbildung 21. Mittelwerte und Standardabweichungen für die Zufriedenheit bei den vier Profilen

Die Lehrpersonen, die zum Profil 1 gehören, unterscheiden sich auch in ihrem selbstberichteten Lernertrag am Ende des zweiten Moduls (t_4) signifikant negativ von den Lehrpersonen der drei anderen Profile ($F [3, 252] = 5,88, p < .01, \eta^2 = .07$; s. Abbildung 22). Das Ergebnis bedeutet, dass die Lehrpersonen, die zu Profil 1 gehören, in einem geringeren Maß den Eindruck haben, aufgrund der Teilnahme am zweiten Modul von *Vielfalt fördern* Neues gelernt und Impulse erhalten zu haben, die für den eigenen schulischen Alltag nützlich sind. Die mittlere Differenz liegt zwischen -1.20 und -1.40 ($p < .01$). Die Lehrpersonen der Profile 2, 3 und 4 unterscheiden sich im Mittel nicht darin, wie hoch sie ihren individuellen Lernertrag aufgrund der Teilnahme am zweiten Fortbildungsmodul zu Fragen der Diagnostik wahrnehmen.

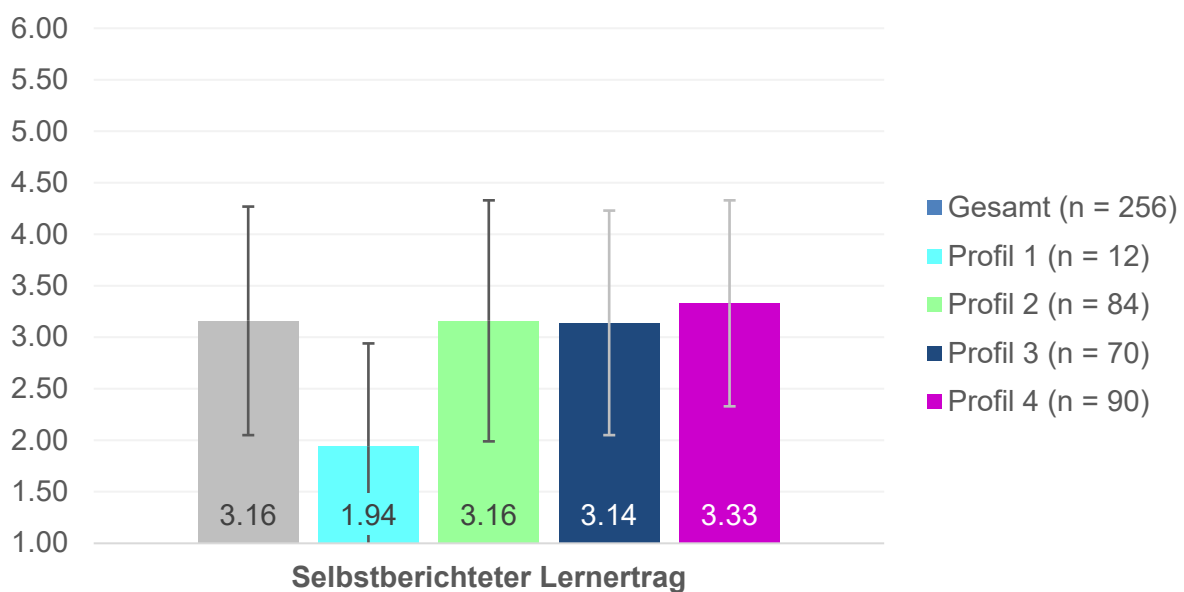


Abbildung 22. Mittelwerte und Standardabweichungen für den selbstberichteten Lernertrag bei den vier Profilen

6.3 Teilnahme aufgrund des schulischen Mehrheitsentscheids

Die am Anfang der Fortbildung gemessene Erwartungsanpassung ist ein Indikator für den Grad einer eher extrinsischen Fortbildungsteilnahmemotivation. Analysen auf Individualebene zeigen, dass es sich um ein erklärungsstarkes Merkmal handelt, das Unterschiede in Wahrnehmungs-, Nutzungs- und Ertragsmerkmalen zu einem späteren Zeitpunkt im Fortbildungsprozess erklären kann (s. Abschnitt 5.2). Auch die in dem vorangehenden Abschnitt berichteten Befunde weisen darauf hin, dass nachteilige Teilnahmevoraussetzungen, zu denen unter anderem auch eine hoch ausgeprägte extrinsische Teilnahmemotivation gezählt werden kann, mit einer ungünstigeren Nutzung und Wahrnehmung des Angebots einhergehen. Wichtige Fragen für Anbieter zukünftiger Lehrerfortbildungen könnten sein, ob es sinnvoll ist, dass (zunächst) nur hoch intrinsisch motivierte Lehrpersonen eines Kollegiums, die sich freiwillig für eine Teilnahme entschieden haben, an einer Fortbildung teilnehmen. Im Folgenden wird daher untersucht, ob und inwieweit der Anteil an extrinsisch motivierten Lehrpersonen in einem Kollegium sich ungünstig auf die Wahrnehmung, Nutzung und den Ertrag der Fortbildung auswirkt.

Um einzuschätzen, wie hoch die extrinsische Motivation des gesamten Kollegiums an einer Schule für die Teilnahme an der Fortbildung ausgeprägt war, wurde das Item „An der Fortbildung „Vielfalt fördern“ nehme ich teil, weil sich die Schule mehrheitlich dafür entschieden hat“ herangezogen.¹¹ Die Lehrpersonen eines Kollegiums wurden zunächst ausgehend von ihrem Antwortverhalten zu diesem Item in zwei Gruppen eingeteilt. Alle Lehrpersonen einer Schule, die angegeben haben, dass das Item *eher nicht, größtenteils nicht* oder *gar nicht* auf sie zutrifft, wurden einer Gruppe zugeordnet. Ebenso wurden alle Lehrpersonen, die angegeben haben, dass das Item *eher, größtenteils* oder *voll und ganz* auf sie zutrifft, zu einer Gruppe zusammengefasst. Anschließend wurde ermittelt, wie hoch der Anteil der Lehrpersonen einer Schule ist, die der letztgenannten Gruppe angehören. Je größer der Anteil der Lehrpersonen innerhalb einer Schule ist, die das Item hoch bewertet haben, desto höher ist die Zurückhaltung der Kolleginnen und Kollegen einer Schule. Hohe Werte für diese Anpassungsquote stehen also für ein Kollegium, in dem viele Lehrpersonen auch deshalb an der Fortbildung teilnahmen, weil es einen Mehrheitsentscheid gab, ohne zwingend von der Relevanz von *Vielfalt fördern* überzeugt und zu einer aktiven Nutzung des Fortbildungsangebots bereit zu sein. Darauf deuten ebenfalls die Korrelationen der Anpassungsquote mit anderen Merkmalen der Ausgangslage auf Schulebene hin (s. Tabelle 8), auch wenn diese nur auf einem 10 %-Niveau signifikant werden.

¹¹ Werte auf diesem Item korrelieren zu $r \geq .60$ ($p < .001$), d. h. stark positiv, mit den beiden anderen Einzelitems der Skala „externale Erwartungsanpassung“, zu $-.19 \leq r \leq -.13$ ($.01 < p < .001$) mit eher intrinsischen Facetten der Fortbildungsteilnahmemotivation (Schulentwicklungsorientierung, soziale Interaktion, Interesse) sowie ebenfalls negativ mit der Relevanz, die eine Lehrperson individuell einer Fortbildung zum Thema individuelle Förderung beimisst ($r = -.20$, $p < .001$). Die Korrelation mit der Anstrengungsbereitschaft am Anfang der Fortbildung ($r = -.23$, $p < .001$) fällt ebenfalls negativ aus. Das Item misst also vor allem eine Facette extrinsischer Motivation der Teilnehmerinnen und Teilnehmer.

Tabelle 8. Korrelation zwischen der Anpassungsquote und anderen Merkmalen der schulischen Ausgangslage

	Anpassungsquote	Anstrengungs- bereitschaft zu t ₁	Fortbildungs- relevanz zu t ₁	(Lern-)Interesse zu t ₁
Anpassungs- quote	1	-.42 [†]	-.40 [†]	-.38 [†]

Anmerkung: [†] = $p < .10$, $n = 21$.

So zeigt sich beispielsweise, dass je höher die Anpassungsquote ausfällt, d. h. je größer der Anteil an Lehrpersonen innerhalb eines Kollegiums ist, die aufgrund des getroffenen Mehrheitsentscheids zur Teilnahme an *Vielfalt fördern* motiviert sind, desto geringer ist tendenziell die selbstberichtete anfängliche Bereitschaft in den Kollegien, sich aktiv in die Fortbildung einzubringen und mitzuarbeiten ($r = -.42$, $p < .10$). Zudem ergibt sich: Je höher die Anpassungsquote einer Schule ist, desto weniger finden die Lehrpersonen des Kollegiums eine Fortbildung zu dem Kernthema von *Vielfalt fördern*, der individuellen Förderung von Schülerinnen und Schülern, relevant ($r = -.40$, $p < .10$) und desto geringer fällt auch das Interesse der Lehrpersonen an neuen Themen aus ($r = -.38$, $p < .10$). Da es sich um Korrelationen handelt, gilt aber auch: Je höher die Anstrengungsbereitschaft, die Relevanzwahrnehmung des Fortbildungsthemas und das (Lern-)Interesse ausgeprägt sind, desto geringer fällt die Anpassungsquote des Kollegiums aus.

Es zeigt sich, dass die Anpassungsquote, d. h. der Anteil der Lehrpersonen, die auch deshalb an der Fortbildung teilnahmen, weil es einen Mehrheitsentscheid im Kollegium hierzu gab, in allen Kollegien bei über 50 % liegt (s. Tabelle 9). Die niedrigsten Quoten (Tabelle 9, grün) weisen die Schule 111 (52.8 %), die Schule 125 (55.0 %) und die Schule 119 (54.2 %) auf. Das Kollegium der Schule 112 (90.2 %), das Kollegium der Schule 128 (86.7 %) und das Kollegium der Schule 124 (85.7 %) weisen die höchsten Quoten auf (Tabelle 9, rot), d. h., dass hier die motivationale Ausgangsbedingung im Vergleich zu den anderen Schulen ungünstiger ist. Ein erheblicher Teil der Lehrpersonen an diesen Schulen passt sich also in gewisser Weise an den Mehrheitsentscheid an und ist unter anderem aus diesem eher extrinsischen Grund stark zur Teilnahme an *Vielfalt fördern* motiviert.

Tabelle 9. Schulquote für die Beteiligung aufgrund des Mehrheitsentscheids

Schule	110	111	112	113	114	115	116	117	118	119	
%	75.0	52.8	90.2	70.6	57.9	83.8	78.3	71.4	75.0	54.2	
Schule Fortsetzung	120	121	122	123	124	125	126	127	128	129	130
%	66.7	70.6	78.6	80.6	85.7	55.0	77.8	76.5	86.7	79.6	75.0

Anmerkung: rote Hervorhebung = besonders hohe Anpassungsquote, grüne Hervorhebung = besonders niedrige Anpassungsquote

In weiterführenden Analysen wurde der Anteil der Lehrpersonen im Kollegium, die diesem Item eher zugestimmt haben, als Prädiktor auf Schulebene genutzt, um Unterschiede im Erleben des

Fortbildungsprozesses (Fortbildungsklima und kooperative Umsetzung) und in Reaktionen (Zufriedenheit und selbstberichteter Lernertrag) auf die Fortbildung zu erklären.

Für die Auswertungen (s. Abschnitte 6.3.1–6.3.4) wurden Mehrebenenmodelle spezifiziert (s. Abschnitt 4.1.3), die ausschließlich die Anpassungsquote als Merkmal auf Schulebene (Ebene 2) enthalten, um Mittelwertunterschiede zwischen den Schulen zu erklären. Die Ergebnisse aus diesen Mehrebenenanalysen deuten darauf hin, dass sich ein höherer Anteil an Lehrpersonen, die sich an den Mehrheitsentscheid im Kollegium anpassen, ungünstig darauf auswirkt, wie Kollegien das Fortbildungsklima und die kooperativen Erprobungen erleben, wie zufrieden sie mit *Vielfalt fördern* sind und wie groß sie den Lernertrag bewerten.

6.3.1 Fortbildungsklima zwischen den Lehrpersonen

Im folgenden Abschnitt wird untersucht, inwieweit sich das von den Teilnehmerinnen und Teilnehmern wahrgenommene Klima im zweiten Modul durch die Anpassungsquote, d. h. den ermittelten Anteil von Lehrpersonen, die auch an der Fortbildung teilnehmen, weil sich die Schule hierfür entschieden hat, vorhersagen lässt.

Erfasst wurde das wahrgenommene Klima im Rahmen der Prozesserhebung von Modul 2 (t_3) mit insgesamt vier Items (Beispielitem: „In den Fortbildungssitzungen findet ein konstruktiver und freundlicher Austausch zwischen den Kollegen statt“, $\alpha = .90$, ICC-2₁₃ = .776). Es fällt auf, dass das Fortbildungsklima von den teilnehmenden Lehrpersonen mit einem Mittelwert von $M = 5.08$ ($SD = 0.82$) sehr positiv bewertet wird. Der ICC-1-Wert des Fortbildungsklimas beträgt 13.9. Dieser Wert ist so zu interpretieren, dass 13.9 % der Varianz im Fortbildungsklima durch Unterschiede zwischen Schulen erklärt werden können.

In Abbildung 23 ist das Ergebnis der Analyse abgebildet, durch die überprüft wurde, ob sich Unterschiede im Fortbildungsklima zwischen den teilnehmenden Schulen durch die Höhe des Anteils der Lehrpersonen im Kollegium erklären lassen, die maßgeblich aufgrund des schulischen Mehrheitsentscheids teilnehmen.

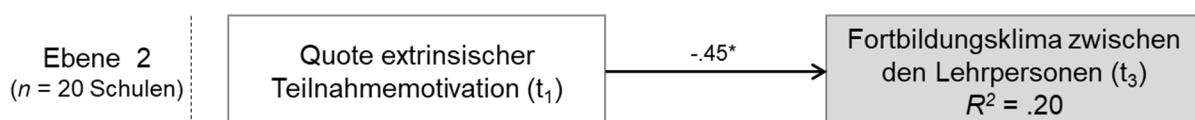


Abbildung 23. Vorhersage des Fortbildungsklimas durch die Anpassungsquote. Anmerkung: * = $p < .05$

Der negative Regressionskoeffizient Beta ($\beta = -.45$, $p < .05$) deutet darauf hin, dass sich ein größerer Anteil von Lehrpersonen im Kollegium, die auch aufgrund des Mehrheitsentscheids der Schule an *Vielfalt fördern* teilnehmen, negativ auf das Klima auswirkt, das die Kollegien in den Fortbildungssitzungen von Modul 2 wahrnehmen. Ungefähr 20 % der Unterschiede in der wahrgenommenen Atmosphäre lassen sich mit Unterschieden in der Anpassungsquote erklären.

6.3.2 Kooperative Erprobungen der Lehrpersonen zwischen den Sitzungen des zweiten und dritten Moduls

Das Merkmal kooperative Umsetzung/Erprobung von Fortbildungsinhalten zwischen den Sitzungen wurde bereits in Abschnitt 6.2 vorgestellt. Die ICC-1-Werte ($ICC-1_{t_3} = .128$; $ICC-1_{t_6} = .063$) deuten darauf hin, dass 12.8 % beziehungsweise 6.3 % der Unterschiede in dem Kriterium durch die Schulzugehörigkeit erklärt werden können. Die ICC-2-Werte zu den beiden Erhebungszeitpunkten zeigen, dass die Lehrpersonen die kooperativen Erprobungen zwischen den Fortbildungssitzungen durchaus kongruent beurteilt haben ($ICC-2_{t_3} = .753$; $ICC-2_{t_6} = .567$). Ein deskriptiver Vergleich der Mittelwerte zu den beiden Erhebungszeitpunkten (t_3 : $M = 3.37$, $SD = 1.16$; t_6 : $M = 3.40$, $SD = 1.13$) deutet auf eine nahezu identische Einschätzung durch die Kolleginnen und Kollegen hin.

Die Analysen zeigen für die kooperativen Erprobungen der Lehrpersonen zwischen Fortbildungssitzungen ein unterschiedliches Ergebnis für die beiden Erhebungszeitpunkte. Auf die gemeinsamen Erprobungen der Teilnehmerinnen und Teilnehmer während des zweiten Moduls wirkt sich die Anpassungsquote nicht signifikant aus ($\beta = -.24$, n. s., s. Abbildung 24), auch wenn das Vorzeichen des Koeffizienten negativ ist und der Effekt damit in die erwartete Richtung weist.

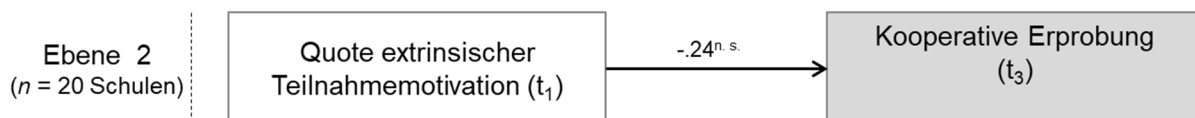


Abbildung 24. Vorhersage der kooperativen Erprobung zu t_3 durch die Anpassungsquote. Anmerkung: n. s. = nicht signifikant

Dagegen zeigt sich ein deutlich negativer Effekt des Anteils der Lehrpersonen, für die der Mehrheitsentscheid des Kollegiums eine wichtige Bedeutung für die eigene Motivation hatte, auf die Intensität der kooperativen Erprobung während der Module 3 und 4 ($\beta = -.71$, $p < .001$, s. Abbildung 25). Über den Anteil an Lehrpersonen, die sich mit ihrer Teilnahme an *Vielfalt fördern* stark nach der Entscheidung des Kollegiums richten, lassen sich ca. 50 % der Unterschiede in der Intensität der kooperativen Umsetzung und Erprobung von Fortbildungsinhalten während des dritten Moduls zwischen den teilnehmenden Schulen erklären. Je höher der Prozentanteil der Lehrpersonen ist, die zu Beginn der Fortbildung angeben, aufgrund des schulischen Mehrheitsentscheids teilzunehmen, desto geringer ist die berichtete Intensität in der Erprobung der Inhalte aus Modul 3 und 4. Der Effekt ist beträchtlich und auch deshalb bemerkenswert, weil zwischen der Messung der Anpassungsquote und der kooperativen Erprobung ein Zeitraum von ca. 18 Monaten¹² liegt.

¹² Der zeitliche Abstand zwischen den beiden Erhebungszeitpunkten t_1 und t_6 variiert je nach Schule zwischen 13 und 31 Monaten, da die Befragungen zeitlich an bestimmte Fortbildungsbausteine angebunden waren (s. Abschnitt 4.1.1).

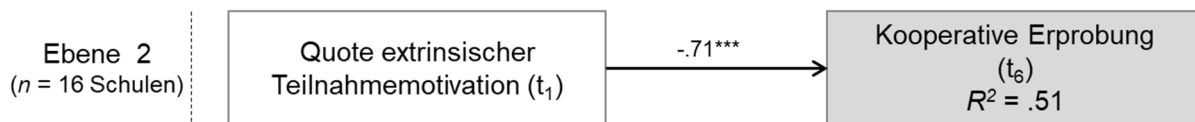


Abbildung 25. Vorhersage der kooperativen Erprobung zu t₆ durch die Anpassungsquote. Anmerkung: *** = $p < .001$

6.3.3 Zufriedenheit mit der Fortbildung

Als Wirksamkeitskriterium der ersten Ebene (s. Abschnitt 2.2) beziehungsweise als spontane Reaktion wurde unter anderem die durchschnittlich erlebte Zufriedenheit der Lehrpersonen mit dem Fortbildungsangebot analysiert. Die Zufriedenheit wurde zu beiden Erhebungszeitpunkten mit vier Items erfasst (Beispielitem: „Mit der bisherigen Fortbildung bin ich zufrieden“, $\alpha_{t_4} = .92$, $\alpha_{t_7} = .91$) und bezog sich generell auf die Fortbildung. Ein hoher Skalenmittelwert ist so zu interpretieren, dass die Teilnehmerinnen und Teilnehmer mit dem Angebot zufrieden sind und gerne teilnehmen, während ein niedriger Wert bedeutet, dass die Lehrpersonen die Fortbildung gerne schon vorzeitig beenden würden und eher unzufrieden sind. Die Zufriedenheit beträgt in der Gesamtstichprobe am Ende von Modul 2 durchschnittlich $M = 3.16$ ($SD = 1.28$) und am Ende des vierten Moduls geben die Lehrpersonen im Mittel $M = 3.01$ ($SD = 1.28$) an. Das bedeutet: Beide Werte liegen unter dem theoretischen Skalenmittel und sind demnach eher als niedrig einzustufen. Die ICC-1-Werte ($ICC-1_{t_4} = .286$; $ICC-1_{t_7} = .243$) zeigen, dass zu t₄ 28.6 % der Varianz in der Zufriedenheit und zu t₇ 24.3 % der Varianz durch Merkmale auf der Schulebene erklärt werden können.

Auch hier zeigt sich (s. Abbildung 26), dass über den Anteil von Lehrpersonen im Kollegium, für die der Mehrheitsentscheid des Kollegiums mit ausschlaggebend dafür war, an der Fortbildung teilzunehmen, Unterschiede zwischen Schulen hinsichtlich der Zufriedenheit erklärt werden können ($\beta = -.42$, $p < .05$). Ein höherer Anteil an Lehrpersonen im Kollegium, die auch wegen des Mehrheitsentscheids teilnehmen, geht mit einer signifikant niedrigeren Fortbildungszufriedenheit am Ende von Modul 2 einher. Hier sei daran erinnert, dass zwischen der Messung der Teilnahmemotivation zu t₁ und der Messung der Zufriedenheit ca. zehn Monate liegen.

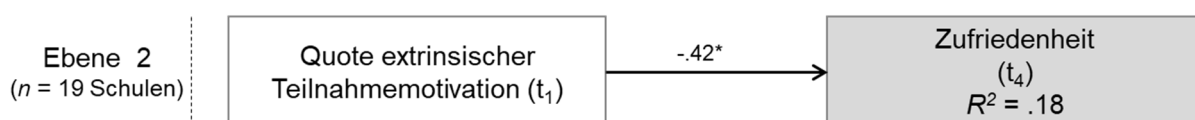


Abbildung 26. Vorhersage der Zufriedenheit zu t₄ durch die Anpassungsquote. Anmerkung: * = $p < .05$

Aus Abbildung 27 geht hervor, dass der anfängliche Anteil der Lehrpersonen, die auch aufgrund des schulischen Mehrheitsentscheids an *Vielfalt fördern* teilnehmen, hingegen keinen signifikanten Einfluss auf die Zufriedenheit am Ende von *Vielfalt fördern* hat ($\beta = -.40$, n. s.).

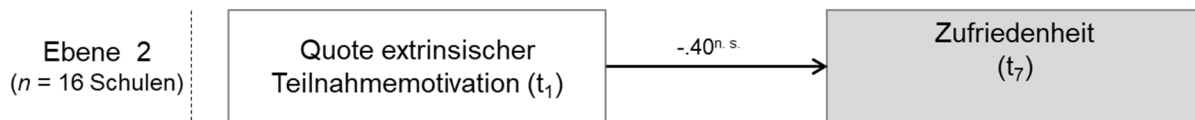


Abbildung 27. Vorhersage der Zufriedenheit zu t₇ durch die Anpassungsquote. Anmerkung: n. s. = nicht signifikant

6.3.4 Selbstberichteter Lernertrag

Der selbstberichtete Lernertrag erfasst, ob die Lehrpersonen den Eindruck haben, durch die Teilnahme an *Vielfalt fördern* Neues gelernt und gewinnbringende Impulse erhalten zu haben. Er wurde zu t₄, also nach dem Modul zur Diagnostik, und zu t₇, also nach dem vierten Modul und somit am Ende der Fortbildung, gemessen. Ein hoher Skalenmittelwert ist als hoher selbstberichteter Lernertrag zu interpretieren und entsprechend weist ein geringer Skalenmittelwert darauf hin, dass die Fortbildungsteilnehmerinnen und -teilnehmer den Eindruck haben, durch die Teilnahme an *Vielfalt fördern* nicht viel Neues oder Hilfreiches für den schulischen Alltag gelernt zu haben. Ein Item zu beiden Erhebungszeitpunkten, d. h. am Ende von Modul 2 und am Ende von Modul 4, lautet beispielsweise: „Durch die Fortbildung habe ich viele Impulse erhalten, die mir etwas für meinen beruflichen Alltag bringen.“ Die Skala, die zu beiden Erhebungszeitpunkten aus vier Items besteht, weist eine gute Reliabilität auf ($\alpha_{t_4} = .87$, $\alpha_{t_7} = .85$). Wie auch die Zufriedenheitswerte liegen die Mittelwerte für den selbstberichteten Lernertrag zu beiden Erhebungszeitpunkten unter dem theoretischen Skalenmittel, sodass die Lehrpersonen eher zu einer kritischen Einschätzung gelangen (t₄: $M = 3.13$, $SD = 1.13$; t₇: $M = 3.08$, $SD = 1.09$). Es lassen sich 15.4 % (t₄) beziehungsweise 13.4 % (t₇) der Unterschiede in der Zufriedenheit durch Merkmale auf Schulebene erklären ($ICC-1_{t_4} = .154$; $ICC-1_{t_7} = .134$).

Die erste Analyse zeigt, dass der Anteil der Lehrpersonen innerhalb eines Kollegiums, die insbesondere an *Vielfalt fördern* teilnehmen, weil es einen entsprechenden schulischen Mehrheitsentscheid gab, den selbstberichteten Lernertrag am Ende von Modul 2 nicht signifikant beeinflusst ($\beta = -.39$, n. s., s. Abbildung 28).

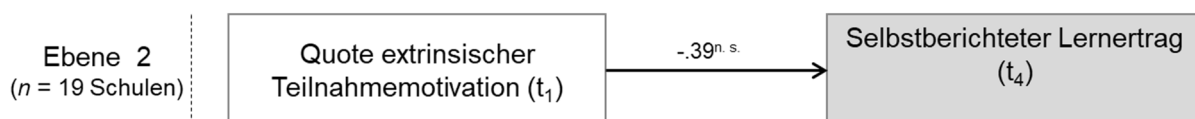


Abbildung 28. Vorhersage des selbstberichteten Lernertrags zu t₄ durch die Anpassungsquote. Anmerkung: n. s. = nicht signifikant

Demgegenüber zeigt sich aber ein Effekt auf den selbstberichteten Lernertrag zu t₇ (s. Abbildung 29). Lehrpersonen schätzen am Ende von Modul 4 ihren Lernertrag und Kompetenzzuwachs geringer ein, wenn zu Beginn der Fortbildung – also ca. 1.5–2 Jahre zuvor – viele Kolleginnen und Kollegen in der eigenen Schule auch deshalb an der Fortbildung teilnahmen, weil es einen Mehrheitsentscheid gab ($\beta = -.58$, $p < .05$). Diese Variable erklärt immerhin 34 % der Unterschiede im mittleren selbstberichteten Lernertrag zwischen den Schulen.

Das ist auch deshalb ein interessanter Befund, weil dieser Zurückhaltung des Kollegiums offenbar vor allem langfristig – wenn über eine lange Zeit ein hohes Maß an Ausdauer und Anstrengung investiert werden muss – eine hohe Bedeutung zukommt.

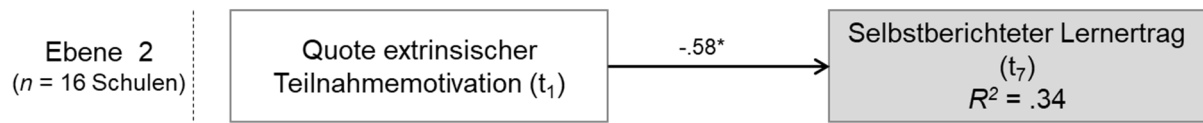


Abbildung 29. Vorhersage des selbstberichteten Lernertrags zu t₇ durch die Anpassungsquote. Anmerkung:
* = $p < .05$

7 Veränderung der Zusammenarbeit im Kollegium

Die professionelle Kooperation im Kollegium kann zugleich eine zentrale *Voraussetzung* für eine gemeinsame und systematische Unterrichtsentwicklung und ein *Ziel* von Schulentwicklungsprozessen sein (Holtappels, 2013). Im Zusammenhang von *Vielfalt fördern* sind beide von Holtappels (2013) genannten Aspekte wiederzufinden: Einerseits soll die Kooperation der Lehrpersonen durch das erste Modul zur Teambildung systematisch und zielgerichtet weiterentwickelt werden, um günstigere Rahmenbedingungen für die gemeinsame Weiterentwicklung von Schule und Unterricht anzubahnen. Durch die kontinuierliche gemeinsame Teilnahme an der Fortbildung in den Modulen 2 – 4 ist außerdem die Entwicklung einer professionellen Kooperation im Sinne eines „Mitnahmeeffektes“ plausibel. Andererseits stellen die Kooperationserfahrung und -praxis von Lehrpersonen eines Kollegiums, die vor *Vielfalt fördern* bestanden und/oder durch das erste Modul *Vielfalt fördern* gewonnen wurden, eine Rahmenbedingung für die Wahrnehmung und Nutzung des Fortbildungsangebots dar.

Im Folgenden werden zunächst die Ergebnisse von Analysen vorgestellt, bei denen einzelne Schulen genauer in den Blick genommen wurden (s. Abschnitt 7.1.1 und Abschnitt 7.1.2). Der Fokus liegt dabei nicht auf der Kooperation als Rahmenbedingung, sondern auf der Kooperation als eine Zielvariable der Fortbildung. Neben zwei Kooperationsfacetten, die dem Konzept der professionellen Lerngemeinschaft zugeordnet werden können, wird auch die kollektive Selbstwirksamkeitserwartung der Schulen untersucht, d. h. deren Überzeugung, herausfordernde Situationen aufgrund der im Kollegium vorhandenen Ressourcen erfolgreich bewältigen zu können.

Anschließend wird in den Abschnitten 7.2.1 und 7.2.2 erläutert, ob sich die Interaktionsfrequenz der Lehrpersonen eines Kollegiums intensiviert hat und ob es zu einer strukturellen Einrichtung spezifischer Formen der Kooperation während des Fortbildungszeitraums kam.

7.1 Facetten der Lehrerkooperation

Bei den folgenden Analysen wurde induktiv vorgegangen. Damit ist gemeint, dass zunächst Schulen ermittelt wurden, die sich im Mittel durch besonders positive Entwicklungen bei bestimmten Merkmalen des Fortbildungsertrags auszeichnen. Diese Schulen, bei denen sich die Teilnahme an *Vielfalt fördern* als wirksam herausgestellt hat, wurden anschließend mit Schulen ohne nachweislichen Lernertrag oder mit einer negativen Entwicklung kontrastiert. Um einen Indikator für die Entwicklung der Schulen zu erhalten, wurde die Differenz aus dem Skalenmittelwert eines Merkmals zu den unterschiedlichen Erhebungszeitpunkten gebildet. Die Skalen, mit denen ein

Merkmal erfasst wurde, bestehen zu den unterschiedlichen Erhebungszeitpunkten aus den gleichen Items.¹³

Als ein Indikator des Fortbildungsertrags wurde die Entwicklung auf der Kooperationsdimension *Fokus auf das Lernen der Schülerinnen und Schüler*, bei der es sich um eine bedeutsame Eigenschaft von professionellen Lerngemeinschaften sowie um ein vielversprechendes Merkmal von wirksamen Lehrerfortbildungen handelt, betrachtet. Da bei der pädagogischen Diagnostik, die Inhalt des zweiten Fortbildungsmoduls ist, ein Blick auf die Lernprozesse von Schülerinnen und Schülern erforderlich ist, wurde als Entwicklungszeitraum die Laufzeit des zweiten Moduls in den Blick genommen, wobei die hier herangezogenen Postwerte aus der t_5 -Erhebung am Anfang von Modul 3 stammen. Die Skala setzt sich aus drei Items zusammen ($\alpha_{t_2} = .62$, $\alpha_{t_5} = .74$). Ein Beispielitem lautet: „*Wir sprechen uns ab, wie wir gemeinsam die Förderung einzelner Schüler unterstützen können.*“ Am Anfang von Modul 2 liegt der Skalenmittelwert bei $M = 4.20$ ($SD = 0.79$) und am Anfang von Modul 3 bei $M = 4.54$ ($SD = 0.84$), was zu beiden Zeitpunkten auf eine tendenziell eher hoch ausgeprägte Zusammenarbeit mit einem Fokus auf das Lernen der Schülerinnen und Schüler hindeutet.

Darüber hinaus wurden die Entwicklung der Deprivatisierung, d. h. die kollegiale Praxis, im Unterricht zu hospitieren und sich gegenseitig Rückmeldungen zum Unterricht zu geben, ab dem Anfang des zweiten Moduls bis zum Ende der Fortbildung (t_2 – t_7) in den Blick genommen. Bonsen und Rolff (2006) gelangten zu dem Ergebnis, dass die Deprivatisierung von Unterricht einen positiven Einfluss darauf hat, wie Lehrpersonen beispielsweise kollegiales Feedback und Anerkennung im eigenen Kollegium wahrnehmen. Im Projekt *LIQUID* wurde die Deprivatisierung mit drei Items gemessen (Beispielitem: „*Gegenseitige Unterrichtsbesuche sind ein selbstverständlicher Teil unserer Arbeit*“, $\alpha_{t_2} = .74$, $\alpha_{t_7} = .83$). Sowohl am Anfang von Modul 2 ($M = 3.25$, $SD = 1.05$) als auch am Ende der Fortbildung ($M = 3.46$, $SD = 1.22$) ist die Deprivatisierungspraxis bei den Lehrpersonen eher gering ausgeprägt, da der Skalenmittelwert der Gesamtstichprobe zu beiden Erhebungszeitpunkten unter dem theoretischen Skalenmittel von 3.50 liegt.

Die kollektive Selbstwirksamkeitserwartung erfasst die subjektive Annahme einer Lehrperson, wie gut das Lehrerkollegium in der Lage ist, mit zukünftigen Anforderungen umzugehen (Schwarzer & Warner, 2014). Es wird davon ausgegangen, dass Erfahrungen aus früheren Situationen die Selbstwirksamkeitserwartungen beeinflussen (Köller & Möller, 2010). Die Selbstwirksamkeitserwartung lässt sich grundsätzlich als Überzeugung begreifen, zukünftige Anforderungen aufgrund vorhandener Fähigkeiten und Ressourcen bewältigen zu können (Fröhlich, 2010; Schaarschmidt, 2001). Häufig wird die Selbstwirksamkeitsüberzeugung als handlungsbezogener

¹³ Ob Konstrukte, die zu mehreren Messzeitpunkten mit identischen Items erfasst wurden – hier die Kooperationsfacetten –, auch wirklich das inhaltlich gleiche Merkmal messen, wird in der Regel mit sogenannten Invarianzanalysen geprüft. Dieses Verfahren konnte für die hier verwendeten Skalen jedoch nicht abschließend durchgeführt werden.

Optimismus umschrieben, der sich nach Schwarzer und Warner (2014) nicht auf Routineaufgaben bezieht, sondern auf Anforderungen, deren Erfüllung mit Anstrengung und Ausdauer verbunden sind. In den Blick genommen werden die kollektive Selbstwirksamkeitserwartung am Anfang von Modul 2 (t_2) und am Anfang von Modul 3 (t_5). Ein Beispielitem der aus insgesamt sechs Items bestehenden Skala zur Erfassung der kollektiven Selbstwirksamkeitserwartung lautet: „Wie sicher sind Sie, dass Sie es als Kollegium schaffen, pädagogisch wertvolle Arbeit zu leisten, auch wenn die Rahmenbedingungen schwierig sind?“ ($\alpha_{t_2} = .87$, $\alpha_{t_7} = .93$). Der Mittelwert für die Gesamtstichprobe beträgt am Anfang von Modul 2 $M = 3.75$ ($SD = 0.84$) und liegt am Ende der Fortbildung etwas höher bei $M = 3.98$ ($SD = 0.99$).

7.1.1 Deskriptive Betrachtung der Veränderungswerte aller Schulen

Die durchschnittlichen Entwicklungen, die sich bei den drei genannten Merkmalen an Teilnahmeschulen von *Vielfalt fördern* zeigen, sind in Tabelle 10 dargestellt. Farblich hervorgehoben sind jeweils die drei Schulen mit den negativsten oder geringsten durchschnittlichen Entwicklungen (*rot*) sowie die drei Schulen mit den positivsten mittleren Entwicklungen (*grün*). Bei negativen oder ungünstigen Entwicklungen liegt ein geringer oder ein negativer Differenzwert zugrunde, welcher dadurch zustande kommt, dass der zweite Messwert geringer ist als der erste Messwert. Bei positiven Entwicklungen ist der zweite Messwert deutlich größer als der erste Wert.

Tabelle 10. Schulische Differenzwerte für den Fokus auf das Lernen der Schülerinnen und Schüler, die Deprivatisierung und die kollektive Selbstwirksamkeitserwartung des Kollegiums

Schule	110	111	112	113	114	115	116	117	118	119	
FLSUS	0.19	0.48	0.51	0.40	0.13	-0.19	0.16	Daten Keine Daten	Keine Daten Keine	0.17	
DEP	-0.10	0.70	Daten Keine	0.11	0.10	0.60	0.37			0.19	
KSWK	0.27	0.19		0.43	0.87	0.02	-0.24			-0.13	
Schule Fortsetzung	120	121	122	123	124	125	126	127	128	129	130
FLSUS	0.32	Daten Keine	0.90	0.42	0.44	0.43	0.43	0.54	Daten Keine	0.20	0.19
DEP	-0.07		0.11	0.08	0.49	-0.37	-0.25	0.35		0.01	0.18
KSWK	0.51		0.25	0.27	0.59	0.27	0.11	0.46		-0.07	0.12

Anmerkung: rote Hervorhebung = besonders negative oder geringe Entwicklung, grüne Hervorhebung = besonders positive Entwicklung, FLSUS = Fokus auf das Lernen der Schülerinnen und Schüler, DEP = Deprivatisierung, KSWK = kollektive Selbstwirksamkeitserwartung

An dem Farbmuster in Tabelle 10 lässt sich erkennen, dass keine der Schulen besonders positiv oder negativ auffällt, indem sie ausschließlich positive oder ungünstige mittlere Entwicklungen aufzeigt. Vielmehr wird deutlich, dass eine Schule bei einem Merkmal besonders günstige Entwicklungen zeigen kann und bei einem anderen Merkmal eher geringe Veränderungen. So weist die Schule 114 die positivste Entwicklung bei der kollektiven Selbstwirksamkeitserwartung auf,

während der Fokus auf das Lernen der Schülerinnen und Schüler, im Vergleich mit den anderen Schulen der UG, in dem Kollegium der Schule 114 mit am geringsten zunimmt.

7.1.2 Deskriptive Betrachtung deutlich hervortretender Schulen

Im Folgenden wird ein Vergleich zwischen drei Schulen, bei denen sich im Mittel keine/geringe Veränderungen zeigten, und drei Schulen, bei denen sich die Mittelwerte über die Zeit hinweg positiv entwickelten, angestellt.

Es wird die Ausgangslage der Schulen in den Blick genommen, um deskriptiv zu betrachten, ob ausbleibende oder geringe Entwicklungen bei Schulen damit zusammenhängen, dass die Kollegien in dem Bereich bereits vor der Teilnahme an *Vielfalt fördern* weit entwickelt sind, d. h. hohe Ausgangswerte aufweisen, oder ob sich eher ein Matthäuseffekt andeutet, d. h. Schulen mit günstigeren Ausgangsbedingungen positivere Entwicklungen zeigen. Neben dem Mittelwert wird auch die Streuung in den Schulen untersucht, um einen Hinweis darauf zu erhalten, ob große Unterschiede in der Wahrnehmung eines Merkmals zwischen den Lehrpersonen einer Schule in einem Zusammenhang mit deren Entwicklung stehen.

Um zu prüfen, welche Merkmale zwischen den Schulen trennen und demnach die Zugehörigkeit zu einer der beiden Gruppen bedingt haben könnten, werden Merkmale der Fortbildungswahrnehmung und -nutzung herangezogen. Betrachtet werden die Mittelwerte für das Fortbildungsklima (t_3) und die kooperative Umsetzung der Fortbildungsinhalte (t_3). Beide Merkmale wurden in vorigen Abschnitten bereits erläutert (s. Abschnitt 6.2 und Abschnitt 6.3). Anzunehmen wäre, dass Schulen, in denen das Klima untereinander positiver erlebt wird und eine intensivere gemeinsame Umsetzung und Erprobung der Fortbildungsinhalte während der einzelnen Sitzungen stattfindet, eine günstigere Entwicklung zeigen.

Zudem wird die mittlere Wahrnehmung des Moderatorenteams (t_3) in den Kollegien der Extremgruppen dargestellt. Mit der Skala *wahrgenommene Moderatorenkompetenz* wurde über 14 Items erfasst, wie die Lehrpersonen den Umgang des Moderatorenteams mit Rückfragen der teilnehmenden Lehrpersonen, die Anschaulichkeit und Beispiele der Moderatorinnen und Moderatoren, deren Sicherheit und Klarheit bei der Präsentation, aber auch die Kooperation und das Zusammenspiel des Moderatorenteams untereinander erlebt haben (Beispielitem: „Die Moderatoren reagieren/antworten auf Fragen kompetent“, $\alpha = .97$). Im Mittel nehmen die Lehrperson ihr Moderatorenteam eher als kompetent im Umgang mit Rückfragen und in der Präsentation der Inhalte wahr. Es scheint den Moderatorenteams demnach zu gelingen, mit geeigneten Beispielen aus der unterrichtlichen Praxis die Umsetzung der Fortbildungsinhalte zu veranschaulichen ($M = 4.23$, $SD = 1.08$). Vor diesem Hintergrund erscheint es plausibel, dass sich Schulen, in denen eine höhere Kompetenz des Moderatorenteams erlebt wird, über den Fortbildungsverlauf hinweg deutlicher weiterentwickeln.

7.1.2.1 Veränderungen beim Fokus auf das Lernen der Schülerinnen und Schüler

Bei dem Kooperationsaspekt Fokus auf das Lernen der Schülerinnen und Schüler zeigten die Kolleginnen und Kollegen der Schulen 114, 115 und 116 eine negative oder sehr geringe Entwicklung. Diese Kollegien werden im Folgenden mit den Schulen 112, 122 und 127 kontrastiert, bei denen eine positive Veränderung bei der Zusammenarbeit unter den Lehrpersonen festgestellt werden konnte (s. Tabelle 10).

Die jeweiligen Mittelwerte und Standardabweichungen der Schulen „ohne Veränderung“ werden in den nachfolgenden Abbildungen (Abbildung 30–41) als graue Balken dargestellt und die Schulen mit „positiven Veränderungen“ sind durch blaue Balken abgebildet.

Betrachtet man bei den Extremgruppenschulen (s. Abbildung 30) die Streuung beziehungsweise den Konsens der Lehrpersonen innerhalb einer Schule, der aus der Länge der Whisker (auch: Antennen) hervorgeht, fallen keine bedeutsamen Unterschiede auf. Die Mittelwerte liegen in den Schulen in beiden Gruppen über dem theoretischen Skalenmittel von $M = 3.50$. Auffallend ist jedoch, dass an zwei Schulen (114, 115), die über die Zeit hinweg keine deutliche Entwicklung beziehungsweise eine leicht negative Entwicklung beim Fokus auf das Lernen der Schülerinnen und Schüler zeigen, das Ausgangsniveau dieser Kooperationsfacette höher ausfällt als an den anderen Schulen.

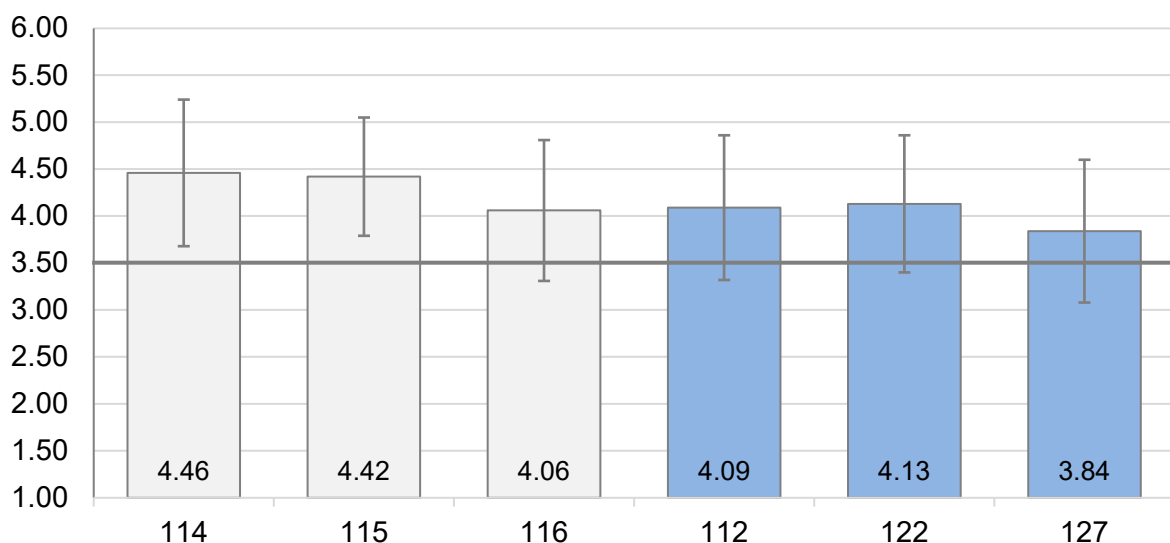


Abbildung 30. Mittelwerte und Standardabweichungen für die Skala Fokus auf das Lernen der Schülerinnen und Schüler zu t_2 (Ausgangswerte). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Das Klima unter den Kolleginnen und Kollegen während der Fortbildungssitzungen von Modul 2 (t_3) wird von den Lehrpersonen aller sechs Schulen sehr positiv beurteilt (s. Abbildung 31). Zwischen den Schulen, die aufgrund ihrer Entwicklungen bei der Kooperationsfacette Fokus auf das Lernen der Schülerinnen und Schüler einer der Extremgruppen zugeordnet wurden, lässt sich kein eindeutiger Unterschied feststellen.

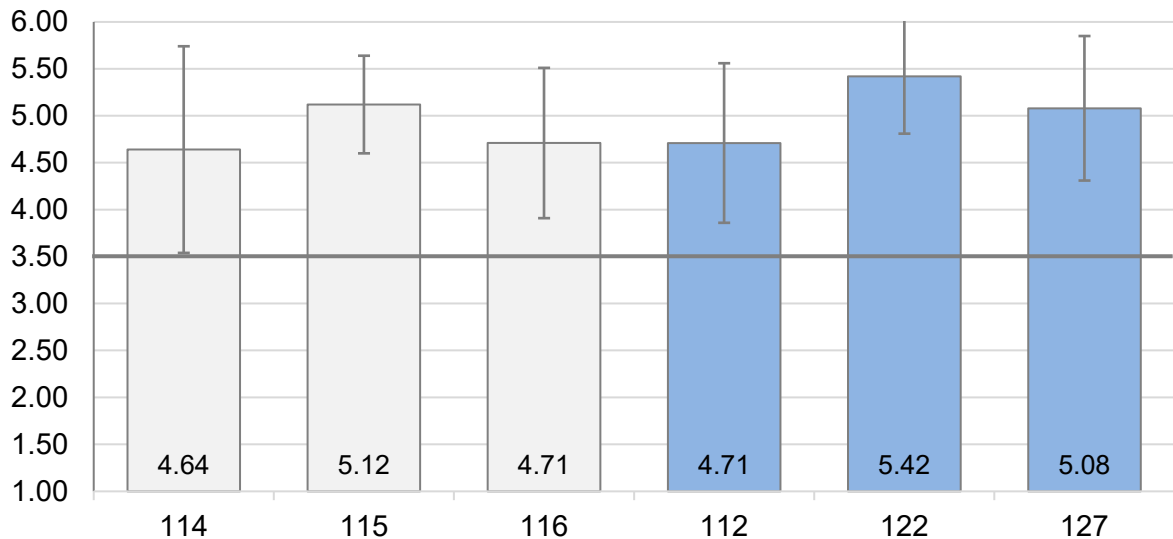


Abbildung 31. Mittelwerte und Standardabweichungen für das Fortbildungsklima zwischen den Lehrpersonen in Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Eine weitere Annahme ist, dass sich die Schulen mit positiven Entwicklungen beim Fokus auf das Lernen der Schülerinnen und Schüler und die Schulen, die keine oder negative Veränderungen aufweisen, darin unterscheiden, wie intensiv die teilnehmenden Lehrpersonen die Fortbildungsinhalte gemeinsam zwischen den Fortbildungssitzungen erprobt hatten. Insgesamt zeigt sich, dass die kooperative Erprobung der Fortbildungsinhalte, die in der Prozesserhebung zu Modul 2 (t_3) erfasst wurde, eher weniger intensiv erfolgte, da die Mittelwerte überwiegend unter dem theoretischen Mittel von $M = 3.50$ liegen (s. Abbildung 32). Es fallen beim Vergleich keine eindeutigen Unterschiede zwischen den Schulen beider Extremgruppen auf. Besonders tritt vielmehr die hohe Streuung innerhalb der verschiedenen Kollegien hervor sowie der besonders niedrige Mittelwert von $M = 2.59$ der Schule 112.

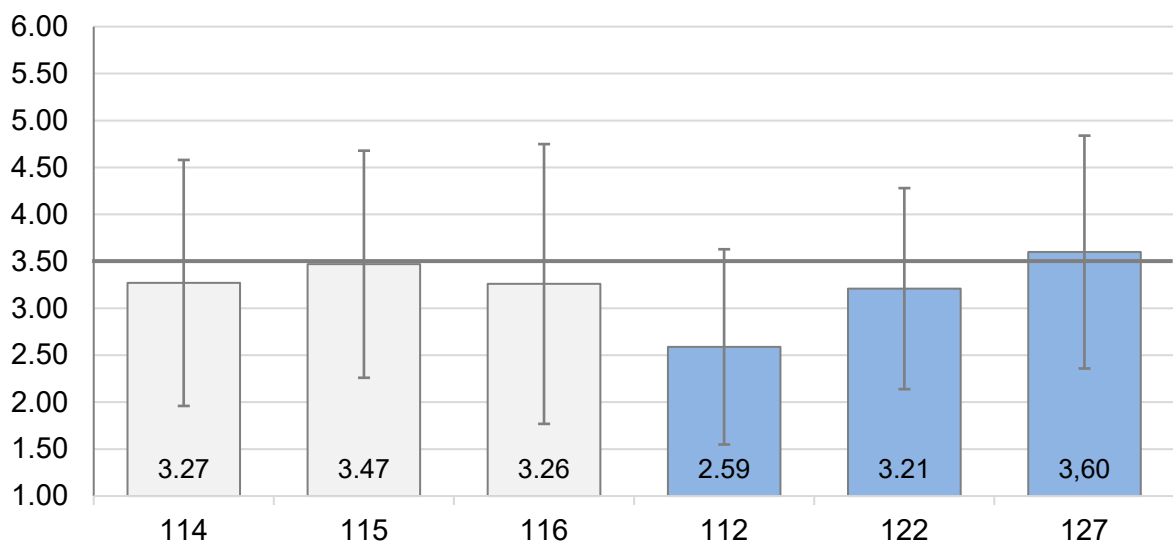


Abbildung 32. Mittelwerte und Standardabweichungen für die kooperative Umsetzung während Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Bei der Wahrnehmung des Moderatorenteams während des zweiten Moduls (t_3) zeigt sich zwischen den Schulen teilweise ein differenziertes Bild (s. Abbildung 33). Die Fortbildungsteilnehmerinnen und -teilnehmer der Schule 115 ($M = 3.54$) und der Schule 112 ($M = 3.29$) beurteilen ihr jeweiliges Moderatorenteam eher kritisch, während die Moderatorenteams an den anderen Schulen eher als kompetent im Umgang mit Rückfragen und als sicher und verständlich bei der Präsentation der Fortbildungsinhalte erlebt werden. Da die Schule 115 allerdings eine negative Entwicklung beim Fokus auf das Lernen der Schülerinnen und Schüler aufweist und die Schule 112 eine positive Veränderung, trennt auch die wahrgenommene Moderatorenkompetenz nicht zwischen beiden Extremgruppen.

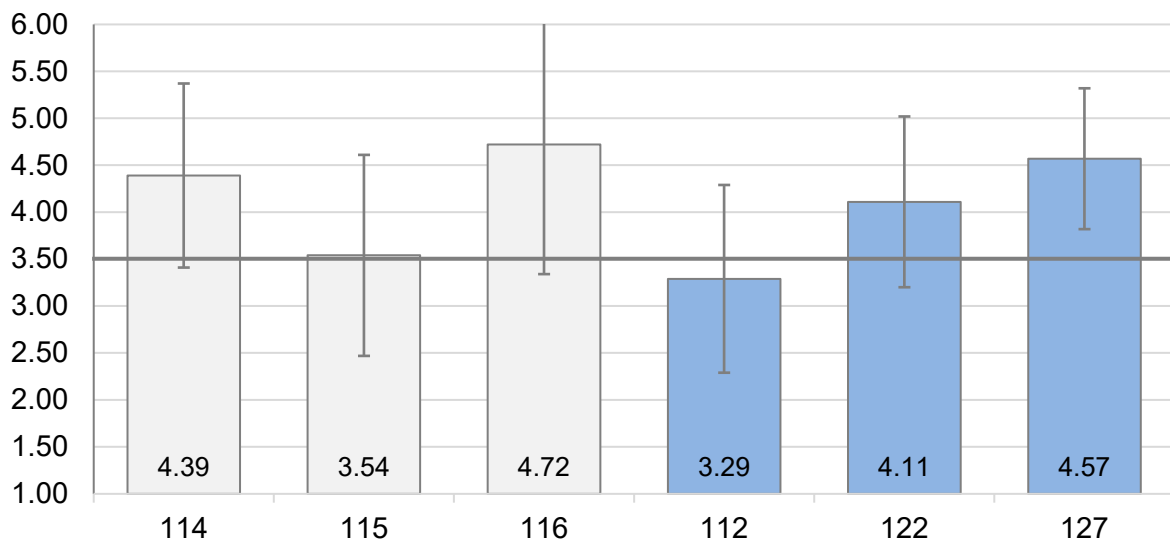


Abbildung 33: Mittelwerte und Standardabweichungen für die wahrgenommene Moderatorenkompetenz in Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

7.1.2.2 Veränderungen der Deprivatisierungspraxis

Bei der Deprivatisierungspraxis zeigen die Schulen 110, 125 und 126 vom Anfang des zweiten Moduls bis zum Ende der Fortbildung keine oder eine negative Entwicklung und die Schulen 111, 115 und 124 eine positive Entwicklung (s. Tabelle 10).

Die Deprivatisierungspraxis ist in den Schulen am Anfang von Modul 2 eher gering ausgeprägt (s. Abbildung 34). Eine Ausnahme bildet das Kollegium der Schule 124 ($M = 4.27$), welches auch eine positive Entwicklung der Deprivatisierungspraxis zeigt. Obwohl sich die Lehrpersonen der Schule 111 ($M = 3.24$) und der Schule 115 ($M = 3.29$) im Mittel eher selten gegenseitig im Unterricht besuchen und Rückmeldungen zur Unterrichtspraxis geben, zählen die Mittelwerte im schulischen Vergleich eher zu den hohen Werten.

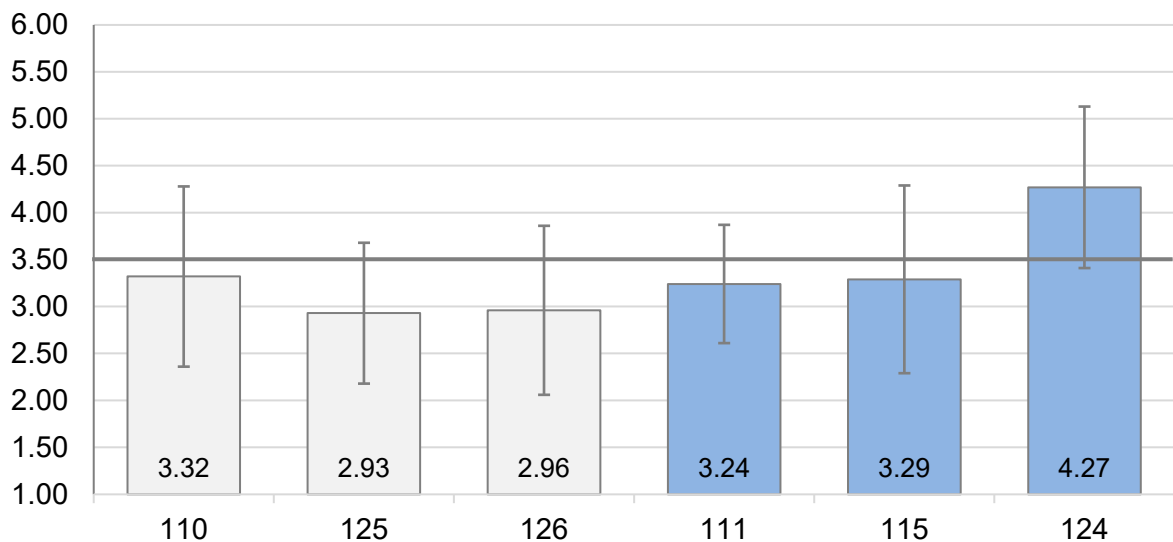


Abbildung 34. Mittelwerte und Standardabweichungen für die Deprivatisierung zu t_2 (Ausgangswerte). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Bei der Wahrnehmung des Fortbildungsklimas innerhalb des eigenen Kollegiums fallen bei den Schulen, die aufgrund ihrer Entwicklungen bei der Deprivatisierung einer der beiden Extremgruppen zugeordnet wurden, keine bedeutsamen Unterschiede auf (s. Abbildung 35). Das Fortbildungsklima wird sowohl in den Schulen (110, 125, 126), bei denen keine Öffnung des Unterrichts für kollegiale Hospitationen und Ratschläge erfolgte, als entspannt und vertrauensvoll erlebt als auch in den Schulen (111, 115, 124), in denen eine zunehmende Deprivatisierung der unterrichtlichen Praxis festgestellt werden konnte.

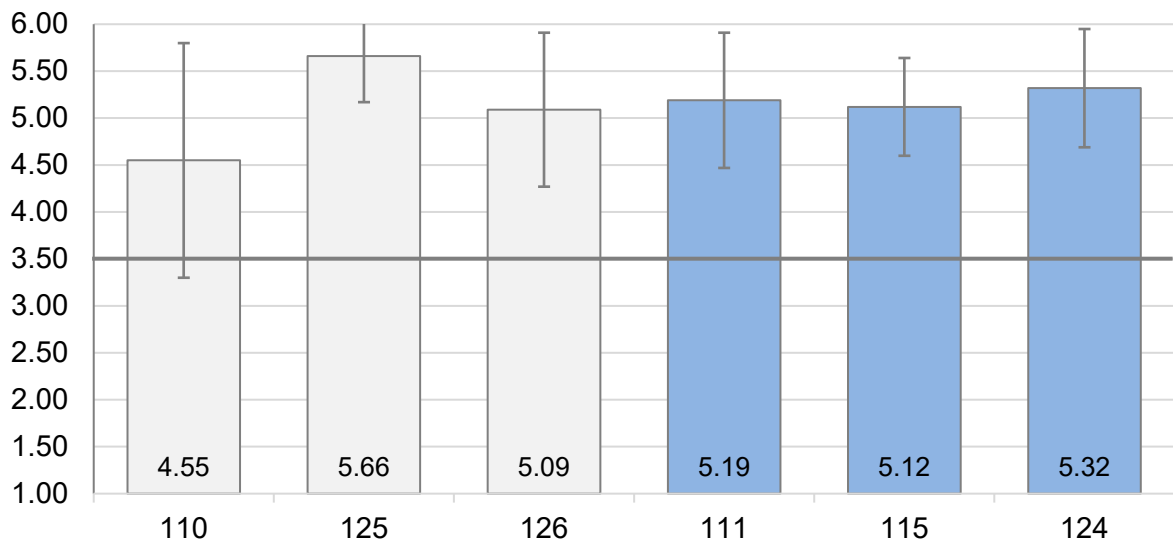


Abbildung 35. Mittelwerte und Standardabweichungen für das Fortbildungsklima zwischen den Lehrpersonen in Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Insgesamt findet eine kooperative Erprobung zwischen den Fortbildungssitzungen an den Schulen eher in einem geringen Maß statt (s. Abbildung 36). Ein eindeutiger Unterschied lässt sich

zwischen den beiden Gruppen (negative/keine Entwicklung vs. positive Entwicklung) deskriptiv nicht ausmachen. Die Intensität der kooperativen Erprobung von Fortbildungsinhalten zwischen den Fortbildungsbausteinen ist an zwei Schulen (110, 126), die eine negative mittlere Entwicklung auf der Skala Deprivatisierung zeigen, im Vergleich mit den anderen Schulen eher gering ausgeprägt. Gleichzeitig berichten die Lehrpersonen der Schule 125, bei der sich die Deprivatisierung ebenfalls negativ entwickelt hat, aber eine durchschnittliche kooperative Umsetzung von Fortbildungsinhalten, die mit der Intensität von Schule 115 (positive Entwicklung) vergleichbar ist und sogar über dem Mittelwert von Schule 111 (positive Entwicklung) liegt.

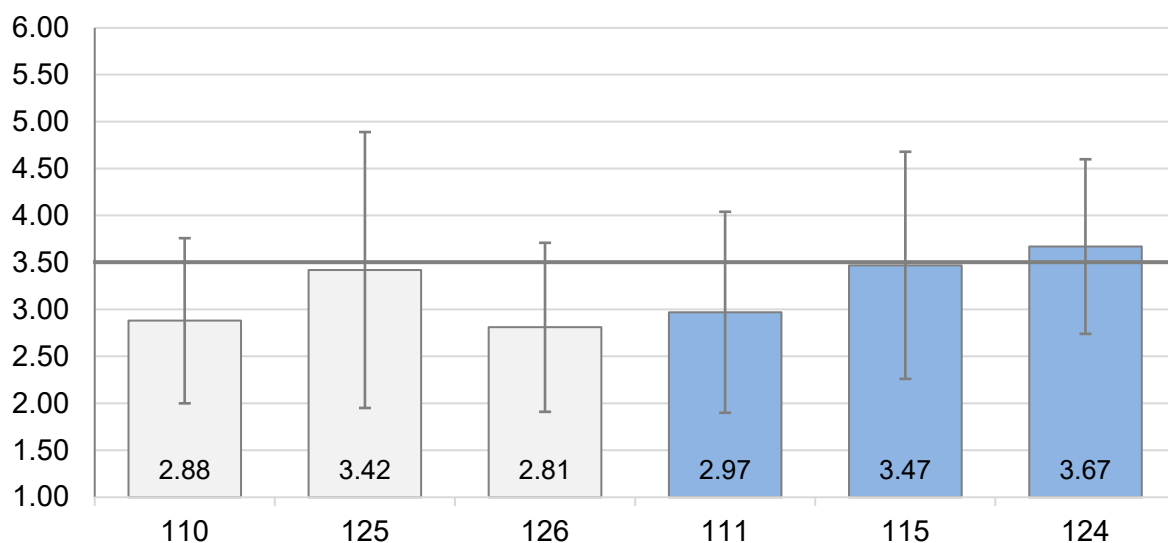


Abbildung 36. Mittelwerte und Standardabweichungen für die kooperative Umsetzung während Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Ein Zusammenhang zwischen der von den Lehrpersonen wahrgenommenen Moderatorenkompetenz und den Entwicklungsverläufen der Deprivatisierung ist ebenfalls nicht ersichtlich (s. Abbildung 37). Sowohl Schulen, die ihre Deprivatisierungspraxis positiv weiterentwickelt haben (blau), als auch Schulen, bei denen keine positiven Veränderungen festgestellt werden konnten (grau), beurteilen ihr Moderatorenteam im Mittel eher sehr positiv. Das Moderatorenteam an der Schule 115 ($M = 3.54$) und an der Schule 110 ($M = 3.48$) wird beispielsweise sehr ähnlich wahrgenommen, aber die Schule 115 hat ihre Deprivatisierungspraxis während *Vielfalt fördern* positiv weiterentwickelt, während für das Kollegium der Schule 110 im Mittel eine negative Entwicklung konstatiert werden musste.

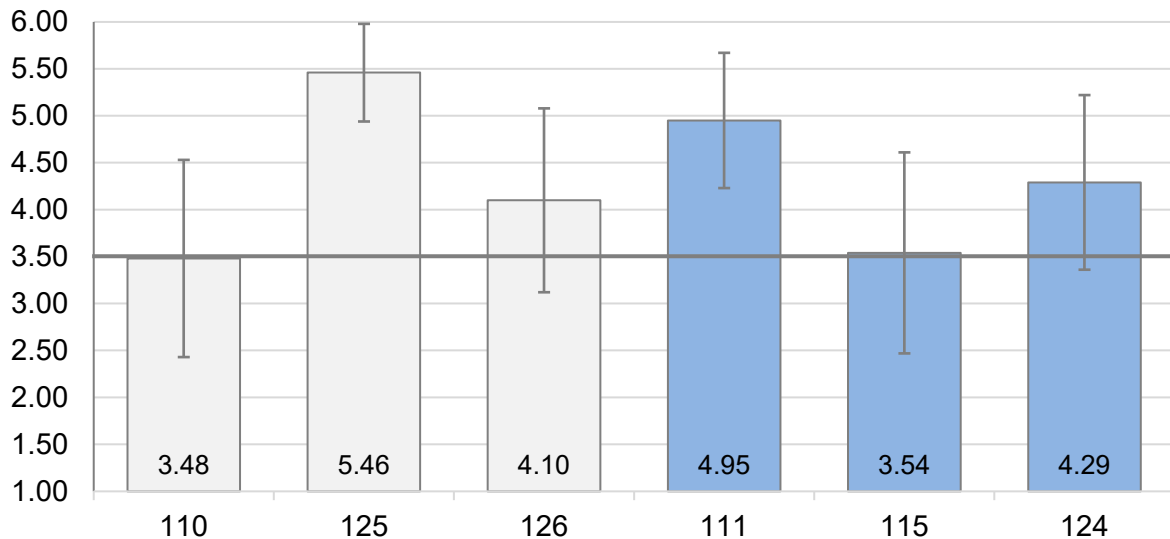


Abbildung 37. Mittelwerte und Standardabweichungen für die wahrgenommene Moderatorenkompetenz in Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Auch für unterschiedliche Verläufe in der Deprivatisierung erweisen sich die hier untersuchten Merkmale – Fortbildungsklima, Intensität der kooperativen Umsetzung und Moderatorenkompetenz – offenbar als nicht relevant.

7.1.2.3 Veränderungen der kollektiven Selbstwirksamkeitserwartung

Im Folgenden wird geprüft, ob das Fortbildungsklima, die Intensität der kooperativen Umsetzung und die wahrgenommene Moderatorenkompetenz zwischen denjenigen Schulen trennen kann, die einen hohen vs. geringen Zuwachs der kollektiven Selbstwirksamkeitserwartung aufweisen. Hierzu werden die Schulen 116, 119 und 129, die bezogen auf die kollektive Selbstwirksamkeitserwartung die Extremgruppe „negative Entwicklung“ repräsentieren, und die Schulen 114, 120 und 124, die zur Gruppe „positive Entwicklung“ gehören, verglichen (s. Tabelle 10).

Die Ausgangslage bei der kollektiven Selbstwirksamkeitserwartung, die am Anfang von Modul 2 (t_2) erfasst wurde, ist zwischen den beiden Extremgruppen vergleichbar, d. h., es zeigen sich deskriptiv keine deutlichen Unterschiede bei den Mittelwerten oder im Ausmaß der Streuung (s. Abbildung 38).

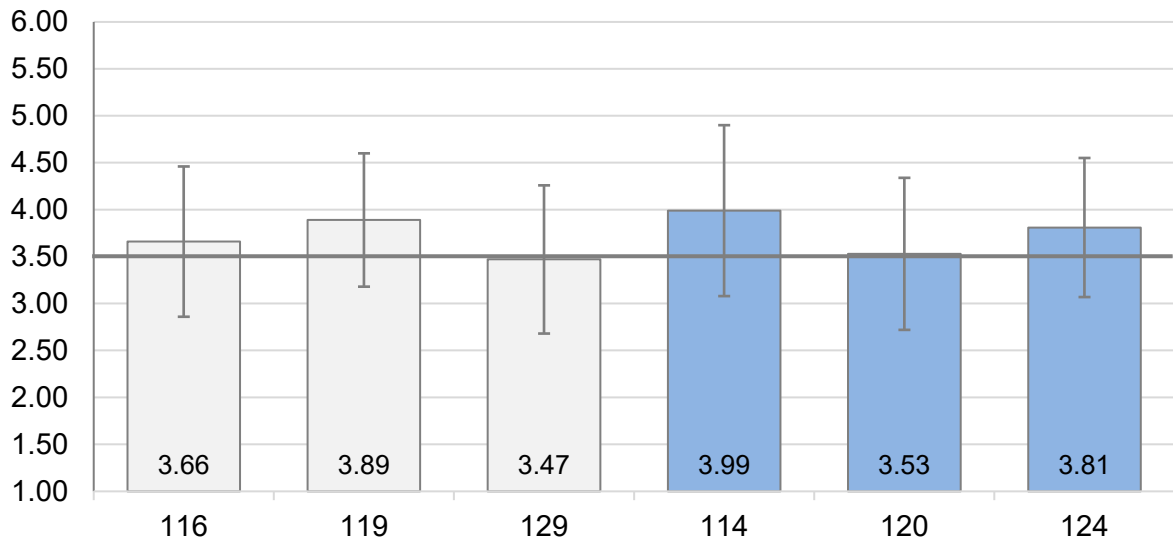


Abbildung 38. Mittelwerte und Standardabweichungen für die kollektive Selbstwirksamkeitserwartung zu t_3 . Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Was das Fortbildungsklima unter den Lehrpersonen anbelangt, zeigen sich erneut sehr positive Bewertungen der sechs Schulen, aber es ist auch hier keine besonders augenfällige Differenz zwischen den beiden Extremgruppen erkennbar (s. Abbildung 39). Sowohl Schulen, bei denen sich die kollektive Selbstwirksamkeitserwartung des Kollegiums negativ entwickelt hat (grau), als auch Schulen, die einen stärkeren handlungsbezogenen Optimismus erworben haben, erleben die Atmosphäre in der Fortbildung als angenehm und vertrauensvoll.

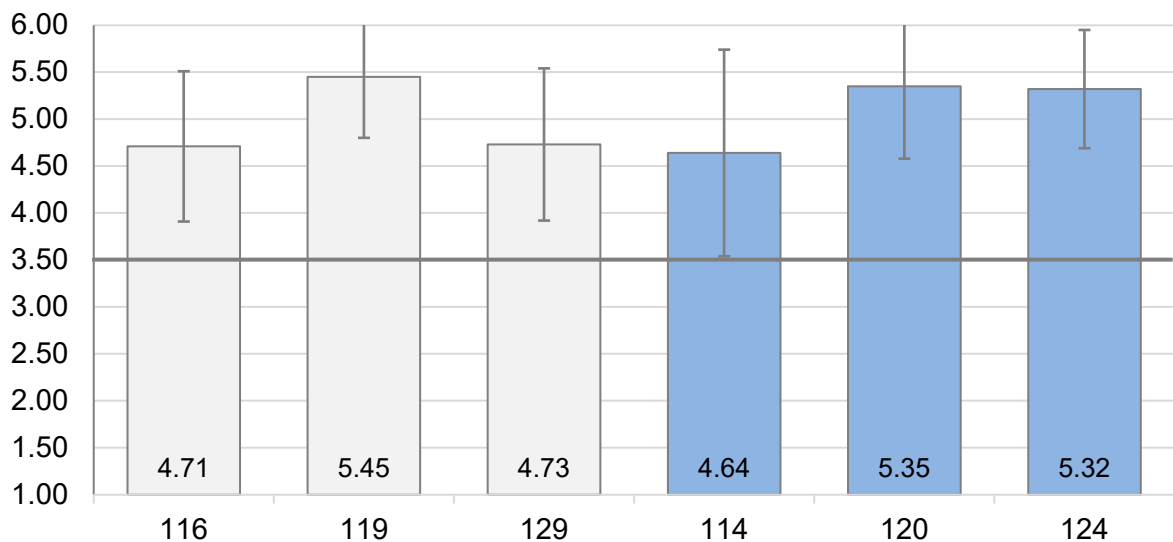


Abbildung 39. Mittelwerte und Standardabweichungen für das Fortbildungsklima zwischen den Lehrpersonen in Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Vergleicht man die Schulen 120 und 124, die zur Gruppe mit positiver Entwicklung bei der Selbstwirksamkeitserwartung gehören, mit der Schule 116 ($M = 3.26$, $SD = 1.49$) und der Schule 129

($M = 2.82$, $SD = 1.19$), in denen sich die kollektive Selbstwirksamkeitserwartung über die Fortbildung hinweg eher ungünstig entwickelte, dann fällt bei der Schule 120 ($M = 3.69$, $SD = 0.80$) und der Schule 124 ($M = 3.67$, $SD = 0.93$) im Mittel eine stärkere kooperative Auseinandersetzung mit den Fortbildungsinhalten und eine geringere Streuung in dieser Einschätzung auf (s. Abbildung 40). Gleichzeitig ist aber zu berücksichtigen, dass die Schule 116 (negative Entwicklung) und das Kollegium der Schule 114 (positive Entwicklung) sich in der durchschnittlichen Beurteilung der kooperativen Erprobung ($M_{116} = 3.26$, $M_{114} = 3.27$) nicht unterscheiden. Zudem weist die Schule 119 den höchsten Mittelwert bei der kooperativen Erprobung mit $M = 3.79$ auf, in der die kollektive Selbstwirksamkeitserwartung am Ende der Fortbildungsteilnahme geringer war als am Anfang von Modul 2. Deskriptiv ist demnach kein eindeutiger Unterschied in der Intensität der kooperativen Umsetzung von Fortbildungsinhalten während des zweiten Moduls zwischen den beiden Extremgruppen erkennbar.

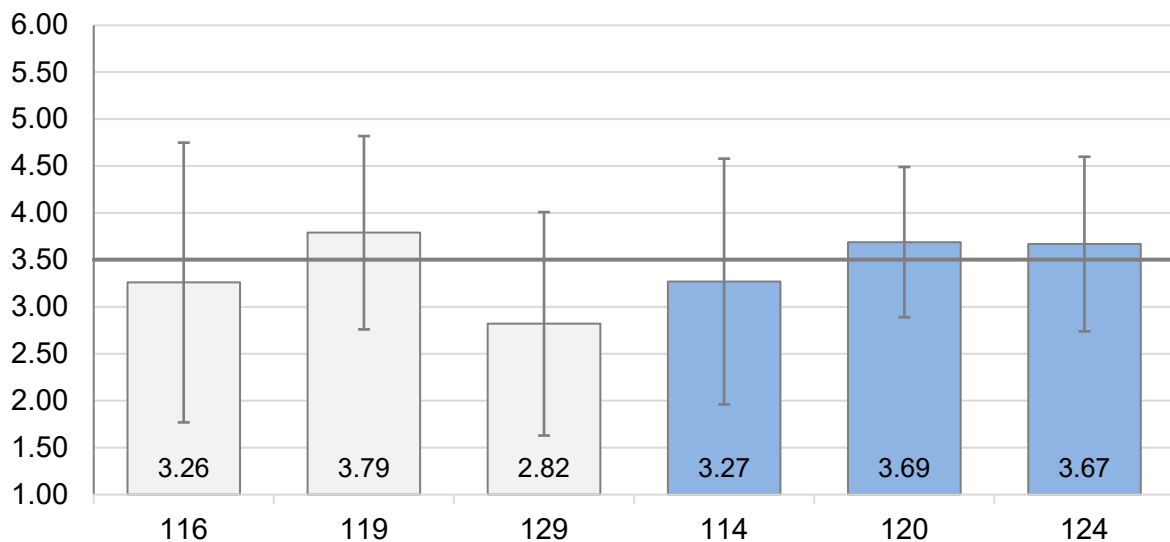


Abbildung 40. Mittelwerte und Standardabweichungen für die kooperative Umsetzung während Modul 2 (t_3). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

In Abbildung 41 sind die Mittelwerte und Standardabweichungen für die wahrgenommene Kompetenz der Moderatorenteams für die sechs Schulen dargestellt, die aufgrund ihrer auffallend negativen oder positiven Entwicklung bei der kollektiven Selbstwirksamkeitserwartung über den Verlauf von *Vielfalt fördern* hinweg miteinander verglichen werden. Deskriptiv zeigt sich allerdings kein merklicher Unterschied zwischen den beiden Extremgruppen. Abgesehen von den Lehrpersonen der Schule 129 ($M = 3.43$) erleben die Kollegien das Moderatorenteam, das die Fortbildung an ihrer Schule durchführt, als kompetent im Umgang mit Rückfragen und als sicher und verständlich bei der Vermittlung der Inhalte.

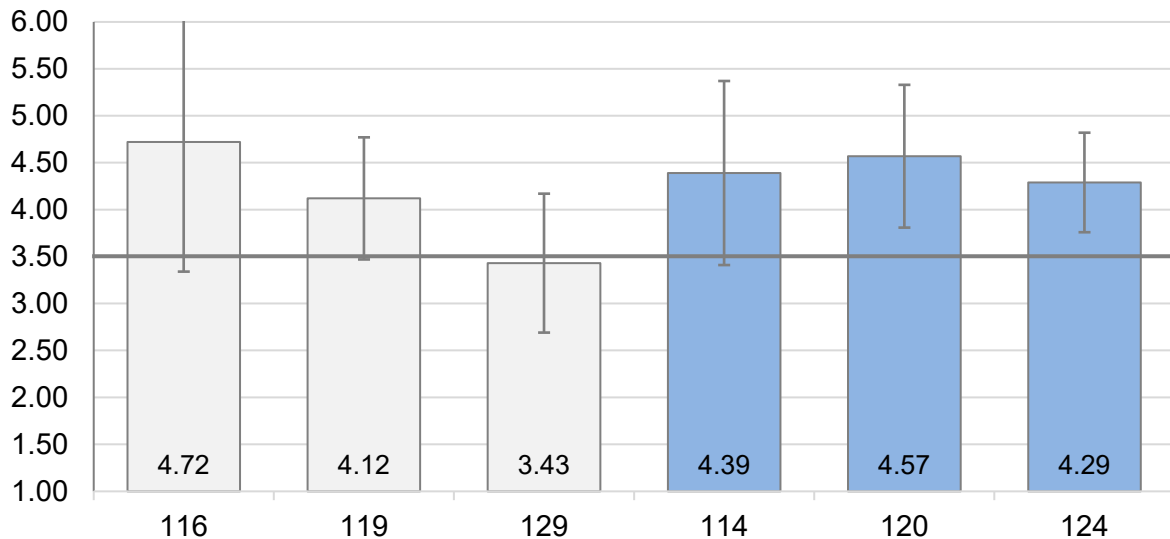


Abbildung 41. Mittelwerte und Standardabweichungen für die wahrgenommene Moderatorenkompetenz in Modul 2 (t₃). Anmerkung: graue Säulen = Schulen mit negativer oder geringer Entwicklung, blaue Säulen = Schulen mit positiver Entwicklung

Die Befunde des Kapitels sind so zusammenzufassen, dass für einzelne Schulen positive Entwicklungen in der Hinwendung zum Lernen der Schülerinnen und Schüler, in der Deprivatisierung des eigenen Unterrichts und in der kollektiven Selbstwirksamkeitserwartung über den Fortbildungszeitraum von *Vielfalt fördern* festgestellt werden können.¹⁴ In weiterführenden deskriptiven Analysen wurden jeweils drei Schulen mit günstigen Entwicklungen und drei Schulen mit ungünstigeren Entwicklungen als Extremgruppen betrachtet und hinsichtlich weiterer Merkmale verglichen. Die Vergleiche zeigen zum einen, dass sich die Ausgangsbedingungen der beiden Schulgruppen nicht merklich unterscheiden. Zum anderen erweisen sich das wahrgenommene Klima, die Intensität und die wahrgenommene Moderatorenkompetenz nicht als relevant für die Zugehörigkeit der Schulen zu den beiden Extremgruppen. Die mittlere Ausgangsbedingung der Schulen wie auch das Fortbildungsklima, die Intensität der kooperativen Umsetzung und die Wahrnehmung des Moderatorenteams sind also offenbar keine guten Indikatoren für die Entwicklung der Kollegien in den drei ausgewählten Zielvariablen.

7.2 Intensität und Struktur der kollegialen Zusammenarbeit

Im vorigen Abschnitt wurde die Entwicklung von zwei Aspekten der Lehrerkooperation (Fokus auf das Lernen der Schülerinnen und Schüler und Deprivatisierung) betrachtet, die eine förderliche Lehrerkooperation kennzeichnen. Es konnte gezeigt werden, dass sich die beiden genannten Kooperationsformen bei einzelnen Schulen der UG während der Teilnahme an *Vielfalt fördern* positiv entwickelten.

¹⁴ Es handelt sich hier um eine deskriptive Betrachtung, die nicht auf statistische Signifikanz geprüft wurde.

Im Folgenden soll zunächst auf die Häufigkeit verschiedener Kooperationshandlungen von Lehrpersonen geblickt werden, um zu überprüfen, ob sich diese im Verlauf von *Vielfalt fördern* intensiviert haben. Anschließend wird untersucht, ob *Vielfalt fördern* strukturelle Veränderungen angeregt hat, die mit einer intensiveren Zusammenarbeit in den Schulen einhergehen sollten.

7.2.1 Veränderung der Kooperationsintensität

Die Kooperationsaktivitäten (s. Tabelle 11) wurden im Projekt *LIQUID* als Einzelitems bezogen auf die letzten drei Monate mit einem fünfstufigen Antwortformat (1 „0 mal“ bis 5 „7 mal oder häufiger“) am Anfang von Modul 3 (t_5) und am Ende der Fortbildung (t_7) erfasst. Höhere Werte sind demnach als eine intensivere Ausübung in den zurückliegenden drei Monaten zu interpretieren. Es fällt auf (s. Tabelle 11), dass kollegiale Gespräche über die konkreten Lernschwierigkeiten von einzelnen Schülerinnen und Schülern (t_5 : $M = 4.33$, $SD = 1.01$; t_7 : $M = 4.40$, $SD = 0.90$) und ein Austausch über die Verbesserung von Unterricht (t_5 : $M = 3.18$, $SD = 1.25$; t_7 : $M = 3.21$, $SD = 1.21$) deutlich häufiger stattfinden als die kollegiale Unterrichtshospitation im eigenen Unterricht (t_5 : $M = 1.76$, $SD = 1.14$; t_7 : $M = 1.74$, $SD = 1.20$) und die Hospitation bei einer Kollegin oder einem Kollegen (t_5 : $M = 1.72$, $SD = 1.03$; t_7 : $M = 1.72$, $SD = 1.15$). Vergleicht man die Mittelwerte der beiden Erhebungszeitpunkte für die fünf verschiedenen Kooperationshandlungen (s. Tabelle 11), dann deuten sich keine bedeutsamen Veränderungen im Zeitverlauf an. Die Standardabweichungen und ICC-Werte deuten jedoch darauf hin, dass durchaus eine Varianz gegeben ist, sodass eine fehlende mittlere Entwicklung in der Kooperationsintensität nicht notwendigerweise bedeutet, dass es in den Schulen bei einzelnen Lehrpersonen nicht doch zu bedeutsamen Veränderungen gekommen ist. Die ICC-1-Werte sind auf vier der fünf Kooperationsaktivitäten zum zweiten Messzeitpunkt am Ende der Fortbildung etwas geringer. Das bedeutet, dass die Unterschiede in der Einschätzung der fünf Kooperationsaktivitäten nicht mehr so stark auf die Schulzugehörigkeit, sondern stärker auf andere Faktoren zurückzuführen sind. Unterschiede in der Intensität der Kooperationsaktivitäten lassen sich am Ende der Fortbildung zu 4.00 % bis 11.5 % durch die Zugehörigkeit zu einer Schule erklären (s. Tabelle 11).

An den ICC-2-Werten, die überwiegend zwischen .561 und .743 liegen, wird deutlich, dass die Lehrpersonen innerhalb einer Schule durchaus zu einer kongruenten Einschätzung gelangen. Einzig der ICC-2-Wert für die kollegialen Gespräche über die Verbesserung von Unterricht ist am Ende der Fortbildung mit .439 als marginal zu bewerten, d. h., die Lehrpersonen geben hier unterschiedliche Einschätzungen ab.

Tabelle 11. Mittelwert, Standardabweichung und ICC-Werte für Kooperationsaktivitäten

Kooperationsaktivitäten (Einzelitems)	t ₅			t ₇		
	M (SD)	ICC 1	ICC 2	M (SD)	ICC 1	ICC 2
Wie häufig haben Sie in den letzten 3 Monaten mit einem Kollegen über konkrete Lernschwierigkeiten einzelner Schüler gesprochen?	4.33 (1.01)	.076	.635	4.40 (0.90)	.065	.569
Wie häufig haben Sie in den letzten 3 Monaten mit einem Kollegen über die Verbesserung von Unterricht in Ihrem gemeinsamen Fach gesprochen?	3.18 (1.25)	.057	.561	3.21 (1.21)	.040	.439
Wie häufig haben Sie in den letzten 3 Monaten gemeinsam mit einem Kollegen eine Unterrichtseinheit durchgeführt (mind. gemeinsame Planung und Reflexion der Erprobung)?	2.78 (1.46)	.121	.743	2.68 (1.42)	.115	.710
Wie häufig sind Sie in den letzten 3 Monaten im Rahmen der kollegialen Unterrichtshospitation von einem Kollegen im Unterricht besucht worden?	1.76 (1.14)	.075	.631	1.74 (1.20)	.091	.654
Wie häufig haben Sie in den letzten 3 Monaten im Rahmen der kollegialen Unterrichtshospitation einen Kollegen im Unterricht besucht?	1.72 (1.03)	.094	.687	1.72 (1.15)	.088	.645

Im Folgenden wird nun untersucht, durch welche Merkmale sich die Intensität der Kooperation am Ende der Fortbildung, welche über die fünf oben genannten Merkmale erfasst wurde, vorher-sagen lässt. Als Prädiktoren, d. h. als unabhängige, erklärende Variablen, wurden die Einstellung der Schulleiterin/des Schulleiters zur Unterrichtsvisitation (t₁), die Einstellung der Schulleite-rin/des Schulleiters zur Kooperation von Lehrpersonen (t₁), die kollektive Innovationsbereitschaft des Kollegiums (t₁) und das wahrgenommene Klima im Kollegium (t₃) untersucht. Die Prädiktoren wurden, abgesehen von dem Ausgangswert des jeweiligen Kooperationsmerkmals, jeweils in separaten Modellen überprüft, in denen die hierarchische Struktur der Daten, also die Gruppierung in Schulen, berücksichtigt wurde.

Als Abbildungen werden im Folgenden ausschließlich die Modelle mit einem signifikanten Effekt dargestellt, wobei selbstverständlich alle Ergebnisse berichtet werden.

7.2.1.1 Einstellung der Schulleiterin/des Schulleiters

Die *Einstellung der Schulleitung zur Unterrichtsvisitation* wurde mit drei Items erfasst ($\alpha = .65$). Ein Beispielitem lautet: „*Unterrichtsbesuche mit anschließendem Feedback sind ein Baustein für die Personalentwicklung einer Schule*“. Hohe Werte auf der Skala deuten darauf hin, dass die Schulleiterinnen und Schulleiter Unterrichtsbesuche als Chance zur Personalentwicklung und Qualitätssicherung verstehen und nicht als Vorgehen, welches die Selbstständigkeit der einzelnen Lehrkraft angreift. Für die Schulleiterinnen und Schulleiter an Schulen der UG beträgt der Mittelwert für die Einstellung zur Unterrichtsvisitation $M = 5.06$ ($SD = 0.84$). Die Schulleiterinnen und Schulleiter betrachten Unterrichtsvisitationen also in einem hohen Maße als ein wichtiges Instrument, um das unterrichtliche Angebot der Schule weiterentwickeln zu können. Es scheint plausibel, dass eine positive und konstruktive Einstellung der Schulleitung gegenüber Unterrichtsvisitationen eine positive Einstellung und eine intensivere kollegiale Hospitationspraxis der Lehrpersonen begünstigen kann.

Die Skala, mit der die *Einstellung von Schulleiterinnen und Schulleitern zur Kooperation von Lehrpersonen* erhoben wurde, besteht aus insgesamt sieben Items (Beispielitem: „*Kollegiale Hilfe bei Problemen ist ein wichtiger Bestandteil der Lehrerarbeit*“, $\alpha = .70$). Werden die Itemausagen von Schulleiterinnen und Schulleitern als zutreffend beurteilt, dann betrachten sie diese Kooperation als einen wichtigen Aspekt des Lehrerberufs und nicht bloß als eine unnötige Mehrarbeit. Auch hier liegt die Annahme nahe, dass eine Schulleiterin/ein Schulleiter, die/der der Kooperation einen großen Wert beimisst, auch eher die Kooperation im Kollegium anregt und unterstützt als eine Schulleitung, die Lehrerkoooperation als weniger bedeutsam erachtet. Der Mittelwert von $M = 5.65$ ($SD = 0.41$) deutet darauf hin, dass die Schulleiterinnen und Schulleiter der UG-Schulen der Kooperation von Lehrkräften eine sehr große Relevanz beimessen. An der Standardabweichung lässt sich außerdem erkennen, dass die Schulleiterinnen und Schulleiter sich in ihrer Einstellung kaum unterscheiden.

Um die Bedeutung der Einstellung der Schulleitung für die Kooperation der Lehrkräfte zu untersuchen, werden zwei Analyseschritte vorgenommen. Zunächst wird geprüft, ob die Einstellungen der Schulleitungen mit der von den Lehrpersonen berichteten Kooperationsintensität zum ersten Messzeitpunkt zusammenhängt. Im jeweils zweiten Schritt und in dem eigentlichen Analysemodell wird dann als abhängige Variable die jeweilige Facette der Kooperationsintensität zum zweiten Messzeitpunkt unter Kontrolle des Wertes zum ersten Messzeitpunkt vorhergesagt. Durch die Kontrolle des ersten Wertes wird demnach untersucht, ob die Einstellungen der Schulleitungen für die Entwicklung der Kooperationsaktivitäten relevant sind.

Die Analysen, die im ersten Schritt durchgeführt wurden, um zu überprüfen, ob die Einstellung der Schulleiterinnen und Schulleiter mit der von den Lehrpersonen berichteten Kooperationshäufigkeit bei der ersten Erhebung, d. h. am Anfang von Modul 3, zusammenhängt, zeigen keine statistisch bedeutsamen Korrelationen. Dieses Ergebnis bedeutet, dass die von den Lehrpersonen angegebene Häufigkeit, mit der die oben genannten fünf Kooperationsaktivitäten durchgeführt werden, weder mit der Einstellung der Schulleitung zu Unterrichtsbesuchen ($.00 \leq r \leq .07$, n. s.) noch mit deren Einstellung zur Bedeutung kollegialer Zusammenarbeit ($-.02 \leq r \leq .12$, n. s.) zusammenhängt.

In den anschließenden Analysemodellen erweisen sich die Einstellungen der Schulleitung zur Unterrichtsvisitation ($-.07 \leq \beta \leq .01$, n. s.) und zur Bedeutung von Kooperation für den Lehrberuf ($-.03 \leq \beta \leq .07$, n. s.) ebenfalls nicht als signifikanter Prädiktor für die Entwicklung der Häufigkeit der unterschiedlichen Kooperationsaktivitäten während der letzten drei Monate. Das bedeutet, dass nach Kontrolle des Eingangswertes die Entwicklung der Kooperationsintensität der Lehrpersonen am Ende der Fortbildung nicht von den Einstellungen der Schulleiterin/des Schulleiters beeinflusst wird, sondern hiervon unabhängig ist.

7.2.1.2 Kollektive Innovationsbereitschaft des Kollegiums

Die *kollektive Innovationsbereitschaft* ist die Wahrnehmung der Lehrpersonen eines Kollegiums darüber, wie aufgeschlossen und bereit das Kollegium an der Weiterentwicklung der Schule und des pädagogischen Konzepts arbeitet. Dieses Merkmal wurde über fünf Items erfasst (Beispielitem: „*Unser Kollegium ist stets um die Arbeit am schuleigenen pädagogischen Konzept bemüht*“, $\alpha = .88$). Im Mittel schätzen die Lehrpersonen die Innovationsbereitschaft des Kollegiums ihrer Schule eher hoch ein ($M = 4.35$, $SD = 0.89$). Der ICC-2-Wert in Höhe von .920 deutet auf eine sehr hohe Kongruenz der Einschätzung von Lehrpersonen eines Kollegiums hin und der Schulmittelwert kann somit als ein zuverlässiger, reliabler Wert betrachtet werden. Erwartet wird, dass sich eine hohe kollektive Innovationsbereitschaft positiv auf die Kooperationsintensität am Ende der Fortbildung auswirkt, d. h., dass an Schulen mit einer hohen Bereitschaft, am eigenen pädagogischen Konzept zu arbeiten, die Lehrpersonen auch häufiger kooperieren.

Unter Kontrolle der Ausgangsintensität (t_5) hat die am Anfang von *Vielfalt fördern* bei den Lehrpersonen erfasste kollektive Innovationsbereitschaft des Kollegiums einen positiven Einfluss auf die Kooperationsintensität von nahezu allen fünf Merkmalen (s. Tabelle 11). Ausschließlich die Häufigkeit, mit der Gespräche zur Verbesserung von Unterricht geführt werden, wird nicht durch die mittlere kollektive Innovationsbereitschaft des Kollegiums beeinflusst ($\beta = .06$, n. s.).

Die durchschnittliche Häufigkeit, mit der die Lehrpersonen mit einer Kollegin oder mit einem Kollegen über eine konkrete Schülerin oder einen konkreten Schüler und deren/dessen Schwierigkeiten und Probleme beim Lernen sprechen, lässt sich unter Kontrolle des Eingangswertes ($\beta = .38$, $p < .001$) signifikant positiv durch die kollektive Innovationsbereitschaft des Kollegiums

erklären ($\beta = .22, p < .001$, s. Abbildung 42). In Kollegien, die von einer höheren Innovationsbereitschaft berichten, finden nach Angaben der Lehrpersonen auch häufiger Gespräche über Lernschwierigkeiten von Schülerinnen und Schülern statt.

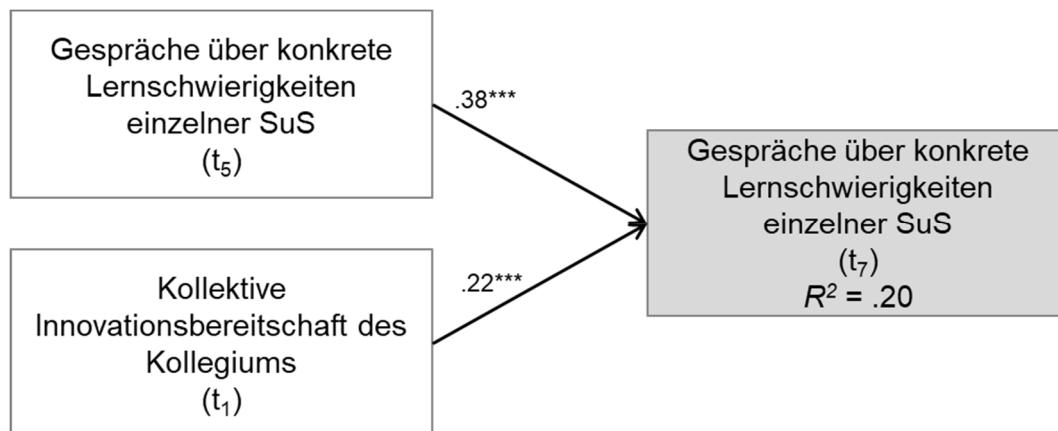


Abbildung 42. Vorhersage der Häufigkeit von Gesprächen über konkrete Lernschwierigkeiten einzelner Schülerinnen und Schüler durch die kollektive Innovationsbereitschaft des Kollegiums. Anmerkung: *** = $p < .001$

Das folgende Modell (s. Abbildung 43) untersucht, ob sich Unterschiede in der Entwicklung der Häufigkeit, mit der Lehrpersonen gemeinsam Unterrichtseinheiten durchführen, durch die kollektive Innovationsbereitschaft des Kollegiums vorhersagen lassen. Es zeigt sich, dass die Ausgangshäufigkeit (t_5), mit der die Lehrpersonen gemeinsam eine Unterrichtseinheit durchführen, ein bedeutsamer Prädiktor für diese Kooperationsaktivität am Ende der Fortbildung ist ($\beta = .51, p < .001$). Darüber hinaus erweist sich die kollektive Innovationsbereitschaft als signifikanter Einflussfaktor. Der Regressionskoeffizient in Höhe von $\beta = .14 (p < .01)$ ist so zu interpretieren, dass mit einer höheren kollektiven Innovationsbereitschaft eines Kollegiums zu Beginn der Fortbildung (t_1) die Kooperation bei der gemeinsamen Durchführung von Unterrichtseinheiten bis zum Ende der Fortbildung (t_7)¹⁵ zunimmt.

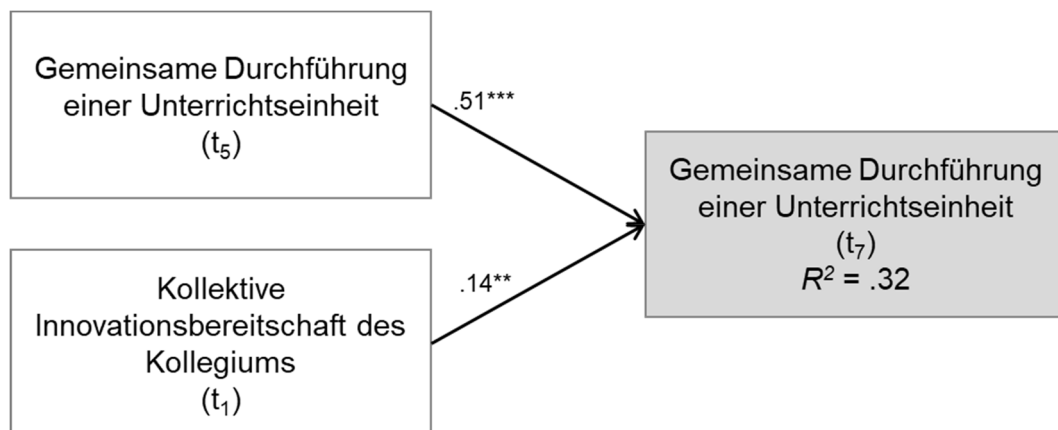


Abbildung 43. Vorhersage der Häufigkeit einer gemeinsamen Durchführung einer Unterrichtseinheit durch die kollektive Innovationsbereitschaft des Kollegiums. Anmerkung: ** = $p < .01$, *** = $p < .001$

¹⁵ Die Häufigkeit der gemeinsamen Durchführung von Unterrichtseinheiten wurde zum Ende der Fortbildung (t_7) mit einem zeitlichen Bezug auf die letzten drei Monate vor der Erhebung erfasst.

In den folgenden Modellen wird analysiert, ob sich die kollektive Innovationsbereitschaft auch auf die Häufigkeit der Hospitation auswirkt. Hierbei wird zwischen der Hospitation von Kolleginnen und Kollegen im eigenen Unterricht und der eigenen Hospitation im Unterricht von Kolleginnen und Kollegen unterschieden. Tatsächlich erweist sich die Innovationsbereitschaft auch in diesen beiden Modellen als signifikant (s. Abbildungen 44 und 45). Da in beiden Modellen die Intensität der entsprechenden Kooperationsform zu Beginn des Moduls 3 kontrolliert wurde, lassen sich die Effekte der Innovationsbereitschaft als Effekte auf die *Entwicklung* der Kooperationsintensität begreifen. Mit anderen Worten: Je höher die Innovationsbereitschaft im Kollegium zu Beginn der Fortbildung ausgeprägt ist, desto stärker nehmen die Hospitationen von Kolleginnen und Kollegen im eigenen Unterricht ($\beta = .18, p < .05$, s. Abbildung 44) und die eigenen Hospitationen im Unterricht von Kolleginnen und Kollegen zu ($\beta = .17, p < .05$, s. Abbildung 45).

Das bedeutet: An Schulen, die um ihre Erneuerung und Entwicklung bemüht und offen gegenüber neuen pädagogischen Ansätzen sind, nimmt die gegenseitige kollegiale Unterrichtshospitation während der Module 3 und 4 stärker zu als an Schulen, in denen die Innovationsbereitschaft gering ausgeprägt ist.

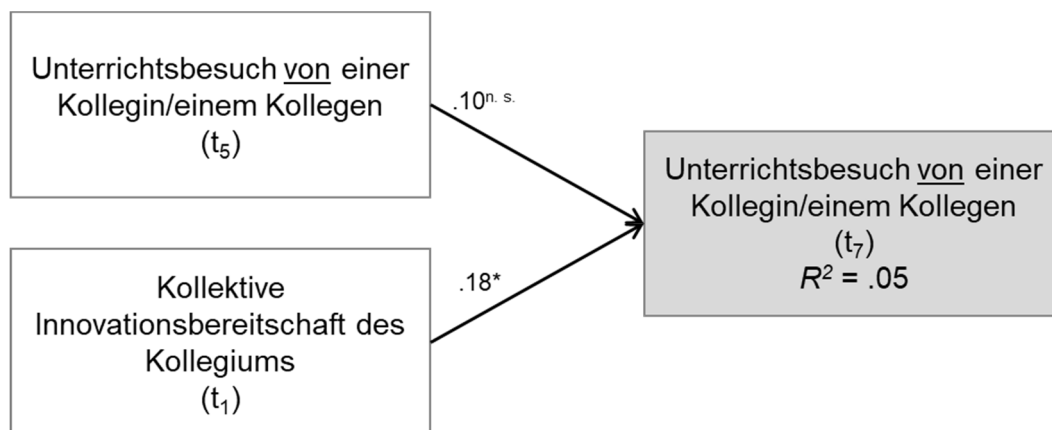


Abbildung 44. Vorhersage der Häufigkeit von Unterrichtsbesuchen von einer Kollegin/einem Kollegen im eigenen Unterricht durch die kollektive Innovationsbereitschaft des Kollegiums. Anmerkung: n. s. = nicht signifikant, * = $p < .05$

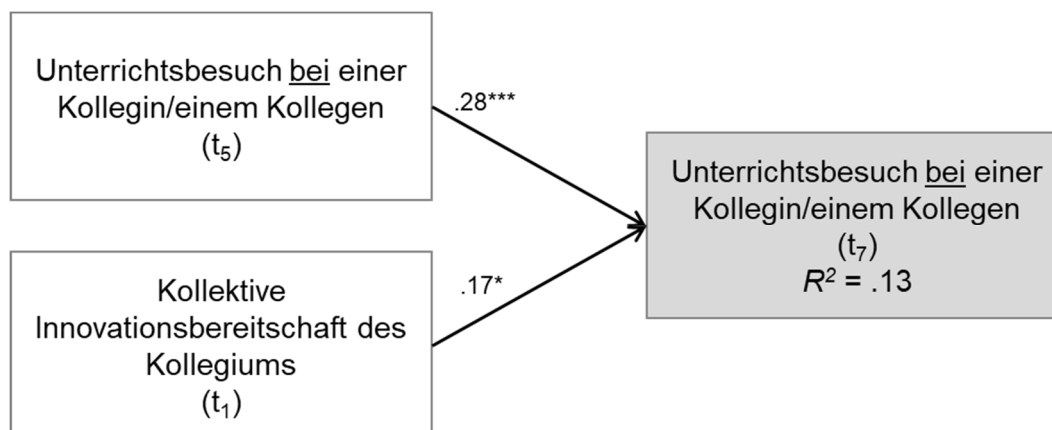


Abbildung 45. Vorhersage der Häufigkeit von Unterrichtsbesuchen bei einer Kollegin/einem Kollegen durch die kollektive Innovationsbereitschaft des Kollegiums. Anmerkung: * = $p < .05$, *** = $p < .001$

7.2.1.3 Klima während der Fortbildung

Die Skala, mit der das Klima während der Fortbildungstreffen erfasst wurde, ist bereits oben vorgestellt worden (s. Abschnitt 6.3). Es wird angenommen, dass ein besseres Fortbildungsklima unter den Lehrpersonen während des zweiten Moduls (t_3) eine intensivere Kooperation am Ende von *Vielfalt fördern* (t_7) begünstigt. Wie oben, wird auch in den nachfolgenden Analysen der Wert der Kooperationsaktivitäten zum ersten Messzeitpunkt kontrolliert, sodass die Modelle die Entwicklung der Kooperationsaktivitäten vom Anfang des dritten Moduls (t_5) bis zum Ende der Fortbildung vorhersagen (t_7).

Das in Modul 2 erfasste Fortbildungsklima erweist sich als prädiktiv für die Entwicklung der selbstberichteten Häufigkeit, mit der die Lehrpersonen einer Schule Gespräche über die Qualität des Unterrichts führen (s. Abbildung 46, $\beta = .09$, $p < .05$). Das Ergebnis bedeutet, dass mit einem günstigeren Klima in der Fortbildung kollegiale Gespräche zur Verbesserung des Unterrichts über den Zeitraum von Modul 3 und 4 hinweg stärker zunehmen als in einem Fortbildungskontext, in dem das Klima ungünstiger ausgeprägt ist.

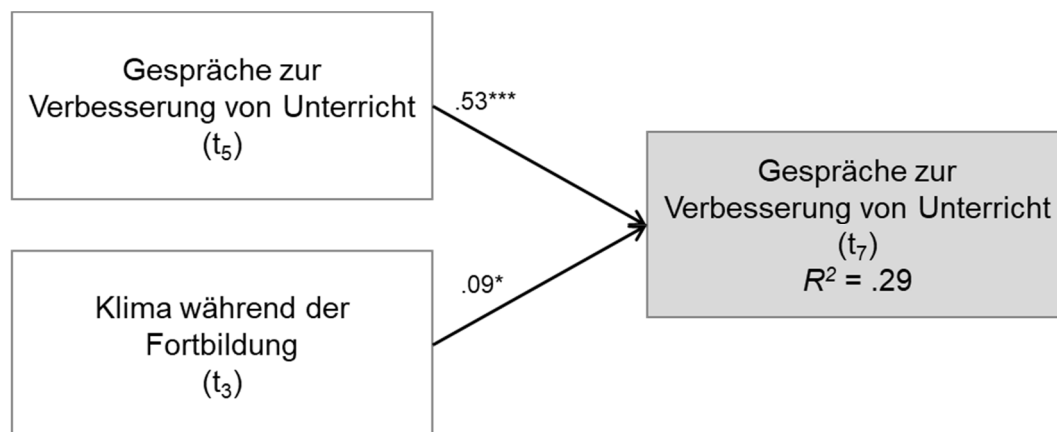


Abbildung 46. Vorhersage der Häufigkeit von Gesprächen zur Verbesserung von Unterricht durch das Klima während der Fortbildung. Anmerkung: * = $p < .05$, *** = $p < .001$

Das Klima unter den Lehrpersonen während der Bausteine von Modul 2 kann zudem Unterschiede in der Entwicklung der Häufigkeit gemeinsam durchgeführter Unterrichtseinheiten erklären ($\beta = .13$, $p < .001$, s. Abbildung 47). Ein freundliches und vertrauensvolles Klima während Modul 2 führt demnach dazu, dass die von den Lehrpersonen berichtete gemeinsame Unterrichtsplanung und -auswertung im Zeitverlauf zunimmt.

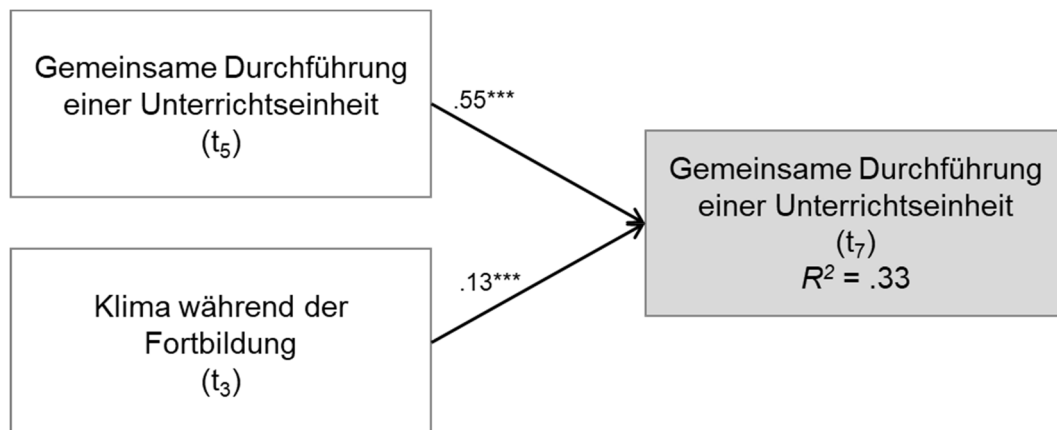


Abbildung 47. Vorhersage der Häufigkeit einer gemeinsamen Durchführung einer Unterrichtseinheit durch das Klima während der Fortbildung. Anmerkung: *** = $p < .001$

Wie häufig die Lehrpersonen der UG-Schulen nach eigenen Angaben am Ende der Fortbildung während der letzten drei Monate von einer Kollegin oder einem Kollegen im Unterricht besucht wurden, hängt in einem gewissen Maße davon ab, wie positiv das Fortbildungsklima empfunden wurde (s. Abbildung 48). Das Klima in den Fortbildungssitzungen hat einen kleinen positiven Einfluss auf die Häufigkeit der Unterrichtsbesuche durch Kolleginnen und Kollegen ($\beta = .15$, $p < .01$). Insgesamt werden jedoch nur 5 % der Unterschiede zwischen den Schulen zu t₇ durch das Modell erklärt ($R^2 = .05$).

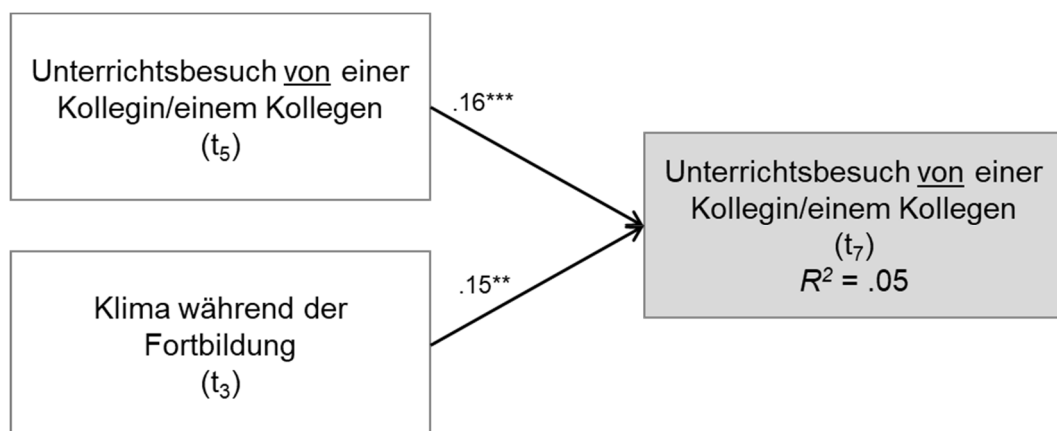


Abbildung 48. Vorhersage der Häufigkeit von Unterrichtsbesuchen von einer Kollegin/einem Kollegen im eigenen Unterricht durch das Klima während der Fortbildung. Anmerkung: ** = $p < .01$, *** = $p < .001$

Auch wie häufig die Lehrpersonen nach eigenen Angaben im Mittel bei einer Kollegin oder bei einem Kollegen hospitieren, hängt vom Klima unter den Lehrpersonen ab (s. Abbildung 49). Der Regressionskoeffizient beträgt $\beta = .11$ ($p < .05$). Demnach handelt es sich um einen schwachen positiven Einfluss. Dies bedeutet: Je positiver das Fortbildungsklima in Modul 2 wahrgenommen wird, desto stärker nimmt die Häufigkeit zu, mit der eine Lehrperson im Unterricht eines Kollegen/einer Kollegin hospitiert.

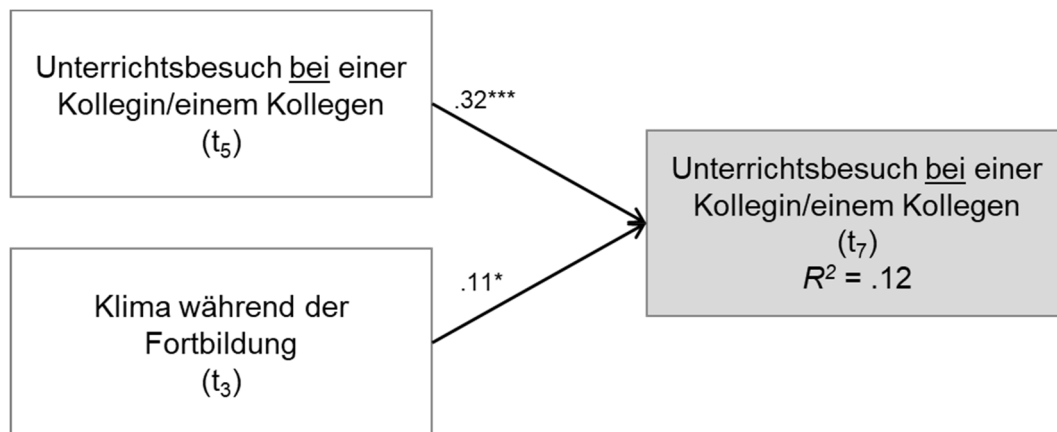


Abbildung 49. Vorhersage der Häufigkeit von Unterrichtsbesuchen bei einer Kollegin/einem Kollegen durch das Klima während der Fortbildung. Anmerkung: * = $p < .05$, *** = $p < .001$

Zusammengefasst zeigt sich in den dargestellten Analysen, dass sich durch das Fortbildungsklima im Modul 2 Unterschiede in der Entwicklung der nachfolgenden Kooperationsintensität in den Modulen 3 und 4 vorhersagen lassen, und zwar bezogen auf die Häufigkeit von Gesprächen zur Verbesserung des Unterrichts sowie auf die Häufigkeit von gemeinsam durchgeführten Unterrichtseinheiten und Unterrichtsbesuchen.

7.2.2 Einrichtung institutionalisierter Formen der Teamarbeit

Vielfalt fördern setzt nicht explizit bei der Veränderung schulischer Strukturen an. Dennoch ist es plausibel anzunehmen, dass Erfahrungen, die ein Kollegium während der Fortbildung mit der unterrichtsbezogenen Kooperation macht, eine feste und verbindliche Einrichtung von Kooperationsstrukturen befördern kann.

7.2.2.1 Deskriptive Beschreibung der Einrichtung fester Teamarbeitsstrukturen

Im folgenden Abschnitt wird zunächst untersucht, ob an den Schulen seit Beginn des Projekts *Vielfalt fördern* nach Angaben der Schulleiterinnen und Schulleiter neue Teamstrukturen eingerichtet wurden. Betrachtet wird die Etablierung fester Arbeits- oder Steuergruppen. Hierzu waren die teilnehmenden Schulen bereits vor *Vielfalt fördern* einmal explizit aufgefordert, da sie eine Steuergruppe für den Fortbildungsprozess bilden sollten. Darüber hinaus sind in Schulen weitere Arbeits- und Steuergruppen für unterschiedliche Schwerpunkte der Schul- und Unterrichtsentwicklung denkbar. Es wird darüber hinaus in den Blick genommen, ob Klassenleitungsteams, Jahrgangsteams und eine kontinuierliche Doppelbesetzung in einzelnen wöchentlich festgelegten Unterrichtsstunden fest in die Arbeitsstrukturen der Schulen integriert wurden. Die Arbeit in Klassen- und/oder Jahrgangsteams sind zwei Aspekte, die in *Vielfalt fördern* eine zentrale Rolle einnehmen und die vielfach an Schulen dominierende fachbezogene Zusammenarbeit ergänzen sollen. Kontinuierliche Doppelbesetzungen in festgelegten Unterrichtsstunden könnten den ge-

meinsamen Blick auf Schülerinnen und Schüler und zudem gegebenenfalls die gegenseitige kollegiale Unterrichtshospitation erleichtern. Die Einrichtung von Teams, die gegenseitige Unterrichtshospitationen durchführen, ist ein weiterer Aspekt, der im Folgenden betrachtet wird.

Aufgrund der geringen Stichprobe wird lediglich eine beschreibende Darstellung der Daten vorgenommen. Um die Angaben von den Schulleiterinnen und Schulleitern der UG besser einordnen zu können, werden an dieser Stelle auch die Daten der Schulleiterinnen und Schulleiter aus den Schulen der KG berichtet.

Es zeigt sich, dass an über der Hälfte aller Schulen der UG und der KG (UG: 53.8 %, KG: 56.2 %) bereits vor dem Evaluationszeitraum feste Arbeits- oder Steuergruppen z. B. für die Entwicklung neuer pädagogischer Konzepte oder die Erarbeitung von neuen Angeboten etabliert waren (s. Abbildung 50). An Schulen, in denen es entsprechende Gruppen innerhalb des Kollegiums noch nicht gab, wurden diese – mit Ausnahme von einer Schule der Kontrollgruppe – neu eingerichtet.

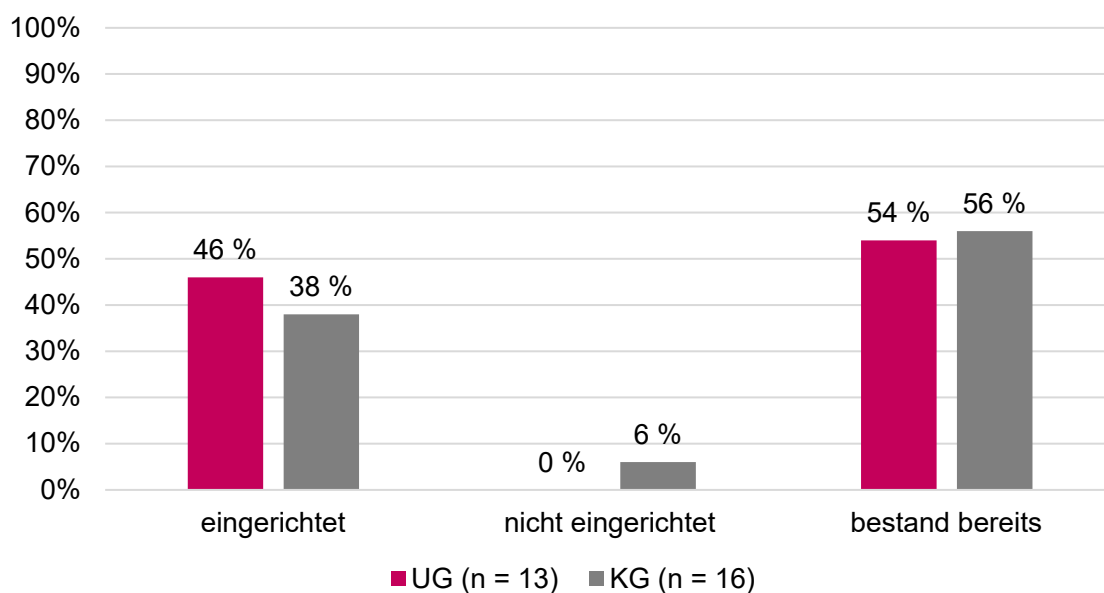


Abbildung 50. Einrichtung fester Arbeits- oder Steuergruppen

In 53.8 % der UG-Schulen und 50.0 % der KG-Schulen gab es bereits zu Beginn von *Vielfalt fördern* eine gemeinsame Klassenführung von zwei Lehrpersonen, denen auch feste Teamzeiten zur Verfügung standen (s. Abbildung 51). Im Evaluationszeitraum von *Vielfalt fördern* wurde an Schulen aus beiden Gruppen, bei denen es noch keine gemeinsame Klassenführung von Lehrpersonen gab, eine entsprechende Arbeitsform eingeführt.

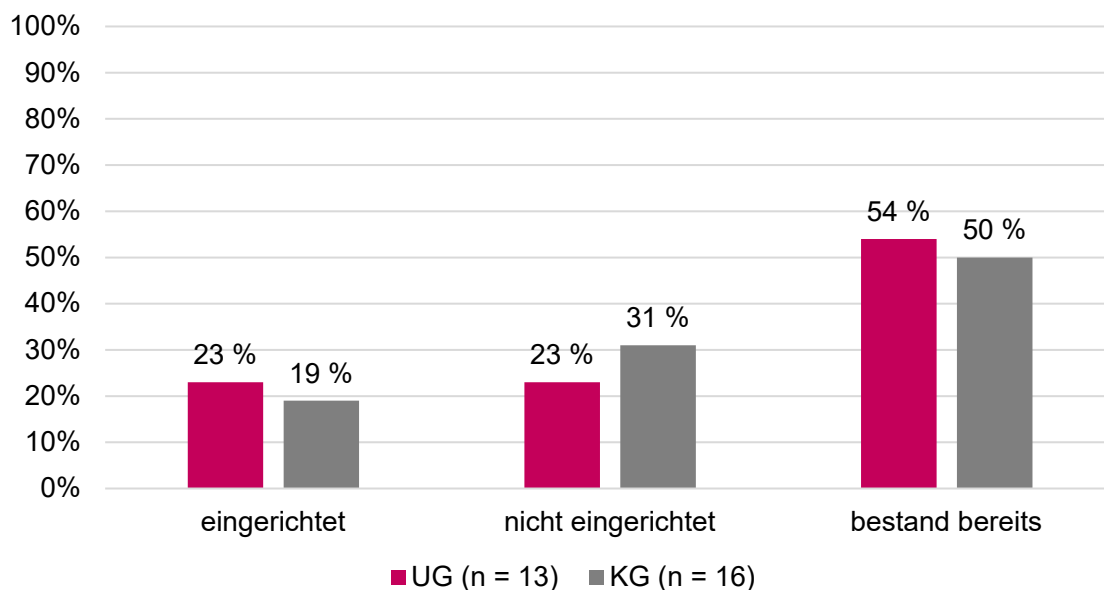


Abbildung 51. Einrichtung einer gemeinsamen Klassenführung durch zwei Lehrpersonen, denen feste Teamzeiten zur Verfügung stehen

Zu Beginn der Projektlaufzeit bestanden in weniger als der Hälfte aller Schulen beider Gruppen Jahrgangsteams, in denen alle Lehrkräfte einer Jahrgangsstufe vertreten sind und die sich regelmäßig zu Teambesprechungen treffen (UG: 38.4 %, KG: 43.7 %). Es zeigt sich (s. Abbildung 52), dass eine entsprechende Teamstruktur nach Angaben der befragten Schulleiterinnen und Schulleiter etwas häufiger an Schulen der UG (46.1 %) neu eingerichtet wurde als an Schulen der KG (31.2 %).

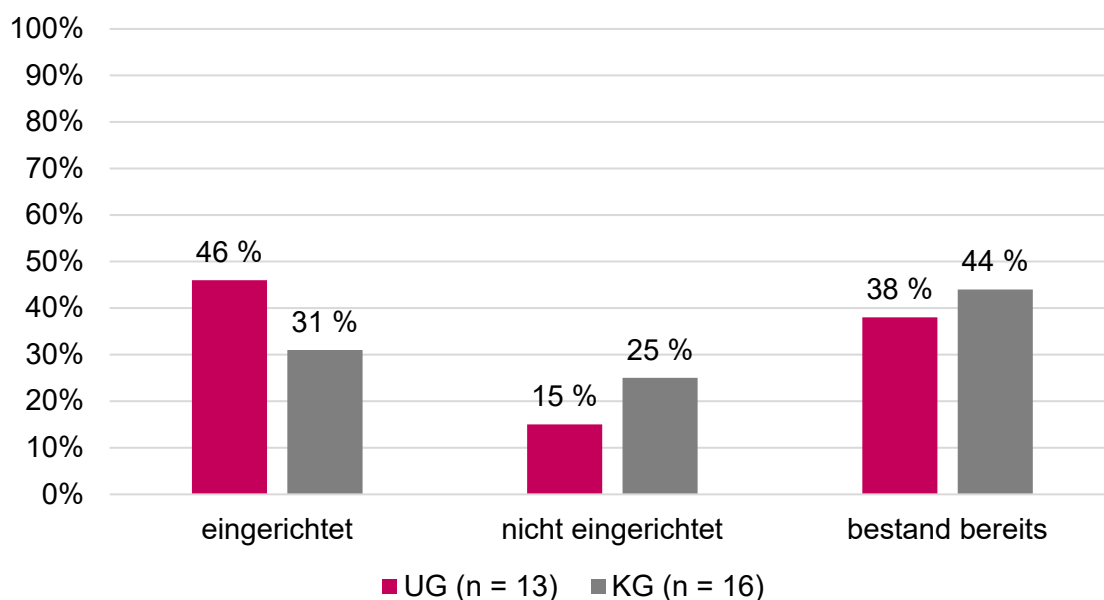


Abbildung 52. Einrichtung von Jahrgangsteams

Auffällig ist, dass deutlich mehr Schulen, die an *Vielfalt fördern* teilnehmen (61.5 %), als Schulen der KG (25.0 %) bereits am Anfang des Fortbildungsprozesses eine kontinuierliche Doppelbesetzung in einzelnen wöchentlich festgelegten Unterrichtsstunden umgesetzt haben (s. Abbildung 53). Mehr Lehrpersonen der UG hatten demnach am Anfang des Fortbildungsprozesses bereits Erfahrungen mit Team-Teaching. Entsprechend ist auch der Anteil an UG-Schulen (15.3 %), an denen ein entsprechendes Lehrkonzept seit dem Beginn von *Vielfalt fördern* eingeführt wurde, geringer als in der KG (37.5 %).

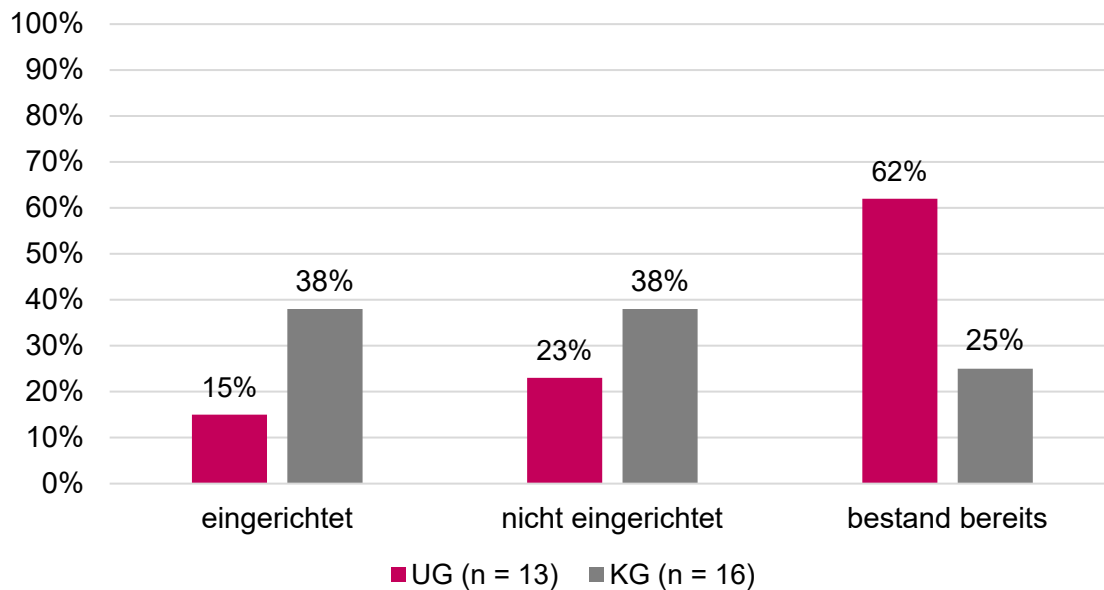


Abbildung 53. Einrichtung einer kontinuierlichen Doppelbesetzung in einzelnen wöchentlich festgelegten Unterrichtsstunden

Bei den bisher betrachteten Teamstrukturen unterschieden sich die Schulen, die an *Vielfalt fördern* teilgenommen haben, kaum von den Schulen, die zur Kontrollgruppe gehören. Ein deutlicher Unterschied zeigt sich jedoch, wenn man die Etablierung von Teams innerhalb des Kollegiums betrachtet, die eine gegenseitige kollegiale Unterrichtshospitation durchführen (s. Abbildung 54). In der UG gab es zu Beginn der Fortbildung keine Schule, die über entsprechende Strukturen verfügte. In der KG waren es 12.5 % aller Schulen, was zwei Schulen entspricht. Seit Projektbeginn wurden an 83.3 % der UG-Schulen Lehrerteams etabliert, die sich gegenseitig in ihrem Unterricht besuchen. Der Anteil an KG-Schulen, in denen eine entsprechende Teamarbeit institutionalisiert wurde, liegt hingegen nur bei 12.5 %. Es liegt daher die Interpretation nahe, dass die Teilnahme an *Vielfalt fördern* zu einer Einrichtung von entsprechenden Teamstrukturen geführt hat.

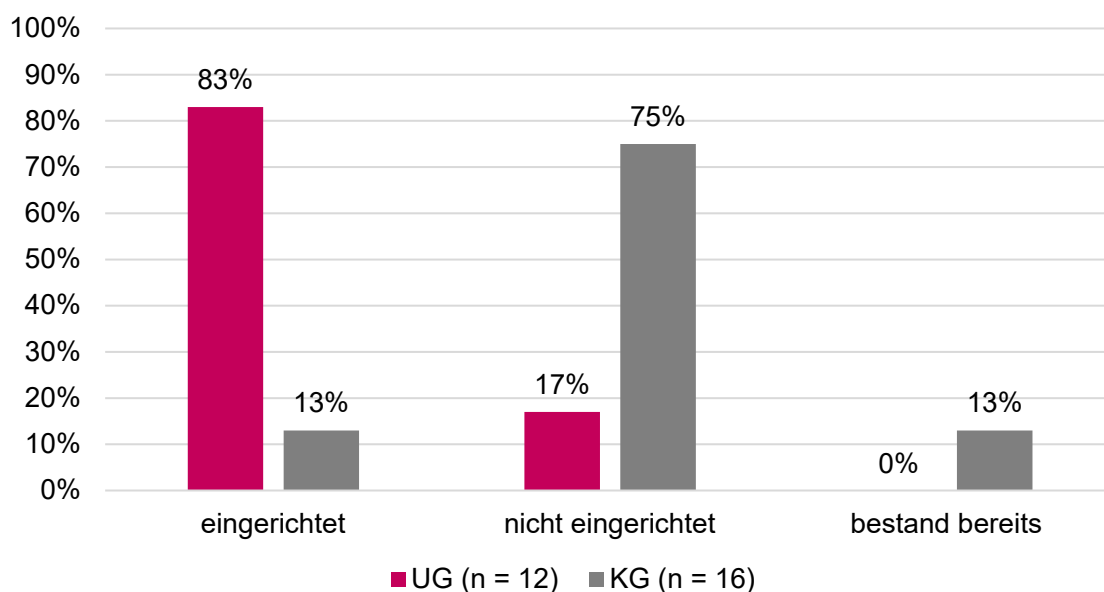


Abbildung 54. Einrichtung von Teams, die gegenseitig Unterrichtshospitation praktizieren

In einer weiterführenden Analyse wurde überprüft, ob an allen Schulen, die an *Vielfalt fördern* teilgenommen haben, mindestens eine der fünf Teamstrukturen neu eingerichtet worden ist oder ob sich einzelne Schulen zeigen, bei denen die Schulleiterin/der Schulleiter über keine teamstrukturellen Veränderungen berichtet. Aus Abbildung 55 geht hervor, dass an allen UG-Schulen eine Etablierung neuer Teamarbeitsformen stattgefunden hat, gleichwohl aber eine gewisse Varianz besteht.

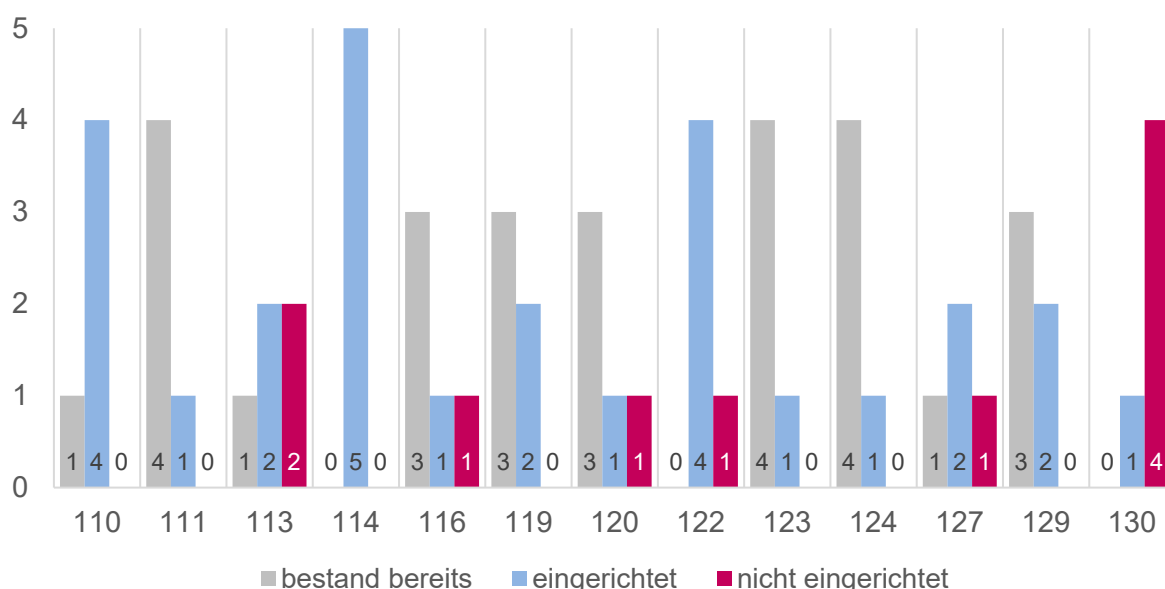


Abbildung 55. Anzahl an eingerichteten Teamstrukturen pro Schule

So bestand beispielsweise an Schule 122 zu Beginn von *Vielfalt fördern* keine der fünf erfassten Teamstrukturen. Die Schulleitung dieser Schule berichtet zum Zeitpunkt der Abschlusserhebung (t_7) eine Einrichtung von vier Formen der Zusammenarbeit. Auch an Schule 130 ist zu Beginn

von *Vielfalt fördern* noch keine der fünf Teamarbeitsformen etabliert. Im Gegensatz zur Schule 122 wurde im Verlauf von *Vielfalt fördern* allerdings nur eine der fünf Teamstrukturen an dieser Schule neu eingeführt. An Schule 114 bestand zu Beginn keine der erfassten Teamstrukturen. Im Verlauf der Fortbildung wurden nach Angaben der Schulleiterin/des Schulleiters alle fünf strukturellen Veränderungen vorgenommen. Aussagen über die Qualität der Umsetzung und der Intensität der neuen kooperativen Arbeitsformen können auf Basis der vorliegenden Daten allerdings nicht getroffen werden.

7.2.2.2 Einstellungen der Schulleitung als erklärendes Merkmal für die Einrichtung fester Teamarbeitsstrukturen

Wenn es um die Etablierung fester Formen und Strukturen für die Teamarbeit an einer Schule geht und nicht nur um die Zusammenarbeit zwischen einzelnen Lehrpersonen, dürfte der Schulleitung eine besondere Bedeutung zukommen.

Im folgenden Abschnitt wird daher überprüft, ob sich die am Anfang von *Vielfalt fördern* erfassten Einstellungen der Schulleiterinnen/Schulleiter zur Unterrichtsvisitation und zur Bedeutung von Kooperation für den Lehrberuf (s. Abschnitt 7.2.1.1) zwischen Schulen unterscheiden, bei denen feste Teamstrukturen bereits vor der Fortbildung bestanden, und solchen Schulen, die entsprechende Strukturen über den Evaluationszeitraum hinweg etabliert haben.

Wie schon im vorigen Abschnitt 7.2.2.1 werden für diese Auswertungen mittels U-Test (s. Abschnitt 4.1.3.1) die Daten von Schulleiterinnen und Schulleitern der UG und KG¹⁶ genutzt. Es zeigt sich, dass sich die Einstellungen von Schulleiterinnen und Schulleitern zur Durchführung von Unterrichtsbesuchen als Mittel zur Weiterentwicklung der Unterrichtsqualität und zur Bedeutung von Kooperation unter Lehrpersonen nicht zwischen Schulen, bei denen die fünf untersuchten Teamstrukturen (feste Arbeits- und Steuergruppen, Klassenteams mit festen Teamzeiten, Jahrgangsteams, kontinuierliche Doppelbesetzung in definierten Unterrichtsstunden, Hospitatonsteams) bereits am Anfang von *Vielfalt fördern* bestanden, und solchen Schulen, an denen entsprechende Strukturen eingeführt wurden, unterscheiden. Das deutet darauf hin, dass die beiden untersuchten Einstellungen der Schulleiterinnen und Schulleiter – Einstellung zur Unterrichtsvisitation und Einstellung zur Bedeutung von Kooperation für den Lehrberuf – nicht bedeutsam für die feste Etablierung von Teamstrukturen sind.

¹⁶ Für die Analysen wurden die Daten der UG und der KG gemeinsam ausgewertet, da die Stichprobe der UG zu klein für eine eigenständige statistische Auswertung ist.

8 Subjektive Bewertung und Verarbeitung durch die Lehrpersonen

In den zuvor berichteten Befunden wurden vielfach Entwicklungen auf schulischer Ebene betrachtet. Um Unterschiede zwischen Schulen zu erklären, wurden Merkmale der Schule als unabhängige Variablen herangezogen, z. B. die schulische Teilnahmequote aus einem extrinsischen Pflichtgefühl heraus, das durchschnittlich eingeschätzte Fortbildungsklima unter den Lehrpersonen, die durchschnittlich wahrgenommene kollegiale Umsetzung der Fortbildungsinhalte und Einstellungen der Schulleitung.

In Abschnitt 6.2 konnte gezeigt werden, dass die von den Lehrpersonen auf verschiedenen Ebenen selbstberichtete Wirksamkeit der Fortbildung auch davon abhängig ist, mit welchen motivationalen Voraussetzungen die Fortbildung begonnen wird. So beeinflussen die motivationalen Voraussetzungen unter anderem, wie die Lernangebote des zweiten Moduls von *Vielfalt fördern* wahrgenommen, genutzt und bewertet werden und wie die zeitliche Belastung aufgrund der Teilnahme an *Vielfalt fördern* eingeschätzt wird.

Im Folgenden wird nun weiterführend untersucht, ob Merkmale der Fortbildungswahrnehmung und -nutzung einen Einfluss auf die Entwicklung der Motivation zur Diagnostik, die Entwicklung von Selbstwirksamkeitserwartungen und auf die Transferabsichten am Ende der Fortbildung haben. Als unabhängige Variablen werden die individuellen Lerngelegenheiten und die zeitliche Belastung der Lehrpersonen in den Blick genommen. Zusätzlich zu diesen beiden Merkmalen der individuellen Nutzung und Wahrnehmung wird die prädiktive Kraft des Erlebens und der subjektiven Bewertung der Fortbildung, gemessen anhand der beiden Dimensionen Challenge Appraisal und Threat Appraisal, sowie der Einfluss der (kognitiven) Verarbeitung der Fortbildungsinhalte untersucht.

In der Evaluation wurden zwei subjektive Bewertungen der Lehrpersonen zu der Fortbildung *Vielfalt fördern* erfasst: die Fortbildung als eine Bedrohung (*Threat Appraisal*, Beispielitem: „Die Anforderungen der Fortbildung beunruhigen mich“, $\alpha = .85$, 6 Items) und die Fortbildung als eine positiv wahrgenommene Herausforderung (*Challenge Appraisal*, Beispielitem: „Die Fortbildung gefällt mir, da sie mich fordert“, $\alpha = .88$, 4 Items). Nach dem kognitiv-affektiven Modell der konzeptuellen Veränderung von Gregoire (2003) ist eine hohe Wahrnehmung an Herausforderungen günstig zu bewerten. Gregoire (2003) geht davon aus, dass eine hohe Ausprägung des Challenge Appraisal in Verbindung mit einer systematischen Verarbeitung der Fortbildungsinhalte konzeptuelle Veränderungen bei den Fortbildungsteilnehmerinnen und -teilnehmern befördert. Die systematische Verarbeitung wurde mit insgesamt sieben Items zu t_3 und mit acht Items zu t_6 erfasst (Beispielitem: „Ich denke über die praktische Anwendung der Fortbildungsinhalte nach“, $\alpha_{t_3} = .91$, $\alpha_{t_6} = .92$). Für die Erfassung der Appraisals und der systematischen Verarbeitung wurden Skalen von Decker (2015) adaptiert. Das Antwortformat ist bei diesen Skalen vierstufig und reicht von 1 „stimmt überhaupt nicht“ bis 4 „stimmt ganz genau“. Niedrige Werte sind demnach

als eine geringe Zustimmung zu der jeweiligen Bewertung oder als Ausdruck einer wenig systematischen Verarbeitung zu interpretieren; höhere Werte bedeuten, dass die Fortbildungsteilnehmerinnen und -teilnehmer den Itemaussagen eher zustimmen. Der Mittelwert für die Skala Threat Appraisal beträgt $M = 1.74$, d. h., die Lehrpersonen fühlen im Mittel kein Unbehagen aufgrund der Anforderungen, die mit der Teilnahme an *Vielfalt fördern* verbunden sind. Die Standardabweichung fällt mit $SD = 0.59$ eher gering aus und zeigt, dass diese Einschätzung bei den Fortbildungsteilnehmerinnen und -teilnehmern ähnlich ausfällt. Allerdings nehmen die Lehrpersonen die Teilnahme an der Fortbildung eher nicht als eine positive Herausforderung wahr ($M = 2.27$, $SD = 0.68$). Vor diesem Hintergrund überrascht es nicht, dass die Mittelwerte für die systematische Verarbeitung der Fortbildungsinhalte während Modul 2 ($M = 2.87$, $SD = 0.60$) und am Anfang von Modul 4 ($M = 2.81$, $SD = 0.63$) ebenfalls eher gering ausgeprägt sind. Die Lehrpersonen denken also nur mäßig intensiv über die praktische Anwendung der Fortbildungsinhalte nach und bringen die Fortbildungsinhalte nach eigenen Angaben kaum mit ihren bisherigen Kenntnissen in Verbindung.

Wie hängen die Nutzung individueller Lernaktivitäten, die erlebte zeitliche Belastung und die Bewertung von *Vielfalt fördern* als eine positive Herausforderung und/oder als eine subjektiv empfundene Bedrohung mit der Entwicklung von motivationalen Merkmalen der teilnehmenden Lehrpersonen über den Fortbildungszeitraum und deren Transferabsichten am Ende der Fortbildung zusammen? Geht mit einer intensiveren kognitiven Verarbeitung der Fortbildungsinhalte eine positivere Entwicklung der Motivation zur Diagnostik und/oder von Selbstwirksamkeitserwartungen der Lehrpersonen einher? Und führt eine elaboriertere kognitive Auseinandersetzung mit den Inhalten der Fortbildung auch zu einer stärkeren Absicht, die Anregungen aus *Vielfalt fördern* im Anschluss an die Fortbildung im eigenen schulischen Alltag anzuwenden? Diese Fragestellungen werden in den folgenden Analysen untersucht.

Als abhängige Variable werden die Motivation zur Diagnostik am Ende des zweiten Moduls (t_4), die Selbstwirksamkeitserwartung zur lernbegleitenden Diagnostik am Ende des zweiten Moduls (t_4), die Selbstwirksamkeitserwartung zur Klassenführung am Ende der Fortbildung (t_7) und die Selbstwirksamkeitserwartung zur adaptiven Unterrichtsgestaltung am Ende von *Vielfalt fördern* (t_7) betrachtet. In den folgenden Modellen wird jeweils der Eingangs-/Präwert des untersuchten Merkmals kontrolliert. Damit bilden die Modelle die Entwicklung der Merkmale ab. Eine Ausnahme stellen hier nur die Transferabsichten dar, die ausschließlich am Ende von Modul 4 (t_7) erfasst wurden.

Die motivationale Haltung der Lehrpersonen gegenüber der Durchführung einer pädagogischen Diagnostik dürfte ein relevanter Faktor für deren tatsächliches diagnostisches Handeln darstellen. Lehrpersonen, die mehr Freude und Interesse am Einsatz unterrichtlicher Diagnostik haben, dürften diagnostische Aktivitäten auch regelmäßiger und elaborierter durchführen. Erfasst wurde

die Motivation zur Übernahme diagnostischer Aufgaben (kurz: Motivation zur Diagnostik) mit einer adaptierten Skala von Klug (2011), die sich aus vier Items zusammensetzt (Beispielitem: *„Die Diagnostik von Lernvoraussetzungen der Schüler ist eine Aufgabe, die mir am Lehrerberuf richtig Spaß macht“*, $\alpha_{t_2} = .70$, $\alpha_{t_4} = .65$). Im Mittel geben die teilnehmenden Lehrpersonen zu beiden Erhebungszeitpunkten an, dass sie durchaus motiviert sind, den Lernstand und die Lernprozesse ihrer Schülerinnen und Schüler zu diagnostizieren (t_2 : $M = 4.11$, $SD = 0.87$; t_4 : $M = 4.03$, $SD = 0.84$).

Die Selbstwirksamkeitserwartung bezogen auf die Umsetzung einer lernbegleitenden Diagnostik erfasst das Zutrauen, Informationen zum Lernstand und -prozess zu erheben und Konsequenzen daraus abzuleiten, auch wenn die Zeit im Unterricht knapp ist oder die Lernenden sehr unterschiedlich sind (Beispielitem: *„Wie sicher sind Sie, dass Sie es im Unterricht in dem angegebenen Fach schaffen, bei der Formulierung von individuellen Lernzielen die Lernprozesse der Schüler zu berücksichtigen, auch wenn diese sehr unterschiedlich sind?“*, $\alpha_{t_2} = .79$, $\alpha_{t_4} = .83$, 4 Items). Das Antwortformat reichte von 1 *„absolut unsicher“* bis 6 *„absolut sicher“*. Zu Beginn von Modul 2 beträgt die mittlere Selbstwirksamkeitserwartung $M = 3.92$ ($SD = .80$) und liegt damit, ebenso wie am Ende des Moduls ($M = 3.77$, $SD = 0.85$), über dem theoretischen Skalenmittel von 3.50. Gleichzeitig deuten die Mittelwerte der beiden Erhebungszeitpunkte darauf hin, dass die Selbstwirksamkeitserwartungen der Lehrpersonen am Ende von Modul 2 tendenziell etwas niedriger ausfallen als zu Beginn des Moduls.

Die Selbstwirksamkeitserwartung bezogen auf die Klassenführung wurde mit vier Items erhoben und bildet das Vertrauen in die eigenen Fähigkeiten und Ressourcen ab, Störungen im Unterricht einzudämmen und einen reibungslosen Ablauf sicherzustellen (Beispielitem: *„Wie sicher sind Sie, dass Sie es im Unterricht schaffen, Störungen einzudämmen, wenn Ihre Schüler im Unterricht Blödsinn machen?“*, $\alpha_{t_5} = .92$, $\alpha_{t_7} = .92$). Auch hier reichte das Antwortformat von 1 *„absolut unsicher“* bis 6 *„absolut sicher“*. Deskriptiv deutet sich bei einem Vergleich der Skalenmittelwerte zu den beiden Erhebungszeitpunkten (t_5 : $M = 4.04$, $SD = 0.96$; t_7 : $M = 4.04$, $SD = 0.91$) keine Veränderung für diese Facette der Selbstwirksamkeitserwartung an.

Der dritte Anforderungsbereich, für den die Selbstwirksamkeitserwartung der Lehrkräfte erfasst wurde, ist die adaptive Unterrichtsgestaltung. Lehrpersonen mit einer hohen Selbstwirksamkeitserwartung trauen sich zu, Schülerinnen und Schüler entsprechend ihrer Voraussetzungen zu fördern und angemessene Aufgaben sowohl für stärkere als auch für schwächere Lernende anbieten zu können (Beispielitem: *„Wie sicher sind Sie, dass Sie es im Unterricht schaffen, den Unterricht so zu organisieren, dass sowohl die schwächeren als auch die stärkeren Schüler Aufgaben entsprechend ihrer Fähigkeiten bearbeiten?“*, $\alpha_{t_5} = .86$, $\alpha_{t_7} = .84$). Das Zutrauen der Lehrpersonen, den eigenen Unterricht adaptiv zu gestalten, fällt im Vergleich mit den beiden zuvor genannten Selbstwirksamkeitserwartungen im Durchschnitt etwas geringer aus, liegt aber zu

beiden Erhebungszeitpunkten über dem theoretischen Skalenmittel und ist damit tendenziell immer noch eher hoch ausgeprägt (t_5 : $M = 3.71$, $SD = 0.90$; t_7 : $M = 3.79$, $SD = 0.85$).

Die Transferabsichten der Fortbildungsteilnehmerinnen und -teilnehmer wurden am Ende des vierten Moduls mit einer von Richter, Diehl, Krüger und Vigerske (2011) adaptierten Skala erfasst, die aus fünf Items gebildet wurde (Beispielitem: „*Ich werde in der Zeit nach der Fortbildung die vermittelten Inhalte in der Schule anwenden*“, $\alpha_{t7} = .96$). Hohe Werte auf der Skala sind so zu interpretieren, dass eine Lehrperson plant, die Fortbildungsinhalte für die Anwendung im eigenen schulischen Kontext anzupassen, weiterzuentwickeln und zu nutzen. Für die Teilnehmerinnen und Teilnehmer der Fortbildung *Vielfalt fördern* wurde eine durchschnittliche Transferabsicht von $M = 3.31$ ($SD = 1.21$) ermittelt. Dieser Mittelwert kann so interpretiert werden, dass die Lehrpersonen eher weniger beabsichtigen, sich über die Anpassung und Anwendung der Fortbildungsinhalte nach dem Abschluss von *Vielfalt fördern* Gedanken zu machen. Die Standardabweichung deutet aber darauf hin, dass durchaus eine Varianz zwischen den Lehrpersonen existiert.

Zwischen den unterschiedlichen Prädiktoren bestehen überwiegend signifikant negative oder signifikant positive Zusammenhänge (s. Tabelle 12). Lehrpersonen, die angeben, dass sie die Lernaktivitäten in Modul 2 intensiver nutzen, berichten beispielsweise auch von einer systematischeren kognitiven Verarbeitung der Fortbildungsinhalte zu t_3 ($r = .64$, $p < .01$). Auch die Zusammenhänge zwischen der Bewertung von *Vielfalt fördern* als eine spannende Herausforderung während des Moduls 2 auf der einen Seite und der gleichzeitig erhobenen systematischen Verarbeitung der Fortbildungsinhalte des Moduls 2 ($r = .49$, $p < .01$) und der Module zur Didaktik ($r = .44$, $p < .01$) auf der anderen Seite fallen signifikant und positiv aus. Hier gilt also: Je stärker *Vielfalt fördern* als Herausforderung erlebt wird, desto intensiver denken die teilnehmenden Lehrpersonen über die Inhalte der Module nach. Das Threat Appraisal zu t_3 , d. h. die Bewertung der Fortbildung als eine Bedrohung, hängt hingegen negativ mit der systematischen Verarbeitung zu t_3 zusammen: Je höher die erlebte Bedrohung durch die Fortbildung ist, desto geringer fällt die systematische Verarbeitung zu t_3 aus ($r = -.10$, $p < .05$). Der Zusammenhang ist aber nur schwach.

Die signifikanten Korrelationen ($-.10 \leq r \leq .64$) fallen teilweise zu hoch aus, um alle Prädiktoren in einem Modell gemeinsam prüfen zu können, denn bei einer hohen Korrelation der einzelnen Prädiktoren kann es aufgrund der geteilten Varianz zu Ungenauigkeiten in der Schätzung der Einflussfaktoren kommen. Daher werden im Folgenden die Nutzung der individuellen Lernaktivitäten aus Modul 2 und die zu dem gleichen Zeitpunkt berichtete Bewertung der Fortbildung als eine positiv erlebte Herausforderung ($r = .53$, $p < .01$) nicht zusammen in einem Modell betrachtet. Auch die selbstberichtete Nutzung des Lernangebots und die systematische Verarbeitung der Inhalte der Fortbildung zu den jeweiligen Erhebungszeitpunkten sollten nicht gemeinsam

analysiert werden (t_3 : $r = .64$, $p < .01$; t_6 : $r = .58$, $p < .01$). Um den Einfluss der Nutzung individueller Lernaktivitäten zu untersuchen, werden daher jeweils separate Modelle betrachtet.

Tabelle 12. Korrelation der Prädiktorvariablen auf Individualebene

	ILA3	ILA6	ZBEL3	ZBEL6	CHAL3	THREA3	SYSVA3	SYSVA6
ILA3	1	.57**	-.38**	-.28**	.53***	-.19**	.64**	.35**
ILA6		1	-.30**	-.37**	.45**	-.09	.49**	.58**
ZBEL3			1	.49**	-.41**	-.19**	-.30**	-.30**
ZBEL6				1	-.36**	.08	-.27**	-.44**
CHAL3					1	-.06	.49**	.44**
THREA3						1	-.10*	-.00
SYSVA3							1	.47**
SYSVA6								1

Anmerkung: ILA3 = Nutzung individueller Lernaktivitäten (t_3), ILA6 = Nutzung individueller Lernaktivitäten (t_6), ZBEL3 = erlebte zeitliche Belastung (t_3), ZBEL6 = erlebte zeitliche Belastung (t_6), CHAL3 = Challenge Appraisal (t_3), THREA3 = Threat Appraisal (t_3), SYSVA3 = systematische Verarbeitung (t_3), SYSVA6 = systematische Verarbeitung (t_6), * = $p < .05$, ** = $p < .01$, *** = $p < .001$, $n = 168-423$

8.1 Entwicklung der Motivation zur Diagnostik über das Modul 2

Hängt die Entwicklung der Motivation, Lernvoraussetzungen, -prozesse und/oder -schwierigkeiten von Schülerinnen und Schülern im eigenen Unterricht zu diagnostizieren, über den Zeitraum des zweiten Moduls (t_2-t_4) hinweg unter anderem davon ab, als wie herausfordernd und/oder bedrohend die teilnehmenden Lehrpersonen die Fortbildungen bewerten? Erklärt die Intensität, mit der die Teilnehmerinnen und Teilnehmer die Inhalte von *Vielfalt fördern* verarbeiten, wie sich deren Motivation zur Diagnostik über den Zeitraum des Diagnostikmoduls hinweg entwickelt? Diese Fragen sollen im folgenden Abschnitt untersucht werden.

Es zeigt sich: Wenn die Fortbildung als Bedrohung erlebt wird, wirkt sich dies – nach Kontrolle der motivationalen Ausgangsbedingungen (Motivation zur Diagnostik zu t_2) – negativ auf die Motivation zur Diagnostik zu t_4 aus ($\beta = -.20$, $p < .01$, Abbildung 56). Mit anderen Worten: Je stärker die Lehrpersonen *Vielfalt fördern* als etwas Bedrohliches empfinden oder sich während der Fortbildung unwohl fühlen, desto geringer fällt ihre Bereitschaft, diagnostische Aufgaben zu übernehmen, am Ende des Moduls 2 aus.

Wie das Modell in Abbildung 56 ebenfalls zeigt, erweist sich die systematische Verarbeitung der Fortbildungsinhalte als positiver Prädiktor für die Motivation zur Diagnostik zu t_4 ($\beta = .15$, $p < .05$). Mit einer elaborierteren Verarbeitung und kognitiven Auseinandersetzung geht demnach eine Zunahme der Motivation zur Diagnostik vom Anfang des zweiten Moduls bis zum Ende des zweiten Moduls einher.

Das Challenge Appraisal, also die Bewertung von *Vielfalt fördern* als eine positiv konnotierte Herausforderung ($\beta = .00$, n. s.), und die erlebte zeitliche Belastung aufgrund der Teilnahme an dem zweiten Fortbildungsmodul ($\beta = -.08$, n. s.) haben aus statistischer Sicht keinen bedeutsamen Einfluss auf die Entwicklung der Motivation zur Diagnostik über den Zeitraum von Modul 2 hinweg.

Insgesamt können durch das Regressionsmodell 44 % der Unterschiede in der Motivation der Diagnostik am Ende von Modul 2 erklärt werden, wobei die motivationalen Voraussetzungen zu t_2 ($r = .59$, $p < .001$) die stärkste Vorhersagekraft haben. Das wiederum bedeutet, dass die Bereitschaft und die Freude daran, regelmäßig den Lernvoraussetzungen und -ständen der eigenen Schülerinnen und Schüler durch diagnostische Aktivitäten nachzugehen, vergleichsweise stabil ist und durch die Variablen des Erlebens und der Verarbeitung der Fortbildungsinhalte kaum beeinflusst werden kann.

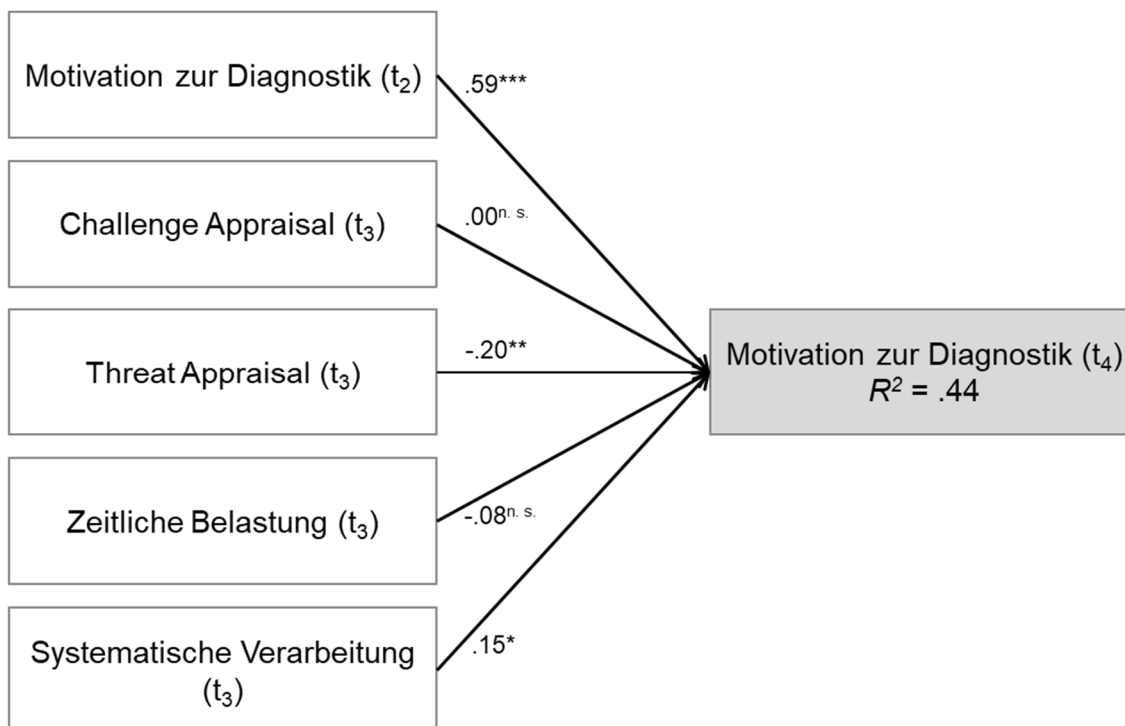


Abbildung 56. Vorhersage der Entwicklung der Motivation zur Diagnostik durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung. Anmerkung: n. s. = nicht signifikant, * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Im nächsten Schritt wurde ein Modell untersucht, das die selbstberichtete Intensität der Nutzung von Lerngelegenheiten aus dem zweiten Modul, die Bewertung von *Vielfalt fördern* als eine Bedrohung und die erlebte zeitliche Belastung als Prädiktoren für die Entwicklung der Motivation zur Diagnostik über den Zeitraum des zweiten Moduls hinweg beinhaltet (s. Abbildung 57).

Auch hier zeigt sich, dass das Empfinden einer zeitlichen Belastung aufgrund von *Vielfalt fördern* unabhängig davon ist, wie sich die Motivation zur Diagnostik bei den Fortbildungsteilnehmerinnen und -teilnehmern entwickelt ($\beta = -.03$, n. s.). Die Bewertung der Fortbildungsteilnahme während des zweiten Moduls als etwas Bedrohliches erweist sich erneut als ein negativer

Einflussfaktor für die Motivationsentwicklung. Mit anderen Worten entwickeln Lehrpersonen weniger Freude und Spaß daran, regelmäßig den Lernvoraussetzungen und Lernständen ihrer Schülerinnen und Schüler nachzugehen, wenn die Fortbildungsteilnahme sie beunruhigt ($r = -.15, p < .05$). Mit einer umfassenderen Nutzung der Lernangebote, die das zweite Modul bereithält, geht hingegen eine positivere Entwicklung der Motivation einher ($r = .17, p < .01$). Das bedeutet also: Lehrpersonen, die die Lerngelegenheiten intensiver nutzen, berichten im Zeitverlauf über eine höhere Motivation zur Diagnostik. Durch das Modell kann insgesamt ein beachtlicher Anteil an Varianz in den Motivationsunterschieden der Lehrkräfte zu t_4 erklärt werden ($R^2 = .51$).

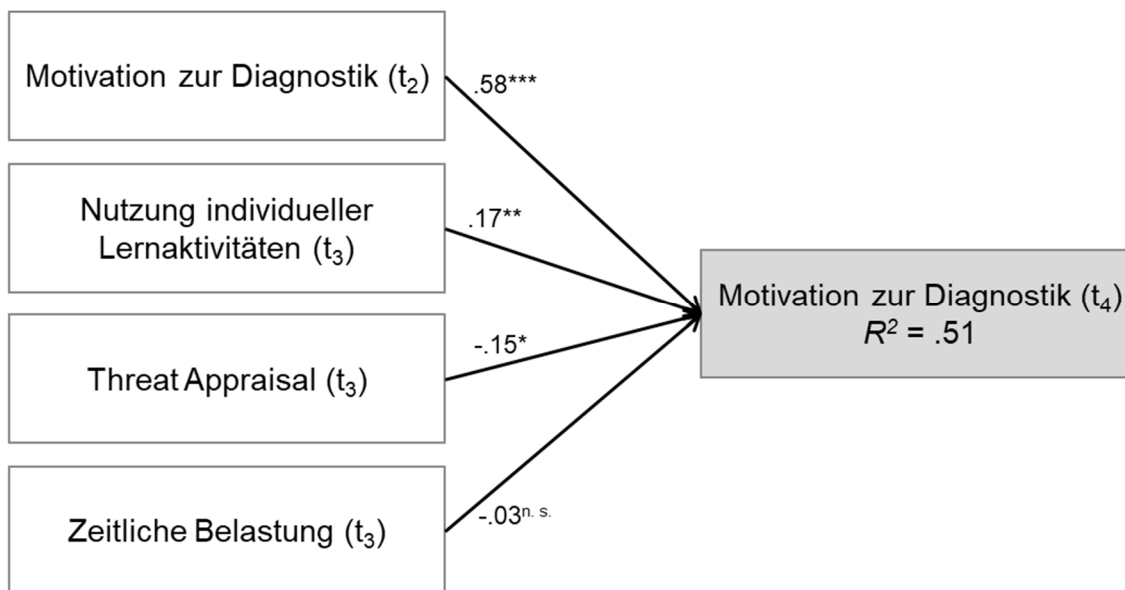


Abbildung 57. Vorhersage der Entwicklung der Motivation zur Diagnostik durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung. Anmerkung: n. s. = nicht signifikant, * = $p < .05$, ** = $p < .01$, *** = $p < .001$

8.2 Entwicklung der Selbstwirksamkeitserwartung im Bereich lernbegleitende Diagnostik im Module 2

Als weiteres Kriterium wurde die Entwicklung der Erwartung der Lehrpersonen betrachtet, eine lernbegleitende Diagnostik in der unterrichtlichen Praxis umsetzen zu können. Darunter wird verstanden, inwieweit sich die Lehrpersonen zutrauen, trotz geringer Ressourcen, zeitlichen Drucks oder einer großen Heterogenität der Schülerinnen und Schüler Lern- und Wissensstände ihrer Lernenden sicher zu beurteilen. Untersucht wird, ob die Entwicklung der Selbstwirksamkeitserwartung von Anfang des Moduls 2 (t_2) bis zum Ende des Moduls (t_4) von der erlebten zeitlichen Belastung durch die Fortbildungsteilnahme und der Bewertung als Bedrohung (Threat Appraisal) beeinflusst wird. In dem einen Modell wird zusätzlich zu den beiden genannten Prädiktoren auch der Einfluss der Bewertung als Herausforderung (Challenge Appraisal) und der systematischen Auseinandersetzung mit den Fortbildungsinhalten untersucht, während in einem zweiten Modell die individuelle Nutzung der Lernaktivitäten von Modul 2 untersucht wird.

In dem ersten Regressionsmodell (s. Abbildung 58) erweist sich, neben dem Ausgangswert der genannten Selbstwirksamkeitserwartung ($r = .41, p < .001$), nur die erlebte zeitliche Belastung als ein signifikanter Prädiktor für die Entwicklung der Selbstwirksamkeitserwartung im Bereich lernbegleitende Diagnostik ($r = .19, p < .01$). Es handelt sich hierbei um einen positiven Einfluss, was so zu interpretieren ist, dass Lehrpersonen, die nach eigenen Angaben aufgrund der Teilnahme an dem zweiten Modul von *Vielfalt fördern* zeitlich belasteter sind, ihre Selbstwirksamkeitserwartung stärker entwickeln.

Die beiden subjektiven Bewertungsmerkmale Challenge und Threat Appraisal und auch das Ausmaß der kognitiven Verarbeitung der Fortbildungsinhalte beeinflussen die Entwicklung der Selbstwirksamkeitserwartung hingegen nicht bedeutsam ($-.08 \leq \beta \leq .09, n. s.$).¹⁷ Zwischen den Lehrpersonen bestehende Unterschiede in der Selbstwirksamkeitserwartung zum zweiten Messzeitpunkt (t_4) können zu 22 % durch das Regressionsmodell erklärt werden.

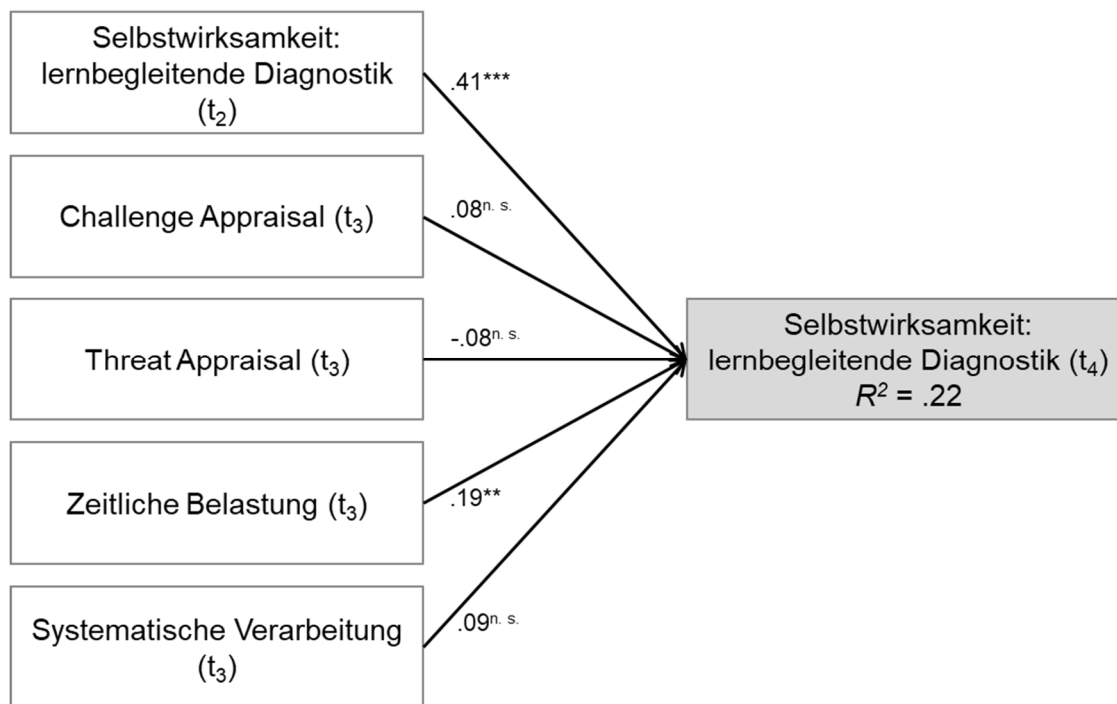


Abbildung 58. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich der lernbegleitenden Diagnostik durch Appraisals, die zeitliche Belastung und die subjektive Verarbeitung. Anmerkung: n. s. = nicht signifikant, ** = $p < .01$, *** = $p < .001$

Welche Einflüsse zeigen sich auf die Entwicklung der Selbstwirksamkeitserwartung, den Lernprozess der Schülerinnen und Schülern angemessen diagnostizieren zu können, wenn anstelle

¹⁷ In einem weiteren Schritt wurde der Einfluss der Bewertung als Herausforderung (Challenge Appraisal), der Bewertung als Bedrohung (Threat Appraisal) und der systematischen Verarbeitung auf die Entwicklung der Selbstwirksamkeitserwartung im Bereich lernbegleitende Diagnostik im Modul 2 jeweils in separaten Regressionsmodellen geprüft. Es zeigt sich, dass die Variablen in diesen Modellen einen signifikanten Einfluss auf die Entwicklung der Selbstwirksamkeitserwartung haben. Das Challenge Appraisal ($\beta = .21, p < .01$) und die systematische Verarbeitung ($\beta = .25, p < .01$) erweisen sich als positive Prädiktoren, während das Threat Appraisal die Entwicklung der Selbstwirksamkeitserwartung negativ beeinflusst ($\beta = -.34, p < .001$).

der beiden Appraisals und der selbstberichteten systematischen Verarbeitung der Fortbildungsinhalte die Intensität der individuellen Nutzung der Lernaktivitäten von Modul 2 betrachtet wird? Aus Abbildung 59 geht hervor, dass die von einer Lehrperson erlebte zeitliche Belastung erneut einen signifikant positiven Einfluss auf die Entwicklung der Selbstwirksamkeitserwartung aufweist ($\beta = .23, p < .01$). Neben der zeitlichen Belastung wirkt sich die Intensität, mit der eine Lehrperson die Lerngelegenheiten nutzt, d. h. beispielsweise wie stark sie über die eigenen Methoden und das eigene Vorgehen bei der Diagnostik nachgedacht und wie intensiv sie diese selbst im Unterricht erprobt hat, positiv darauf aus, wie sich die Selbstwirksamkeit über den Zeitraum des zweiten Moduls hinweg verändert ($\beta = .22, p < .01$). Ob die Lehrpersonen der Teilnahme am zweiten Modul eher mit Sorgen begegnen und diese als etwas Beunruhigendes erleben, wirkt sich hingegen nicht auf die Selbstwirksamkeitserwartung aus ($\beta = -.07, n. s.$).

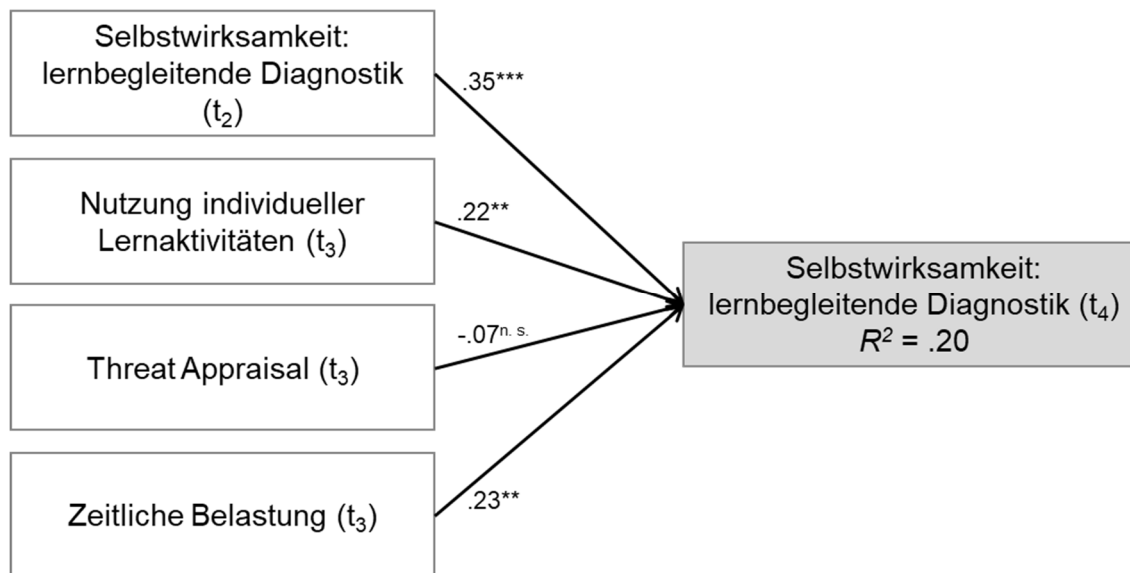


Abbildung 59. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich der lernbegleitenden Diagnostik durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung. Anmerkung: n. s. = nicht signifikant, ** = $p < .01$, *** = $p < .001$

8.3 Entwicklung der Selbstwirksamkeitserwartung im Bereich Klassenführung im Rahmen der Module 3 und 4

Im Folgenden wird die Selbstwirksamkeitserwartung bezogen auf die Klassenführung untersucht. Zur Erinnerung: Lehrpersonen, die über eine hohe Selbstwirksamkeitserwartung in diesem Bereich verfügen, sind davon überzeugt, dass sie einen reibungslosen Unterrichtsablauf gewährleisten und Störungen gering halten können.

Die Selbstwirksamkeitserwartung der Lehrpersonen bezogen auf die Klassenführung erweist sich als sehr stabil ($\beta = .56, p < .001$, s. Abbildung 60). Weder das Erleben der Fortbildung als eine interessante und mit Freude verbundene Herausforderung ($\beta = -.01, n. s.$) noch die Verarbeitung der Fortbildungsinhalte ($\beta = -.05, n. s.$) können etwas zur Erklärung der Selbstwirksamkeitserwartung am Ende der Fortbildung beitragen, wenn man die entsprechende Einschätzung

zum Messzeitpunkt t_5 kontrolliert. Das Threat Appraisal, d. h. die zu t_3 erfasste Sorge bezüglich der Fortbildungsanforderungen und der Konsequenzen, die mit *Vielfalt fördern* verbunden sein können, hat einen negativen Einfluss auf die Entwicklung der Selbstwirksamkeitserwartung im Bereich Klassenführung über den Verlauf der Module 3 und 4. Der Regressionskoeffizient wird aber nur auf dem 10 %-Niveau signifikant ($\beta = -.16, p < .10$). Tendenziell geht eine geringere Bedrohung also mit einer günstigeren Selbstwirksamkeitserwartung im Bereich der Klassenführung über den Zeitraum von Modul 3 und Modul 4 (t_5-t_7) hinweg – also über ein Jahr – einher. Ebenfalls einen auf 10 %-Niveau signifikanten Effekt übt die systematische Verarbeitung der Fortbildungsinhalte, die am Anfang von Modul 4 (t_6) erfasst wurde, auf die Entwicklung der Selbstwirksamkeitserwartung bezogen auf die Klassenführung aus. Mit anderen Worten: Die Selbstwirksamkeitserwartungen entwickeln sich bei denjenigen Lehrpersonen günstiger, die intensiver über die Fortbildungsinhalte und deren Anwendung im eigenen Unterricht nachdenken.

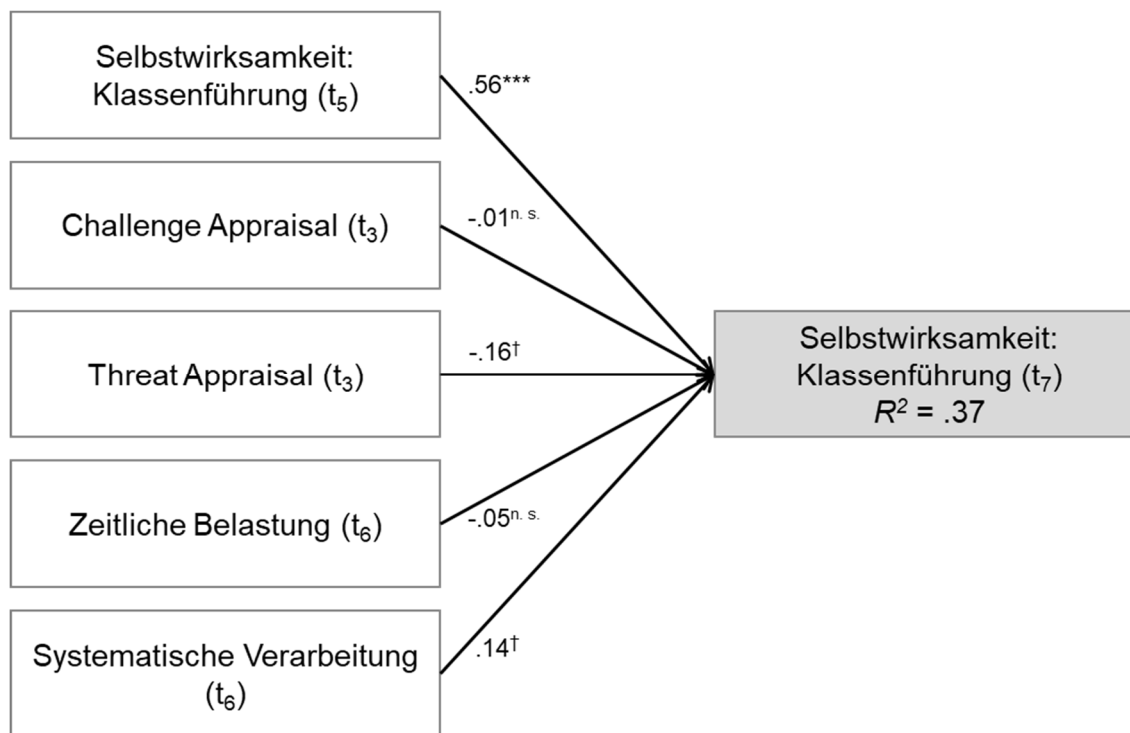


Abbildung 60. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich Klassenführung durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung. Anmerkung: n. s. = nicht signifikant, † = $p < .10$, *** = $p < .001$

Durch das Modell, in dem die Intensität der individuellen Nutzung von Lernaktivitäten untersucht wurde (s. Abbildung 61), können insgesamt 44 % der Unterschiede in der Selbstwirksamkeitserwartung bezogen auf ihre Klassenführung am Ende der Fortbildung (t_7) erklärt werden. Den stärksten Prädiktor stellt, wie schon in dem zuvor erläuterten Modell (s. Abbildung 60), die Ausgangslage der Lehrpersonen dar. Die erlebte zeitliche Belastung erweist sich nicht als statistisch bedeutsamer Einflussfaktor für die Entwicklung der Selbstwirksamkeitserwartung ($\beta = .02, n. s.$). Das Threat Appraisal wird auf dem 10 %-Niveau signifikant ($\beta = -.14, p < .10$). Einen sta-

tistisch bedeutsamen Anteil an der Erklärung von Unterschieden in der Selbstwirksamkeitserwartung der Lehrpersonen am Ende von *Vielfalt fördern* (t_7) besitzt die selbstberichtete Nutzung der Lerngelegenheiten aus den Modulen zum Lernen und Lehren ($\beta = .15, p < .01$). Der Einfluss ist positiv, was bedeutet, dass sich die Selbstwirksamkeitserwartung der Lehrpersonen, einen störungsarmen Unterricht sicherzustellen, positiver entwickelt, wenn die Lehrpersonen während der Module 3 und 4 ihre eigene unterrichtliche Praxis überdenken und hinterfragen sowie unterrichtsbezogene Fortbildungsinhalte erarbeitet und im eigenen Unterricht erprobt haben.

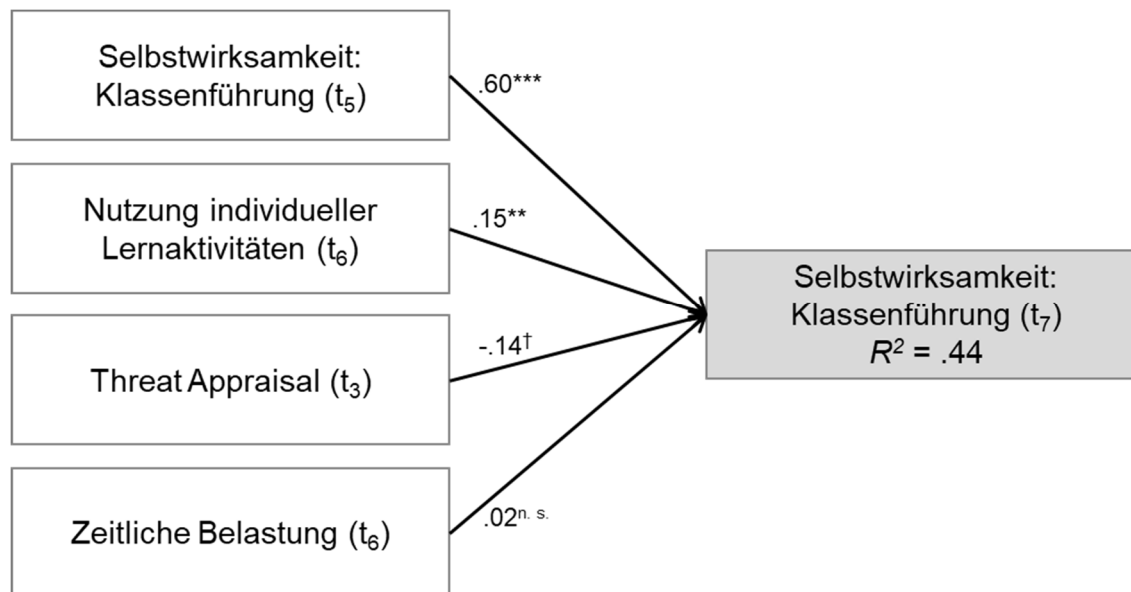


Abbildung 61. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich Klassenführung durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung. Anmerkung: n. s. = nicht signifikant, † = $p < .10$, ** = $p < .01$, *** = $p < .001$

8.4 Entwicklung der Selbstwirksamkeitserwartung im Bereich adaptive Unterrichtsgestaltung im Rahmen der Module 3 und 4

Wodurch lässt sich erklären, dass sich das individuelle Zutrauen von Lehrpersonen, die eigenen Schülerinnen und Schüler entsprechend ihrer Voraussetzungen individuell fördern zu können, über den Zeitraum von Modul 3 und Modul 4 hinweg unterschiedlich entwickelt? Haben das individuelle Erleben und die Nutzung der Fortbildung einen Einfluss auf die Entwicklung der Selbstwirksamkeitserwartung bezogen auf eine adaptive Unterrichtsgestaltung?

Das Ergebnis des Regressionsmodells (s. Abbildung 62) zeigt, dass das Threat Appraisal, also die Bewertung als Bedrohung, einen signifikant negativen Prädiktor darstellt ($\beta = -.30, p < .01$). Anders ausgedrückt: Sorgen sich die Lehrpersonen darum, ob sie die Inhalte des dritten und vierten Moduls richtig verstehen, oder beunruhigt sie die Konfrontation mit den neuen Inhalten, dann entwickelt sich ihr Zutrauen in die eigenen Fähigkeiten und Ressourcen, ihren Unterricht adaptiv zu gestalten, eher ungünstig. Die systematische Verarbeitung der Fortbildungsinhalte in den Didaktikmodulen, d. h. intensivere Überlegungen dazu, wie das neue Wissen mit dem eigenen bisherigen Wissen und der eigenen Unterrichtspraxis in Verbindung gebracht werden kann,

und Überlegungen dazu, welche Konsequenzen sich aus den neuen Kenntnissen für die eigene schulische Praxis ergeben, erweist sich nicht als ein signifikanter Prädiktor für die Selbstwirksamkeitsentwicklung zu t_7 ($\beta = .12$, n. s.). Auch die Bewertung von *Vielfalt fördern* als eine interessante Herausforderung ($\beta = -.06$, n. s.) und das Erleben einer zeitlichen Beanspruchung aufgrund der Fortbildungsteilnahme ($\beta = -.14$, n. s.) beeinflussen die Selbstwirksamkeitserwartung der Lehrkräfte am Ende der Fortbildung (t_7) nicht.

Insgesamt werden durch die Prädiktoren des Modells 34 % der Varianz in der Selbstwirksamkeitserwartung zur adaptiven Unterrichtsgestaltung zu t_7 erklärt. Die Koeffizienten machen aber auch hier deutlich, dass dies im Wesentlichen auf die Selbstwirksamkeitserwartung zum ersten Messzeitpunkt zurückgeht.

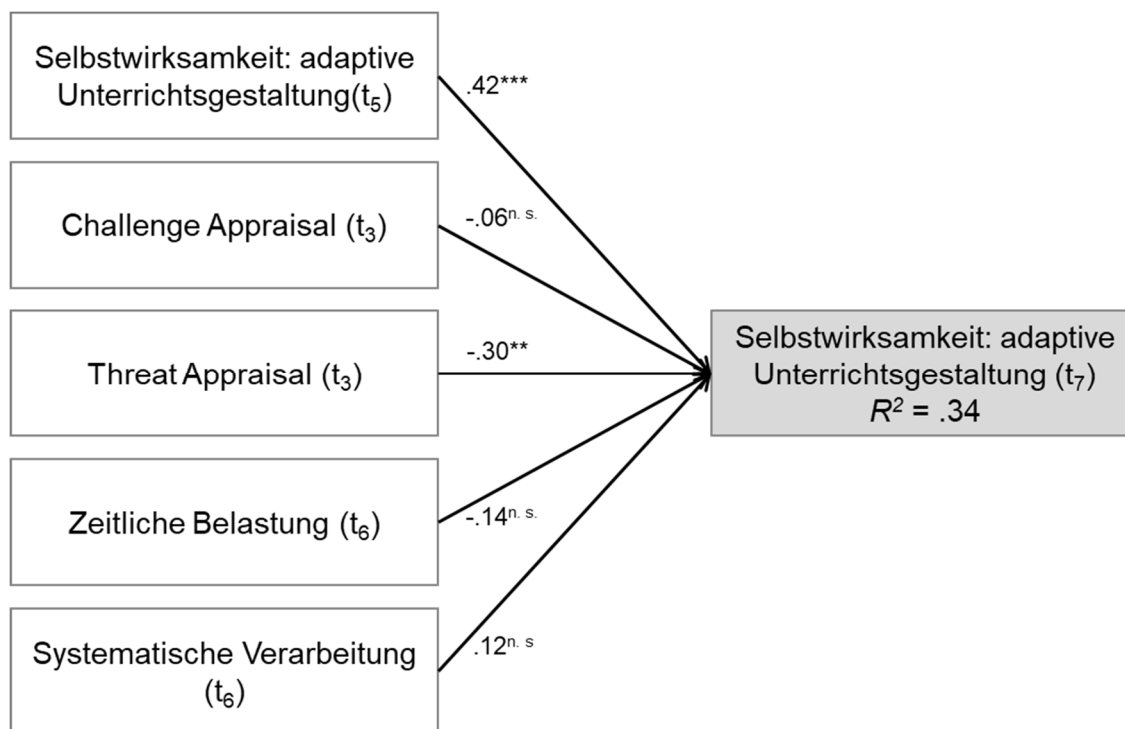


Abbildung 62. Vorhersage der Selbstwirksamkeitserwartung im Bereich adaptive Unterrichtsgestaltung durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung. Anmerkung: n. s. = nicht signifikant, ** = $p < .01$, *** = $p < .001$

Das Modell unter Einbezug der selbstberichteten Nutzungsintensität der Lerngelegenheiten, die am Anfang des vierten Moduls (t_6) erfasst wurde, zeigt einen positiven Einfluss dieser Nutzung von Lerngelegenheiten auf die Entwicklung der Selbstwirksamkeitserwartung im Bereich adaptive Unterrichtsgestaltung (s. Abbildung 63, $\beta = .14$, $p < .05$) und einen negativen Effekt für das Threat Appraisal ($\beta = -.26$, $p < .01$). Hingegen erweist sich die aufgrund der Teilnahme an der Fortbildung subjektiv erlebte zeitliche Belastung auch in diesem Modell nicht als ein signifikanter Prädiktor für die Selbstwirksamkeitserwartung zu t_7 ($\beta = -.08$, n. s.).

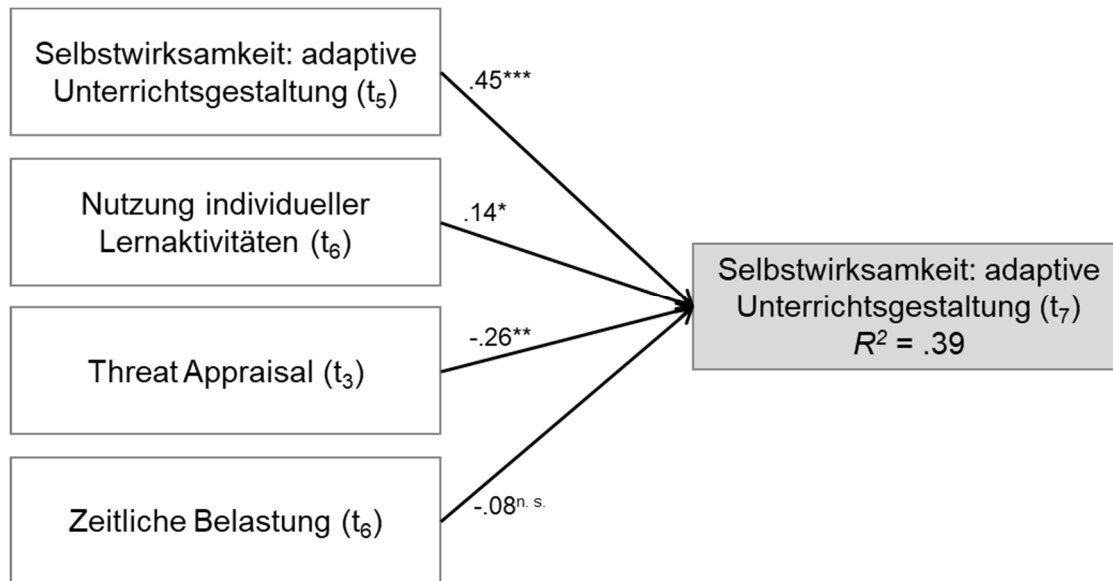


Abbildung 63. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich adaptive Unterrichtsgestaltung durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung. Anmerkung: n. s. = nicht signifikant, * = $p < .05$, ** = $p < .01$, *** = $p < .001$

8.5 Transferabsichten am Ende der Fortbildung

Die Lehrpersonen wurden in der Abschlusserhebung am Ende von Modul 4 (t₇) gefragt, inwieweit sie beabsichtigen, die während des Fortbildungsprozesses kennengelernten Inhalte als Ausgangspunkt zur Entwicklung eigener Ideen zu nutzen und/oder für den eigenen beruflichen Kontext anzupassen und dort anzuwenden. Im Mittel geben die Fortbildungsteilnehmerinnen und -teilnehmer eher eine geringe Absicht zum Transfer der Inhalte an (s. Einleitung von Abschnitt 8).

Um Unterschiede in den Transferabsichten zu erklären, wurden in einem ersten Modell erneut die subjektiven Bewertungen als Herausforderung und Bedrohung (Challenge & Threat Appraisal) zu t₃ sowie die erlebte zeitliche Belastung und die individuelle Auseinandersetzung und inhaltliche Verarbeitung der Fortbildungsinhalte zu zwei Erhebungszeitpunkten (t₃, t₆) als Prädiktoren in einem multiplen Regressionsmodell untersucht.¹⁸ In einem zweiten Modell wurde der Einfluss der individuellen Nutzung der Lernaktivitäten, der Bewertung als Bedrohung (Threat Appraisal) und der erlebten zeitlichen Belastung analysiert.

Aus Abbildung 64 geht das Ergebnis des ersten Modells hervor: Insgesamt werden 45 % der Unterschiede in den selbstberichteten Transferabsichten durch das Vorhersagemodell erklärt. Die stärkste Vorhersagekraft haben das Challenge Appraisal, also die über ein Jahr vor dem Fortbildungsabschluss erfasste Bewertung von *Vielfalt fördern* als eine positive Herausforderung ($\beta = .35$, $p < .001$), und die am Anfang von Modul 4 (t₆) erfasste systematische Verarbeitung der

¹⁸ Es handelt sich hierbei um vergleichsweise viele Prädiktoren, weshalb alternative Modelle mit weniger Prädiktoren untersucht wurden. Es zeigten sich aber keine Unterschiede in der grundlegenden Bedeutung der einzelnen Prädiktoren in den sparsameren Modellen, weshalb hier das umfassendere Regressionsmodell berichtet wird.

Fortbildungsinhalte ($\beta = .34, p < .01$). Auch die systematische Verarbeitung der Fortbildungsinhalte des zweiten Moduls (t_3) erweist sich als ein positiver, wenngleich schwächerer Prädiktor für die Transferabsichten ($\beta = .21, p < .05$). Da es sich um positive Prädiktoren handelt, gilt: Je stärker die Lehrpersonen die Fortbildungsteilnahme als eine anspruchsvolle und spannende Herausforderung betrachten und je elaborierter sie sich mit den Inhalten auseinandersetzen, desto stärker sind ihre Absichten ausgeprägt, die Fortbildungsinhalte nach der Fortbildung zu nutzen und anzuwenden. Das Erleben der zeitlichen Belastung aufgrund der Fortbildungsteilnahme während des zweiten Moduls ($\beta = .21, p < .10$)¹⁹ hat einen auf dem 10 %-Niveau signifikant positiven Einfluss auf die Transferabsichten, während sich die in der zweiten Prozesserhebung (t_6) erfasste zeitliche Belastung ($\beta = -.03, n. s.$)²⁰ – bezogen auf die Didaktikmodule – nicht signifikant auswirkt. Die Bewertung der Fortbildung als Bedrohung (Threat Appraisal) beeinflusst die selbstberichteten Transferabsichten der Lehrpersonen am Ende von *Vielfalt fördern* nicht ($\beta = .01, n. s.$).²¹

¹⁹ Untersucht man den Einfluss der erlebten zeitlichen Belastung zu t_3 auf die Transferabsichten am Ende von *Vielfalt fördern* in einem separaten Modell (einfache lineare Regression), zeigt sich ein signifikant negativer Effekt ($\beta = -.23, p < .01$). Die Variation des Einflusses der erlebten zeitlichen Belastung auf die Transferabsichten zwischen den Modellen ist möglicherweise mit Multikollinearitäten oder Supressoreffekten zu erklären.

²⁰ Der Einfluss der erlebten zeitlichen Belastung zu t_6 bleibt weiterhin nicht signifikant, wenn man die erlebte zeitliche Belastung zu t_3 aus dem Modell entfernt. In einem linearen Regressionsmodell, in dem ausschließlich der Einfluss der erlebten zeitlichen Belastung zu t_6 als Prädiktor für die Transferabsichten untersucht wird, erweist sich die zeitliche Belastung hingegen als ein signifikant negativer Prädiktor ($\beta = -.31, p < .001$). Dies bedeutet, dass eine Lehrperson, die sich am Anfang von Modul 4 stärker von der Fortbildungsteilnahme zeitlich belastet erlebt, geringere Absichten hat, die Fortbildungsinhalte nach der Fortbildung für die eigene schulische Praxis anzupassen und dort anzuwenden.

²¹ Die Bewertung als Bedrohung (Threat Appraisal) stellt sich auch in einem linearen Regressionsmodell nicht als ein signifikanter Prädiktor für die Transferabsichten heraus.

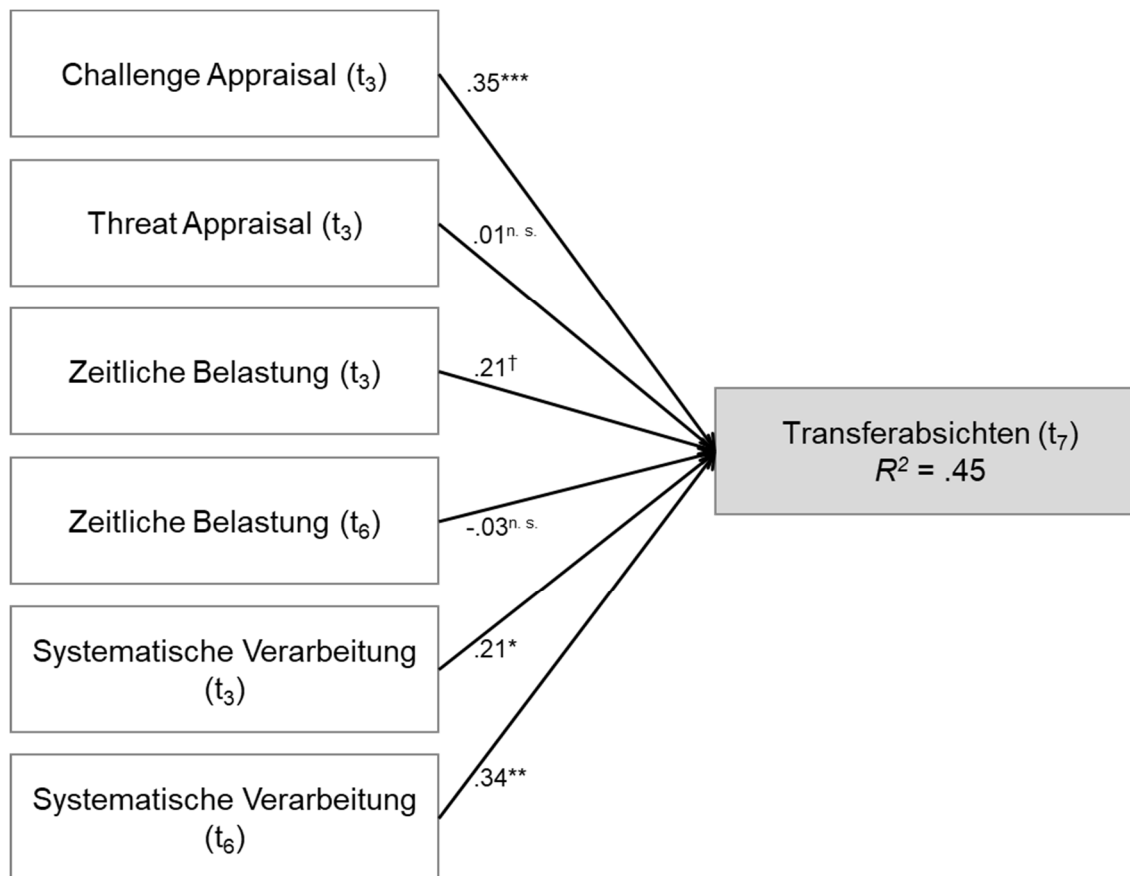


Abbildung 64. Vorhersage der Transferabsichten durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung. Anmerkung: n. s. = nicht signifikant, † = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Durch das zweite Modell (s. Abbildung 65), in dem statt der Bewertung als Herausforderung (Challenge Appraisal) und der systematischen Verarbeitung die selbstberichtete Intensität der Nutzung der Lerngelegenheiten des zweiten Moduls (t₃) untersucht wurde, können 24 % der Varianz in den Transferabsichten am Ende von *Vielfalt fördern* erklärt werden. Die selbstberichteten Transferabsichten werden allerdings ausschließlich davon beeinflusst, wie intensiv die Lehrpersonen die Lernaktivitäten nach eigenen Angaben während des zweiten Moduls genutzt haben ($\beta = .48, p < .001$).²² Ein intensiveres selbstberichtetes Nutzungsverhalten geht mit stärkeren Absichten einher, die Fortbildungsinhalte nach dem Abschluss von *Vielfalt fördern* weiterzuentwickeln und anzuwenden. Weder die Bewertung als Bedrohung noch die erlebte zeitliche Belastung zu t₃²³ und zu t₆ erweisen sich als signifikante Prädiktoren ($-.11 \leq \beta \leq .06$, n. s.).

²² Das maßgebliche Ergebnis des Modells ändert sich nicht, wenn man anstelle der selbstberichteten Intensität der Nutzung der Lernaktivitäten des zweiten Moduls (t₃) die Angaben zu t₆, d. h. am Anfang von Modul 4 betrachtet. Es ändert sich lediglich die Stärke dieses Prädiktors ($\beta = .56, p < .001$).

²³ Dass die zeitliche Belastung zu t₃ die Transferabsichten nicht erneut (wie in Abbildung 64) beeinflusst, deutet darauf hin, dass das zeitliche Belastungserleben in seiner Vorhersagekraft nicht stabil ist, da der Einfluss je nach Modellspezifikation variiert. Als Erklärung kommen auch hier möglicherweise Multikollinearitäten oder Supressoreffekte in Frage.

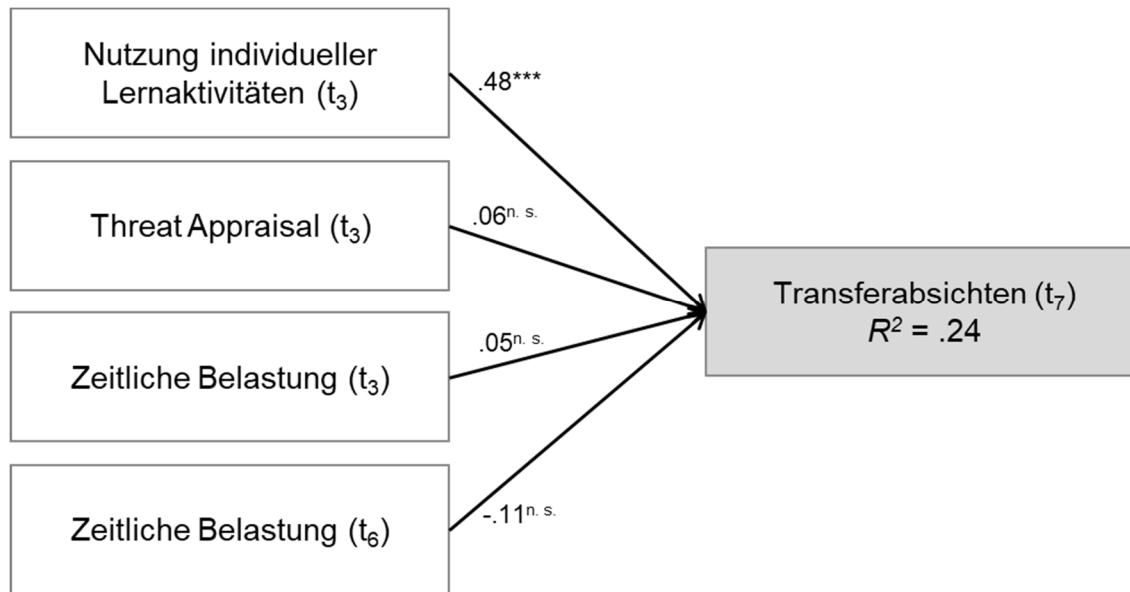


Abbildung 65. Vorhersage der Transferabsichten durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung. Anmerkung: n. s. = nicht signifikant, *** = $p < .001$

9 Lerngelegenheiten der Fortbildung

In den folgenden Abschnitten geht es um die Inhalte des Curriculums von *Vielfalt fördern*. Zunächst wird deskriptiv berichtet, wie die Moderatorinnen und Moderatoren den verfügbaren zeitlichen Rahmen für die Fortbildungsinhalte beurteilen und die Freiräume für eine Orientierung an den Bedürfnissen der Teilnehmerinnen und Teilnehmer bewerten (s. Abschnitt 9.1). In Abschnitt 9.2 wird anschließend aus der Perspektive der teilnehmenden Lehrpersonen und der Moderatorinnen und Moderatoren dargestellt, wie intensiv – im Sinne von ausführlich – unterschiedliche Themen behandelt wurden. Abschließend wird untersucht, ob die Intensität der Thematisierung von unterschiedlichen Inhalten die Entwicklung der beiden im vorigen Abschnitt (s. Abschnitt 8) vorgestellten Merkmale – Motivation zur Diagnostik während des Moduls 2 (t_2 – t_4) und die Selbstwirksamkeitserwartung in Bezug auf eine adaptive Unterrichtsgestaltung über den Verlauf der Module 3 und 4 (t_5 – t_7) – vorhersagt.

9.1 Zeit für die Fortbildungsinhalte und Freiräume für Teilnehmerorientierung

Während der abschließenden Erhebung (t_7) wurden die Moderatorinnen und Moderatoren der zweiten Kohorte um eine globale Einschätzung darüber gebeten, ob ausreichend Zeit für die Behandlung der geplanten Inhalte gegeben war (Beispielitem: „*Es bestand ausreichend Zeit, um alle geplanten Inhalte in den Fortbildungsveranstaltungen unterzubringen*“). Insgesamt setzt sich die Skala aus vier Items zusammen ($\alpha = .92$).

Zudem haben die Moderatorinnen und Moderatoren beurteilt, inwieweit während der Fortbildung genügend Freiräume vorhanden waren, um sich an den Wünschen der Schulen und den Bedürfnissen der Lehrpersonen zu orientieren. Die Skala, mit der die wahrgenommenen Freiräume erfasst wurden, besteht aus fünf Items (Beispielitem: „*Das Fortbildungscurriculum ließ ausreichend Spielraum, um sich an den Bedürfnissen der Teilnehmer zu orientieren*“, $\alpha = .76$).

Es liegt nahe, dass Moderatorinnen und Moderatoren, die einen hohen Zeitdruck bei der Vermittlung der curricularen Inhalte erleben, auch über geringere Freiräume berichten, die Bedürfnisse der Lehrpersonen zu berücksichtigen.

Der Mittelwert für die eingeschätzte Zeit für die Fortbildungsinhalte von $M = 2.85$ ($SD = 1.28$) liegt unter dem theoretischen Skalenmittel ($M = 3.50$) und deutet somit darauf hin, dass das Fortbildungscurriculum von *Vielfalt fördern* eher (zu) viele Themen und Inhalte für die verfügbare Fortbildungszeit vorsieht. Der Mittelwert für die Skala Freiräume zur Teilnehmerorientierung liegt bei $M = 3.85$ ($SD = 0.53$). Diese mittlere Einschätzung indiziert, dass die Moderatorinnen und Moderatoren das Fortbildungskonzept nicht als zu starr empfunden haben, sondern es durchaus für möglich halten, sich an den Bedürfnissen der Teilnehmerinnen und Teilnehmer zu orientieren. Zwischen den beiden Merkmalen besteht zwar – wie erwartet – ein negativer Zusammenhang ($r = -.23$, n. s.), der jedoch nicht signifikant wird.

9.2 Intensität der Bearbeitung von Fortbildungsthemen

Der vorliegende Abschnitt untergliedert sich erneut in zwei Teile. Zunächst wird die Intensität der Behandlung unterschiedlicher Fortbildungsthemen anhand deskriptiver Daten dargestellt. Dabei werden die Einschätzungen der Lehrpersonen und die der Moderatorinnen und Moderatoren einander gegenübergestellt. Anschließend wird die Behandlung von Fortbildungsinhalten als Prädiktor genutzt, um die Entwicklung der Motivation zur Diagnostik während des Moduls 2 und der Selbstwirksamkeitserwartung in Bezug auf eine adaptive Unterrichtsgestaltung während der Module 3 und 4 (s. Abschnitt 8) zu erklären. Die pädagogische Diagnostik und die adaptive Unterrichtsgestaltung stellen wesentliche Aspekte der Fortbildung *Vielfalt fördern* dar, sodass Effekte der Fortbildungsteilnahme auf motivationale Merkmale der Lehrpersonen in diesem Bereich plausibel erscheinen.

9.2.1 Deskriptive Betrachtung der Thematisierungsintensität

Aus der Perspektive der teilnehmenden Lehrpersonen und aus der Perspektive der Moderatorinnen und Moderatoren wurde erfasst, wie intensiv unterschiedliche Inhalte im Laufe der Module thematisiert wurden. Das Antwortformat war vierstufig von 1 „gar nicht“ bis 4 „ausführlich“, sodass höhere Werte eine intensivere Behandlung bedeuten.

In Tabelle 13 sind die durchschnittlichen Einschätzungen der Lehrpersonen und der Moderatorinnen und Moderatoren für die verschiedenen erfassten Themen sowie die Differenz, die zwischen den Urteilen der beiden Personengruppen vorliegt, dargestellt. Es zeigt sich, dass die Intensität, mit der die verschiedenen Lerninhalte behandelt wurden, von den Moderatorinnen und Moderatoren in den meisten Fällen höher bewertet wurde als von den Lehrpersonen. Eine Ausnahme stellen in den Modulen zum Lernen und Lehren die Behandlung von empirischen Befunden zur Lerndokumentation und die Leseförderung mit den Textdetektiven dar. Hier geben die Lehrpersonen im Mittel eine intensivere Thematisierung an als die Moderatorinnen und Moderatoren (Tabelle 13, orange hervorgehoben) und gleichzeitig handelt es sich um die beiden Inhalte, in denen sich die geringste Differenz ($0.03 < M\text{-Diff} < 0.07$) zwischen den Einschätzungen der Lehrpersonen und der Moderatorinnen und Moderatoren zeigt.

Geht man von den Mittelwerten aus, dann beurteilen die teilnehmenden Lehrpersonen und die Moderatorinnen und Moderatoren übereinstimmend, dass das Thema *Hypothesenbildung/-formulierung* am ausführlichsten (Tabelle 13, grün hervorgehoben) während des Diagnostikmoduls behandelt wurde ($M_{LPS} = 3.03$, $SD_{LPS} = 0.76$; $M_{Mod} = 3.63$, $SD_{Mod} = 0.77$). In den Modulen zum Lernen und Lehren divergieren die Einschätzungen der beiden Personengruppen. Während die Lehrpersonen im Mittel angegeben haben, dass *Instrumente, die die Lernplanung unterstützen* (z. B. *Wochenpläne, Lernlandkarten*), am umfassendsten thematisiert wurden ($M = 2.71$, $SD = 0.85$), geben die Moderatorinnen und Moderatoren an, das Thema *Merkmale guter (kognitiv aktivierender, kompetenzorientierter) Aufgaben* am stärksten bearbeitet zu haben ($M = 3.68$,

$SD = 0.48$). Hier zeigt sich mit einer Differenz von -1.10 auch die größte Abweichung zwischen der Wahrnehmung der Lehrpersonen und der Wahrnehmung der Moderatorinnen und Moderatoren.

Unterschiedlich wahrgenommen wird auch, welche Themen in den betrachteten Modulen am kürzesten aufgegriffen (Tabelle 13, lila hervorgehoben) wurden. Im Durchschnitt geben die Lehrpersonen für das Modul 2 an, dass zentrale Aspekte, die bei der Datenerhebung zu beachten sind, nur kurz besprochen wurden ($M = 1.88$, $SD = 0.79$). Die Moderatorinnen und Moderatoren haben nach eigenen Angaben dem Verhältnis von Individualisierung und Standardisierung bei der Diagnostik am wenigsten Zeit gewidmet ($M = 2.40$, $SD = 0.84$). Bei den Didaktikmodulen geben beide Personengruppen an, dass die geringste Zeit für die Besprechung konkreter Ansätze der Leseförderung verwendet wurde, wobei die Lehrpersonen das Konzept „Lesen(d) lernen“ (Leseschule NRW) nennen ($M = 1.56$, $SD = 0.89$) und die Moderatorengruppe die Leseförderung mit den Textdetektiven ($M = 1.59$, $SD = 0.85$).

Betrachtet man die Korrelationen zwischen den Wahrnehmungen der Kollegien und ihres Moderatorenteams für die einzelnen Themen (s. Tabelle 13), fällt auf, dass sich über alle 34 Themen hinweg insgesamt nur drei signifikant positive Zusammenhänge zeigen. Diese Korrelationen der beiden Perspektiven beziehen sich darauf, wie ausführlich Lernlandkarten ($r = .83$, $p < .001$) und Klassenarbeiten ($r = .56$, $p < .05$) als Beobachtungs- beziehungsweise Testinstrument behandelt wurden, sowie auf den Umfang, mit dem auf das Konzept „Lesen(d) lernen“ (Leseschule NRW) eingegangen wurde ($r = .86$, $p < .001$). Dieses Ergebnis mag auf den ersten Blick erstaunen, da die Lehrpersonen und Moderatorinnen und Moderatoren doch an den gleichen Veranstaltungen teilgenommen haben. In gewisser Weise steht es aber durchaus im Einklang mit Befunden aus der Unterrichtsforschung. Dort zeigt sich, dass auch die Wahrnehmungen von Lehrkräften und ihren Schülerinnen und Schülern auf den Unterricht nicht konvergieren (Clausen, 2002).

Tabelle 13. Mittelwerte, Standardabweichungen, Differenzen und Korrelationen für die Lerngelegenheiten aus Sicht der Lehrpersonen und Moderatorinnen und Moderatoren

Fortbildungsinhalt	Lehrpersonen		Moderatoren		M-Diff	r
	M	SD	M	SD		
<i>Modul 2</i>						
Typische Beurteilungstendenzen	2.28	0.70	2.54	0.67	-0.26	-.08
Verschiedene Beurteilungsfehler	2.26	0.70	2.49	0.75	-0.23	-.09
Unterschiede zwischen impliziter und expliziter Diagnostik	2.07	0.76	2.85	0.73	-0.78	.37
Anwendungsfelder pädagogischer Diagnostik (z. B. Einschulung, Leistungsbeurteilung, Differenzierung)	2.35	0.77	3.02	0.76	-0.67	.03
Bereiche pädagogischer Diagnostik (z. B. Lernvoraussetzungen, Lernkontexte, Unterrichtsqualität)	2.66	0.78	3.46	0.55	-0.80	-.35
(Teil-)Schritte im diagnostischen Prozess	2.74	0.80	3.44	0.63	-0.70	.04
Kennzeichen einer diagnostischen Fragestellung	2.85	0.80	3.44	0.55	-0.58	.24
Hypothesenbildung/-formulierung	3.03	0.76	3.63	0.77	-0.60	.24
Diagnostische Grundmethoden (Beobachten, Befragen, Testen)	2.71	0.73	3.46	0.67	-0.75	.27
Die Auswahl von Diagnoseinstrumenten	2.65	0.79	3.25	0.74	-0.60	.24
Die Lernlandkarte als Instrument der Befragung	2.43	0.86	3.15	0.82	-0.71	.83***
Beobachtungsbögen für die kollegiale Unterrichtshospitation	2.96	0.82	3.51	0.60	-0.55	.43†
Klassenarbeiten als eine Form des Testens	2.32	0.84	3.05	0.92	-0.73	.56*
Das Verhältnis von Individualisierung und Standardisierung bei der Diagnostik	2.07	0.82	2.40	0.84	-0.33	.22
Zentrale Aspekte, die bei der Datenerhebung zu beachten sind	1.88	0.79	2.45	0.90	-0.57	.42†
<i>Modul 3–4</i>						
Selbstgesteuertes Lernen	2.64	0.74	3.27	0.70	-0.63	-.24
Lehrerrollen bei verschiedenen Unterrichtsformen	2.26	0.78	2.68	0.72	-0.42	.29
Faktoren, die das Lernen beeinflussen (z. B. Lernstrategien, Motivation)	2.60	0.80	3.18	0.85	-0.59	.02
Instrumente, die die Lernplanung unterstützen (z. B. Wochenpläne, Lernlandkarten)	2.71	0.85	3.64	0.66	-0.92	-.33
Helfersysteme (Schüler helfen Schülern) im Unterricht	2.23	0.78	2.59	0.80	-0.36	.01
Maßnahmen zur Förderung des Selbstvertrauens	1.91	0.80	2.27	0.63	-0.36	-.22
Attributionsmuster (Typen von Ursachenzuschreibungen)	1.86	0.79	2.05	1.00	-0.19	.21

Fortbildungsinhalt	Lehrpersonen		Moderatoren			
Lern- und Arbeitstechniken von Schülern	2.27	0.76	3.09	0.68	-0.82	-.16
Leseförderung mit den Textdetektiven	1.66	0.82	1.59	0.85	0.07	.48
Leseförderung nach dem Konzept „Lesen(d) lernen“ (Leseschule NRW)	1.56	0.89	1.77	1.07	-0.22	.86***
Förderung der Selbsteinschätzungskompetenz von Schülern	2.34	0.79	2.91	0.61	-0.57	.14
Instrumente zur Lernprozessreflexion	2.45	0.72	2.86	0.77	-0.41	-.35
Aufgaben eines Lerncoaches/Mentors	2.12	0.80	2.73	0.70	-0.61	.10
Merkmale guter (kognitiv aktivierender, kompetenzorientierter) Aufgaben	2.58	0.84	3.68	0.48	-1.10	-.34
Aufgabentypen zur Differenzierung (z. B. Zusatzaufgaben, Blütenaufgaben)	2.67	0.87	3.45	0.67	-0.78	-.09
Kriterien zur Differenzierung	2.62	0.78	3.36	0.66	-0.74	-.24
Empirische Befunde zur Lerndokumentation	2.13	0.79	2.10	0.72	0.03	.32
Formen der Lerndokumentation (z. B. Lerntagebuch, Portfolio)	2.58	0.85	2.90	0.72	-0.32	.46
Empirische Befunde zu Merkmalen lernwirksamen Unterrichts	2.11	0.68	2.41	0.85	-0.30	.15

Anmerkung: grüne Hervorhebung = das am ausführlichsten behandelte Thema eines Moduls je nach Perspektive, lila Hervorhebung = das am kürzesten behandelte Thema eines Moduls je nach Perspektive, orange Hervorhebung = ausführlichere Thematisierung nach Angaben der Lehrpersonen als nach Angaben der Moderatorinnen und Moderatoren. *** = $p < .001$, ** = $p < .01$, * = $p < .05$, † = $p < .10$

9.2.2 Die Thematisierungsintensität als erklärendes Merkmal

Im anschließenden Analyseschritt wurde mittels Regressionsanalysen²⁴ untersucht, ob über die Intensität der Lerngelegenheiten – wahrgenommen von den Lehrpersonen – die Entwicklung der Motivation zur Diagnostik im Verlauf des Moduls 2 (t_2-t_4) und die Entwicklung der Selbstwirksamkeitserwartung in Bezug auf eine adaptive Unterrichtsgestaltung im Verlauf der Module 3 und 4 (t_5-t_7) vorhergesagt werden kann. Als Prädiktoren werden nicht alle Lerngelegenheiten betrachtet, sondern es werden jene fünf Lerngelegenheiten ausgewählt, bei denen eine besondere inhaltliche Nähe zu dem Kriterium besteht. Die einzelnen Lerngelegenheiten der Module wurden zu dem jeweiligen Erhebungszeitpunkt in einem ersten Schritt zu einem Summenscore zusammengefasst. Da das Antwortformat von 1 „gar nicht“ bis 4 „ausführlich“ reichte und jeweils fünf Items ausgewählt wurden, kann der Summenscore maximal den Wert 20 erreichen. Der Summenscore ist folgendermaßen zu interpretieren: Je höher der Summenscore, desto intensiver wurden die fünf ausgewählten Themen behandelt. Vice versa gilt: Je niedriger der Summenscore, desto weniger wurden die Inhalte thematisiert.

²⁴ Die Regressionsanalysen wurden auf Individualebene durchgeführt. Bei der Modellierung wurde jedoch die Gruppierung der Lehrpersonen in Schulen durch den in Mplus eingebetteten Analysebefehl „complex“ berücksichtigt.

Für das Modul 2, d. h. zur Vorhersage von Unterschieden in der Entwicklung der Motivation zur Diagnostik über den Zeitraum des zweiten Moduls hinweg wurden folgende fünf Lerngelegenheiten ausgewählt, die zu einem Summenscore zusammengefasst wurden: *Bereiche pädagogischer Diagnostik, (Teil-)Schritte im diagnostischen Prozess, diagnostische Grundmethoden, die Auswahl von Diagnoseinstrumenten und Beobachtungsbögen für die kollegiale Unterrichtshospitation*. Die Werte reichen von $Min = 5$ bis $Max = 20$ ($M = 13.5$, $SD = 3.0$).²⁵

Zur Erklärung von Unterschieden in der Selbstwirksamkeitserwartung bezüglich der adaptiven Unterrichtsgestaltung werden folgende Lerngelegenheiten zum Summenscore zusammengefasst: *selbstgesteuertes Lernen, Lern- und Arbeitstechniken von Schülern, Instrumente zur Lernprozessreflexion, Aufgabentypen zur Differenzierung und Formen der Lerndokumentation*. Der Maximalwert des entsprechenden Summenscores beträgt zu t_6 $Max = 19$ ($M = 12.4$, $SD = 2.9$).

Um Unterschiede in den Voraussetzungen der Lehrpersonen zu berücksichtigen, werden für die folgenden Modelle jeweils die Ausgangswerte zum ersten Messzeitpunkt einbezogen. Das bedeutet: In den Modellen zur Vorhersage der Motivation zur Diagnostik wird der Wert, den die Lehrpersonen vor dem Modul 2 (t_2) berichtet haben, kontrollierend einbezogen. So prüfen die entsprechenden Modelle, ob die Lerngelegenheiten aus Modul 2 (t_3) die Motivation am Ende des Moduls 2 (t_4) nach Kontrolle der individuellen Voraussetzungen der Lehrpersonen vorhersagen können. Letztlich wird damit also die Entwicklung der Motivation zur Diagnostik vorhergesagt. Analog wird mit der zweiten Kriteriumsvariable, der Selbstwirksamkeitserwartung zur adaptiven Unterrichtsgestaltung, verfahren. Hier wird der Ausgangswert zu Beginn von Modul 3 (t_5) kontrolliert und damit die Entwicklung über die Module 3 und 4 untersucht.

Wie bereits in dem Modell zur Erklärung der Motivationsentwicklung durch Merkmale der subjektiven Wahrnehmung und Verarbeitung der Fortbildungsinhalte (s. Abschnitt 8.1, Abbildung 56) sichtbar wurde, erweisen sich auch in der hier durchgeführten Analyse die motivationalen Voraussetzungen zu Beginn des zweiten Moduls als wirkmächtige Prädiktoren für die Motivation am Ende des Moduls (Abbildung 66, $\beta = .66$, $p < .001$), sodass von einer vergleichsweise hohen Stabilität der Motivation zur Diagnostik ausgegangen werden muss. Die Intensität, mit der unterschiedliche Themen aus dem Bereich der pädagogischen Diagnostik behandelt werden, wirkt sich hingegen nicht auf die Entwicklung der Motivation der Lehrpersonen aus, pädagogische Diagnostik im eigenen Unterricht betreiben zu wollen ($\beta = .08$, n. s.).²⁶

²⁵ Dass die Werte insgesamt so stark variieren, erstaunt auf den ersten Blick, da das Curriculum von *Vielfalt fördern* festgelegt war und insofern vergleichbar sein sollte. Zu beachten ist aber, dass die Angaben der Lehrpersonen auch aufgrund der Nichtanwesenheit bei einigen Fortbildungsterminen und aufgrund unterschiedlich intensiver Nutzung der Lerngelegenheiten bei der schulischen Erprobung so stark variieren können.

²⁶ In einem zweiten Schritt wurden die fünf Lerngelegenheiten in separaten Modellen untersucht. Es zeigt sich, dass eins der fünf Themen – die Thematisierung von (Teil-)Schritten im diagnostischen Prozess – einen signifikant positiven Einfluss auf die Entwicklung der Motivation zur Diagnostik besitzt ($\beta = .12$, $p < .05$). Eine intensivere Auseinandersetzung damit, in welchen Schritten eine systematische pädagogische Diagnostik abläuft, führt demnach dazu, dass die Lehrpersonen eine höhere Bereitschaft entwickeln, die Lernvoraussetzungen der Schülerinnen und Schüler zu diagnostizieren und den Ursachen von Lernschwierigkeiten auf den Grund zu gehen.

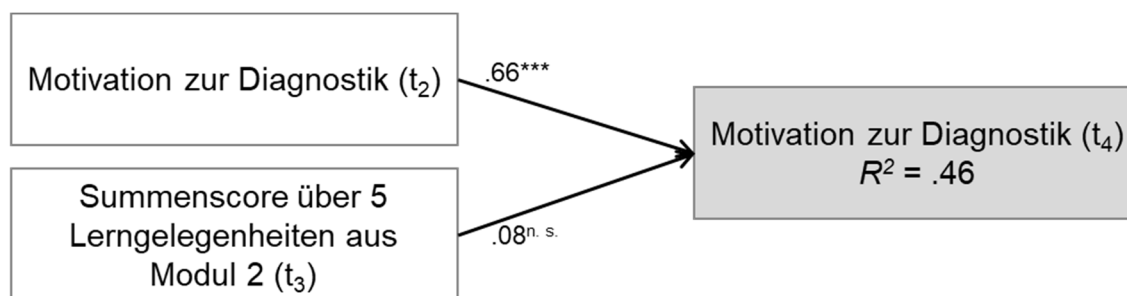


Abbildung 66. Vorhersage der Entwicklung der Motivation zur Diagnostik durch die Intensität der Thematisierung von Lerngelegenheiten aus den Modul Diagnostik. Anmerkung: n. s. = nicht signifikant, *** = $p < .001$

Das Modell zur Erklärung von Varianzanteilen in der Selbstwirksamkeitserwartung, den eigenen Unterricht adaptiv zu gestalten (s. Abbildung 67), zeigt, dass neben der anfänglichen Selbstwirksamkeitserwartung am Anfang von Modul 3 ($\beta = .52, p < .001$) auch der Summenscore, der über fünf Themen aus den Modulen Lernen und Lehren gebildet wurde, einen signifikant positiven Einfluss auf die Entwicklung der Selbstwirksamkeitserwartung der Lehrpersonen besitzt ($\beta = .23, p < .05$). Eine ausführlichere Behandlung der Modulinhalte *selbstgesteuertes Lernen, Lern- und Arbeitstechniken von Schülern, Instrumente zur Lernprozessreflexion, Aufgabentypen zur Differenzierung* und *Formen der Lerndokumentation* geht demnach mit einer günstigeren Entwicklung der Selbstwirksamkeitserwartung bei der adaptiven Unterrichtsgestaltung über den Zeitraum der Module 3 und 4 einher.²⁷

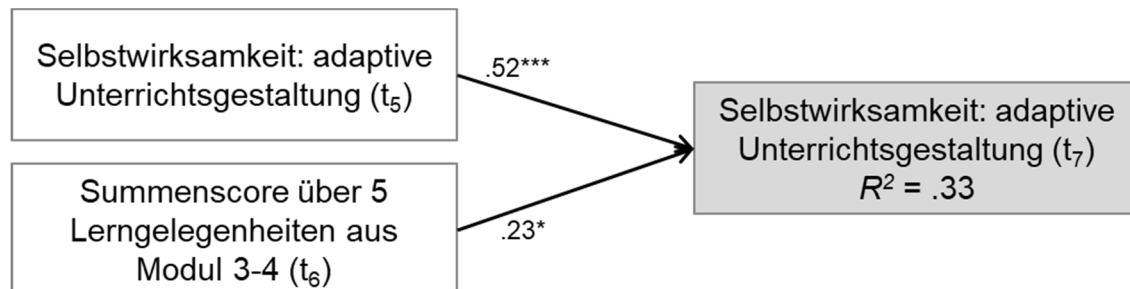


Abbildung 67. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich der adaptiven Unterrichtsgestaltung durch die Intensität der Thematisierung von Lerngelegenheiten aus den Modulen zum Lernen und Lehren. Anmerkung: * = $p < .05$, *** = $p < .001$

²⁷ Auch hier wurde im zweiten Schritt der Einfluss der einzelnen Lerngelegenheiten auf die Entwicklung der Selbstwirksamkeitserwartung untersucht. Es zeigt sich, dass die Intensität der Thematisierung von vier Inhalten (Lern- und Arbeitstechniken von Schülern, Instrumente zur Lernprozessreflexion, Aufgabentypen zur Differenzierung und Formen der Lerndokumentation) einen signifikant positiven Einfluss auf die Entwicklung der Selbstwirksamkeitserwartung hat ($.16 \leq \beta \leq .29, p < .01-.05$). Ausschließlich die Intensität, mit der das Thema selbstgesteuertes Lernen behandelt wurde, beeinflusst die Entwicklung der Selbstwirksamkeitserwartung nicht bedeutsam ($\beta = .09, n. s.$).

10 Enthusiasmus des Moderatorenteams

Die Forschungslage zum Einfluss von Moderatorinnen und Moderatoren auf den Prozess und Ertrag von Lehrerfortbildungen ist insgesamt eher dünn. Ein Grund hierfür liegt auch darin, dass selten eine größere Anzahl an Fortbildnerinnen und Fortbildnern annähernd zeitgleich dasselbe Fortbildungsprogramm anbietet. Die Evaluation von *Vielfalt fördern* kann hierzu wichtige Erkenntnisse beisteuern, da es möglich ist, eine Gruppe von Moderatorinnen und Moderatoren, die alle eine gemeinsame Qualifizierung durchlaufen haben, zu untersuchen und zusätzlich mit Angaben der teilnehmenden Lehrpersonen in Verbindung zu bringen.

Die von den Lehrpersonen wahrgenommene Moderationskompetenz des Moderatorenteams an ihrer Schule wurde in Abschnitt 7.1.2 bereits einmal für einen deskriptiven Vergleich von Schulen mit und ohne sichtbare Entwicklungen während des Fortbildungszeitraumes herangezogen. In der beschreibenden Gegenüberstellung zeigte sich kein auffälliger Unterschied in der Einschätzung des Moderatorenteams zwischen den Schulen der beiden Extremgruppen.

Die Bedeutung von Merkmalen des Moderatorenteams soll daher im folgenden Abschnitt anhand der vorliegenden Daten aus der zweiten Kohorte von *Vielfalt fördern* vertiefend untersucht werden. In einer Studie zum Lehrerenthusiasmus konnte gezeigt werden, dass der Lehrerenthusiasmus einen positiven Einfluss auf die Lernfreude von Schülerinnen und Schülern im Unterricht hat (Bleck, 2018). Es erscheint plausibel, dass ein positives affektiv-motivationales Erleben nicht nur von Lehrpersonen auf Schülerinnen und Schüler „überschwappen“ kann, sondern dass sich auch erwachsene Lernerinnen und Lerner von einem enthusiastischen Moderatorenteam „anstecken“ lassen.

Der Enthusiasmus des Moderatorenteams wurde im Projekt *LIQUID* aus der Perspektive der teilnehmenden Lehrpersonen und als selbstberichteter Enthusiasmus der Moderatorinnen und Moderatoren erfasst. Die Erhebung erfolgte bei den Lehrpersonen ausschließlich im Rahmen der ersten Prozesserhebung in Modul 2 (t_3) mit einer Skala aus sechs Items (Beispielitems: „Die Moderatoren zeigen, dass sie die Arbeit mit Lehrern begeistert“ und „Die Moderatoren vertreten die Inhalte der Fortbildung mit Leidenschaft“, $\alpha = .94$). Über die Items werden der Enthusiasmus für das Fortbildungsthema und der Enthusiasmus für die Tätigkeit des Fortbildens abgebildet. Theoretisch lassen sich diese beiden Facetten des Enthusiasmus unterscheiden, empirisch bilden die Items jedoch einen Faktor.

Bei den Moderatorinnen und Moderatoren wurde für den Enthusiasmus für die Tätigkeit des Fortbildens und den Enthusiasmus für das Fortbildungsthema jeweils eine Skala gebildet. Der Enthusiasmus für das Fortbilden wurde ausschließlich in der ersten Prozesserhebung (t_3) mit fünf Items erhoben (Beispielitem: „Ich bin glücklich, wenn ich Fortbildungen geben kann“, $\alpha = .72$). Der Enthusiasmus für das Fortbildungsthema wurde in der Prozesserhebung zum Diagnostikmodul (t_3) und in der Prozesserhebung zu den Modulen zum Lernen und Lehren (t_6) mit fünf

Items (Beispielitem: „Die Inhalte der Fortbildung vertrete ich leidenschaftlich gerne“, $\alpha_{t3} = .82$; $\alpha_{t6} = .82$) erfasst.

Der Mittelwert für den von den Lehrpersonen eingeschätzte Enthusiasmus der Moderatorinnen und Moderatoren liegt über dem theoretischen Skalenmittel von 3.50 ($M = 3.90$, $SD = 1.13$), d. h., die Moderatorinnen und Moderatoren werden eher als begeistert erlebt. Der von den Moderatorinnen und Moderatoren selbstberichtete Enthusiasmus für das Fortbilden von Lehrpersonen liegt bei $M = 4.65$ ($SD = 0.64$) und damit deutlich über dem theoretischen Mittel. Die Begeisterung für das Thema Diagnostik ($M = 4.14$, $SD = 0.77$) ist geringer als die Begeisterung für die Durchführung von Lehrerfortbildungen und auch geringer als der Enthusiasmus für die Themen der Module zum Lernen und Lehren ($M = 4.73$, $SD = 0.78$). Im Mittel schätzen sich die Moderatorinnen und Moderatoren demnach enthusiastischer ein, als sie auf die Fortbildungsteilnehmerinnen und -teilnehmer wirken. Zwischen der Wahrnehmung des Enthusiasmus durch die Kollegien und dem von den Moderatorenteams selbstberichteten Enthusiasmus besteht zu beiden Erhebungszeitpunkten, d. h. in der Mitte von Modul 2 (t_3) und am Anfang von Modul 4 (t_6), kein Zusammenhang ($r_{\text{Fortbilden}} = -.03$, n. s.; $r_{\text{Thema-Modul}_2} = -.37$, n. s.; $r_{\text{Thema-Modul}_{3-4}} = -.50$, n. s.).

Einzelne Moderatorenteams haben *Vielfalt fördern* in mehreren Schulen durchgeführt. Aus Abbildung 68 lässt sich erkennen, dass die Moderatorenteams, die an mehreren Schulen waren, von den teilnehmenden Lehrpersonen überwiegend vergleichbar wahrgenommen wurden. Eine Ausnahme stellen lediglich das Team C und das Team D dar, deren Enthusiasmus in jeweils einer der beiden Schulen geringer beurteilt wird.

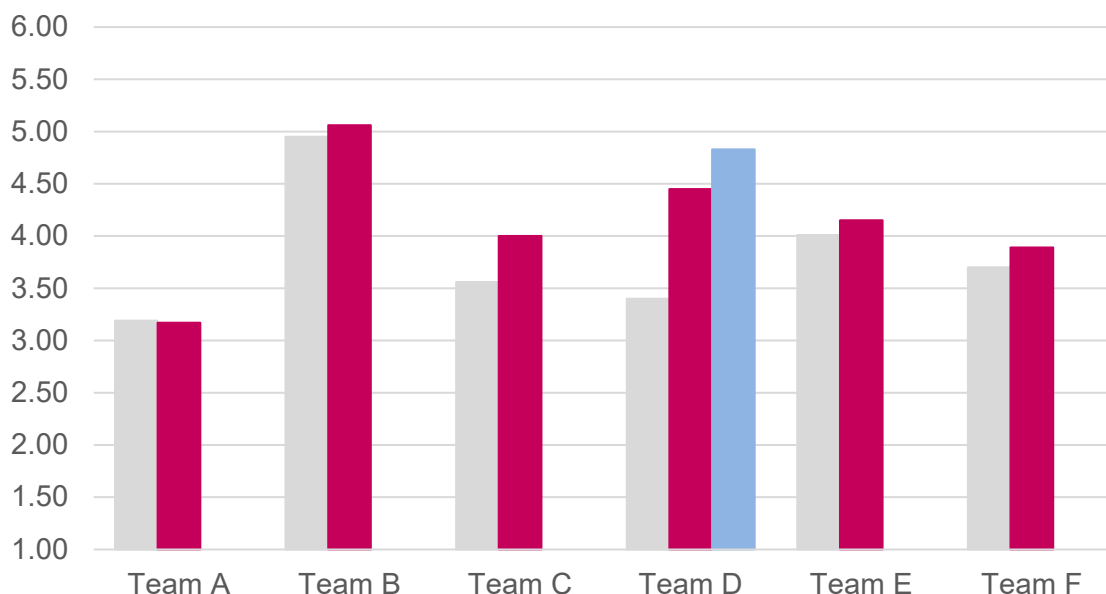


Abbildung 68. Durchschnittlich wahrgenommene Moderatorenkompetenz eines Moderatorenteams an unterschiedlichen Schulen

10.1 Erklärung der Fortbildungsnutzung und -bewertung durch den Enthusiasmus des Moderatorenteams

Beeinflusst der Enthusiasmus des Moderatorenteams, wie die teilnehmenden Lehrpersonen *Vielfalt fördern* wahrnehmen und nutzen? Diese Frage wurde mittels verschiedener Regressionsmodelle unter Berücksichtigung der hierarchischen Datenstruktur untersucht. In einem ersten Schritt wurde überprüft, ob der von den Moderatorenteams berichtete Enthusiasmus, der als aggregiertes Merkmal²⁸ auf Schulebene behandelt wurde, die subjektive Wahrnehmung und die individuelle Nutzung des Lernangebots von *Vielfalt fördern* beeinflusst. Da der von den Moderatorinnen und Moderatoren angegebene Enthusiasmus für das Fortbilden und der Enthusiasmus für das Thema bei der Prozesserhebung des zweiten Moduls (t_3) stark miteinander zusammenhängen ($r = .78, p < .001$), wurden diese beiden Merkmale in separaten Modellen überprüft. Im zweiten Schritt wurde zusätzlich der von den Lehrpersonen erlebte Enthusiasmus des Moderatorenteams in die Modelle aufgenommen, um zu untersuchen, inwiefern dieses individuelle Erleben zusätzlich beeinflusst, wie die Lehrpersonen *Vielfalt fördern* wahrnehmen und nutzen.

10.1.1 Vorhersage der systematischen Verarbeitung von Fortbildungsinhalten während Modul 2

Es zeigt sich in den linearen Regressionsanalysen auf Individualebene, in denen die Gruppierung der Lehrpersonen in Schulen berücksichtigt wurde (s.

Abbildung 69 und Abbildung 70), dass weder der selbstberichtete Enthusiasmus der Moderatorinnen und Moderatoren für das *Fortbilden* ($\beta = .08, n. s.$) noch der Enthusiasmus für das *Thema Diagnostik* ($\beta = .05, n. s.$) sich darauf auswirken, wie intensiv und systematisch die Lehrpersonen die Fortbildungsinhalte verarbeiten.

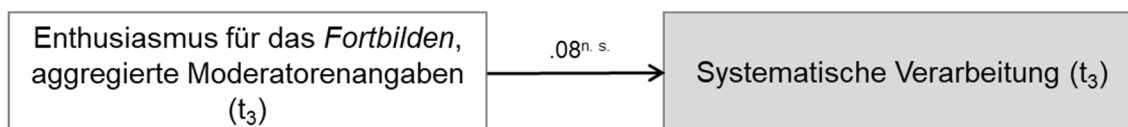


Abbildung 69. Vorhersage der systematischen Verarbeitung der Fortbildungsinhalte durch den Enthusiasmus des Moderatorenteams für das Fortbilden aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant

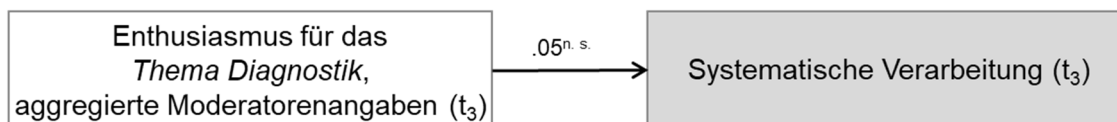


Abbildung 70. Vorhersage der systematischen Verarbeitung der Fortbildungsinhalte durch den Enthusiasmus des Moderatorenteams für das Thema Diagnostik aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant

²⁸ Hierzu wurden die Werte der Moderatorinnen und Moderatoren auf Schulebene gemittelt.

Da der von den Moderatorinnen und Moderatoren selbstberichtete Enthusiasmus die systematische Verarbeitung der Lehrpersonen nicht signifikant beeinflusst, wird in das nachfolgende Modell ausschließlich der von den Lehrkräften berichtete Enthusiasmus des Moderatorenteams einbezogen. Die individuelle Wahrnehmung des Enthusiasmus der Moderatorinnen und Moderatoren durch die Lehrpersonen, die zum gleichen Zeitpunkt wie die systematische Verarbeitung der Fortbildungsinhalte, d. h. nach etwa der Hälfte des zweiten Moduls (t_3), erfasst wurde, übt – im Unterschied zum von den Moderatorinnen und Moderatoren eingeschätzten Enthusiasmus – einen signifikant positiven Einfluss auf die systematische Verarbeitung aus (s. Abbildung 71, $\beta = .37$, $p < .001$). Mit anderen Worten: Wenn eine Lehrperson das Moderatorenteam als sehr begeistert erlebt, dann geht dies mit einer elaborierteren Auseinandersetzung mit den Fortbildungsinhalten einher. Durch den wahrgenommenen Enthusiasmus des Moderatorenteams können 14 % der Unterschiede in der systematischen Verarbeitung der Fortbildungsinhalte von Modul 2 erklärt werden.

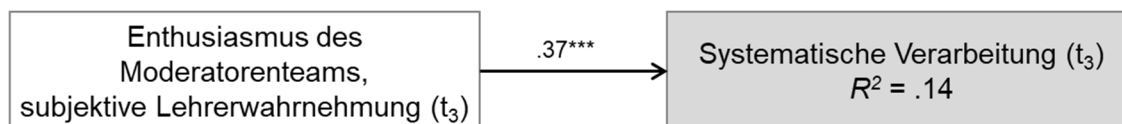


Abbildung 71. Vorhersage der systematischen Verarbeitung der Fortbildungsinhalte durch den Enthusiasmus des Moderatorenteams aus Lehrerperspektive. Anmerkung: *** = $p < .001$

10.1.2 Vorhersage der kooperativen Erprobungen zwischen Fortbildungssitzungen von Modul 2

Betrachtet man als Kriterium die kooperative Umsetzung und Erprobung der Fortbildungsinhalte zwischen den Fortbildungstreffen, also den gemeinsamen Austausch über die Fortbildungsinhalte und/oder eine gemeinsame Anpassung der Fortbildungsinhalte an die eigenen schulischen Rahmenbedingungen, dann zeigt sich ein ähnliches Bild wie für die systematische Verarbeitung (s. Abschnitt 10.1.1). Der Enthusiasmus für das *Fortbilden* (s. Abbildung 72, $\beta = .10$, n. s.) und der Enthusiasmus für das *Thema Diagnostik* (s. Abbildung 72, $\beta = .06$, n. s.), den die Moderatorinnen und Moderatoren selbst berichteten, haben keinen Einfluss auf die kooperative Erprobung der Fortbildungsinhalte zwischen den ersten Fortbildungssitzungen des zweiten Moduls.

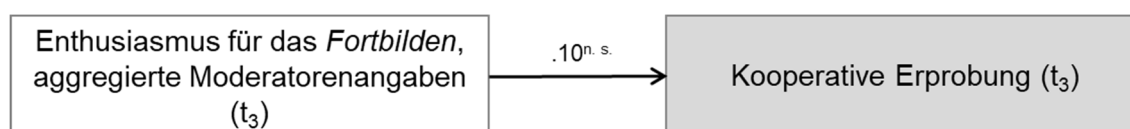


Abbildung 72. Vorhersage der kooperativen Umsetzung von Fortbildungsinhalten zwischen den Bausteinen durch den Enthusiasmus des Moderatorenteams für das Fortbilden aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant

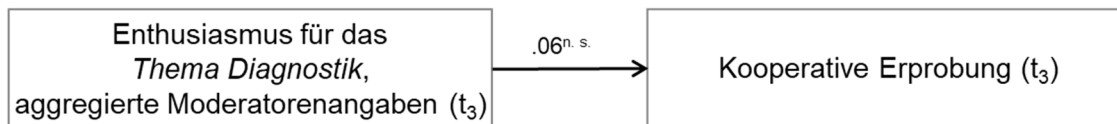


Abbildung 73. Vorhersage der kooperativen Umsetzung von Fortbildungsinhalten zwischen den Bausteinen durch den Enthusiasmus des Moderatorenteams für das Thema Diagnostik aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant

Die Wahrnehmung des Moderatorenenthusiasmus durch die Lehrpersonen wirkt sich erneut signifikant positiv auf die gemeinsame Anwendung und Erprobung der Fortbildungsinhalte mit Kolleginnen und Kollegen aus (s. Abbildung 74, $\beta = .37$, $p < .001$, $R^2 = .14$). Lehrpersonen, die ihr Moderatorenteam enthusiastischer erlebt haben, geben demnach an, dass sie in den Erprobungsphasen zwischen den Modulbausteinen intensiver mit Kolleginnen und Kollegen planen, wie die Fortbildungsinhalte umgesetzt werden können.

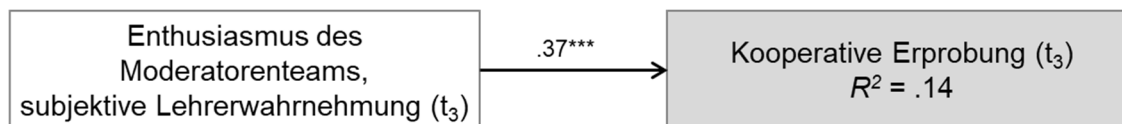


Abbildung 74. Vorhersage der kooperativen Umsetzung von Fortbildungsinhalten zwischen den Bausteinen durch den Enthusiasmus des Moderatorenteams aus Lehrerperspektive. Anmerkung: *** = $p < .001$

10.1.3 Vorhersage des selbstberichteten Lernertrags der Lehrpersonen am Ende von Modul 2

Mit dem selbstberichteten Lernertrag zu t_4 wird das subjektive Empfinden einer Lehrperson erfasst, aufgrund der Teilnahme an dem Modul 2 hilfreiche Impulse erhalten zu haben, Neues gelernt zu haben und sich kompetenter zu fühlen als vor dem Modul.

Unterschiede im selbstberichteten Lernertrag am Ende von Modul 2 lassen sich zu 5 % durch das Regressionsmodell (s. Abbildung 75) erklären, in dem ausschließlich der von den Moderatorinnen und Moderatoren selbstberichtete Enthusiasmus für das Fortbilden während Modul 2 als Prädiktor berücksichtigt ist ($\beta = .23$, $p < .001$). Es handelt sich um einen schwachen Effekt, der bedeutet, dass die Begeisterung, die das Moderatorenteam während der Arbeit mit Lehrpersonen im Rahmen der Fortbildung selbst erlebt, dazu beiträgt, dass die teilnehmenden Lehrpersonen in einem stärkeren Maße den Eindruck haben, während der Fortbildung hilfreiche neue Impulse erhalten zu haben und sich kompetenter zu fühlen.

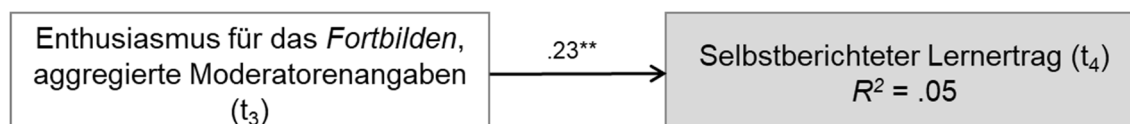


Abbildung 75. Vorhersage des selbstberichteten Lernertrags durch den Enthusiasmus des Moderatorenteams für das Fortbilden aus Moderatorenperspektive. Anmerkung: ** = $p < .01$

Aus dem Regressionsmodell, das den Einfluss des von den Moderatorinnen und Moderatoren berichteten Enthusiasmus für das Thema Diagnostik untersucht (s. Abbildung 76), geht kein statistisch bedeutsamer Einfluss auf den selbstberichteten Lernertrag der Lehrpersonen am Ende von Modul 2 hervor ($\beta = .11$, n. s.).

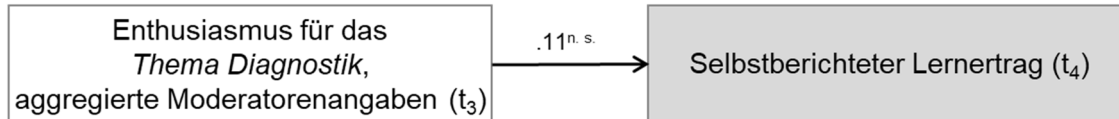


Abbildung 76. Vorhersage des selbstberichteten Lernertrags durch den Enthusiasmus des Moderatorenteams für das Thema Diagnostik aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant

Da sich der aus der Perspektive des Moderatorenteams erfasste Enthusiasmus für das Fortbilden als ein signifikant positiver Prädiktor für den selbstberichteten Lernertrag herausstellte (s. Abbildung 75), wird dieser auch in dem Modell, welches im zweiten Schritt untersucht wird, aufgenommen. Das Modell beinhaltet als zweiten Prädiktor zur Erklärung von Varianzanteilen im selbstberichteten Lernertrag den von den Lehrpersonen subjektiv erlebten Enthusiasmus des Moderatorenteams (s. Abbildung 77). Insgesamt werden durch das Modell 25 % der Unterschiede im selbstberichteten Lernertrag zwischen Lehrpersonen erklärt. Sowohl der von den Lehrpersonen wahrgenommene Enthusiasmus des Moderatorenteams ($\beta = .45$, $p < .001$) als auch der von den Moderatorinnen und Moderatoren selbstberichtete Enthusiasmus für die Arbeit mit Lehrpersonen ($\beta = .20$, $p < .01$) erweisen sich als signifikant positive Prädiktoren. Je enthusiastischer die Moderatorenteams nach eigenen Angaben während des zweiten Moduls sind und je enthusiastischer das Moderatorenteam von den Fortbildungsteilnehmerinnen und Fortbildungsteilnehmern erlebt wird, desto eher haben die Lehrpersonen den Eindruck, durch das zweite Modul hilfreiche Anregungen für die eigene schulische Praxis erhalten und Neues gelernt zu haben.

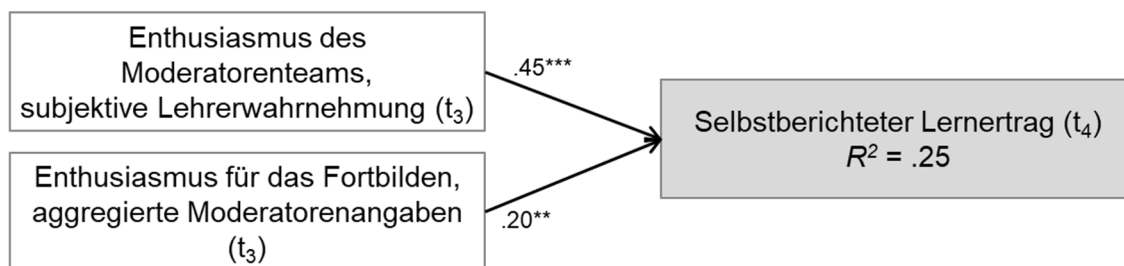


Abbildung 77. Vorhersage des selbstberichteten Lernertrags durch den Enthusiasmus des Moderatorenteams aus Lehrer- und Moderatorenperspektive. Anmerkung: ** = $p < .01$, *** = $p < .001$

Insgesamt ergaben die Analysen, deren Ergebnisse in den vorangehenden Abschnitten 10.1.1 bis 10.1.3 dargestellt wurden, dass der selbstberichtete Enthusiasmus des Moderatorenteams die Nutzung und Verarbeitung der Fortbildungsinhalte durch die teilnehmenden Lehrpersonen nicht bedeutsam beeinflusst. Das subjektive Erleben der Lehrpersonen erwies sich hingegen als bedeutsam dafür, wie elaboriert die Lehrpersonen die Fortbildungsinhalte verarbeiten und wie

umfassend die Lehrpersonen auch zwischen den einzelnen Modulbausteinen an der Erprobung der Fortbildungsinhalte gemeinsam mit Kolleginnen und Kollegen arbeiten. Die von den Lehrpersonen wahrgenommene Freude und Begeisterung des Moderatorenteams während des zweiten Moduls fungiert auch als ein positiver Prädiktor dafür, wie hoch die Lehrpersonen ihren Lernertrag am Ende des Moduls bewerten. Die subjektive Bewertung der Lehrpersonen, durch das zweite Modul nützliche Impulse erhalten zu haben und kompetenter geworden zu sein, wird zudem von dem durch die Moderatorinnen und Moderatoren selbstberichteten Enthusiasmus für das Fortbilden positiv vorhergesagt.

10.2 Beziehung zwischen der Fortbildungsmotivation der Lehrpersonen und dem Enthusiasmus des Moderatorenteams

Die Teilnahmemotivation der Fortbildungsteilnehmerinnen und -teilnehmer wirkt sich darauf aus, wie die Lehrpersonen die Fortbildung erleben und bewerten (s. Abschnitte 5 und 6). Vor diesem Hintergrund erscheint die Annahme plausibel, dass auch die Wahrnehmung des Enthusiasmus der Moderatorinnen und Moderatoren durch die Teilnehmerinnen und Teilnehmer nicht unabhängig davon ist, mit welcher Motivation die Lehrpersonen die Teilnahme an *Vielfalt fördern* beginnen.

Im Abschnitt 10.1 wurde deutlich, dass der von den Lehrpersonen erlebte Enthusiasmus ihres Moderatorenteams ebenfalls einen bedeutsamen Einfluss auf die Fortbildungsnutzung und -bewertung ausübt. Demnach geht ein stärker wahrgenommener Enthusiasmus bei den Moderatorinnen und Moderatoren aus Sicht der Teilnehmerinnen und Teilnehmer mit tieferehenden Verarbeitungsprozessen, einer intensiveren kooperativen Erprobung und einem höheren selbstberichteten Lernertrag am Ende des zweiten Moduls von *Vielfalt fördern* einher.

Ausgehend von den bisherigen Befunden zur Teilnahmemotivation und zum von den Lehrpersonen wahrgenommenen Enthusiasmus des Moderatorenteams wird angenommen, dass sich diese Merkmale auch darauf auswirken, wie motiviert die Lehrpersonen am Anfang von Modul 4 (t_6) sind, aktiv in der Fortbildung mitzuarbeiten.

Die in der zweiten Prozesserhebung (t_6) erfasste Motivation wird als *Nutzungsmotivation* bezeichnet. Wie bei der Teilnahmemotivation (s. Abschnitt 6.1) wurden bei der Nutzungsmotivation unterschiedliche Facetten der Motivation – *externale Erwartungsanpassung*, *Schulentwicklungsorientierung* und *soziale Interaktion* – unterschieden.

Die Skala zur Erfassung der externalen Erwartungsanpassung besteht aus zwei Items (Beispielitem: „*Ich arbeite in den Fortbildungstreffen von Vielfalt fördern mit, weil ich dazu verpflichtet bin*“, $\alpha = .87$). Ein hoher Mittelwert bedeutet, dass die Lehrpersonen sich stark aufgrund eines Pflicht-

gefühls in die Fortbildung einbringen und mitarbeiten. Für die Stichprobe der zweiten Fortbildungskohorte beträgt der Mittelwert $M = 4.50$ ($SD = 1.51$), sodass von einer hohen extrinsisch motivierten Mitarbeit der Teilnehmerinnen und Teilnehmer zu sprechen ist.

Die Nutzungsmotivationsfacetten Schulentwicklungsorientierung und soziale Interaktion setzen sich jeweils aus drei Items zusammen. Wenn die Lehrpersonen vorwiegend zur aktiven Teilnahme während der Fortbildungssitzungen motiviert sind, weil sie denken, dass das Projekt für die Schule wichtig ist und deren Entwicklung voranbringt, käme dies in einem hohen Skalenmittelwert bei der Schulentwicklungsorientierung zum Ausdruck (Beispielitem: „*Ich arbeite in den Fortbildungstreffen von Vielfalt fördern mit, weil ich so die Entwicklung der Schule voranbringen kann*“, $\alpha = .92$). Im Mittel geben die Lehrpersonen einen Wert von $M = 3.68$ ($SD = 1.33$) an, der knapp über dem theoretischen Skalenmittel von 3.50 liegt. Damit arbeiten die Lehrpersonen während der Fortbildung durchaus auch aus dem Beweggrund mit, auf diese Weise einen positiven Beitrag für die Weiterentwicklung ihrer Schule zu leisten.

Ein hoher Mittelwert auf der Skala soziale Interaktion ist so zu interpretieren, dass die Möglichkeit zum Austausch mit Kolleginnen und Kollegen und die Freude am kollegialen Lernen wesentliche Beweggründe für die Nutzung des Fortbildungsangebots darstellen (Beispielitem: „*Ich arbeite in den Fortbildungstreffen von Vielfalt fördern mit, weil ich so berufliche Kontakte zu Kollegen pflegen kann*“, $\alpha = .75$). Der Mittelwert für diese Facette der Nutzungsmotivation liegt bei $M = 4.18$ ($SD = 1.11$) und ist damit höher als der Wert für die Schulentwicklungsorientierung, aber niedriger als der Mittelwert auf der Skala externale Erwartungsanpassung.

Um zu überprüfen, ob sich die Teilnahmemotivation der Lehrpersonen (t_1) darauf auswirkt, wie diese den Enthusiasmus ihres Moderatorenteams während Modul 2 (t_3) wahrnehmen und ob diese Wahrnehmung des Enthusiasmus der Moderatorinnen und Moderatoren wiederum einen Einfluss auf die Nutzungsmotivation der Lehrpersonen (t_6) besitzt, werden Mediatormodelle berechnet (s. Abschnitt 4.1.3). Den Mediator, d. h. die vermittelnde Variable, bildet in den Modellen der von den Lehrpersonen wahrgenommene Enthusiasmus des Moderatorenteams. Hierbei wird angenommen, dass es sich um eine partielle Mediation handelt, die *Teilnahmemotivation* also neben dem vermittelten Effekt über den Enthusiasmus auch einen direkten Effekt auf die *Nutzungsmotivation* besitzt.

10.2.1 Motivationsfacette: Externale Erwartungsanpassung

Im Folgenden wird untersucht, inwiefern die externale Erwartungsanpassung am Anfang der Fortbildung (t_1), vermittelt über den von den Lehrpersonen wahrgenommenen Enthusiasmus des Moderatorenteams während Modul 2 (t_3), die externale Erwartungsanpassung als eine Facette der Nutzungsmotivation am Anfang von Modul 4 (t_6) vorhersagt.

Der Abbildung 78 können die direkten Effekte entnommen werden: Wenn eine Lehrperson unter anderem deshalb stark zur Teilnahme an *Vielfalt fördern* motiviert ist, weil sie der mehrheitlichen Entscheidung des Kollegiums folgt und es als eine notwendige Pflicht empfindet, führt dies dazu, dass das Moderatorenteam in Modul 2 weniger begeistert wahrgenommen wird ($\beta = -.21$, $p < .001$). Der Effekt ist aber nur schwach ($R^2 = .04$). Je enthusiastischer das Moderatorenteam während Modul 2 auf eine Lehrperson wirkt, desto geringer ist die aktive Mitarbeit der Lehrperson im weiteren Verlauf der Fortbildung durch ein Pflichtgefühl motiviert ($\beta = -.21$, $p < .01$). Die von den Lehrpersonen berichtete externale Erwartungsanpassung am Anfang von Modul 4 (t_6) wird zudem positiv davon beeinflusst, ob die Lehrperson bereits am Anfang von *Vielfalt fördern* angegeben hat, aufgrund des schulischen Mehrheitsentscheids an der Fortbildung teilzunehmen ($\beta = .46$, $p < .001$). Hierin deutet sich an, dass die extrinsische Fortbildungsmotivation bei den Fortbildungsteilnehmerinnen und -teilnehmern eher konstant ist.

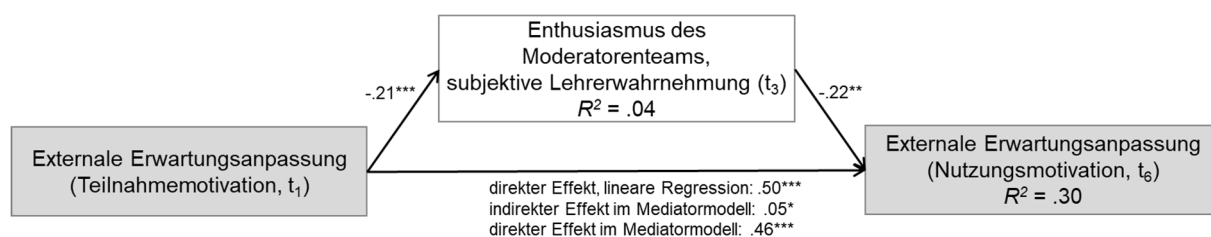


Abbildung 78. Mediatormodell mit der Teilnahmemotivation externe Erwartungsanpassung als Prädiktor, dem Enthusiasmus des Moderatorenteam aus Lehrerperspektive als Mediator und der Nutzungsmotivation externe Erwartungsanpassung als Kriterium. Anmerkung: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Die Mediation, d. h. der indirekte Effekt von der Teilnahmemotivation externe Erwartungsanpassung über den von den Lehrpersonen wahrgenommenen Enthusiasmus des Moderatorenteam, erweist sich als signifikant, aber schwach (s. Abbildung 78, $\beta = .05$, $p < .05$). Da der direkte Einfluss von der Teilnahmemotivation zu t_1 auf die Nutzungsmotivation zu t_6 weiterhin stark und signifikant bleibt ($\beta = .46$, $p < .001$), handelt es sich um eine partielle Mediation.

10.2.2 Motivationsfacette: Schulentwicklungsorientierung

In diesem Abschnitt wird die Motivationsfacette Schulentwicklungsorientierung untersucht. Lehrpersonen, bei denen diese Motivationsfacette hoch ausgeprägt ist, motiviert die Aussicht, dass die Schule als Ganzes von *Vielfalt fördern* profitieren und sich weiterentwickeln kann. Diese Facette der Motivation wurde sowohl als Motivation für die Teilnahme an *Vielfalt fördern* zu t_1 als auch als Nutzungsmotivation am Anfang von Modul 4 (t_6) erfasst.

Es wird angenommen, dass die Teilnahmemotivation die zu einem späteren Zeitpunkt erfasste Nutzungsmotivation positiv vorhersagt. Eine hohe Schulentwicklungsorientierung am Anfang von *Vielfalt fördern* sollte nach dieser Annahme dazu führen, dass auch die aktive Mitarbeit während der Fortbildung stark dadurch motiviert ist, einen als wichtig erachteten Beitrag zur Weiterentwicklung der eigenen Schule zu leisten. Hierbei wird angenommen, dass der Einfluss nicht nur

direkter Natur ist, sondern in Teilen darüber vermittelt wird, wie enthusiastisch das eigene Moderatorenteam während der Fortbildung erlebt wird. Es wird also eine partielle Mediation angenommen.

Das Ergebnis der Mediatoranalyse zeigt, dass 10 % der Unterschiede in dem Enthusiasmus des Moderatorenteam, der aus Perspektive der Lehrpersonen erfasst wurde, durch die Teilnahmemotivation Schulentwicklungsorientierung erklärt werden kann ($\beta = .32, p < .001$). Dies ist so zu interpretieren, dass Lehrpersonen, die insbesondere an *Vielfalt fördern* teilnehmen, weil sie den darin liegenden Vorteil der eigenen Schule und deren Weiterentwicklung vor Augen haben, eher den Eindruck haben, dass ihrem Moderatorenteam die Arbeit mit Lehrpersonen Spaß macht.

Durch das Modell können außerdem 28 % der Unterschiede zwischen Lehrpersonen in der Nutzungsmotivationsfacette Schulentwicklungsorientierung erklärt werden (s. Abbildung 79). Die am Anfang von *Vielfalt fördern* erfasste Teilnahmemotivation (t_1) ($\beta = .28, p < .001$) und der wahrgenommene Enthusiasmus des eigenen Moderatorenteam während Modul 2 (t_3) ($\beta = .37, p < .001$) wirken sich im Mediationsmodell positiv auf die spätere Nutzungsmotivation (t_6) aus. Wie sind diese positiven Effekte inhaltlich zu interpretieren? Die Effekte bedeuten, dass sich die Nutzungsmotivation einer Lehrperson nicht nur aus der Schulentwicklungsorientierung zu Beginn der Fortbildung, sondern auch durch den von der Lehrkraft erlebten Enthusiasmus der Moderatorinnen und Moderatoren speist und dass der Einfluss der Teilnahmemotivation zum ersten Messzeitpunkt partiell erklärbar ist durch den als höher erlebten Enthusiasmus. Während das Modell ohne Mediation nur 16 % der Nutzungsmotivation aufklärt, ist die Varianzaufklärung im Mediationsmodell mit 28 % deutlich höher.

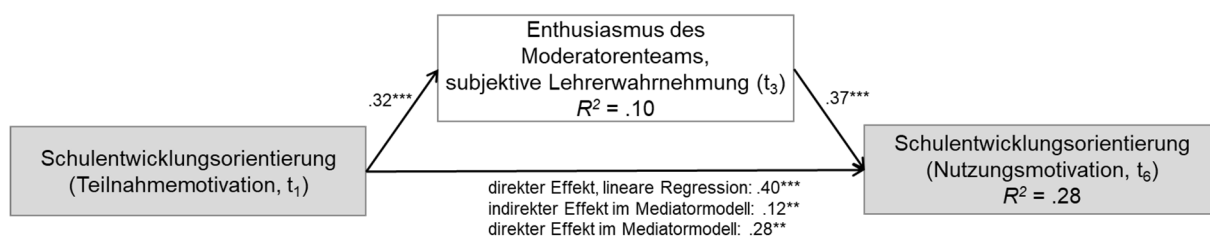


Abbildung 79. Mediatormodell mit der Teilnahmemotivation Schulentwicklungsorientierung als Prädiktor, dem Enthusiasmus des Moderatorenteam aus Lehrerperspektive als Mediator und der Nutzungsmotivation Schulentwicklungsorientierung als Kriterium. Anmerkung: $** = p < .01$, $*** = p < .001$

Der indirekte Effekt, respektive der Einfluss der Teilnahmemotivation, der vermittelt über den von den Lehrpersonen wahrgenommenen Enthusiasmus des Moderatorenteam auf die Nutzungsmotivation wirkt, ist signifikant positiv ($\beta = .12, p < .01$). Es handelt sich um eine partielle Mediation, da die Teilnahmemotivation zu Beginn von *Vielfalt fördern* zudem noch einen direkten Einfluss auf die über ein Jahr später erfasste Motivation hat, die Lernangebote aktiv zu nutzen ($\beta = .28, p < .01$).

10.2.3 Motivationsfacette: Soziale Interaktion

Im Folgenden wird untersucht, inwiefern die soziale Interaktion, die als eine Facette der Teilnahmemotivation am Anfang von *Vielfalt fördern* (t_1) erfasst wurde, die Motivation der Lehrpersonen vorhersagt, während der Fortbildungsbausteine (t_6) aktiv mitzuarbeiten, weil sie sich dadurch einen intensiveren Austausch mit Kolleginnen und Kollegen erhoffen und an deren beruflichen Erfahrungen teilhaben möchten. Wie bereits in den beiden vorherigen Abschnitten wird angenommen, dass die Teilnahmemotivation nicht nur einen direkten Einfluss auf die über ein Jahr später erfasste Nutzungsmotivation hat, sondern der Einfluss teilweise darüber vermittelt wird, wie enthusiastisch die Lehrpersonen ihr Moderatorenteam während des zweiten Moduls erlebt haben.

Wie Abbildung 80 zeigt, wirkt sich die Teilnahmemotivation positiv auf die Nutzungsmotivationsfacette soziale Interaktion aus ($\beta = .40$, $p < .001$). Mit anderen Worten: Eine höher ausgeprägte Teilnahmemotivation führt über ein Jahr später zu einer höheren Nutzungsmotivation. Lehrpersonen, die an der Fortbildung teilnehmen, weil ihnen der soziale Austausch wichtig ist, sind auch während der Fortbildungsbausteine zur Mitarbeit motiviert, weil sie gerne mit Kolleginnen und Kollegen zusammenarbeiten und um berufliche Erfahrungen mit Kolleginnen und Kollegen zu teilen. Das Mediationsmodell verdeutlicht, dass darüber hinaus die Teilnahmemotivation einen positiven Einfluss auf den von den Lehrpersonen wahrgenommenen Enthusiasmus während Modul 2 ausübt ($\beta = .19$, $p < .01$), wobei es sich um einen schwachen Effekt handelt ($R^2 = .04$).

Von dem aus der Lehrerperspektive wahrgenommenen Enthusiasmus des Moderatorenteams geht hingegen nur ein auf dem 10 %-Niveau signifikanter Einfluss darauf aus, wie stark die Lehrpersonen motiviert sind, bei der Nutzung der Fortbildungsinhalte in einen intensiveren fachlichen Austausch mit Kolleginnen und Kollegen zu treten.

Ein weiteres Ergebnis der Analyse ist, dass der Einfluss der Teilnahmemotivation nicht über den wahrgenommenen Enthusiasmus des Moderatorenteams aus Sicht der Lehrpersonen vermittelt wird. Der indirekte Effekt ist nicht signifikant ($\beta = .03$, n. s.).

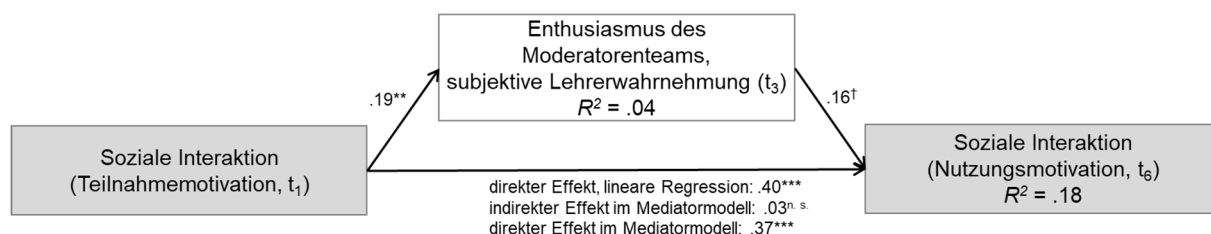


Abbildung 80. Mediatormodell mit der Teilnahmemotivation soziale Interaktion als Prädiktor, dem Enthusiasmus des Moderatorenteams aus Lehrerperspektive als Mediator und der Nutzungsmotivation soziale Interaktion als Kriterium. Anmerkung: n. s. = nicht signifikant, $\dagger = p < .10$, $** = p < .01$, $*** = p < .001$

11 Einbezug von Erfahrungen der teilnehmenden Lehrpersonen

In Abschnitt 10.1 wurde untersucht, ob sich Unterschiede in der Nutzung und Bewertung der Lerngelegenheiten des zweiten Fortbildungsmoduls von *Vielfalt fördern* durch den von den Moderatorinnen und Moderatoren selbstberichteten Enthusiasmus und durch den von den Lehrpersonen wahrgenommenen Enthusiasmus ihres Moderatorenteams erklären lassen. Hierbei erwies sich das subjektive Erleben der Lehrperson gegenüber den von den Moderatorinnen und Moderatoren berichteten Angaben als bedeutungsvoller.

Im daran anschließenden Abschnitt 10.2 wurde daher ausschließlich der von den Lehrpersonen erlebte Enthusiasmus der Moderatorinnen und Moderatoren näher untersucht. Es wurde analysiert, ob der Einfluss der zu t_1 erfassten Teilnahmemotivation der Lehrpersonen auf die Nutzungsmotivation am Anfang von Modul 4 (t_6) darüber vermittelt wird, wie enthusiastisch die Lehrpersonen ihr Moderatorenteam während Modul 2 (t_3) erlebt haben. Die Teilnahmemotivation erwies sich in allen drei Modellen als ein signifikant positiver Prädiktor für die Motivation, die Lernangebote von *Vielfalt fördern* aktiv zu nutzen und sich in die Fortbildung einzubringen. Darüber hinaus wurde der Einfluss der externalen Erwartungsanpassung und der Schulentwicklungsorientierung aber partiell darüber vermittelt, wie die Lehrpersonen die Freude und Begeisterung des Moderatorenteams ihrer Schule wahrnehmen.

Durch die nachfolgenden Analysen sollen erneut die drei Facetten der Nutzungsmotivation der Fortbildungsteilnehmerinnen und -teilnehmer (*externale Erwartungsanpassung*, *Schulentwicklungsorientierung* und *soziale Interaktion*) in den Blick genommen werden. Als Prädiktor wird ein Merkmal der Fortbildungsgestaltung, der Einbezug von Teilnehmererfahrungen, untersucht. Dieses Merkmal wurde, wie auch der Enthusiasmus, aus Sicht des Moderatorenteams und aus Sicht der teilnehmenden Lehrpersonen erhoben. Bei den Moderatorinnen und Moderatoren wurde der Einbezug von Teilnehmererfahrungen während der zweiten Prozesserhebung (t_6) mit vier Items erfasst ($\alpha = .91$). Ein Item der Skala lautet beispielsweise „*Ich beziehe die unterrichtlichen Erfahrungen der Teilnehmer ein*“. Bei den Lehrpersonen besteht die Skala zu t_6 ebenfalls aus vier Items (Beispielitem: „*Meine Erfahrungen aus dem schulischen Alltag wurden aktiv in die Fortbildung einbezogen*“, $\alpha = .84$). Bei beiden Personengruppen ist ein höherer Skalenmittelwert als ein umfangreicherer Einbezug von Erfahrungen der Teilnehmerinnen und Teilnehmer zu interpretieren. Ein deskriptiver Vergleich zeigt, dass die Moderatorenteams angeben, die Beispiele und Erfahrungen der teilnehmenden Lehrpersonen deutlich stärker zu berücksichtigen ($M = 5.40$, $SD = 0.53$), als die Lehrpersonen dies selbst erleben ($M = 4.11$, $SD = 1.08$). Interessant ist, dass zwischen der Wahrnehmung der Moderatorinnen und Moderatoren und der Kollegien kein statistisch bedeutsamer Zusammenhang besteht ($r = .14$, n. s.).

Die Analysen werden erneut in zwei Schritten vorgenommen. Im ersten Schritt wird der von den Moderatorinnen und Moderatoren berichtete Einbezug von Teilnehmererfahrungen als Prädiktor

zur Erklärung von Varianz in der Motivation zur aktiven Nutzung der Lernangebote während des Moduls zum Lernen und Lehren untersucht. Im zweiten Schritt wird zusätzlich der von den Lehrpersonen wahrgenommene Einbezug der Erfahrungen von Teilnehmerinnen und Teilnehmern in die Fortbildungssitzungen in das Modell aufgenommen.

Zusätzlich zu dem Einbezug der Teilnehmererfahrungen wird in allen Modellen die anfängliche Teilnahmemotivation kontrolliert. Hierbei wird jeweils nur die Teilnahmemotivationsfacette in das Modell aufgenommen, die auch als Nutzungsmotivation untersucht wird. Mit anderen Worten: Wenn Unterschiede in dem Bedürfnis nach sozialer Interaktion als Beweggrund für die aktive Mitarbeit während der Fortbildungssitzungen als Kriterium untersucht werden, dann wird zusätzlich zu dem Einbezug der Teilnehmererfahrungen auch die Teilnahmemotivationsfacette soziale Interaktion in das Modell aufgenommen.

11.1 Motivationsfacette: Externale Erwartungsanpassung

Im Folgenden wird untersucht, ob sich die Nutzungsmotivation externe Erwartungsanpassung oder mit anderen Worten die Motivation, in der Fortbildung mitzuarbeiten, weil man dies als Pflicht empfindet, durch den Einbezug von Erfahrungen der teilnehmenden Lehrpersonen erklären lässt.

Unter Kontrolle der externalen Erwartungsanpassung, die am Anfang von *Vielfalt fördern* erfasst wurde, zeigt sich (s. Abbildung 81), dass die Nutzungsmotivation externe Erwartungsanpassung zu Beginn von Modul 4 nicht davon beeinflusst wird, wie stark die Erfahrungen der teilnehmenden Lehrpersonen nach der Einschätzung des Moderatorenteams in die Fortbildungssitzungen einbezogen wurden ($\beta = -.04$, n. s.).

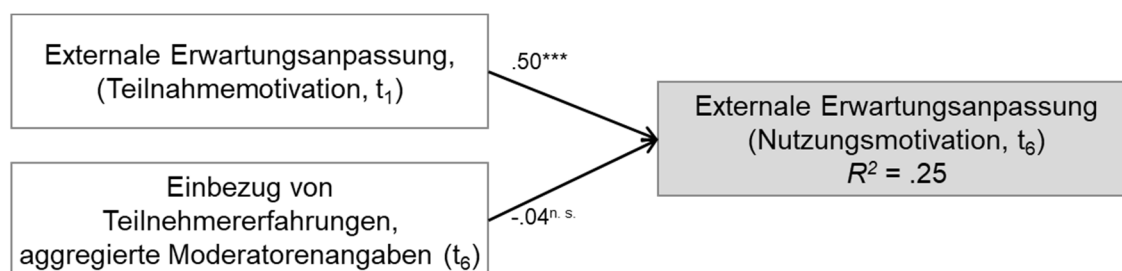


Abbildung 81. Vorhersage der Fortbildungsnutzungsmotivationsfacette externe Erwartungsanpassung durch den Einbezug von Teilnehmererfahrungen aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant, *** = $p < .001$

In das Modell, das im zweiten Schritt analysiert wird und den Einbezug der Teilnehmererfahrungen aus Sicht der Lehrpersonen als Prädiktor überprüft, wird das Aufgreifen von Teilnehmererfahrungen aus Sicht des Moderatorenteams aufgrund des Ergebnisses aus dem ersten Analyseschritt (s. Abbildung 81) nicht mehr einbezogen.

Es zeigt sich (s. Abbildung 82), dass die anfängliche Motivation erneut einen stark positiven Prädiktor für die Nutzungsmotivation am Anfang von Modul 4 darstellt ($\beta = .43$, $p < .001$). Je stärker

die Motivation von Lehrpersonen am Anfang von *Vielfalt fördern* unter anderem daraus resultiert, dass sich die Schule mehrheitlich dafür entschieden hat, desto stärker arbeiten die Lehrpersonen auch in den Modulen – erfasst am Anfang von Modul 4 – aufgrund einer empfundenen Pflicht mit. Darüber hinaus zeigt sich, dass der von den Lehrpersonen wahrgenommene Einbezug eigener Erfahrungen ein signifikant negativer Prädiktor für die externe Erwartungsanpassung zu t_6 ist ($\beta = -.27, p < .001$).

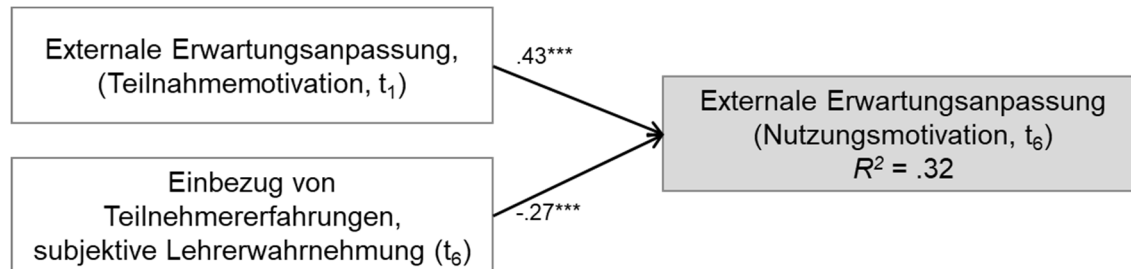


Abbildung 82. Vorhersage der Fortbildungsnutzungsmotivationsfacette externe Erwartungsanpassung durch den Einbezug von Teilnehmererfahrungen aus Lehrerperspektive. Anmerkung: *** = $p < .001$

Das negative Vorzeichen des Betakoeffizienten des lehrerberichteten Einbezugs von Teilnehmererfahrungen bedeutet: Wenn die teilnehmenden Lehrpersonen den Eindruck haben, dass die eigenen Erfahrungen von den Moderatorinnen und Moderatoren berücksichtigt werden und sie eigene Beispiele einbringen können, dann fallen die Werte auf der extrinsischen Motivationsfacette externe Erwartungsanpassung geringer aus. Zur Erinnerung: Mit dieser Facette wurde gemessen, inwieweit die Lehrpersonen während den Bausteinen von *Vielfalt fördern* aktiv mitarbeiten, weil sie dies als ihre Pflicht betrachteten. Ein Beispielitem lautet: „*Ich arbeite in den Fortbildungstreffen von Vielfalt fördern mit, weil ich dazu verpflichtet bin.*“

11.2 Motivationsfacette: Schulentwicklungsorientierung

In dem folgenden Abschnitt wird die Schulentwicklungsorientierung als eine Facette der Nutzungsmotivation näher untersucht. Eine hohe Ausprägung der Schulentwicklungsorientierung, die am Anfang von Modul 4 erfasst wurde, bedeutet, dass eine Lehrperson zur aktiven Mitarbeit in den Bausteinen von *Vielfalt fördern* motiviert ist, weil sie dadurch die Weiterentwicklung der Schule als Ganzes befördern möchte.

Mit den Regressionsmodellen auf Individualebene wird geprüft, ob sich nach Kontrolle der Teilnahmemotivation zu Beginn von *Vielfalt fördern* (t_1) die über ein Jahr später erfasste Nutzungsmotivation auch aus dem Einbezug von Teilnehmererfahrungen speist. Der Einbezug der Teilnehmererfahrungen, also unter anderem von Beispielen aus dem schulischen und unterrichtlichen Alltag der teilnehmenden Lehrpersonen, wurde ebenfalls am Anfang von Modul 4 (t_6) aus der Perspektive der Moderatorinnen und Moderatoren und aus der Perspektive der Lehrpersonen erfasst.

Im ersten Modell zeigt sich, dass der aus Moderatorenperspektive erfasste Einbezug von Erfahrungen der teilnehmenden Lehrpersonen keine Unterschiede zwischen Lehrpersonen in der Schulentwicklungsorientierung am Anfang von Modul 4 erklärt (s. Abbildung 83, $\beta = -.01$, n. s.).

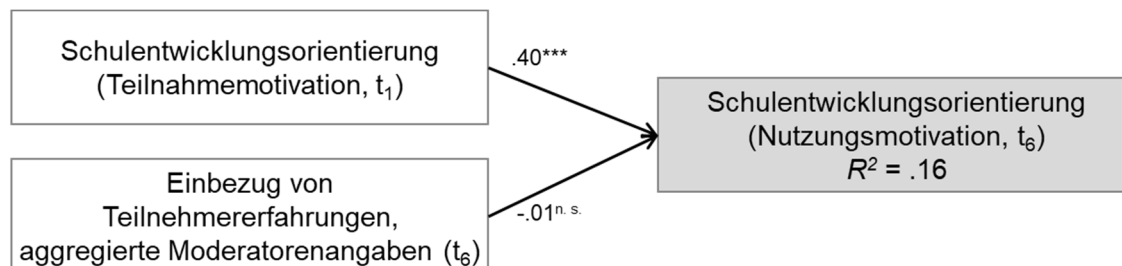


Abbildung 83. Vorhersage der Fortbildungsnutzungsmotivationsfacette Schulentwicklungsorientierung durch den Einbezug von Teilnehmererfahrungen aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant, *** = $p < .001$

Da der von den Moderatorinnen und Moderatoren angegebene Einbezug von Erfahrungen und Beispielen der teilnehmenden Lehrpersonen die Nutzungsmotivation nicht beeinflusst, wird diese Skala im zweiten Analyseschritt nicht erneut einbezogen. Als Prädiktoren zur Vorhersage der Schulentwicklungsorientierung am Anfang von Modul 4 werden die Ausprägung der anfänglichen Teilnahmemotivation und der von den Lehrpersonen wahrgenommene Einbezug von Teilnehmererfahrungen berücksichtigt (s. Abbildung 84).

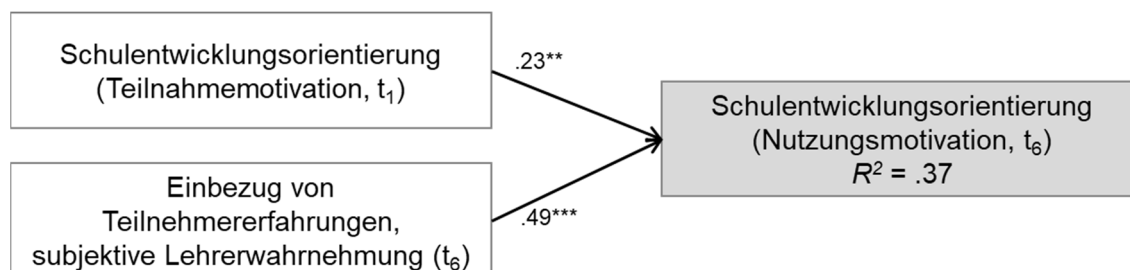


Abbildung 84 Vorhersage der Fortbildungsnutzungsmotivationsfacette Schulentwicklungsorientierung durch den Einbezug von Teilnehmererfahrungen aus Lehrerperspektive. Anmerkung: ** = $p < .01$, *** = $p < .001$

Tatsächlich erweisen sich sowohl die Teilnahmemotivation ($\beta = .23$, $p < .01$) als auch der Einbezug der Erfahrungen aus der Perspektive der Lehrpersonen ($\beta = .49$, $p < .001$) als vorhersage-relevant. Je stärker eine Lehrperson am Anfang von *Vielfalt fördern* zur Teilnahme motiviert ist, weil die Schule davon profitieren kann, und je stärker eine Lehrperson den Eindruck hat, dass die eigenen Erfahrungen und Beispiele der Teilnehmerinnen und Teilnehmer in den Fortbildungssitzungen einen Raum erhalten, desto stärker ist sie motiviert, die Fortbildungsinhalte zum Wohle der gesamten Schule zu nutzen.

11.3 Motivationsfacette: Soziale Interaktion

Abschließend wird das Ergebnis für das Modell erläutert, in dem die soziale Interaktion als eine Facette der Fortbildungsnutzungsmotivation als Kriterium untersucht wurde (s. Abbildung 85).

Hohe Werte auf dieser Skala drücken aus, dass eine Lehrperson in den Fortbildungssitzungen aktiv mitarbeitet, weil sie dadurch in einen intensiveren Kontakt mit Kolleginnen und Kollegen kommt und an deren beruflichen Erfahrungen teilhaben kann.

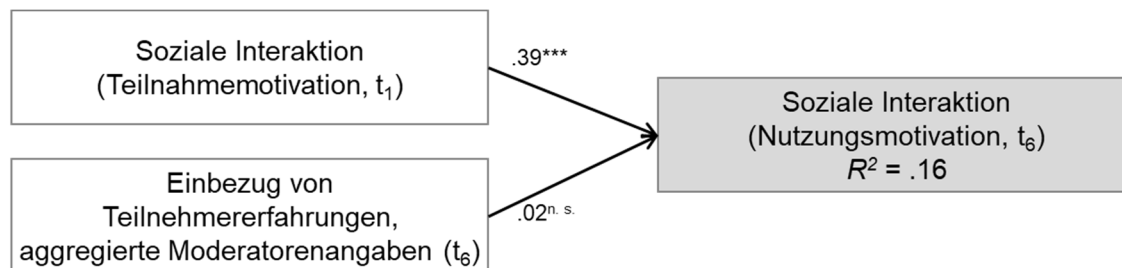


Abbildung 85. Vorhersage der Fortbildungsnutzungsmotivationsfacette soziale Interaktion durch den Einbezug von Teilnehmererfahrungen aus Moderatorenperspektive. Anmerkung: n. s. = nicht signifikant, *** = $p < .001$

Ebenso wie für die Motivationsfacette externale Erwartungsanpassung und Schulentwicklungsorientierung (s. Abschnitte 11.1 und 11.2) zeigt sich für die Motivationsfacette soziale Interaktion, dass die Einschätzung der Moderatorinnen und Moderatoren dazu, in welchem Maße die unterrichtlichen Erfahrungen der teilnehmenden Lehrpersonen in der Fortbildung aufgegriffen wurden, keine Unterschiede in der Nutzungsmotivation der teilnehmenden Lehrpersonen erklärt ($\beta = .02$, n. s.).

Hingegen erweist sich der lehrerberichtete Einbezug von Teilnehmererfahrungen erneut als ein statistisch bedeutsamer positiver Prädiktor neben der anfänglichen Teilnahmemotivation (s. Abbildung 86, $\beta = .40$, $p < .001$). Durch die beiden Prädiktoren können 32 % der Unterschiede in der Motivation der Lehrpersonen, deshalb an der Fortbildung mitzuarbeiten, weil sich damit Kontakte zu Kolleginnen und Kollegen vertiefen lassen, erklärt werden.

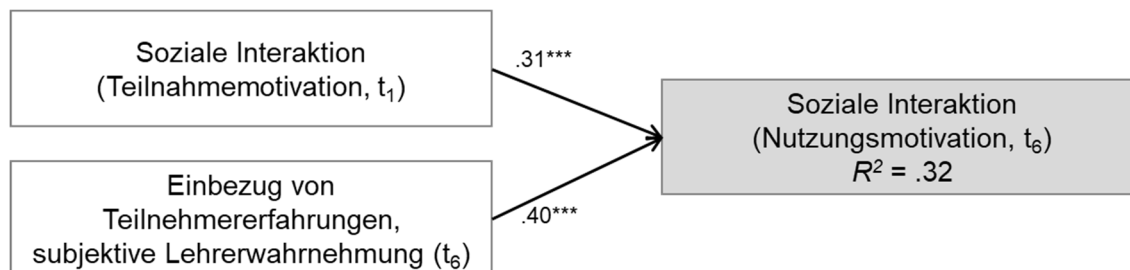


Abbildung 86. Vorhersage der Fortbildungsnutzungsmotivationsfacette soziale Interaktion durch den Einbezug von Teilnehmererfahrungen aus Lehrerperspektive. Anmerkung: *** = $p < .001$

Insgesamt haben die drei Modelle gezeigt, dass es günstig für die Nutzungsmotivation der teilnehmenden Lehrpersonen ist, wenn diese den Eindruck gewinnen, dass an ihren Erfahrungen angeknüpft wird und Beispiele aus der eigenen unterrichtlichen Praxis aufgegriffen werden. Dieser Einbezug von Teilnehmererfahrungen aus der Perspektive der Lehrpersonen führt zu einer geringen extrinsisch motivierten Nutzung des Fortbildungsangebots (externale Erwartungsanpassung) sowie zu einer stärkeren Motivation, in der Fortbildung aktiv mitzuarbeiten, um die Schule voranzubringen und den Austausch mit Kolleginnen und Kollegen zu intensivieren. Die

eigenen Wahrnehmungen der Lehrpersonen erweisen sich hierbei als deutlich vorhersagerelevanter als die Einschätzungen der Moderatorinnen und Moderatoren.

12 Moderatorenschulung und Fortbildungsmaterial

Ausgehend von dem Angebots-Nutzungs-Modell der Lehrerfortbildung (s. Abschnitt 2.1) dürften sich nicht nur Merkmale der Lehrpersonen und deren beruflicher Rahmenbedingungen auf den Erfolg einer Fortbildung auswirken, sondern auch Merkmale des Fortbildungsangebots und Merkmale derjenigen Personen, die die Fortbildung leiten, hier also die Gruppe der Moderatorinnen und Moderatoren.

Eine Besonderheit des Fortbildungsangebots *Vielfalt fördern* ist, dass die Moderatorinnen und Moderatoren eine spezifische Qualifizierung für die Durchführung der Fortbildung durchlaufen und die Möglichkeit haben, hierfür ein Zertifikat zu erwerben. Bei der ersten und zweiten Kohorte von *Vielfalt fördern* fand die Moderatorenschulung zu den einzelnen Modulen stets ein halbes Jahr vor der eigentlichen Fortbildungsdurchführung statt. Es zeigte sich jedoch im Prozess, dass dieser „Vorsprung“ der Moderatorinnen und Moderatoren gegenüber den teilnehmenden Lehrpersonen zu knapp ist. Ab der vierten Kohorte von *Vielfalt fördern* wurde die Qualifikation mit ca. einem Jahr Abstand zur Fortbildung durchgeführt. Entsprechend systematische und langfristige Moderatorenqualifikationen sind in der Fortbildungslandschaft bisher kaum anzutreffen.

Um eine Rückmeldung zur Moderatorenqualifizierung anbieten zu können, wurden die Moderatorinnen und Moderatoren am Ende der Fortbildung um eine Beurteilung ihrer eigenen Qualifizierung gebeten. Die Skala setzt sich aus insgesamt acht Items zusammen. Ein Skalenmittelwert über $M = 3.50$ ist so zu interpretieren, dass sich die Moderatorinnen und Moderatoren aufgrund ihrer Schulung optimal auf die Fortbildungsanforderungen vorbereitet fühlten und die Qualifizierung als ausreichend lang, systematisch und insgesamt von hoher Qualität erlebt wurde (Beispielitem: „Die Schulungen waren insgesamt von hoher Qualität“, $\alpha = .83$). Im Mittel bewerten die Moderatorinnen und Moderatoren der zweiten Kohorte die angebotene Qualifizierung für die Durchführung von *Vielfalt fördern* mit $M = 2.56$ ($SD = .80$). Der Wert liegt unter dem theoretischen Skalenmittel. Das bedeutet: Die Moderatorinnen und Moderatoren beurteilen die eigenen Schulungen, durch die sie auf die Aufgaben und Anforderungen von *Vielfalt fördern* vorbereitet wurden, eher kritisch. Wie bereits angedeutet, erfolgte an dieser Stelle bereits eine Nachsteuerung. Durch eine frühere Qualifizierung zu den Fortbildungsinhalten erhalten die Moderatorinnen und Moderatoren einen größeren zeitlichen Spielraum, die Ausbildungsinhalte zunächst im eigenen Unterricht anzuwenden und zu erproben, bevor sie mit der Fortbildung beginnen.

Ergänzt werden soll die Rückmeldung aus Sicht der Moderatorinnen und Moderatoren an dieser Stelle um eine Beurteilung der Fortbildungsmaterialien. Den Moderatorinnen und Moderatoren wird für die Durchführung von *Vielfalt fördern* ein umfangreiches Materialpaket bereitgestellt. Es stellt sich daher die Frage, ob diese Unterlagen als ansprechend und motivierend, als konkret genug und nicht zu theorieelastig und abstrakt empfunden wurden, ob eine gute Einbindung in die

Fortbildungsdurchführung möglich war und ob die Materialien immer rechtzeitig vorlagen. Entsprechende Aspekte wurden mit einer Skala aus sieben Items erfasst (Beispielitem: *„Die Materialien (Präsentationen, Inputs, Unterlagen, Lektüren ...) für die Fortbildungsbausteine waren ansprechend und motivierend gestaltet“*, $\alpha = .71$).

Auch hier spiegelt sich in der mittleren Bewertung der Qualität des Fortbildungsmaterials ein eher kritisches Urteil der Moderatorinnen und Moderatoren wider. Im Mittel ergibt sich ein Wert, der mit $M = 2.59$ ($SD = .74$) unter dem theoretischen Skalenmittel liegt. Diese Einschätzung ist so zu interpretieren, dass die Materialien eher als zu theorielastig und abstrakt und demzufolge als überarbeitungswürdig beurteilt wurden.

13 Unterrichtliche Anregungen zur Reflexion über das Lernen

Unter dem Stichwort „*Selbstgesteuertes Lernen: Schülerinnen und Schüler werden Experten für das eigene Lernen*“ sollen die Fortbildungsteilnehmerinnen und -teilnehmer im zweiten Baustein des dritten Moduls lernen und sicherer darin werden, ihre Schülerinnen und Schüler zu Experten für das eigene Lernen werden zu lassen (Müncher & Sturm-Schubert, 2015). Dazu, Expertin/Experte für das eigene Lernen zu sein und selbstgesteuert zu lernen, gehört auch, dass Schülerinnen und Schüler metakognitive Strategien an die Hand bekommen, um ihre Lernprozesse zu überwachen, und dass sie während des Unterrichts dazu veranlasst werden, diese anzuwenden.

Eine solche metakognitive Förderung und metakognitives Lernen beziehen sich auf den Erwerb von Wissen über das eigene Lernverhalten sowie auf die gezielte Planung, Überwachung, Reflexion und Anpassung des Lernverhaltens in Bezug auf das Lernziel. Entsprechende Anregungen der Lernenden durch die Lehrperson können nicht nur das Wissen der Lernenden über metakognitive Strategien erweitern, sondern sich auch positiv auf deren Lernerfolge auswirken (Lipowsky, 2015).

Im Rahmen der Unterrichtsstudie wurden die Lernenden zu beiden Erhebungszeitpunkten (ust₁, ust₂) danach gefragt, inwiefern sie im Mathematik- beziehungsweise Deutschunterricht eine Anregung zur Reflexion über das Lernen durch ihre jeweiligen Lehrpersonen wahrnehmen. Dieses Konstrukt wurde in den beiden Unterrichtsfächern über insgesamt fünf Items erfasst, die unterschiedliche fachbezogene Formulierungen aufwiesen. Ein Item²⁹ im Fach Mathematik lautet beispielsweise: „*Mein Mathematiklehrer fordert mich immer wieder dazu auf, selbst zu erkennen, was ich in Mathematik gut kann und was noch nicht.*“ Im Fach Deutsch sind die Items³⁰ mit Bezug auf die Leseförderung formuliert, z. B.: „*Mein Deutschlehrer fordert mich immer wieder dazu auf, über meine Fortschritte im Lesen nachzudenken.*“ Das Antwortformat reichte von 1 „*stimmt gar nicht*“ bis zu 4 „*stimmt genau*“. In beiden Fächern sind höhere Werte auf der Skala daher als eine stärkere Anregung zur Reflexion über das Lernen während des Unterrichts zu interpretieren, während niedrigere Werte darauf hindeuten, dass die Schülerinnen und Schüler im Unterricht weniger dazu angeregt werden, über ihr Lernen zu reflektieren.

Das Strategiewissen der Schülerinnen und Schüler wurde in beiden Fächern mit erprobten Tests erfasst (s. Abschnitt 4.1.2.4). Den Lernenden werden in beiden Verfahren mehrere fachspezifische Lern- und Leistungsszenarien zusammen mit einer Auswahl von jeweils fünf bis sechs strategischen Handlungsalternativen präsentiert. Diese Handlungsvorschläge müssen die

²⁹ Die Items beginnen bei der Posterhebung am Anfang des siebten Schuljahrs (ust₂) stets mit: „*Mein Mathematiklehrer in der 6. Klasse forderte mich immer wieder dazu auf, ...*“, um den Bezug auf das zurückliegende sechste Schuljahr herzustellen.

³⁰ Auch im Fach Deutsch werden die Schülerinnen und Schüler bei der Erhebung am Anfang des siebten Schuljahres (ust₂) durch die Items angeregt, sich auf das zurückliegende sechste Schuljahr zu beziehen. Die Items beginnen alle mit: „*Mein Deutschlehrer in der 6. Klasse forderte mich immer wieder dazu auf, ...*“.

Lernenden hinsichtlich ihrer Qualität und Angemessenheit beurteilen. Zu beiden Erhebungszeitpunkten können im Fach Mathematik maximal 34 Gesamtpunkte erreicht werden und im Fach Deutsch 41 Punkte.

Im globalen Vergleich (s. Abschnitt 5.2.3) zeigten sich für die Lernenden der UG und die Lernenden der KG mehrebenanalytisch im Mittel keine unterschiedlichen Einschätzungen der unterrichtlichen Anregungen zur Reflexion über den eigenen Lernprozess sowie in der Entwicklung des Strategiewissens. Hängen aber eine intensivere Unterstützung des selbstgesteuerten Lernens der Schülerinnen und Schüler und spezifisch die unterrichtliche Anregung zur Reflexion über das eigene Lernen damit zusammen, wie Lernende ihr fachbezogenes Strategiewissen entwickeln? Mit den folgenden Korrelationsanalysen auf Klassenebene wird der Zusammenhang zwischen der unterrichtlichen Anregung zur Reflexion über das Lernen aus der Perspektive der Lernenden und der mittlere Zuwachs in den UG-Klassen im strategiebezogenen Wissen in den Fächern Mathematik und Deutsch untersucht.

13.1 Fach Mathematik

Für das Fach Mathematik weist der Mittelwert für die Skala *Anregungen zur Reflexion über das Lernen* ($M = 2.81$, $SD = 0.59$), der über dem theoretischen Skalenmittel von $M = 2.50$ liegt, darauf hin, dass sich die Lernenden von ihrer Mathematiklehrkraft am Ende des fünften Schuljahres eher dazu angeregt fühlen, über den eigenen Lernprozess nachzudenken und zu reflektieren. Bei der zweiten Erhebung wurde ein Mittelwert von $M = 2.74$ ($SD = 0.61$) ermittelt. Das bedeutet, die Schülerinnen und Schüler geben auch zu diesem Zeitpunkt an, dass ihre Mathematiklehrkraft sie eher zur Reflexion über das eigene Lernen anregt.

Im mathematikbezogenen Strategiewissenstest erreichen die Schülerinnen und Schüler der UG am Ende der 5. Klasse zwischen 1 und 29 Punkten. Im Mittel wird die Angemessenheit von $M = 14.06$ ($SD = 5.74$) Handlungsalternativen richtig beurteilt. Etwa ein Jahr später, d. h. zu ust_2 am Anfang der 7. Klasse, werden maximal 32 Punkte von den Schülerinnen und Schülern der UG erreicht. Der Mittelwert beträgt $M = 15.42$ ($SD = 6.36$). Auf Klassenebene betrachtet, hängt das mathematikbezogene Strategiewissen, das am Ende der 5. Klasse erfasst wurde, positiv mit dem Strategiewissen zusammen, das am Anfang der 7. Klasse ermittelt wurde ($r = .53$, $p < .01$). Demnach gehen höhere Werte in dem Strategiewissenstest zu dem einen Erhebungszeitpunkt mit höheren Werten zum anderen Erhebungszeitpunkt einher.

Wie hängt nun die mittlere Entwicklung des Strategiewissens der UG-Klassen damit zusammen, wie die Klassen die Anregung zur Reflexion über das eigene Lernen durch ihre Mathematiklehrkraft während des Unterrichts wahrnehmen? Es zeigt sich, dass der mittlere Zuwachs im mathematikbezogenen Strategiewissen in den UG-Klassen unabhängig davon ist, wie intensiv die Schülerinnen und Schüler sich von ihrer Mathematiklehrkraft im Unterricht dazu angeregt fühlen,

über das eigene Vorgehen und den eigenen Lernprozess nachzudenken ($r_{ust1} = .22$, $r_{ust2} = .08$, n. s.). Betrachtet man nicht die Wahrnehmung der Schülerinnen und Schüler einer Klasse zu den jeweiligen Erhebungszeitpunkten, sondern die Veränderung darin, wie sehr sich die Lernenden durch ihre Lehrkraft aufgefordert und angeregt fühlen, über das eigene Lernen zu reflektieren, zeigt sich, dass diese Entwicklung der Wahrnehmungen unabhängig von der mittleren Entwicklung im mathematikbezogenen Strategiewissen der Lernenden ist ($r = .00$, n. s.).

13.2 Fach Deutsch

Die Anregung zur Reflexion über den eigenen Lernprozess wird von den Schülerinnen und Schülern der UG zum ersten Messzeitpunkt eher mäßig beurteilt. Der Mittelwert von $M = 2.50$ ($SD = 0.71$) entspricht exakt dem theoretischen Skalenmittel. Bei der zweiten Erhebung am Anfang der 7. Klassen wird die Anregung zur Reflexion über das eigene Lernen durch die Deutschlehrkraft von den Schülerinnen und Schülern der UG tendenziell etwas geringer beurteilt ($M = 2.29$, $SD = 0.70$).

Im metakognitiven Strategiewissenstest für den Bereich erreichen die Lernenden der UG am Ende der 5. Klasse zwischen 1 und 23 Punkte. Durchschnittlich wird etwa ein Viertel der Handlungsalternativen von den Lernenden korrekt beurteilt ($M = 9.82$, $SD = 4.67$). Bei der zweiten Erhebung erreicht die schwächste Schülerin/der schwächste Schüler ebenfalls nur einen Punkt. Die höchste Punktzahl, die am Anfang der 7. Klasse erreicht wird, beträgt 22 Punkte ($M = 7.91$, $SD = 4.43$). Das Strategiewissen der Schülerinnen und Schüler der UG hängt auf Klassenebene zu den beiden Erhebungszeitpunkten positiv zusammen ($r = .56$, $p < .01$). Mit anderen Worten: Schülerinnen und Schüler in Klassen, die zum ersten Erhebungszeitpunkt über ein umfassenderes Strategiewissen im Bereich Lesen verfügen, besitzen auch zum zweiten Erhebungszeitpunkt ein größeres Strategiewissen. Gleichzeitig deutet die Korrelationsstärke darauf hin, dass es sich um ein Merkmal handelt, das zwar relativ stabil ist, aber dennoch Potenzial für den Einfluss weiterer Variablen bietet.

Im nächsten Schritt wurde untersucht, inwiefern die mittlere Entwicklung von Klassen im strategiebezogenen Wissen damit zusammenhängt, wie die Klassen die Anregung zur Reflexion über den eigenen Lernprozess durch ihre Deutschlehrkraft wahrnehmen. Es zeigt sich, dass die mittlere Entwicklung im metakognitiven Strategiewissen im Bereich Lesen weder zum ersten Erhebungszeitpunkt ($r = .01$, n. s.) damit zusammenhängt, wie umfangreich sich die Schülerinnen und Schüler einer Klasse durch ihre Lehrkraft dazu angeregt fühlten, über das eigene Vorgehen und die eigene Entwicklung im Bereich Lesen nachzudenken, noch zum zweiten Erhebungszeitpunkt, d. h. am Anfang der 7. Klasse ($r = .17$, n. s.). Auch wenn man auf Klassenebene den Zusammenhang zwischen der mittleren Entwicklung im Strategiewissen und der mittleren Veränderung in der unterrichtlichen Anregung zur Reflexion über das eigene Lernen untersucht, zeigt sich keine statistisch bedeutsame Korrelation ($r = .15$, n. s.). Die Ergebnisse deuten darauf

hin, dass die Entwicklung von metakognitivem Wissen im Bereich Lesen unabhängig davon ist, wie die Klassen in ihrem Deutschunterricht nach eigenen Angaben dazu angeregt werden, bewusst über das eigene Vorgehen und den eigenen Lernprozess nachzudenken.

14 Befunde der Vertiefungsstudie

Die Ergebnisdarstellung des folgenden Abschnitts basiert auf den Daten aus den Interviews, die mit sieben Schulen aus der UG im Rahmen der Vertiefungsstudie durchgeführt wurden. Die Schulen stammen aus allen Regionen, die im Rahmen der zweiten Kohorte in das Projekt *LIQUID* einbezogen wurden.

Die Schulen A, B und G sind Schulen, die bei den schulbezogenen Auswertungen in Abschnitt 7.1.1 weder durch eine negative noch durch eine positive Entwicklung aufgefallen sind. Die Schule F ist durch eine marginal negative Entwicklung bei der kollektiven Selbstwirksamkeitserwartung aufgefallen. Eine eher ungünstige Entwicklung der kollektiven Selbstwirksamkeitserwartung weist auch Schule D auf. Neben der kollektiven Selbstwirksamkeitserwartung hat sich in diesem Kollegium allerdings auch die Kooperationsfacette Fokus auf das Lernen der Schülerinnen und Schüler ungünstig entwickelt. Die Schulen C und E fielen durch eine positive Entwicklung der kollektiven Selbstwirksamkeitserwartung auf. Schule C weist außerdem eine günstige Entwicklung bei der Deprivatisierung auf.

An dem Interview in den Schulen B, D und G haben nur die Schulleitung und Steuergruppenmitglieder teilgenommen, während in den Interviews mit den Schulen A, C, E und F auch Lehrpersonen ohne besondere Funktion im Zusammenhang mit der Durchführung von *Vielfalt fördern* vertreten waren. Eine Besonderheit des Personenkreises an Schule F ist, dass eine Lehrperson selbst auch die Moderation von *Vielfalt fördern* an anderen Schulen durchgeführt hat und somit sowohl die teilnehmende als auch die fortbildungsleitende Rolle kennt. Die Interviews mit den Schulen haben überwiegend zwischen eineinhalb und zweieinhalb Stunden ($M = 107$ Minuten, $SD = 20$ Minuten) gedauert.

Im Folgenden wird inhaltlich herausgearbeitet, wie die Interviewteilnehmerinnen und -teilnehmer ihre motivationale Ausgangslage und Entwicklung einschätzen und welche Themen von *Vielfalt fördern* als besonders wichtig empfunden wurden. Es werden darüber hinaus Aussagen zu den Moderatorinnen und Moderatoren dargestellt. Durch diese thematische Schwerpunktsetzung sind Illustrierungen, Ergänzungen und/oder auch Kontrastierungen zu den Ergebnissen aus dem quantitativ angelegten Evaluationsteil möglich. In Fällen, in denen unter bestimmten Voraussetzungen eine Identifikation einzelner Schulen aufgrund von Angaben potenziell möglich wäre, wurde auf die Angabe der Schule verzichtet. Bei längeren Beispielsequenzen aus dem Interviewmaterial werden sprachliche Glättungen vorgenommen, indem unter anderem bejahende Sprecherelemente und/oder Ergänzungen anderer Personen, die im Gespräch nicht weitergeführt wurden, herausgelassen werden.

14.1 Motivation

In Abschnitt 6 wurde dargestellt, dass die Fortbildungsteilnehmerinnen und -teilnehmer ausgehend von ihren Voraussetzungen, zu denen unter anderem Facetten der Fortbildungsteilnahmemotivation zählen, unterschiedliche Cluster (Gruppen) bilden. Dabei zeigen sich Gruppen mit eher ungünstigeren Voraussetzungen und Gruppen mit günstigeren Voraussetzungen. Zwischen den Lehrpersonen dieser Gruppen wurden bedeutsame Unterschiede in der Nutzung und Wahrnehmung des Fortbildungsangebotes ermittelt.

Mit dem quantitativen Erhebungsinstrument des Projekts *LIQUID* wurden unterschiedliche Facetten der Motivation erfasst, von denen manche eher eine extrinsische Motivationsqualität (z. B. externale Erwartungsanpassung) abbilden und manche eher eine intrinsische Motivationsqualität (z. B. Entwicklungsorientierung). Im Fokus der quantitativen Auswertungen standen drei Facetten der Motivation: die externale Erwartungsanpassung, die Schulentwicklungsorientierung und die soziale Interaktion.

Diese drei Facetten der Fortbildungsteilnahmemotivation zeigen sich auch in den qualitativen Daten neben weiteren Beweggründen, wie etwa einer allgemeinen Fortbildungsfreude sowie dem Wunsch nach neuem „Handwerkszeug“ und Methoden zum Umgang mit der zunehmenden Heterogenität der Schülerinnen und Schüler.

14.1.1 Ausdruck extrinsischer Motivation

Nicht von allen Schulen wird in dem Interview ausdrücklich auf die teilweise fehlende Teilnahmemotivation oder auf eine auffindbare extrinsische Motivierung am Anfang der Fortbildung *Vielfalt fördern* eingegangen, jedoch finden sich deutliche Aussagen bei den Schulen A, C, D und F.

So stellt beispielsweise die Schulleiterin/der Schulleiter von Schule A klar, dass die Frage der Freiwilligkeit nicht im Raum stand beziehungsweise die Fortbildungsteilnahme nicht auf freiwilliger Basis erfolgte, sondern einzelne Lehrpersonen gezielt ausgewählt wurden.

Wir haben nicht gefragt, wer möchte daran teilnehmen, sondern wir haben in der Vorbereitungsgruppe überlegt, was macht Sinn und haben dann gesagt aus allen Hauptfächern sollen, Fachlehrer der Hauptfächer der Parallelklassen sollen daran teilnehmen und haben gefragt, wer dagegen sei. Und es war keiner, der es dann nicht wollte. (Schule A)

Die Lehrpersonen hätten also theoretisch die Option gehabt, sich gegen die Beteiligung auszusprechen – „[...] und haben gefragt, wer dagegen sei“ –, allerdings hat sich niemand dagegen ausgesprochen. Ob es von den Lehrpersonen an Schule A jedoch als real gegebene Möglichkeit wahrgenommen wurde, sich gegen die Fortbildungsteilnahme auszusprechen,

bleibt offen. Eine Interviewsequenz, die aus dem Interview an der Schule F stammt, wirft ebenfalls eher Zweifel daran auf, ob einzelne Lehrpersonen dieser Schule im Zusammenhang mit der Teilnahme an *Vielfalt fördern* eine reale Option gehabt hätten, nicht teilzunehmen.

L1: Hätt' ich sagen können, ich komm nicht? Nein.

L2: Ne, also das ist ja schon institutionell. (Schule F)

Die extrinsische Motivierung von einzelnen Kolleginnen und Kollegen wird auch im Interview mit Schule D benannt, allerdings wird dies nicht als ein Spezifikum für die Teilnahme an *Vielfalt fördern* verstanden, sondern eher als eine verschlossene Grundhaltung der betreffenden Lehrpersonen.

[...] manchmal ergibt sich dann so 'ne Art Zwangsläufigkeit über ein Team zu schicken. Dann kommt auch jemand mit ins Boot, der jetzt im ersten Schritt nicht intrinsisch gebrannt hat dafür. Das funktioniert dann manchmal. (...) ³¹ Jetzt waren ja alle im Boot. Und es gibt Leute, die, ob ich die jetzt irgendwo hinschicke oder ob die da sitzen, innerlich genauso verschlossen bleiben, wie sie eben sind. (Schule D)

Die Interviewteilnehmerinnen und -teilnehmer beschreiben die Motivationslage im Kollegium auch an anderen Stellen erneut als durchwachsen und problematisch: „*Also ich glaub die Motivation war insgesamt sehr durchwachsen*“ und „[...] *Schulleitung hat uns das aufgedrückt*“. Der gemeinsam getroffene Beschluss für die Teilnahme an *Vielfalt fördern* wurde aufgrund der Anforderungen und Herausforderungen von den Lehrpersonen im Laufe des Fortbildungsprozesses zum Teil vergessen: „[...] *Also, da wird dann auch gerne vergessen, dass es mal 'n gemeinsam herbeigeführter Beschluss war. Einfach, weil es anstrengend war.*“ Insgesamt wird von einer Interviewpartnerin/einem Interviewpartner aber geschätzt: „*Also es gab einen relativ kleinen Teil des Kollegiums, wo man sagen konnte, die Motivation war von Anfang bis Ende durchgehend nicht vorhanden.*“ Diese Einschätzung, die von der Lehrperson bezogen auf das eigene Kollegium getroffen wurde, kann in einen Zusammenhang mit den quantitativen Daten bezogen auf die gesamte Stichprobe gebracht werden. Denn auch dort zeigte sich, dass die Lehrpersonen, die überwiegend eine extrinsische Motivation besitzen, einen vergleichsweise geringen Anteil ausmachen (s. Abschnitt 6.1).

Eine Lehrperson von Schule C begründet ihre Skepsis oder geringe Motivation am Anfang von *Vielfalt fördern* mit dem Eindruck, dass die Schule in vielen Bereichen schon weit entwickelt ist, sodass ein geringer Mehrwert durch die Teilnahme an *Vielfalt fördern* erwartet wurde.

³¹ In runde Klammern gesetzte Punkte kennzeichnen eine Pause während eines Gesprächsbeitrags, wobei ein Punkt für eine kurze Pause steht und mehrere Punkte eine längere Pause bedeuten.

Aber ich weiß, dass ich damals auch sehr skeptisch war und ich glaube, mich auch in der Abstimmung dagegen entschieden hätte, weil ich genau das, auch das Gefühl hatte, dass wir in vielen Dingen schon so weit sind, weil wir uns so, mit so vielen Ideen auf den Weg gemacht haben, dass ich da schon direkt irgendwo das Gefühl hatte, ob das jetzt der Mehrwert, den man erwartet, ob der für uns jetzt überhaupt trägt, weil wir schon in vielen Dingen glaub ich, das hat sich dann nachher auch bewahrheitet, einfach weiter waren als das, was uns vermittelt werden sollte. (Schule C)

Obwohl von den Lehrpersonen in den Interviews eine extrinsische Motivation zum Ausdruck gebracht wird, sollte dies nicht ohne Weiteres als ungünstige Basis für den Fortbildungsprozess an den Schulen interpretiert werden. Beispielsweise beschreiben sich auch die Lehrpersonen der Schulen C und F als fortbildungsaffin und offen für Entwicklungsprozesse.

Die Offenheit, die allgemein da gewesen ist im Kollegium, die war ja auch sichtbar, sonst hätten wir uns ja sicherlich in dieser Form nicht entschieden. Wir wollten alle einfach mal neue Techniken kennenlernen, neue Aspekte kennenlernen. [...], aber grundsätzlich die Haltung des Kollegiums und auch meine Haltung ist schon, dass wir sagen, wir möchten uns fortbilden und wir möchten da weiter kommen in dem Bereich. (Schule C)

Wir machen jetzt nicht zwei Jahre 'ne Fortbildung und danach ist Ende, sondern danach geht's erst weiter, wir knüpfen daran an und gucken, wo können wir daran aufbauen an einzelnen Punkten. Und ich denke, das war wohl das, was uns natürlich auch alle motiviert hat. (Schule F)

14.1.2 Ausdruck einer Schulentwicklungsorientierung als Motivationsfacette

Für die Schule D war die Entscheidung zur Teilnahme an *Vielfalt fördern* unter anderem dadurch motiviert, dass schulische Entwicklungen nur gemeinsam funktionieren können und die Fortbildung das ganze Kollegium mit „ins Boot“ geholt hat: „*Unterrichtsentwicklung und Schulentwicklung kann ja nur gemeinsam gehen. Und das war jetzt 'ne gute Möglichkeit, das gesamte Kollegium ins Boot zu bekommen und einmal im Boot zu haben.*“

Durch *Vielfalt fördern* wurde aufgrund des zeitlichen Rahmens die Chance gesehen, dass Lehrpersonen gemeinsam Unterricht planen, durchführen und nachbesprechen können, um so das Kollegium weiterzubringen: „[...] *Zeit mit Kollegen gemeinsam Unterricht zu planen, vorzubereiten, durchzuführen, zu evaluieren. (.) Das war eigentlich so der, der Schritt wo ich gedacht habe, das ist was, was unser Kollegium auch vorwärtsbringt, gemeinsame Zeit zu haben.*“

Auch im Interview mit der Schule E wird darauf hingewiesen, dass *Vielfalt fördern* als eine Möglichkeit betrachtet wurde, die Schule weiterzuentwickeln, wie aus der folgenden Aussage deutlich wird: „[...] *als Chance gesehen von unserer Schule, sich da weiterzuentwickeln*“. Die Schulleiterin/der Schulleiter der Schule E begründet ihre/seine Motivation ebenfalls damit, einen langfristigen Prozess für alle Kollegen zu erhalten.

[...] hab gedacht, naja, dann kriegst du, dann kriegst du 'n zweijährigen Prozess, mit dem du deine Schule oder mit dem die Kollegen, die hier, die hier arbeiten, auf der - mit auf 'n Weg nehmen kannst und die kriegen dann nach und nach das Handwerkszeug. Das war so die Motivation, sich auf diesen Prozess einzulassen.
(Schule E)

Ähnlich erklärt auch die Schulleiterin/der Schulleiter von Schule B ihre/seine Motivation für die Teilnahme an *Vielfalt fördern* damit, viele Fortbildungstage und viel Unterstützung beim Aufbau der Schule zu bekommen.

[...] meine große Hoffnung als aufbauende Schule [war] ganz viel aus dieser Fortbildung mitnehmen zu dürfen, mitnehmen zu können, auch die Genehmigung von Seiten der Bezirksregierung, das sehe ich jetzt als Schulleiterin/Schulleiter einfach mal so, so viele Fortbildungstage für eine Schule zu bekommen, die man sonst als Schule ja überhaupt nicht hat. Einen Tag im Halbjahr hat man sonst zur Verfügung und das war's dann auch. Das war 'ne echte Herausforderung, 'n echter Wunsch und quasi so 'ne Motivation, die damit verbunden war. (Schule B)

Auch die Schule C, die sich – wie im vorigen Abschnitt dargestellt wurde – selbst als weit fortgeschritten beschrieben hat, weist an anderer Stelle im Interview darauf hin, dass sie sich als Schule gerade erst auf den Weg gemacht hat und daher hoch motiviert war, bei der Weiterentwicklung unterstützt zu werden: „*Wir war'n eben sehr jung, wir hatten uns grade erst auf den Weg gemacht und da möchte man natürlich von allen Seiten her sowas wie 'ne Struktur bekommen, unterstützende Hilfe bekommen und deswegen ist man natürlich hoch motiviert, sowas mitzumachen.*“

14.1.3 Ausdruck von sozialer Interaktion als Motivationsfacette

Es ist eingangs bereits darauf hingewiesen worden, dass die Lehrpersonen unter anderem der Wunsch nach einer sozialen Interaktion mit ihren Kolleginnen und Kollegen dazu motiviert hat, an *Vielfalt fördern* teilzunehmen.

Eine Lehrperson von Schule A beschreibt diesen Beweggrund folgendermaßen: „*Für mich jetzt die Hoffnung und die Aussicht auf einen prozentweisen Türöffner in der Kooperation mit den Kollegen, auf die ich angewiesen bin, um arbeiten zu können.*“ Im weiteren Verlauf des Interviews mit Schule A wird deutlich, dass das Team und der Teamgedanke zudem eine wichtige

Rolle für die Motivation der Lehrpersonen während *Vielfalt fördern* einnahmen, da man sich gegenseitig motivieren konnte: „*Aber wieso, wir waren doch ein Team und im Team haben wir uns gegenseitig motiviert.*“

Im Interview mit der Schule E formuliert eine Lehrperson, dass sie dadurch motiviert wurde, dass sie während der Fortbildung die Möglichkeit erhält, die Kolleginnen und Kollegen auf einer anderen Ebene als im unterrichtlichen Alltag kennenzulernen und von der Vielfalt innerhalb des Lehrerkollegiums zu erfahren: „*Man lernt auch gut die Vielfalt im Kollegium kennen. Ne, ist ja auch ganz unterschiedlich, was die Leute mitbringen. Ich finde so durch die Zusammenarbeit auf Fortbildungsebene, also nicht so im Schulalltag, sondern auf Fortbildungsebene lernt man die Leute auch noch mal anders kennen.*“

Ein deutlicher Hinweis darauf, dass die Aussicht auf eine Entwicklung und/oder Veränderung der schulischen Teamstrukturen motivierend wirkte, findet sich zudem in dem Interview mit der Schule F. Hier waren vor *Vielfalt fördern* noch keine festen Teamstrukturen etabliert, die Schule F hatte sich dies aber als Entwicklungsaspekt definiert und sah *Vielfalt fördern* daher als ein passendes Angebot: „*Ja, wir müssten daran was tun an dieser Kollegiumsstruktur. Also wir war'n vorher keine Teamschule, das ist ganz wichtig, wir war'n keine Teamschule. (.) Da hab' ich gesagt, das passt, das wär' 'ne richtig tolle Sache.*“

14.1.4 Motivation im Verlauf des Fortbildungsprozesses

Abschließend soll noch die Motivation im Verlauf der Fortbildung dargestellt werden, da nicht nur in den quantitativen Daten ein gewisser Unterschied zwischen der Teilnahme- und der Nutzungsmotivation zu erkennen war, sondern auch aufgrund der Interviewaussagen von einer Veränderung auszugehen ist.

Sehr drastisch wird dies von einer Schule beschrieben, die kurz vor der Entscheidung stand, die Fortbildung vorzeitig zu beenden, weil das Angebot als zu praxisfern empfunden wurde und die Motivation dadurch im ersten Fortbildungsjahr stark gesunken war.

Der Großteil war wirklich enttäuscht und dann haben wir auch auf 'ner Lehrerkonferenz tatsächlich sehr ausführlich darüber diskutiert, ob wir weiterhin an der Fortbildung teilnehmen oder nicht. Die Steuergruppe hat sich eben zum großen Teil, also drei der ehemaligen vier Steuerungsmitglieder haben sich ham gesagt, sie möchten es nicht weiter in der Steuergruppe betreuen. Dann haben wir eben die Steuergruppe noch mal neu besetzt, ich bin als Einzige geblieben und ham halt im Endeffekt mit dem Kollegium und auch in Absprachen mit den Moderatoren, unter der Voraussetzung, dass es praxisnäher wird, dass es mehr auf unsere Schule zugeschnitten wird, tatsächlich dann noch beschlossen, dass wir das Modul abschließen werden, aber es war bis zum Ende einfach so, dass es sehr schwierig

war mit dem gesamten Kollegium, weil die Motivation einfach irgendwie nach dem ersten Jahr eigentlich wirklich geraubt war im Großen und Ganzen.

Die Kollegen hatten irgendwann auch die Faxen dicke. Ja, um das mal ganz auf 'n Punkt zu bringen. Irgendwann war da Schluss, die wollten auch nichts mehr davon annehmen (..) aufgrund dieser ganzen Sachen, dass es nicht passgenau war, dass es nicht individualisiert war, dass es an der Stelle nicht zu uns passte.

Es wird aber auch beschrieben, dass es gelungen sei, die Motivation wiederaufzubauen, indem sich die Schule selbst stärker in die Fortbildungsgestaltung eingebracht hat, was bis zur Übernahme von Fortbildungstagen durch Lehrpersonen der Steuergruppe reichte.

Also ich hatte den Eindruck, genauso auch und dass das dann so im Laufe des dritten Moduls bei einigen Kollegen zumindest dann noch mal 'n bisschen besser wurde, weil, wenn ich das so sagen darf, wir uns dann auch viel eingeklinkt haben, weil wir dann das gemerkt haben, oje oje, das geht jetzt hier total schief, also die ham überhaupt keine Lust mehr, die Kollegen war'n jetzt nur noch so, oh ja, Vielfalt fördern, möcht ich gar nicht hin morgen, dass wir uns dann eingeklinkt haben und gesagt haben, das möchten wir nicht machen, das passt nicht zu uns, wir möchten lieber das 'n bisschen intensiver machen und dann teilweise eben auch Strecken des Fortbildungstages übernommen haben. Da wurd's dann wieder 'n bisschen besser.

Deutlich wird in den Interviewsequenzen, dass eine wahrgenommene Nicht-Passung und eine gefühlte Praxisferne dazu geführt haben, dass eine eigentlich vorhandene Motivation der Teilnehmerinnen und Teilnehmer abgenommen hat.

Eine ähnliche Beschreibung des Motivationsabfalls – bei einer generell offenen Grundhaltung gegenüber Fortbildungen – geben auch die Interviewteilnehmerinnen und -teilnehmer von Schule G an. Von einer Lehrperson aus Schule G wird die nachlassende Motivation unter anderem damit begründet, dass der Eindruck entstanden sei, kritische Rückmeldungen zum Fortbildungsprozess würden nicht berücksichtigt und auf die Erfahrungen von Erprobungen, die zwischen den einzelnen Bausteinen durchgeführt wurden, werde nicht ausreichend eingegangen.

Und die Kollegen haben es mitgemacht. Zum Teil eben, weil sie gesagt haben, gut dann machen wir es halt bis zum Ende und gucken dann, ob wir was rausziehen können. Andere haben vielleicht auch noch drauf gewartet, dass tatsächlich mal was Neues kommt, was wir noch gut gebrauchen können. Von daher eben diese

positive Einstellung war grundsätzlich vorhanden. Aber man hat mit jeder Fortbildung gemerkt, dass dann die Motivation dann auch weiter in den Keller gegangen ist, definitiv. (Schule G)

Ja, das Problem war halt immer, wenn was gesagt worden ist, wenn Kritik gesagt worden ist oder Ähnliches, dann war es auch, dass es eben überhaupt nicht, also wir hatten teilweise das Gefühl, dass das überhaupt nicht wahrgenommen wird. Dass da so'n Weg durchgelaufen wird und dann ist natürlich auch die Motivation im Nachhinein ein bisschen abgesunken. Wenn dann Sachen gemacht worden sind und es ist gar nicht mehr drauf eingegangen oder man hat gemerkt, man hat jetzt was Neues probiert, das funktioniert für uns nicht. Dann hat man es eben auch gelassen. Aber es war ja auch im Nachhinein auch im Verlauf eben nachher nicht mehr viel Neues dabei. Das ist ja so das, der Haken. Also so im Verlauf hat die Motivation einfach deshalb nachgelassen, weil eben nicht mehr auf unsere individuellen Bedürfnisse eingegangen worden ist. (Schule G)

Das empfundene Defizit an neuen und nützlichen Inhalten führte dazu, dass Aufgaben, die während der Fortbildung gestellt wurden, lediglich noch „abgearbeitet“ wurden und die Teilnahme an *Vielfalt fördern* zu einem Unmut führte.

Wir hatten diese Ideen praktisch alle schon im Kopf, (teilweise) schon verwirklicht an unserer Schule und wurden jetzt gezwungen, uns damit noch mal zu beschäftigen, und das schürte natürlich den Unmut auch im Kollegium, die dann sagten, warum sitzen wir jetzt hier nachmittags und machen etwas, was wir bereits schon kennen. (Schule C)

Ich denke schon, dass [...] diese Angebote am Anfang mit Sicherheit sehr angenommen wurden und auch mit 'ner anderen Motivation, als das zum Ende war. Es ist ja, wir haben natürlich trotzdem noch das, was zu erledigen war, aber da merkt man schon an dem Begriff erledigen, was zu erledigen war, ham wir natürlich erledigt, aber am Anfang war das kein Erledigen, sondern da war das eher so'n, ja euphorisch ist zu viel, aber da war 'n freudiges Handeln und Tun und es hat sich halt so abgeschwächt hin, dass es 'n Erledigen war, Haken, ich hab's jetzt gemacht, morgen ist die Fortbildung, ich hab' das, was ich machen musste, getan, Haken dran, erledigt und das ist ja irgendwie schade, weil das ist irgendwie da für mich jetzt eher 'n bisschen negativ behaftet gewesen. (Schule C)

L1: Ich sag mal, bestimmte Dinge wurden zum Schluss kollegial abgearbeitet. [...] Ohne dass man sich jetzt drauf freut, neue Informationen, neue Methoden kennenzulernen, mit denen man den Unterricht gewinnbringend verbessern kann.

L2: Wenn es die [neue Informationen, neue Methoden] gegeben hätte, wäre die Motivation wieder gestiegen. (Schule G)

Ein weiterer Grund für einen Motivationsabfall und/oder Unmut, der von einer Lehrperson von Schule D gesehen wurde, ist die Art und Weise, mit der die Inhalte während der Fortbildungsveranstaltungen vermittelt wurden: *„Und es war auch, das war auch der Widerstand glaub ich, der an verschiedenen Stellen, der vielleicht unabsichtlich erzeugt wird, indem da jemand kommt, 'ne PowerPoint aufklappt und für den ein oder anderen oberflächliche Inhalte präsentiert.“*

Umgekehrt findet man auch positive Formulierungen und Hinweise in den Interviewaussagen, was als motivierend erlebt wurde. So äußert eine Lehrperson von Schule D: *„Bei mir persönlich war es ganz stark auch vom Inhalt abhängig. Also die Praxisphasen fand ich super und da war die intrinsische Motivation auch da“*, und eine Lehrperson von Schule G antizipiert, dass ein praktischeres Arbeiten während der Fortbildungssitzungen auch dazu geführt hätte, dass zwischen den Bausteinen eine intensivere Auseinandersetzung mit den Inhalten stattgefunden hätte.

Ich glaube, die Motivation wäre größer gewesen, wenn wir wirklich in diesen Fortbildungen viel praktischer gearbeitet hätten. Dass dann die Motivation wirklich da gewesen wäre, zu sagen: So, und morgen oder einmal die Woche setzen wir uns auf jeden Fall auch noch mal eine Stunde zusammen und arbeiten daran weiter und probieren das aus. (Schule G)

Eine Lehrperson von Schule E vergleicht das eigene Lernverhalten beziehungsweise das von Lehrpersonen im Allgemeinen mit dem von Schülerinnen und Schülern. Dabei gelangt die Person zu dem Schluss, dass Arbeitsphasen motivierend wirken, wenn man den Eindruck hat, dass etwas entsteht und die Arbeit mit Spaß verbunden ist.

Okay (...) Ich kann da nur für mich persönlich sprechen. Bei mir entsteht die Motivation dann eigentlich immer in der Arbeitsphase. Ne, also, wenn dann irgendeine Arbeitsphase eingeleitet wird, dann bin ich glaub ich schon eher so'n bisschen Schüler und denk, ja gut, okay, zwei Stunden soll's dauern. Ist ja nicht so lange, aber dann, wenn, wenn man dann in dem Prozess drinsteckt und merkt, jetzt läuft da auch tatsächlich was an, dann, dann entsteht die Motivation dabei. Also ich glaube grundsätzlich, dass man, dass wir Lehrer da in solchen Fortbildungen nicht groß anders sind als die Schüler, die wir vor uns sitzen haben, wenn wir denen

Arbeitsaufträge erteilen. Ne (.) wenn man sieht, das macht Spaß und daraus entsteht etwas, dann ist man auch motiviert, das zu machen. (Schule E)

Zwei weitere Interviewpassagen aus dem Gruppeninterview mit der Schule E untermauern die Einschätzung, dass insbesondere die Phasen der Fortbildung als motivierend und zielführend empfunden wurden, in denen die Teilnehmerinnen und Teilnehmer selbst aktiv Dinge erarbeitet und entwickelt haben: *„Ich fand’s motivierend, dass man während des laufenden Prozesses in der Fortbildung in den Arbeitsphasen immer wieder Zeit hatte, Sachen zu entwickeln, handfeste Sachen, die man dann auch wirklich nutzen kann“* und *„Also das war nicht nur Blabla, was da in den Arbeitsphasen rauskommt, sondern etwas, das man tatsächlich auch nutzen kann (.) kurz- oder mittelfristig. (...) Und dann macht’s halt auch Spaß, ne, wenn man sieht, man entwickelt da etwas, das man auch verwenden kann später.“*

Eine Lehrperson von Schule B beschreibt, dass sie am Ende eines Fortbildungstages dann mit einem guten, glücklichen Gefühl nach Hause gegangen sei, wenn sie den Eindruck gehabt habe, dass Ziele erreicht worden seien und neben einem Input auch die Gelegenheit zur eigenständigen Arbeit an den Inhalten bestanden habe.

Da hatte ich auch das Gefühl nach so ‘nem Tag, wenn wir nach Hause gehen. Jetzt ham wir was geschafft. Wir ham wirklich Ziele erreicht. Wir ham, wir sind weitergebildet worden. Ham vielleicht Input bekommen, aber eben auch selber was daran gearbeitet, dass das zu einem Erfolg führen kann. Das war so ‘ne innere Motivation, die mich da an den bestimmten Stellen eben auch weitergebracht hat. (Schule B)

14.2 Relevante Themen

In den Interviews wurde deutlich, dass von den Lehrpersonen insbesondere diejenigen Fortbildungsphasen als motivierend und gewinnbringend wahrgenommen wurden, in denen eine eigenständige Erarbeitung stattfand. Nicht deutlich wurde bisher, welche der thematischen Inhalte von *Vielfalt fördern* als besonders relevant wahrgenommen wurden. Die quantitativ angelegte Studie liefert hierzu keine Hinweise, sondern konnte nur aufzeigen, wie intensiv unterschiedliche Themen behandelt wurden und dass die Intensität der Thematisierung teilweise einen schwachen Einfluss auf die Fortbildungswirksamkeit hat (s. Abschnitt 9.2).

Die Frage, welche Inhalte von *Vielfalt fördern* als besonders relevant erlebt wurden, wird im Folgenden anhand der Interviewdaten betrachtet. Im Verlauf der Gruppeninterviews wurden die Interviewpartnerinnen und -partner gebeten, die drei bedeutungsvollsten Inhalte zu benennen, aber ihre Wahl auch zu diskutieren und zu begründen.

Tabelle 14. Die drei relevantesten Themen/Inhalte von *Vielfalt fördern* aus Sicht der Vertiefungsschulen

	Aspekt 1	Aspekt 2	Aspekt 3
Schule A	Zeit	Deprivatisierung	Diagnoseinstrumente
Schule B	Teambildung	Kollegiale Hospitation	Arbeit am Logbuch
Schule C	Kollegiale Unterrichtsbesuche	Diagnostik mit Blick auf den einzelnen Schüler	Diagnosebögen erstellen
Schule D	Zeit für gemeinsame Planung und Reflexion	Hospitation	Klassenmanagement
Schule E	Hospitation	Fallbeispiele, Blick auf einzelne Schüler	Diagnostik
Schule F	Teambildung	Abläufe von Teamsitzungen	individualisierte Unterrichtsvorhaben
Schule G	Rituale		

Anmerkung: Die Reihenfolge der Angaben ist nicht im Sinne einer Wertigkeit zu interpretieren.

Betrachtet man die Aspekte, die in den Interviews genannt wurden (s. Tabelle 14), können sowohl Gemeinsamkeiten als auch Unterschiede zwischen den Schulen festgestellt werden. So nennt beispielsweise die Schule A den Aspekt „Zeit“, was im eigentlichen Sinne gar keinen Inhalt darstellt. Die Schule E gibt sehr allgemein den Bereich „Diagnostik“ als besonders relevant an, während die Schule C deutlich spezifischer eine Diagnostik mit Blick auf den einzelnen Schüler/die einzelne Schülerin und die Erstellung von Diagnosebögen angibt. Besonders auffällig ist die Schule G, die lediglich einen der Inhalte von *Vielfalt fördern* retrospektiv als besonders relevant betrachtet. Es handelt sich dabei um das Thema Rituale, wofür die Interviewpartnerinnen und -partner als Beispiele das Aufstellen der Schülerinnen und Schüler für den Gang in die Pause und eine regelmäßige Form der Kontrolle, ob benötigte Materialien von den Schülerinnen und Schülern mitgebracht wurden, angeben.

Ja die Rituale. Dieses Aufstellen, Schülerbuch kontrollieren, rausgucken. Also die Rituale, die wir aufgestellt haben. Wir haben uns überlegt, wo hakt es im Moment. Eine Sache war es halt, dass es in den Fluren recht tumultartig stellenweise zugeht. Dass wir halt geguckt haben, dass nicht ständig die Lehrer alles immer ziehen müssen, sondern wir haben uns überlegt, dass sich die jüngeren Klassen halt aufstellen. Und zweites Problem war zum Beispiel Hausaufgaben. Dass wir gesagt haben, das Schülerbuch muss zum Beispiel immer mit auf dem Tisch liegen, die sollen das eintragen, die sollen am Ende der Stunde zum Kontrollieren kommen. Solche Sachen, die wir dann halt als Ritual für die Klassen festgelegt haben. Und das funktioniert auch nach wie vor. (Schule G)

Andere Aspekte von *Vielfalt fördern* wurden von dieser Schule, zumindest unter dem „Neuigkeitsaspekt“, nicht als besonders relevant und nützlich empfunden: *„Ja, unter diesem Neuigkeitsaspekt nicht. Die kooperativen Lernformen, die haben wir gehabt. Wir haben kein Problem mit Team-Teaching und da kommt jemand an sagt ‚hey‘, da geht der Kollege da vorne rein, da wird keiner nervös.“*

Anders war die Interviewsituation an Schule D: Hier wurden zunächst mehr als drei relevante Inhalte benannt und statt einer eindeutigen Auswahl der drei Aspekte haben die Interviewteilerinnen und -teilnehmer quasi eine Clusterung vorgenommen und Inhalte inhaltlich miteinander verbunden.

L1: Ja, ich finde, irgendwie gehört das ja zusammen: Phasen der Reflexion und Zeit für gemeinsame Planung. [...] Also überhaupt dieser, dieser Austausch, dass man ins Gespräch kommt und auch am gemeinsamen Gegenstand. Daraus hat sich dann die Hospitation irgendwie so ergeben. [...] Und das Klassenmanagement.

L2: Und die individualisierenden Aufgabenformate sind ja in der Zeit für gemeinsame Planung eigentlich mit drin, weil das ist ja eigentlich der Auftrag gewesen, das zu tun. (Schule D)

Die Relevanz des Themas effektive Klassenführung wurde damit begründet, dass ein großer Bedarf darin bestand, gemeinsame Regeln aufzustellen, an denen sich dann alle Kolleginnen und Kollegen beim Unterrichten in einer Klasse orientieren. Als konkretes Beispiel bringen die Lehrpersonen von Schule D – ebenso wie die Lehrpersonen der Schule G – die Arbeit an Ritualen ein.

Ja und Classroom Management würde ich jetzt einfach noch mal sagen, dass da einfach ganz viel Bedarf bestanden hat, auch noch mal so gemeinsame Regeln aufzustellen, ne. Also, dass man in einem Jahrgang irgendwie noch mal besonders guckt, wo brennt's da grade und wie kann man da vielleicht mit einfachen Mitteln irgendwie ein bisschen dran schrauben, wenn alle Kollegen in der Klasse das so und so machen, dass es dann einfach insgesamt leichter wird. Da bestand glaub ich relativ viel Bedarf auch, sich noch mal so'n bisschen zurückzubedenken auf so einfache Maßnahmen. (Schule D)

Also zum Beispiel so'n ritualisierten Einstieg. Also wir haben so 'nen Schülerkalender und dann haben wir gesagt, der soll. Also wir sagen: „Alle Materialien auf den Tisch.“ So und dann haben wir aber beim Hospitieren gemerkt: Aha, bei einigen liegt das aufgeschlagen dann schon da. Die Schüler sehen, dass einfach dann

bestimmte Rituale dann so da sind und dass, wenn eben Arbeitsmaterial nicht da ist, also ich glaube, ganz viele sind mittlerweile so, dann gibt's Ersatzmöglichkeiten im Raum irgendwie so: 'ne Schere, 'nen Bleistift, 'nen Stift und solche Sachen. Und das eben auch gleiche Ansagen gemacht werden. (Schule D)

Weshalb die Schule D die gemeinsame Planung und Reflexion von Unterricht sowie die Hospitation als bedeutsam wahrgenommen hat, verdeutlicht das folgende Zitat.

Wir haben gelernt, dass wenn wir planen und wenn wir es gemeinsam machen und wenn wir das auch hospitieren und besprechen, dass trotzdem unter Umständen manchmal andere Dinge dabei rauskommen. Also wir haben viel zu wenig Rückkopplung im Alltag, sodass manchmal dann auch einem bewusst wird, wir besprechen das zwar immer, aber das kommt eben unterschiedlich an, auch bei Kollegen, so wie bei Schülern eben auch. Und ich glaube, dass das auch so 'ne Einsicht ist, wo man auch noch mal sieht, wie wie nötig eigentlich 'n Austausch für uns ist oder so. (Schule D)

Die kollegiale Hospitation wird aber nicht nur von der Schule D als ein besonders relevanter Aspekt im Verlauf von *Vielfalt fördern* betrachtet, sondern noch von vier weiteren Schulen (A, B, C und E).

Von einer Lehrperson der Schule B wird die Relevanz der kollegialen Unterrichtshospitationen, die durch die Teilnahme an der Fortbildung *Vielfalt fördern* angeregt wurden, damit begründet, dass auf diese Weise Transparenz unter den Lehrkräften entstehen kann und auch Kolleginnen und Kollegen miteinander zusammengearbeitet haben, die dies sonst eher nicht tun. Eine andere Lehrperson weist ergänzend dazu darauf hin, dass Lehrpersonen, die der kollegialen Hospitation zunächst skeptisch gegenüberstanden, am Ende darin einen Mehrwert erkennen konnten.

Bei der kollegialen Unterrichtshospitation war's tatsächlich so, dass Kollegen gegenseitig auch sich in den Unterrichtsveranstaltungen besuchten. Was wir ja eigentlich wollen, offene Klassentüren zu haben, auch Transparenz herzustellen, im kollegialen Miteinander keine Ressentiments verfestigen, sich verfestigen lassen. Das führte wirklich dazu, dass sich Kollegen besuchten, die vielleicht sonst gar nicht so miteinander gearbeitet hätten. (Schule B)

Was ich wirklich positiv an der kollegialen Unterrichtshospitation fand, war, dass wir wirklich einige Kollegen hatten, die dem ganzen wirklich sehr skeptisch gegenübergestanden sind und wirklich, ja so'n bisschen skeptisch war'n, dass da jemand jetzt in' Unterricht kommt und dies als sehr sehr positiv eben erfahren haben. Dass,

es ging ja bei diesen Hospitationen immer nur um einen konkreten Schwerpunkt, den man vorher sich auch überlegt hatte, und dass tatsächlich nur darauf geachtet wurde und dass das eben nicht persönlich wurde oder in irgendeiner Art, ja weiß ich nicht, die Unterrichtsqualität da berücksichtigt wurde oder besprochen wurde, sondern dass es wirklich, ja hilfreich war und man aufgrund dieser konkreten Schwerpunktsetzung wirklich 'n Mehrwert davon hatte. (Schule B)

Ein konkretes Beispiel, aus dem die Relevanz der kollegialen Hospitation hervorgeht, wird auch in dem Interview mit Schule C gegeben. Die Interviewpartnerin/der Interviewpartner beschreibt, dass eine kollegiale Unterrichtshospitation durchgeführt wurde, weil ein Schüler ein sehr auffälliges Sozialverhalten im Unterricht zeigte und die unterrichtende Lehrperson eine Rückmeldung zum eigenen Verhalten in solch einer kritischen Situation erhalten wollte.

Also mir fällt da spontan ein, dass wir ein- bei einem Schüler mal untersucht haben, das Sozialverhalten, weil er da wirklich extrem auffällig gewesen ist und der betroffene Lehrer hat halt den Auftrag, du schau mal ganz genau auf diesen Jungen und schau auch mal darauf, wie ich als Lehrperson mit diesem Jungen umgehe. An welchem Punkt sieht man mir an, dass es in mir brodelt, und was mach ich in dem Augenblick. Und die Stunde wurde dann durchgeführt und hinterher kam halt dieses Auswertungsgespräch und dann über dieses Gespräch selbst wurde einem dann schon bewusst, an welchen Stellschrauben man noch hätte drehen können, um das Sozialverhalten des Schülers besser zu steuern, vielleicht es nicht zur Eskalation kommen zu lassen. Und insofern hatte man diesen Blick dann noch mal geschärft auf diesen einzelnen Schüler und konnte in der nächsten Situation dann vielleicht ganz anders damit umgehen und das war natürlich schon sehr hilfreich. (Schule C)

Das Zitat enthält darüber hinaus einen Hinweis auf den zweiten Aspekt, der von der Schule C als besonders relevanter Inhalt von *Vielfalt fördern* benannt wurde, nämlich den Blick auf die einzelne Schülerin/den einzelnen Schüler beziehungsweise eine Diagnostik, die die einzelne Lernende/den einzelnen Lernenden betrachtet. Auch hierfür geben die Lehrpersonen noch ein konkretes Beispiel im Rahmen des Interviews. Aus diesem wird deutlich, dass der Blick auf einen Schüler und die gemeinsame Erstellung eines Diagnoseinstrumentes die Perspektiven der Lehrpersonen erweitert hat und ihnen beim Umgang mit dem Schüler geholfen hat.

L1: Wir ham, also zumindest in der Gruppe, wo ich war, ham wir die Diagnosebögen nicht nur auf nur auf den Unterricht allein bezogen, sondern wir ham auch bei einem Schüler, wo ich mich erinnere – ging es wirklich darum- irgendwie, dass wir uns zusammengesetzt haben und so 'ne praktisch 'n Diagnosebogen erstellt haben, wo wir gesagt haben, was müssen wir jetzt eigentlich erstmal außenrum, also

wie sieht's im Elternhaus aus, wie sieht's in den Freunden aus, wie mal 'n Gespräch mit ihm führen, was bewegt den Schüler. Dass wir sowas auch, so'n Diagnosebogen erstellt haben, den wir dann halt abgearbeitet haben, um einfach außerhalb des Unterrichts auch den besser kennenzulernen, zu wissen, irgendwie, irgendwo ist ja 'n Grund für sein Verhalten und den ham wir auch abgearbeitet und ich meine schon, dass- da ist zumindest heute ist der, den ich jetzt im Kopf hab, unauffällig. Inwiefern das jetzt, war ja jetzt keine spontane (Heilung), aber es hat doch in vielen, wenn man dann irgendwie weiß, wie's dem Schüler so außerhalb der Schule geht, das hilft, wenn man das selber weiß, einem auch sehr weiter, auch im Unterricht dann letztendlich natürlich weiter, wie man mit dem umgehen kann. War nicht nur auf Unterricht selber bezogen.

L2: Ich denke da- möchte dich unterstützen. In der Phase ist glaube ich dem Kollegium noch mal ganz deutlich geworden, dass Fehlverhalten aus verschiedenen Blickwinkeln halt betrachtet werden muss. [...] Und in der Hinsicht haben wir dann auch entsprechende Diagnoseverfahren dann versucht zu entwickeln und ham uns mal überlegt, welche Blickwinkel können wir eigentlich einnehmen. Und das war jetzt auch so im Nachhinein, wenn ich mir das so überlege, das wird ja immer noch hier praktiziert, das war sehr, sehr hilfreich.

Von der Schule E wird ebenfalls, wenn auch im Vergleich deutlich allgemeiner, das Thema Diagnostik als ein wichtiger Inhalt der Fortbildung benannt. Es wird deutlich, dass an der Schule ein weites Verständnis von Diagnostik besteht. Bereits ein Gespräch mit Kolleginnen und/oder Kollegen über eine Schülerin/einen Schüler wird als Diagnostik definiert und aufgrund der leichten Durchführbarkeit als wichtig empfunden.

Die einfachste Form der Diagnostik, die ich hier an der Schule kennengelernt habe, ist mit 'nem Kollegen zu reden. Beispielsweise die Frau S. anzusprechen und zu sagen, hör mal zu, die Schülerin M. und die handelt so und so und so, ich weiß nicht, was ist mit diesem Mädels los punkt punkt punkt. Und dann erzählt die Frau S. mir, die Schülerin M. hat Folgendes zu Hause dies und das und jenes. Dann ist es die einfachste Form der Diagnostik, die ich für mich umsetzen kann, 'n gewisses Verständnis für dieses Kind zu haben und erst mal zu wissen, warum die so ist. Deshalb kann ich die noch nicht ändern, aber dann kann ich vielleicht schon mal mein Verhalten ein wenig umstellen. Und ich hab' dann einfach Erklärungsmuster. Ist für mich die einfachste Form der Diagnostik und von daher find ich die sehr wichtig. (Schule E)

Aufgrund einer anderen Aussage im Interview lässt sich allerdings annehmen, dass mit diesen Gesprächen nicht nur Tür-und-Angel-Gespräche zwischen Kolleginnen und Kollegen gemeint

sind. Eine Lehrperson erklärt, dass aufgrund von *Vielfalt fördern* Teams gebildet worden seien und man Zeit und Handwerkszeug erhalten habe, um in kollegialen Gesprächen über das „Lamentieren“ hinauszukommen und genauer auf einzelne Schülerinnen und Schüler zu blicken.

Da hat man oft auch im Alltag - nimmt man sich nicht die Zeit dafür. Oder, oder wir, wir haben sie manchmal ja auch nicht (..) Und das hängt ja zusammen, ob manche Fallballbeispiele in der Diagnostik – dass, dass wir dann auch über das (...) Lamentieren sag ich mal, nicht hinauskommen, ne. Also erst mal, nicht das sich gegenseitig zu erzählen, worüber man sich geärgert hat, aber wir haben da was an der Hand, wo ich mich dann auch mal hinsetzen kann und wenn wir jetzt, wir haben jetzt ja Teams gebildet, die, ich hab' jetzt zum Glück auch 'n Team, wo wir so personell ganz gut zusammengesetzt sind so, die alle in den gleichen Klassen arbeiten. Da kann man sich auch mal hinsetzen und wirklich mal, noch mal beim ganz schwierigen Schüler dann gucken und versuchen, da sich runterzuholen, also aus den Emotionen rauszuholen und so und wir haben da Handwerkszeug, um sich das noch mal in Ruhe zu durchleuchten (Schule E).

Insgesamt deutet sich in den Interviews an, dass konsensuell vornehmlich die Inhalte des ersten und des zweiten Moduls von den Schulen als besonders relevant wahrgenommen wurden, z. B. die Zeit für den kollegialen Austausch und die Teambildung im Allgemeinen, aber auch speziell die Durchführung kollegialer Unterrichtshospitationen. Im Modul Diagnostik wurde insbesondere das Einnehmen eines speziellen Blicks auf das Lernen und/oder Verhalten einer Schülerin/eines Schülers sowie das Arbeiten an und mit Diagnoseinstrumenten von den Lehrpersonen als bedeutsam betont.

14.3 Lernförderliche Eigenschaften von Moderatorinnen und Moderatoren

In dem Teil des Interviews, in dem es um die Moderatorinnen und Moderatoren von *Vielfalt fördern* geht, äußert sich eine Interviewpartnerin/ein Interviewpartner aus der Schule A folgendermaßen: „Also ich glaube, diese Fortbildung steht und fällt mit den Moderatoren [...]“. Den Moderatorenteams wird damit eine ausgesprochen hohe Bedeutung und Verantwortung für das Gelingen der Fortbildung durch die Lehrperson beigemessen.

Durch die Ergebnisse der quantitativen Auswertungen konnte tatsächlich gezeigt werden, dass der von den Lehrpersonen wahrgenommene Enthusiasmus des Moderatorenteams einen positiven Einfluss auf die Nutzung des Fortbildungsangebots und den selbstberichteten Lernertrag hat (s. Abschnitt 10). Im Folgenden werden Interviewaussagen gebündelt, die die besondere Bedeutung der Moderatorenteams aus Sicht der Interviewpartnerinnen und -partner unterstreichen.

In den Interviews wurden zunächst lernförderliche und unterstützende Merkmale einer Moderatorin/eines Moderators gesammelt. Anschließend sollten die drei wichtigsten Eigenschaften genannt werden. Es zeigte sich jedoch vielfach, dass den Interviewteilnehmerinnen und -teilnehmern eine Festlegung auf drei Merkmale schwerfiel beziehungsweise nicht gelang. Die Zusammenfassung der verschiedenen Eigenschaften variiert hierbei durchaus zwischen den Schulen (s. Tabelle 15). Eine Lehrperson von Schule C fasst das „Dilemma“ folgendermaßen zusammen: *„Okay, wir können uns nicht für drei entscheiden, wir finden alles wichtig. Das ist dann die eierlegende Wollmilchsau.“*

Tabelle 15. Lernförderliche Eigenschaften einer Moderatorin/eines Moderators aus Sicht der Vertiefungsschulen

	Aspekt 1	Aspekt 2	Aspekt 3
Schule A	Passung mit Schulform/ Fallbezug	Präsentationskompetenz/ Fachwissen/ Durchsetzungsvermögen	Flexibilität
Schule B	Fachliche Kompetenz	Präsentationskompetenz/ Authentizität/ Enthusiasmus	Flexibilität
Schule C	Humor/Enthusiasmus	Fachkompetenz/ Moderationskompetenz/ Souveränität	Flexibilität/Schul- und Fallbezug/ Systemerfahrungen
Schule D	Systemerfahrung	Fachwissen/ Flexibilität	Authentizität/ Vermittlung auf Augenhöhe
Schule E	Fachwissen	Authentizität	Flexibilität
Schule F	Teilnehmerorientierung/ Flexibilität	Humor/Authentizität	Souveränität
Schule G	Fachwissen/Erfahrung	Authentizität	Flexibilität

Anmerkung: Die Reihenfolge der Angaben ist nicht im Sinne einer Wertigkeit zu interpretieren.

Der Enthusiasmus einer Moderatorin/eines Moderators wird von den Interviewteilnehmerinnen und -teilnehmern selbst nur an zwei Schulen (B, C) zu den Top-Eigenschaften gezählt. Eine zeitliche Flexibilität, aber auch die Flexibilität, auf die Rahmenbedingungen, Voraussetzungen und Bedarfe der Fortbildungsteilnehmerinnen und -teilnehmer einzugehen, wird hingegen ausnahmslos als eine Eigenschaft benannt, die für das Gelingen und den Erfolg einer Fortbildung von Bedeutung sind. Die zeitliche Flexibilität wird beispielsweise in einer Aussage aus dem Interview mit Schule B deutlich: *„Schulorganisatorisch vielleicht tatsächlich die zeitliche Flexibilität, um die Fortbildung eben über so'n langen Zeitraum auch in den Unterrichtsalltag einbauen zu können“* Als Hinweis auf die inhaltliche Flexibilität können folgende zwei

Interviewsequenzen angeführt werden, wobei die erste aus dem Interview mit Schule C entnommen wurde und die zweite aus dem Interview mit Schule E: *„Die Situation eines Kollegiums erkennen (...) und daraus dann halt flexibel handeln“* und *„Flexibilität echt auch wichtig (...) sonst wird das schwierig, wenn die da nicht abweichen können vom Plan.“*

Der Aspekt der Flexibilität wurde in den Interviews beispielsweise in den Zusammenhang mit dem Fachwissen, der Authentizität und den Systemerfahrungen gestellt, drei Merkmale, welche für die Interviewten eindeutig zu den wichtigen Eigenschaften einer Moderatorin/eines Moderators zählen. Die geforderte Systemerfahrung rührt dabei unter anderem aus der Einstellung, sich nichts von einer Person „sagen lassen“ zu wollen, die an einer anderen Schulform als Lehrkraft tätig ist, wie es sich im Interview mit Schule A ausdrückt: *„Die Bindung an unsere Schulform. Ich möchte mir nicht gerne von Gymnasialen sagen lassen, wie ich das in der Realschule machen muss.“* Im Gespräch mit Schule D wird die Notwendigkeit einer Passung der Schulform damit begründet, dass eventuellen Widerständen – die sich aus einer Einstellung wie derjenigen von der Lehrperson von Schule A ergeben könnten – auf diese Weise präventiv vorgebeugt werden könnte: *„Also es macht immer Sinn, jemanden zu nehmen, der aus dem System kommt. Das hat sich jetzt so hier nicht bewahrheitet und wir sind trotzdem prima mit klargeworden. Aber ich glaube, dass das so erste Widerstände, die entstehen (...) abfangen kann, wenn einfach klar ist, das ist einer, der arbeitet wie ich.“* Ferner werden Erfahrungen in dem gleichen System wie die fortgebildete Schule von den Interviewpartnerinnen und -partnern an Schule D als eine Voraussetzung für die fachliche Kompetenz einer Moderatorin/eines Moderators betrachtet: *„Also sie muss fachkompetent sein, sie muss wissen, wovon sie spricht. (...) Und das war eigentlich an manchen Stellen im ersten Schritt für Moderatorin A. schwierig, weil die [...] von einem anderen System kommt.“* Ähnlich wird die benötigte Systemerfahrung von der Schule C begründet. Aus dem Zitat wird deutlich, dass die Erfahrung in einer Schulform als Voraussetzung dafür betrachtet wird, dass sich eine Moderatorin/ein Moderator in angemessener Weise in die Situation und Rahmenbedingungen der Fortbildungsschule hineinversetzen kann.

[...] Systemerfahrung, weil das ist für mich was anderes, ob jemand als Moderator von 'ner Schulform klassischer Art kommt, von gegliederten Schulsystemen, Realschule, Gymnasium oder aus solchen Systemen wie „längeres gemeinsames Lernen“. [...] also Systemerfahrung, um überhaupt auch sich einzudenken [...]
(Schule C)

Sehr kritisch werden die Interviewpartnerinnen und -partner von Schule C an der Stelle, wo sie um eine Einschätzung darüber gebeten werden, ob das Moderatorenteam, das die Fortbildung an ihrer Schule durchgeführt hat, die von ihnen als besonders wichtig erachteten Moderatoreigenschaften besaßen: *„Ein ideales Moderatorenteam hätte das alles mitgebracht, unser Moderatorenteam hat (...) das alles nicht mitgebracht.“*

Losgelöst von den „drei“ eigentlich genannten lernförderlichen Eigenschaften einer Moderatorin/eines Moderators ist einer Lehrkraft wichtig, einen anderen Aspekt – die mangelnde Klarheit von Arbeitsaufträgen – anzusprechen: *„Ich weiß gar nicht, wo das jetzt hinpasst. Ich hatte am Anfang gedacht, das passt zu Moderation, deswegen bin ich mir gar nicht sicher. Was ich aber noch gerne loswerden wollte. Ich hab’, also das ist wirklich wieder ‘ne blöde Wertung, aber ich hab’ selten so unklare und schwammige Arbeitsanweisungen erlebt, wie da.“*

Allerdings betrachten die Interviewten von Schule C die beiden Personen des Moderatorenteams auch differenziert und haben Unterschiede wahrgenommen. So wird beispielsweise eine Moderatorin/ein Moderator als authentisch erlebt, während dies bei der anderen Moderatorin/dem anderen Moderator eher vermisst wurde.

Ich finde, da müsste man trennen zwischen den beiden Moderatoren. Also ich fand sie/ihn, fand ich authentisch. Bei ihr/ihm hatte ich manchmal so den Eindruck, dass sie/er in dieser Rolle der Moderatorin/des Moderators nicht immer authentisch sein konnte. Aber das ist nur ‘n Gefühl, ich kenne die ja nur jetzt aus dieser Umgebung, aber manchmal hat man so das Gefühl, ob das jetzt authentisch ist oder nicht, also das war jetzt mein persönliches Gefühl. (Schule C)

Abschließend sprechen sie dem Moderatorenteam zusätzlich allerdings ein Bemühen und die Intention zu, engagiert mit dem Kollegium arbeiten zu wollen.

L1: Das Bemühen darum mitgebracht.

S: Ja, hat, also, das muss man ihnen sagen. Sie war'[n]

L2: [Sie] haben auf jeden Fall versucht, engagiert mit uns zu arbeiten. (Schule C)

Die Aussage einer Lehrperson einer anderen Schule bringt ebenfalls zum Ausdruck, dass die Moderatorinnen und Moderatoren eines Teams teilweise unterschiedlich authentisch erlebt wurden: *„Also ich hatte halt teilweise wirklich das Gefühl, da steht, zumindest bei der Frau/dem Herrn A., ‘ne Schauspielerin/nen Schauspieler vor mir, die/der jetzt vorgibt, etwas zu sein, was sie/er einfach nicht war.“* Dieser unterschiedliche Eindruck zwischen den beiden Personen des Moderatorenteams setzt sich auch bei dem wahrgenommenen Enthusiasmus fort. Die/Der als authentisch erlebte Moderatorin/Moderator vermittelt den Eindruck, Freude an der Moderation zu haben, während sich bei der anderen Moderatorin/dem anderen Moderator eher gefragt wird, was der Beweggrund für die Moderationstätigkeit ist: *„Dass sie/er das gemacht hat, weil sie/er das gerne machen wollte (.) und bei der Frau/dem Herrn A., weiß ich nicht. [...] Dass sie/er das gemacht hat, um aus dem Unterricht rauszukommen [...]. Da hat man dann deutlich gemerkt, dass sie/er dann überhaupt keinen Bock mehr hatte eigentlich.“*

Darüber hinaus wird von der Schule B angemerkt, dass das Moderatorenteam nicht über ausreichend fachliche Kompetenz verfügte, um Fragen, die über die auf den Präsentationsfolien angegebenen Informationen hinausgehen, beantworten zu können.

Das klang ja jetzt eben auch schon 'n bisschen an, also die fachliche Kompetenz, die war halt wenig vorhanden. Also sie konnten das ablesen, was auf der Präsentation stand, was jeder von uns auch ablesen konnte. Teilweise haben sie's noch nicht mal in eigenen Worten formuliert, wo wir dann auch schon beim Präsentieren sind und sobald Nachfragen kamen, war halt (.) die Luft raus im wahrsten Sinne des Wortes. Da war'n die halt tatsächlich im wahrsten Sinne des Wortes überfragt. (Schule B)

Dennoch resümieren auch die Interviewpartnerinnen und -partner von Schule B positiv: „*Das Bemühen der Moderatoren war da.*“

Eine unzureichende fachliche Expertise des Moderatorenteam und/oder kaum eigene Erfahrungen in der Anwendung der Fortbildungsinhalte in der eigenen unterrichtlichen Praxis wurde auch von den Schulen A und G empfunden. Die Lehrpersonen an Schule A vermissten unter anderem eine Zielklarheit des Moderatorenteam, wobei sie die ungenügende Vorbereitung nicht in die Verantwortung der Moderatorinnen/Moderatoren selbst legten.

Durch diese schwammigen Antworten hatte man auch eher den Eindruck, ich muss da meiner Kollegin zustimmen, dass (.) den Moderatoren vielleicht teilweise auch noch nicht ganz klar war, was das Ziel dieser Fortbildung war, da sie die auch einfach noch nie durchgeführt hatten. (Schule A)

L1: Die hatten diese Fortbildung noch nie durchgeführt, ich glaube, diese Fortbildung war selber in sich auch noch nicht ganz schlüssig. Das war für die beiden, ich persönlich fand es unmöglich, wie man die dahin, es tut mir leid, wie man die dahin schicken konnte. Verstehe ich nicht.

L2: Das vermittelte etwas den Eindruck, dass sie nur zwei Seiten weiter waren als wir. Noch mal, nicht als böswillige Beschreibung des Verhaltens. (Schule A)

Im Interview mit der Schule G formulieren die Lehrpersonen gleichermaßen, dass das Moderatorenteam kein deutlich größeres Fachwissen und mehr Erfahrungen bezüglich der Fortbildungsinhalte mitgebracht hatte als die Lehrpersonen der eigenen Schule, sodass es ihm nicht möglich war, eine Verbindung zwischen dem fachlichen und/oder theoretischen Wissen und Beispielen aus der schulischen Praxis herzustellen. Infolge des fehlenden Wissens oder der fehlenden Erfahrung des Moderatorenteam nahmen die Lehrpersonen bei diesem eine starke

Orientierung an den vorgegebenen PowerPoint-Folien wahr, was seine Flexibilität eingeschränkt habe.

Ja, also ich hatte den Eindruck, die wissen das, die können das. Aber so wie wir im Prinzip. Und die stehen halt auch noch am Anfang und haben (.) die haben halt nicht so das, was wir uns eigentlich gewünscht haben. Zu profitieren, dass die wirklich diese Differenzierung vielleicht schon mal in sich gemacht haben, dass die uns Beispiele für so was geben. (Schule G)

L1: Es wurde nicht geschafft, das Fachwissen aufgrund der Erfahrungen flexibel hier miteinzubringen, wenn ich da mal so den Bogen schlage da. Also das fehlte so'n bisschen.

L2: Ja und die Flexibilität ist natürlich extrem eingeschränkt, wenn die sich da an ihren PowerPoint-Präsentationen festhalten, die sie von außerhalb bekommen. Da kann man dann natürlich auch nicht so flexibel sein, wie wir es vielleicht uns gewünscht hätten. (Schule G)

Bei Schule E entstand ebenfalls der Eindruck, dass die Moderatorinnen und Moderatoren keinen vollständigen Überblick über die Inhalte und den Prozess von *Vielfalt fördern* hatten, was als nachteilig beziehungsweise als „*nicht so ganz der Optimalfall*“ beschrieben wird, wenn es um die Flexibilität der Moderatorinnen und Moderatoren geht.

Also 'n Moderator muss auch wissen, wovon er spricht. Also er muss das Ganze im Blick haben und nicht nur so'n Teilwissen haben. Bis dahin weiß ich jetzt und kann aber auch nicht abweichen, weil ich nicht weiß, was noch kommt. Also ich kann nur in der Gesamtschau sagen, das kann ich jetzt weglassen oder verändern, dann verlier ich nichts Schlimmes; und das war da eben nicht gegeben. Ne, also weil die die Gesamtschau der Fortbildung nicht hatten. (Schule E)

Darüber hinaus wird es von den Interviewteilerinnen und -teilnehmern von Schule E als Problem für die Authentizität der Moderatorinnen und Moderatoren betrachtet, wenn diese nicht genügend eigene Erfahrungen mit der unterrichtlichen Anwendung der Fortbildungsinhalte sammeln konnten.

Authentizität, hab' ich da nicht, hab' ich vermisst. Weil, [...] die hatten ja auch selbst nicht, da muss man, glaub ich, um authentisch zu sein für diese Sache, muss man da eigentlich schon 'n bisschen eigene Erfahrungen haben und ich hatte das Gefühl, die haben das aus dem, was sie da theoretisch gelernt haben, weitergegeben und sie/er auch. Sie/Er hatte ja auch, das glaub ich, ich hatte nicht das Gefühl, dass die/der jetzt da selber da schon unheimlich viel gearbeitet hat, sondern dass sie/er

sich das angeeignet hat, um das vermitteln zu können, aber nicht selber vorher gelebt hat. (Schule E)

Obwohl in den Interviews, die mit den Schulen geführt wurden, durchaus kritische Einschätzungen dazu abgegeben wurden, inwiefern die Moderatorinnen und Moderatoren von *Vielfalt fördern* die als lernförderlich betrachteten Eigenschaften einer Moderatorin/eines Moderators mitbringen, beschreiben doch alle Schulen eine positive Wahrnehmung ihres Moderatorenteams.

In Schule A enthält das Interview eine Aussage, die das Bemühen des Moderatorenteams unterstreicht, Transparenz herzustellen und flexibel zu reagieren: *„Also das lag sicherlich auch an unseren Moderatoren, die wir dann hatten. Die haben das sehr transparent gemacht und haben uns eigentlich auch immer auf dem Laufenden gehalten und haben eben auch flexibel reagiert, wenn bei uns mal irgendwie was war. Also das war eigentlich in Ordnung.“* Des Weiteren wird in dem Interview mit Schule A das folgende globale Urteil gefällt: *„Ich glaube da haben wir wirklich Glück gehabt mit unserem Moderatorenteam.“*

Vorsichtiger, aber dennoch positiv anmutende Äußerungen findet man in dem Interview mit Schule F, durch die günstige Entwicklungen beim Moderatorenteam im Verlauf der Fortbildung deutlich werden: *„Ich denke mal, einige Aspekte haben sich entwickelt im Laufe des Projektes.“* Spezifiziert wird die positive Entwicklung des Moderatorenteams dann in Bezug auf die Orientierung an den Bedarfen der Lehrpersonen: *„Teilnehmerorientierung, das war von Anfang an nicht immer gegeben, weil die Moderatoren eben sehr stark an den Vorgaben gebunden waren. Wurde dann aber im Laufe des Prozesses ja immer besser, würde ich sagen.“*

Eine vergleichsweise positive Wahrnehmung ihrer Moderatorinnen/Moderatoren vermittelt das Interview mit Schule D: *„Und wir hatten auch eben ganz kompetente Moderatoren.“* Neben der allgemeinen Einschätzung, ein kompetentes Moderatorenteam gehabt zu haben, wird die spürbare Authentizität benannt: *„Also ich hab’ ihr/ihm das, was sie/er da vorne macht, immer abgenommen.“* Darüber hinaus betont eine Lehrperson das Fachwissen, die Flexibilität und einen Kontakt auf Augenhöhe, gibt aber gleichzeitig zu bedenken, dass andere Kolleginnen und Kollegen dies eventuell anders empfunden haben könnten.

Fachwissen hat sie/er aber in jedem Fall gehabt. Flexibel war sie/er. Ich hab’ sie/ihn als auf Augenhöhe erlebt. Nun ist aber meine Rolle auch eine andere als die der Kollegen. [...] Ich glaube, dass einzelne Kollegen sagen würden, nein. [...] Und auch vor allen Dingen auch Kollegen, die selber vielleicht in Fachleitungen sind oder so. Die sich dann eigentlich so’n bisschen (.) herabgewürdigt gefühlt haben, so wie sie da behandelt worden sind. (Schule D)

15 Resümee und Empfehlungen

Dieser Bericht fasst Ergebnisse des Projekts LIQUID zusammen. LIQUID evaluierte die Fortbildung *Vielfalt fördern*, indem Moderatorinnen und Moderatoren, Schulleiterinnen und Schulleiter, Lehrpersonen und deren Schülerinnen und Schüler befragt wurden. Hierbei wurden nicht nur Personen einbezogen, die an der Fortbildung *Vielfalt fördern* teilgenommen haben, sondern auch eine Kontrollgruppe. Diese setzt sich aus Schulleiterinnen und Schulleitern, Lehrpersonen und deren Schülerinnen und Schülern zusammen, die nicht an der Fortbildung *Vielfalt fördern* teilgenommen haben.

Im statistischen Globalvergleich zwischen der UG und KG deuteten sich, abgesehen von einem signifikant positiven Effekt bei der Deprivatisierungspraxis der Lehrpersonen der ersten Kohorte, keine Entwicklungsvorteile für die fortgebildeten Lehrpersonen, für die seitens der Schülerinnen und Schüler wahrgenommene Unterrichtsqualität, für die Entwicklung von Leistungen der Schülerinnen und Schüler in den Fächern Mathematik und Deutsch und für die Entwicklung im metakognitiven Wissen der Schülerinnen und Schüler an.

Da angenommen werden konnte, dass es sich bei der Gruppe der fortgebildeten Schulen und Lehrpersonen nicht um eine homogene, vergleichbar motivierte Gruppe handelte, wurden vertiefte Analysen zu differenziellen Effekten der Fortbildung durchgeführt. Im Mittelpunkt des vorliegenden Berichts stand daher die Frage, welche Faktoren und Bedingungen innerhalb der Gruppe der fortgebildeten Lehrpersonen für positive und negative Effekte verantwortlich sein könnten.

Zunächst wurden dabei die Voraussetzungen der teilnehmenden Lehrpersonen betrachtet. Die Schulen beschreiben in den Interviews eine hohe Motivation für die Teilnahme an *Vielfalt fördern*. Hierbei konnten Indikatoren für die drei Motivationsfacetten externale Erwartungsanpassung, Schulentwicklungsorientierung und soziale Interaktion rekonstruiert werden. Die Daten, die aus der schriftlichen Befragung der Fortbildungsteilnehmerinnen und -teilnehmer am Anfang von *Vielfalt fördern* stammen, sprechen ebenfalls dafür, dass die Gesamtstichprobe der UG-Lehrpersonen im Mittel eine eher günstige motivationale Ausgangslage für die Teilnahme an *Vielfalt fördern* aufweist. Mittels einer Profilanalyse konnten im zweiten Schritt vier disjunkte Gruppen von Lehrpersonen ermittelt werden, die sich hinsichtlich ihrer motivationalen Voraussetzungen, d. h. im vorliegenden Fall hinsichtlich ihrer Lernzielorientierung, ihrer individuellen Innovationsbereitschaft und hinsichtlich drei Facetten der Fortbildungsteilnahmemotivation (externale Erwartungsanpassung, Schulentwicklungsorientierung, soziale Interaktion), bedeutsam unterscheiden (s. Abschnitt 6.1).

Die Lehrpersonen von Cluster 1 besitzen im Vergleich mit den drei anderen Clustern die geringste Lernzielorientierung und Innovationsbereitschaft und sind im geringsten Maße zur Teilnahme an *Vielfalt fördern* motiviert, weil sie die Schule weiterentwickeln und/oder einen

intensiveren Kontakt mit Kolleginnen und Kollegen aufbauen wollen. Hingegen sind diese Lehrpersonen (Cluster 1) am stärksten durch eine Anpassung an externale Erwartungen motiviert. Cluster 2 kennzeichnet sich dadurch, dass die Lehrpersonen eher lernzielorientiert und offen gegenüber Innovationen sind, die die Schule und den Unterricht betreffen, die Lehrpersonen gleichzeitig aber auch eher stark durch den schulischen Mehrheitsentscheid und das Empfinden einer Pflicht zur Fortbildungsteilnahme motiviert sind. Beim Cluster 3 sind die Lehrpersonen lernzielorientiert und offen für Innovationen, die Schule und Unterricht betreffen. Die Lehrpersonen, die diesem Cluster (Cluster 3) zugeordnet sind, sind stark zur Teilnahme an *Vielfalt fördern* motiviert, weil die Schule als Ganzes davon profitieren kann, gleichzeitig ist aber die externale Erwartungsanpassung ebenfalls hoch ausgeprägt. Im Cluster 4 sind Lehrpersonen, die eine hohe Lernzielorientierung und Innovationsbereitschaft besitzen und die stark zur Teilnahme an *Vielfalt fördern* motiviert sind, weil sie auf diesem Wege die Schule weiterentwickeln können, während extrinsische Aspekte wie ein Pflichtgefühl und ein schulischer Mehrheitsentscheid keine bedeutsamen Beweggründe für die Fortbildungsteilnahme dieser Lehrpersonen bilden.

Was die Häufigkeit der vorzufindenden Profile anbelangt, sind die Cluster 2, 3 und 4 mit ca. 30 %–35 % in etwa gleich häufig vertreten. Cluster 1, welches sich aus Lehrpersonen mit deutlich ungünstigen motivationalen Voraussetzungen zusammensetzt, ist das kleinste Cluster und besteht aus ca. 6 % der Lehrpersonen der zweiten Kohorte.

Anschließende Untersuchungen führten zu dem Ergebnis, dass sich die Lehrpersonen der vier Profile teilweise signifikant in ihrer Nutzung und Bewertung des Fortbildungsangebots von *Vielfalt fördern* unterscheiden. So berichteten beispielsweise die Lehrpersonen, die dem Profil mit den günstigsten motivationalen Voraussetzungen zugeordnet wurden (Profil 4), eine intensivere Nutzung des Fortbildungsangebots, eine höhere Zufriedenheit und einen größeren Lernertrag als Lehrpersonen der anderen Profile. Die Effektstärken liegen zwischen $\eta^2 = .07$ und $\eta^2 = .19$ zugunsten der Lehrpersonen, die am Anfang ihrer Teilnahme an *Vielfalt fördern* über förderlichere motivationale Voraussetzungen verfügen.

Diese Ergebnisse sind insofern relevant, als dass sie zum einen die Bedeutung der individuellen Voraussetzungen der Fortbildungsteilnehmerinnen und -teilnehmer für den Fortbildungsprozess unterstreichen und zeigen, dass die motivationalen Voraussetzungen der Teilnehmerinnen und Teilnehmer die Wahrnehmung und Wirksamkeit der Fortbildung (z. B. Nutzungsverhalten, zeitliche Belastung, Zufriedenheit, selbstberichteter Lernertrag) bedeutsam beeinflussen. Zum anderen bestätigen die Befunde, dass es sich bei der Gruppe der fortgebildeten Lehrpersonen um eine motivational eher heterogene Gruppe von Lehrpersonen handelt.

In die gleiche Richtung weisen diejenigen Befunde, nach denen die Anpassungsquote, d. h. der Anteil an Lehrpersonen einer Schule, die auch oder vorwiegend deshalb an der Fortbildung teilnehmen, weil sich die Schule als Ganzes dafür entschieden hat, eine prädiktive Kraft für einige Erfolgskriterien hat (s. Abschnitt 6.3). Je höher die so bezeichnete Anpassungsquote ist, je höher also der Anteil jener Lehrpersonen an der jeweiligen Schule ist, die an der Fortbildung teilnehmen, weil sich die Schule als Ganzes dafür entschieden hat, desto ungünstiger fällt die Einschätzung und Bewertung der Fortbildung aus, und dies, obgleich zwischen der Messung der anfänglichen Motivation und den Erfolgskriterien teilweise bis zu 1,5–2 Jahre liegen. Es zeigte sich etwa, dass mit einer höheren Anpassungsquote innerhalb eines Kollegiums das Fortbildungsklima, die kooperative Erprobung durch die teilnehmenden Lehrpersonen sowie deren Zufriedenheit mit der Fortbildung und deren selbstberichteter Lernertrag ungünstiger bewertet werden. Das Ergebnis kann so interpretiert werden, dass sich ein höherer Anteil an extrinsisch motivierten Lehrpersonen innerhalb eines Kollegiums ungünstig auf den Fortbildungsprozess auswirkt. Damit bestätigen sich auch Annahmen der internationalen Forschung, wonach insbesondere bei langen und verpflichtenden Fortbildungsmaßnahmen mit erheblichen Effekten motivationaler Voraussetzungen von Lehrpersonen gerechnet werden muss (Kennedy, 2016).

Dieses Ergebnis wirft die Frage auf, wer die Adressaten einer Fortbildung sein sollten. Ist es für die Durchführung einer Fortbildung vorteilhafter, den Fortbildungsprozess mit einer stärker intrinsisch motivierten, aufgeschlossenen Teilgruppe eines Lehrerkollegiums zu beginnen, anstatt von Anfang an eine Einbindung des gesamten Kollegiums anzustreben und damit auch skeptische und eher extrinsisch motivierte Lehrpersonen in den Prozess einzubeziehen? Wenn mit einer Teilgruppe begonnen wird: Sollte es sich um eine Fachgruppe oder besser um ein Jahrgangsteam handeln?

Der Start einer Fortbildung mit einem gesamten Kollegium birgt andere Herausforderungen als der Start mit einer (motivierten) homogenen Teilgruppe eines Kollegiums. Hierzu gehört etwa die Frage nach dem Transfer und der Multiplikation der Fortbildungsinhalte in das Kollegium. Der Einbezug des gesamten Kollegiums verspricht hingegen, dass man den Großteil des Kollegiums zusammen anspricht und erreicht. Andererseits dürften Implementationswege in den eigenen Unterricht kürzer ausfallen und die Implementationsprozesse leichter gelingen, wenn man zunächst z. B. mit Fachkollegien oder mit einer motivierten Teilgruppe von Kolleginnen und Kollegen beginnt. In diesem Fall ergeben sich jedoch spätestens mit der Dissemination und Multiplikation der Prozesse in das erweiterte Kollegium neue Herausforderungen.

Zu den überraschenden Befunden der Evaluation gehört, dass die Einstellungen der Schulleiterin/des Schulleiters zu Unterrichtsbesuchen und zur Kooperation im Lehrerberuf Veränderungen in der Kooperation der Lehrpersonen nicht erklären können (s. Abschnitt 7.2). Das

bedeutet: Auch wenn Schulleitungen berichten, dass sie der Kooperation von Lehrkräften aufgeschlossen gegenüberstehen und es wichtig finden, dass eine Lehrkraft mit Kolleginnen und/oder Kollegen zusammenarbeitet, hat dies keine Zunahme der Kooperationsintensität aus Sicht der Lehrkräfte zur Folge.

Dazu passt auch, dass sich jene Schulen, in denen im Verlauf des Projekts bestimmte Formen der Zusammenarbeit (z. B. feste Arbeits- und Steuergruppen, kontinuierliche Doppelbesetzung in wöchentlich festgelegten Unterrichtsstunden) neu etabliert wurden, nicht auffällig in den am Anfang der Fortbildung erfassten Einstellungen ihrer Schulleitungen von jenen Schulen unterscheiden, in denen entsprechende Teamstrukturen schon bestanden. Gezeigt werden konnte aber, dass an fast allen Schulen der UG feste Teamarbeitsstrukturen eingerichtet wurden. Lediglich eine Schulleiterin/ein Schulleiter gibt keine Veränderung an, da alle erfassten Teamstrukturen bereits vor *Vielfalt fördern* an der Schule bestanden haben.

Entscheidend sind dagegen die kollektive Innovationsbereitschaft des Kollegiums und das Fortbildungsklima (s. Abschnitt 7.2.1): In einfachen linearen Regressionen zeigte sich für die am Anfang der Fortbildung im Kollegium erfasste Offenheit gegenüber innovativen schulischen Neuerungen und für ein freundliches und vertrauensvolles Klima während der Fortbildung ein positiver Einfluss auf die Kooperationshäufigkeit (Gespräche über konkrete Lernschwierigkeiten einzelner Schülerinnen und Schüler, gemeinsame Durchführung einer Unterrichtseinheit, Unterrichtshospitation) am Ende der Fortbildung. Die anfängliche kollektive Innovationsbereitschaft eines Kollegiums stellt eine schulische Rahmenbedingung dar, die sich einer Beeinflussbarkeit seitens der Moderatorinnen und Moderatoren entzieht. Hingegen ist das Fortbildungsklima aber ein Aspekt, der zu gewissen Teilen auch durch ein Moderatorenteam positiv mitgestaltet werden kann, sodass hierin eine Stellschraube zu sehen ist, über die ein Fortbildungsprozess in gewissem Maße positiv beeinflusst werden könnte. Der positive Effekt des Fortbildungsklimas stimmt mit bisherigen Erkenntnissen überein, die nahelegen, dass sich ein konstruktiver und offener Umgang mit anderen Ansichten und Herangehensweisen positiv auf die professionelle Weiterentwicklung von Lehrerteams auswirkt, da hierdurch kognitive Irritationen ausgelöst werden können, die zur Überprüfung, Erweiterung und/oder Anpassung der eigenen Ansichten führen können (Pröbstel, 2008).

Neben kontextuellen Bedingungen wurde anknüpfend an das Angebots-Nutzungs-Modell der Fortbildungsforschung untersucht, ob die individuelle Wahrnehmung der Fortbildung als Herausforderung und Bedrohung, die erlebte zeitliche Belastung aufgrund der Fortbildungsteilnahme, die Intensität der Nutzung von Lernaktivitäten der Fortbildungsmodule und die systematische (kognitive) Verarbeitung der Fortbildungsinhalte einen Einfluss auf affektiv-motivationale Variablen der Lehrpersonen und deren Transferabsichten am Ende der Fortbildung nehmen (s. Abschnitt 8). Tatsächlich konnten diese Merkmale teilweise Unterschiede in den

Zielvariablen (Motivation zur Diagnostik, Facetten der Selbstwirksamkeitserwartung, Transferabsichten) erklären. So erwies sich die selbstberichtete Nutzung des Lernangebots der Module in allen Modellen als ein signifikant positiver Prädiktor zur Vorhersage von Entwicklungen der Selbstwirksamkeitserwartung, die für Anforderungen in den Bereichen *lernbegleitende Diagnostik*, *Klassenführung* und *adaptive Unterrichtsgestaltung* untersucht wurde.

Die am Ende der Fortbildung geäußerten Transferabsichten lassen sich durch die Bewertung von *Vielfalt fördern* als eine interessante Herausforderung, durch die selbstberichtete systematische Verarbeitung der Fortbildungsinhalte und durch die selbstberichtete Intensität der Nutzung des Lernangebots positiv vorhersagen. Beispielsweise gilt daher: Je stärker die Fortbildung als positive Herausforderung wahrgenommen wird und je intensiver die Lehrpersonen die Inhalte systematisch verarbeiten, desto stärker fallen deren Absichten aus, die Fortbildungsinhalte für die eigene Praxis anzupassen und anzuwenden.

Auch die motivationalen Merkmale der Teilnehmerinnen und Teilnehmer lassen sich durch einzelne Wahrnehmungs- beziehungsweise Bewertungsmerkmale und die kognitive Verarbeitung oder Intensität der Nutzung erklären. So wird die Motivation dazu, die Lernstände und den Lernprozess der eigenen Schülerinnen und Schüler zu erfassen, von der systematischen kognitiven Auseinandersetzung mit den Fortbildungsinhalten und der Intensität der individuellen Nutzung der Lernaktivitäten positiv beeinflusst. Da die beiden Merkmale – systematische kognitive Auseinandersetzung mit den Fortbildungsinhalten und Intensität der individuellen Nutzung von Lernaktivitäten – stark miteinander zusammenhängen, wurden diese Merkmale in separaten Modellen untersucht. In beiden Modellen erwies sich zusätzlich das Threat Appraisal als ein bedeutsamer Prädiktor. Der Einfluss des Threat Appraisal ist negativ, d. h., je stärker die Lehrpersonen *Vielfalt fördern* als eine Bedrohung bewerten, desto geringer fällt die Entwicklung in der Motivation zur Diagnostik aus.

Das selbstberichtete Ausmaß, mit dem eine Lehrperson von den Lernangeboten von *Vielfalt fördern* Gebrauch macht, wirkt sich auch auf die Entwicklung der anderen motivationalen Merkmale, d. h. auf die Selbstwirksamkeitserwartungen im Bereich der lernbegleitenden Diagnostik, der Klassenführung und der adaptiven Unterrichtsgestaltung positiv aus. Das bedeutet, dass mit einer intensiveren individuellen Nutzung des Lernangebots von *Vielfalt fördern* eine positivere Entwicklung der Selbstwirksamkeitserwartungen verbunden ist. Für die systematische Verarbeitung der Fortbildungsinhalte wurde hingegen nur ein auf dem 10 %-Niveau signifikanter, positiver Einfluss auf die Entwicklung der Selbstwirksamkeitserwartung im Bereich der Klassenführung ermittelt. Das bedeutet, dass sich eine elaboriertere kognitive Auseinandersetzung mit den Fortbildungsinhalten tendenziell positiv darauf auswirkt, wie sich das Vertrauen einer Lehrperson entwickelt, einen störungsarmen Unterricht gestalten zu können.

Grundsätzlich ist zu beachten, dass insbesondere bei so langen Fortbildungsmaßnahmen wie *Vielfalt fördern* damit gerechnet werden muss, dass die Wahrnehmung und Nutzung der Fortbildungsinhalte durch die Teilnehmerinnen und Teilnehmer stärker von deren motivationalen Voraussetzungen beeinflusst werden, als dies bei kürzeren Maßnahmen der Fall ist (Kennedy, 2016). Mit anderen Worten: Je länger die Fortbildung dauert und je anstrengender und aufwendiger die Fortbildung für die teilnehmenden Lehrpersonen wird, desto größer dürfte die Wahrscheinlichkeit werden, dass Lehrpersonen mit ungünstigen motivationalen Voraussetzungen gedanklich aus der Fortbildung „aussteigen“. Aus dieser Perspektive könnte es für eine Weiterführung von *Vielfalt fördern* wichtig sein, die Motivationsförderung der Lehrpersonen im Auge zu behalten oder aber Lehrpersonen z. B. die Teilnahme an einzelnen Modulen zu ermöglichen, um die Fortbildungszeit zu verringern und die Motivation hoch zu halten.

Gleichzeitig sollten Moderatorinnen und Moderatoren sensibel für ein Unbehagen und für eventuelle Sorgen der Teilnehmerinnen und Teilnehmer sein, dass sie Inhalte und Ansprüche der Fortbildung überfordern. Ein Gefühl der Besorgnis und Überforderung aufgrund der Fortbildungsanforderungen dürfte sich vermeiden und/oder verringern lassen, indem nicht zu viele und nicht zu komplexe neue Themen und Konzepte auf einmal behandelt werden, sondern beispielsweise zunächst mit der Weiterentwicklung von sehr spezifischen und konkreten unterrichtlichen Verhaltensweisen begonnen wird, durch die es den teilnehmenden Lehrpersonen vergleichsweise schnell möglich ist, die Umsetzbarkeit der Fortbildungsinhalte und einen eigenen Erfolg zu erleben (Lipowsky & Rzejak, in Vorbereitung).

Im nächsten Schritt (s. Abschnitt 9) wurde der Einfluss der inhaltlichen Lerngelegenheiten untersucht. Hierfür wurden jeweils fünf Lerngelegenheiten aus dem Modul 2 beziehungsweise aus dem Modul 3–4 zu einem Summenscore zusammengefasst. Es konnte gezeigt werden, dass der Summenscore aus Lerngelegenheiten von Modul 2 keinen Einfluss auf die Entwicklung der Motivation zur Diagnostik besitzt. Hieraus sollte jedoch nicht der Schluss gezogen werden, dass die Behandlung dieser Fortbildungsinhalte bedeutungslos ist. Hierzu wären weitere Analysen erforderlich, die neben den Lerngelegenheiten selbst auch die Qualität der Behandlung dieser Lerngelegenheiten einbeziehen. Entsprechende Angaben zur Qualität konnten allerdings nicht erhoben werden.

Für die Lerngelegenheiten in den Modulen zum Lernen und Lehren ließ sich nachweisen, dass mit der ausführlicheren Behandlung von fünf ausgewählten Themen (selbstgesteuertes Lernen, Lern- und Arbeitstechniken von Schülern, Instrumente zur Lernprozessreflexion, Aufgabentypen zur Differenzierung, Formen der Lerndokumentation), die ebenfalls zu einem Summenscore zusammengefasst wurden, eine günstigere Entwicklung der Selbstwirksamkeitserwartung bezogen auf die adaptive Unterrichtsgestaltung einhergeht. Der Effekt ist jedoch nur schwach (s. Abschnitt 9.2.2).

Eine weitere Komponente im Angebots-Nutzungs-Modell, die die Wirkung einer Fortbildung beeinflussen und insofern Unterschiede in der Wirksamkeit der Fortbildung *Vielfalt fördern* bei den teilnehmenden Lehrpersonen erklären kann, stellt die Gruppe der Moderatorinnen und Moderatoren dar. Bisher liegen nur wenige Studien vor, die sich systematisch mit dieser Personengruppe und insbesondere mit dem Einfluss von Moderatoreneigenschaften auf den Fortbildungsprozess auseinandersetzen. Die Datenlage des Projekts *LIQUID* ermöglichte hier wichtige erste Untersuchungen und erbrachte einige interessante Ergebnisse. Ausgehend von Befunden, die sich in Studien mit Lehrpersonen und deren Schülerinnen und Schülern zeigten, wurde untersucht, ob die Begeisterung und der Enthusiasmus der Moderatorinnen und Moderatoren Einfluss auf die Wahrnehmungen der Lehrpersonen sowie auf die Verarbeitung und Umsetzung der Inhalte nehmen. Es zeigte sich, dass weniger der von den Moderatorinnen und Moderatoren selbst berichtete Enthusiasmus als vielmehr die von den Lehrpersonen wahrgenommene Begeisterung der Moderatorinnen und Moderatoren für einige der untersuchten Zielvariablen vorhersagerelevant war (s. Abschnitt 10). So berichteten die Lehrpersonen beispielsweise von einer systematischeren Verarbeitung der Fortbildungsinhalte, wenn sie das Moderatorenteam als begeisterter und motivierter erlebten. Auch die Intensität, mit der die Lehrkräfte die Fortbildungsinhalte zwischen den einzelnen Bausteinen mit Kolleginnen und Kollegen erproben, und das Gefühl, durch die Fortbildungsteilnahme etwas zu lernen und für den eigenen unterrichtlichen Alltag mitnehmen zu können, fallen umso stärker aus, je enthusiastischer die Moderatorinnen und Moderatoren von den Lehrpersonen wahrgenommen wurden.

Auch wenn man erwarten kann, dass eine Person, die sich für die Moderatorentätigkeit entscheidet, die Freude und die Begeisterung für diese Tätigkeit und für die Inhalte einer Fortbildung mitbringen sollte, deuten die Interviewdaten darauf hin, dass diese Eigenschaften nicht immer von den teilnehmenden Lehrpersonen in gleicher Weise wahrgenommen wurden. Die Interviews verdeutlichen aber nicht nur Unterschiede in affektiv-motivationalen Variablen der Moderatorinnen und Moderatoren, sondern machen auch deutlich, dass die Expertise der Moderatorinnen und Moderatoren in Teilen kritisch eingeschätzt wurde. In mehreren Interviews kommt zum Ausdruck, dass insbesondere die Fähigkeit der Moderatorinnen und Moderatoren vermisst wurde, die Bedeutsamkeit und Relevanz der Fortbildungsinhalte aufzuzeigen. Diese Fähigkeit, Fortbildungsinhalte bedeutsam zu machen und die Relevanz des Fortbildungsthemas anhand von Fällen und Beispielen nachvollziehbar zu veranschaulichen und erlebbar zu machen, wird in der Literatur als eine wichtige Qualifikation von Fortbildnerinnen und Fortbildnern bezeichnet (Timperley & Parr, 2009).

Bedeutungsvoll können die Inhalte einer Fortbildung etwa dadurch gemacht werden, dass die Erfahrungen und die Rahmenbedingungen der teilnehmenden Lehrpersonen aufgegriffen werden und dass die Lehrpersonen die Gelegenheit erhalten, Fortbildungsinhalte für eine Anwendung im eigenen schulischen Kontext anzupassen und weiterzuentwickeln.

In einigen Interviews, die im Anschluss an die Fortbildung geführt wurden, kam zum Ausdruck, dass bei den Lehrpersonen die Motivation im Verlauf des Fortbildungsprozesses teilweise nachließ, da das Angebot zu wenig Spielraum für eine Anpassung an die bestehenden Voraussetzungen und Bedarfe der Schulen ließ und da die Phasen, in denen die Lehrpersonen in Teams konkret an Inhalten arbeiten konnten, als zu kurz empfunden wurden. Gleichzeitig wird in den Interviews aber auch deutlich, dass eine Flexibilisierung und stärkere Teilnehmerorientierung mit dem Fortschreiten von *Vielfalt fördern* an einigen Schulen gelang, was als motivationsförderlich beschrieben wurde. In den quantitativen Daten, die aus den schriftlichen Befragungen stammen, zeigte sich zudem, dass der von den Lehrpersonen wahrgenommene Einbezug eigener schulischer und/oder unterrichtlicher Erfahrungen in die Fortbildungsveranstaltungen Unterschiede in der Motivation zur aktiven Mitarbeit während der Bausteine erklären kann und für die Nutzungsmotivation bedeutungsvoller ist als der Einbezug von Teilnehmererfahrungen aus der Perspektive der Moderatorinnen und Moderatoren (s. Abschnitt 11). Wenn die teilnehmenden Lehrpersonen den Eindruck haben, dass ihre Erfahrungen und Beispiele systematisch in die Veranstaltungen integriert werden, reduziert dies die extrinsische Teilnahmemotivation. Hingegen werden intrinsischere Motivationsqualitäten, wie beispielsweise die Bereitschaft zur Mitarbeit an der Weiterentwicklung der Schule oder der Wunsch, mit Kolleginnen und Kollegen zusammenzuarbeiten, dadurch gefördert, dass Erfahrungen der Lehrpersonen aufgegriffen werden.

Damit Moderatorinnen und Moderatoren flexibel auf die vorzufindenden Erfahrungen und Beispiele reagieren können, ist eine hohe Expertise und ein umfassendes Wissen über die Inhalte der Fortbildung und deren Relevanz für den schulischen Alltag erforderlich. Hierzu gehört auch, dass die Moderatorinnen und Moderatoren innerhalb des thematischen Rahmens flexibel auf die Bedürfnisse von Kollegien eingehen und die Bedeutung der Fortbildungsinhalte an multiplen Beispielen und Fällen illustrieren können.

Aus den Interviews der Vertiefungsstudie wird deutlich, dass einige Lehrpersonen eigene Erfahrungen der Moderatorinnen und Moderatoren mit der Umsetzung der Fortbildungsinhalte im eigenen schulischen Alltag und eine tiefergehende fachliche Expertise vermisst haben. Auch die Moderatorinnen und Moderatoren bewerteten die eigene Qualifizierung eher kritisch. Die Moderatorinnen und Moderatoren haben sich also nur bedingt auf ihre Aufgabe der Fortbildungsdurchführung vorbereitet gefühlt.

Dass der zeitliche Abstand zwischen der Qualifizierung der Moderatorinnen und Moderatoren und deren Leitung von *Vielfalt fördern* an den Schulen im Anschluss an die dritte Kohorte von einem halben Jahr auf ein ganzes Jahr erweitert wurde, erscheint vor diesem Hintergrund wichtig. Auf diese Weise erhalten die angehenden Moderatorinnen und Moderatoren mehr Zeit, um die Inhalte der Fortbildung im eigenen schulischen Alltag anzuwenden, Erfahrungen damit zu sammeln und diese dann in der Fortbildung authentisch an die teilnehmenden Lehrpersonen weiterzugeben.

Neben einer längeren Qualifizierungsphase für die Moderatorinnen und Moderatoren erscheint auch wichtig, dass wissenschaftliche Konzepte (z. B. Merkmale der Tiefenstruktur von Unterricht, wie kognitive Aktivierung, metakognitive Förderung, lernförderliches Feedback) mit Inhalten aus der Praxis verknüpft werden und die Moderatorinnen und Moderatoren das Handwerkszeug erhalten, um den teilnehmenden Lehrpersonen diese Verbindungen, z. B. anhand authentischer Fälle, Beispiele und Materialien, deutlich und nachvollziehbar machen zu können. Die vorhandenen Materialien und Unterlagen, die für den Fortbildungsprozess von *Vielfalt fördern* erarbeitet wurden, wurden von den Moderatorinnen und Moderatoren der zweiten Kohorte eher als zu theoretisch und abstrakt empfunden.

Die Analyse des Fortbildungsmaterials mit einem zusätzlichen Implementationscheck, durch den der Einsatz der Materialien und die Durchführung der Fortbildung an den einzelnen Schulen überprüft werden, könnte an dieser Stelle Optimierungsbedarfe aufzeigen und zur Begründung von Wirkungen und/oder deren Ausbleiben beitragen.

Aus didaktischer Sicht erscheint es lohnenswert, den teilnehmenden Lehrpersonen über die Moderatorinnen und Moderatoren Materialpakete für den eigenen Unterricht zur Verfügung zu stellen, mit denen a) Fortbildungsinhalte in ihrer Relevanz von den Lehrpersonen selbst erfahren werden können, b) die Lehrpersonen die Fortbildungsinhalte selbst erproben können, c) entlastende Funktionen, was die alltägliche Unterrichtsvorbereitung anbelangt, verbunden sind und d) die Lehrpersonen eher in die Rolle von *reflective practitioners* versetzt werden, d. h. zur Analyse und Reflexion des eigenen Unterrichts angeregt werden.

Insgesamt wird aus der Evaluation der Fortbildung *Vielfalt fördern* deutlich, dass für den Erfolg einer Maßnahme nicht allein die Entwicklung eines durchdachten Angebots notwendig ist. Um nachhaltige Entwicklungsprozesse von Lehrpersonen und Schulen anzuregen, bedarf es auch Moderatorinnen und Moderatoren, die frühzeitig und systematisch für die Durchführung eines solch komplexen Fortbildungsprogramms qualifiziert werden. Mehrfach erwähnt wurde von den teilnehmenden Lehrkräften die Bedeutung der Authentizität der Moderatorinnen und Moderatoren. Daher erscheint es wichtig, dass Moderatorinnen und Moderatoren entweder über umfassende Erfahrungen zu dem Fortbildungsthema verfügen oder aber weiterhin als Lehrpersonen tätig sind, um sich in der Fortbildung auch auf eigene Erfahrungen und Beispiele

stützen zu können. Darüber hinaus erscheint es zentral, dass die Gruppe der Durchführenden über umfassende diagnostische Kompetenzen und Empathie verfügt, um die Kognitionen, Einstellungen, Motivationen und/oder Emotionen der teilnehmenden Lehrpersonen im Professionalisierungsprozess wahrnehmen und berücksichtigen zu können. Angesichts der Heterogenität, welche nicht nur die Schülerschaft an Schulen kennzeichnet, sondern auch die an Schulen tätige Lehrerschaft, ist dies sicherlich keine triviale Aufgabe, sondern eine große Herausforderung für Moderatorinnen und Moderatoren.

Das Projekt *Vielfalt fördern* hat daher mit der systematischen Moderatorenqualifizierung eine wichtige und richtungsweisende Maßnahme unternommen. Auch wenn der Erfolg dieser Maßnahme bei den ersten beiden Kohorten noch Optimierungspotenzial zeigte, wurden bereits positive Konsequenzen derart gezogen, dass die Moderatorenqualifizierung inzwischen mit einem größeren zeitlichen Abstand zur eigentlichen Fortbildungsdurchführung erfolgt.

Die Bedeutung von Merkmalen der Moderatorinnen und Moderatoren für den Fortbildungsprozess und -erfolg sollte auch in zukünftigen Studien weiter untersucht werden. Der Umstand, dass die von den Moderatorinnen und Moderatoren berichteten Merkmale kaum Wirkungen auf die teilnehmenden Lehrkräfte zeigten, bedeutet nicht, dass sie unwichtig sind. So kann es sein, dass sie nicht direkt, sondern über weitere Prozessvariablen Wirkungen entfalten, die im Rahmen dieser Evaluation nicht erfasst werden konnten. Denkbar ist zudem, dass weniger der hier gebildete Mittelwert der Moderatorenangaben von Belang ist, sondern die Angaben jenes Moderators, jener Moderatorin des jeweiligen Moderatorenteams, der/die von den teilnehmenden Lehrkräften am überzeugendsten eingeschätzt wurde. Dies konnte im Rahmen dieser Evaluation jedoch nicht untersucht werden.

Eine weitere Empfehlung, welche aufgrund der Ergebnisse für *Vielfalt fördern* und zukünftige Fortbildungsmaßnahmen ausgesprochen werden kann, betrifft die Flexibilisierung von Arbeitsformen im Rahmen von Fortbildungen. Insbesondere die Interviewdaten führten vor Augen, dass für Lehrpersonen die Möglichkeit zur praktischen Erarbeitung von Inhalten einen hohen Stellenwert einnimmt und grundlegend für deren Motivation in der Fortbildung ist. Damit es zu einer unmittelbaren Anwendung von Fortbildungsinhalten in der schulischen Praxis von Lehrpersonen kommt, müssen die Lehrpersonen den Nutzen und Mehrwert der Inhalte nachvollziehen können. Vor diesem Hintergrund ist die vielfach geforderte Teilnehmerorientierung ernst zu nehmen, ohne dass dies eine willkürliche Ausrichtung und Anpassung des Fortbildungsangebots an die Wünsche der teilnehmenden Lehrpersonen nach sich ziehen beziehungsweise zu einem gänzlichen Verzicht auf die Thematisierung wissenschaftlich abgesicherter Sachverhalte führen sollte.

Neben Möglichkeiten zur stärkeren Teilnehmerorientierung sollte bei der Weiterentwicklung von *Vielfalt fördern* gegebenenfalls darüber nachgedacht werden, wie sich in dem Fortbildungskonzept ein stärkerer fachlicher Fokus umsetzen lässt. Viele Fortbildungen, die sich positiv auf das unterrichtliche Handeln von Lehrpersonen und auf das Lernen der Schülerinnen und Schüler auswirken, zeichnen sich durch einen vergleichsweise engen fachlichen Fokus aus (zsf. z. B. Darling-Hammond, Hylar & Gardner, 2017; Lipowsky & Rzejak, 2017). Eine mögliche Erklärung hierfür ist, dass Fortbildnerinnen und Fortbildner so inhaltlich stärker in die Tiefe gehen können und eher die Chance haben, die teilnehmenden Lehrpersonen zu einer genauen Analyse von Vorstellungen, Präkonzepten, Lösungsprozessen und Schwierigkeiten ihrer Schülerinnen und Schülern anzuregen, aus denen diese dann Konsequenzen für den Unterricht und die Förderung ihrer Schülerinnen und Schüler ableiten können. Daher könnte bei einer Weiterentwicklung von *Vielfalt fördern* auf umfassendere fachliche Lerngelegenheiten Wert gelegt werden, durch die diagnostische und fachdidaktische Kompetenzen der Fortbildungsteilnehmerinnen und -teilnehmer weiterentwickelt werden.

Um während der Präsenzveranstaltungen von *Vielfalt fördern* das gemeinsame Betrachten und die Gespräche über Unterricht zu strukturieren und zu fokussieren, aber auch, um die kollegiale Weiterentwicklung von Unterricht bei den teilnehmenden Lehrpersonen zu befördern, wird empfohlen, dass ein klares, geteiltes und wissenschaftlich abgesichertes Verständnis von lernwirksamem und motivationsförderlichem Unterricht zugrunde liegt, welches als Orientierungsrahmen für die Fortbildung fungiert. Eine entsprechende Orientierung an evidenzbasierten Merkmalen lernwirksamen Unterrichts ist ein weiteres Merkmal wirksamer Fortbildungen und hilft dabei, sich auf wesentliche und erfolgversprechende Aspekte des Lehrens und Lernens zu konzentrieren (z. B. Gore et al., 2017; Kyriakides & Antoniou, 2011, 2013). Diese Orientierung an evidenzbasierten Merkmalen wirksamen Unterrichts sollte damit einhergehen, dass die an der Fortbildung teilnehmenden Lehrpersonen unmittelbar erleben können, dass die intendierten Veränderungen auch in ihrem eigenen Unterricht funktionieren und „Früchte tragen“, denn dann ist die Bereitschaft größer, die Fortbildungsinhalte nachhaltig im eigenen Unterricht zu implementieren (Guskey, 1985; Timperley, Wilson, Barrar und Fung, 2007).

Literaturverzeichnis

- Antoniou, P., & Kyriakides, L. (2011). The impact of a dynamic approach to professional development on teacher instruction and student learning: Results from an experimental study. *School Effectiveness and School Improvement*, 22(3), 291–311.
- Antoniou, P., & Kyriakides, L. (2013). A Dynamic Integrated Approach to teacher professional development: Impact and sustainability of the effects on improving teacher behaviour and student outcomes. *Teaching and Teacher Education*, 29, 1–12.
- Artelt, C., & Neuenhaus, N. (2010). Metakognition und Leistung. In W. Bos, O. Köller & E. Klieme (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung* (S. 127–146). Münster: Waxmann.
- Bäuerlein, K., Lenhard, W., & Schneider, W. (2012). *Lesetestbatterie für die Klassenstufen 6–7. Verfahren zur Erfassung der basalen Lesekompetenz und des Textverständnisses. Manual*. Göttingen: Hogrefe.
- Behr, F. B. (2017). *Lernhabitus und Weiterbildung. Determinanten des Weiterbildungsverhaltens von Lehrerinnen und Lehrern*. Bad Heilbrunn: Klinkhardt.
- Bleck, V. (2018). *Lehrerenthusiasmus: Entwicklung, Determinanten, Wirkungen*. Wiesbaden: Springer.
- Bonsen, M., & Rolff, H.-G. (2006). Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern. *Zeitschrift für Pädagogik*, 52(2), 167–184.
- Bortz, J., & Weber, R. (2005). *Statistik für Human- und Sozialwissenschaftler* (6., vollst. überarb. und aktualisierte Aufl.). Berlin: Springer.
- Christ, O., & Schlüter, E. (2012). *Strukturgleichungsmodelle mit Mplus. Eine praktische Einführung*. München: Oldenbourg.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* Münster: Waxmann.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Palo Alto: Learning Policy Institute. Verfügbar unter <https://learningpolicyinstitute.org/product/teacher-prof-dev>
- Deci, E. L., & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–239.
- Decker, A.-T. (2015). *Veränderung berufsbezogener Überzeugungen bei Lehrkräften*. Dissertation, Johann Wolfgang Goethe-Universität. Frankfurt am Main.

- Dickhäuser, O., Butler, R., & Tönjes, B. (2007). Das zeigt doch nur, dass ich's nicht kann. Zielorientierung und Einstellung gegenüber Hilfe bei Lehramtsanwärtern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39(3), 120–127.
- Diedrich, M., Abs, H. J., & Klieme, E. (2004). *Evaluation im BLK-Modellprogramm Demokratie lernen und leben: Skalen zur Befragung von Schüler/-innen, Lehrer/-innen und Schulleitungen. Dokumentation der Erhebungsinstrumente*. Frankfurt am Main: DIPF 2004.
- Fröhlich, W. D. (2010). *Wörterbuch Psychologie* (27., Aufl.). München: dtv.
- Gebhardt, M., Schwab, S., Nusser, L., & Hessels, M. G. P. (2015). Einstellungen und Selbstwirksamkeit von Lehrerinnen und Lehrern zur schulischen Inklusion in Deutschland. Eine Analyse mit Daten des Nationalen Bildungspanels Deutschland (NEPS). *Empirische Pädagogik*, 29(2), 211–229.
- Geiser, C. (2010). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung*. Wiesbaden: VS.
- Göb, N. (2018). *Wirkungen von Lehrerfortbildung. Eine explorative Betrachtung von Fortbildungstypen und deren Effekte auf die Teilnehmenden am Beispiel des Pädagogischen Landesinstituts Rheinland-Pfalz*. Weinheim: Beltz Juventa.
- Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H., & Lubans, D. (2017). Effects of professional development on the quality of teaching: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education*, 68, 99–113.
- Gorozidis, G., & Papaioannou, A. G. (2014). Teachers' motivation to participate in training and to implement innovations. *Teaching and Teacher Education*, 39, 1–11.
- Götz, L., Lingel, K., & Schneider, W. (2013a). *Deutscher Mathematiktest für fünfte Klassen. DEMAT 5+*. Göttingen: Hogrefe.
- Götz, L., Lingel, K., & Schneider, W. (2013b). *Deutscher Mathematiktest für sechste Klassen. DEMAT 6+* (Hogrefe Schultests). Göttingen: Hogrefe.
- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review*, 15(2), 147–179.
- Guskey, T. P. (1985). Staff development and teacher change. *Educational Leadership*, 42(7), 57–60.
- Haehnel, C. (2011). *Emotionen bei Buying Center-Entscheidungen*. Wiesbaden: Gabler.

- Hoffmann, L., & Richter, D. (2016). Aspekte der Aus- und Fortbildung von Deutsch- und Englischlehrkräften im Ländervergleich. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend: Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (S. 481–501). Münster: Waxmann.
- Holtappels, H. G. (2013). Schulentwicklung und Lehrerkooperation. In H. G. Holtappels & N. McElvany (Hrsg.), *Empirische Bildungsforschung. Theorien, Methoden, Befunde und Perspektiven. Festschrift für Wilfried Bos* (S. 35–61). Münster: Waxmann. Verfügbar unter <http://www.ciando.com/ebook/bid-889702>
- Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (2nd ed.). New York: Routledge.
- Kao, C.-P., Wu, Y.-T., & Tsai, C.-C. (2011). Elementary school teachers' motivation toward web-based professional development, and the relationship with Internet self-efficacy and belief about web-based learning. *Teaching and Teacher Education*, 27(2), 406–415.
- Kennedy, M. M. (2016). How does professional development improve teaching?. *Review of Educational Research*, 86(4), 945–980.
- Kirkpatrick, D. L. (1994). *Evaluating training programs. The four levels*. San Francisco: Berrett-Koehler.
- Klug, J. (2011). *Modeling and training a new concept of teachers' diagnostic competence*. Dissertation, Technische Universität Darmstadt. Darmstadt.
- Köller, O., & Möller, J. (2010). Selbstwirksamkeit. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4., überarb. und erw. Aufl., S. 767–774). Weinheim: Beltz.
- Lingel, K. (2016). *Metakognitives Wissen Mathematik. Entwicklung und Zusammenhang mit der Mathematikleistung in der Sekundarstufe I*. Dissertation, Julius-Maximilians-Universität Würzburg. Würzburg.
- Lingel, K., Götz, L., Artelt, C., & Schneider, W. (2014). *MAESTRA 5–6+ – Mathematisches Strategiewissen für fünfte und sechste Klassen: PSYNDEX Tests Info*.
- Lipowsky, F. (2014). Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und -weiterbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. Aufl., S. 511–541). Münster: Waxmann.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (2., vollst. überarb. und aktualisierte Aufl., S. 69–105). Berlin: Springer.
- Lipowsky, F., & Rzejak, D. (2015). Key features of effective professional development programmes for teachers. *Ricercazione*, 7(2), 27–51.

- Lipowsky, F., & Rzejak, D. (2017). Fortbildungen für Lehrkräfte wirksam gestalten – Erfolgversprechende Wege und Konzepte aus Sicht der empirischen Bildungsforschung. *Bildung und Erziehung*, 70(4), 379–399.
- Lipowsky, F., & Rzejak, D. (in Vorbereitung). Empirische Befunde zur Wirksamkeit von Lehrerfortbildungen – Ein Update. In W. Böttcher, K. Heinemann, C. Kubina & B. Priebe (Hrsg.), *Qualitätsentwicklung und Wirksamkeit der Lehrkräftefortbildung im Kontext des staatlichen schulischen Unterstützungssystems und der Lehrerinnen- und Lehrerbildung*. Seelze: Friedrich.
- Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments. How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, 34(2), 120–131.
- Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1(3), 85–91.
- Ministerium für Schule und Unterricht, & Bertelsmann Stiftung. (2014). *Übersicht der Projektregionen*. Zugriff am 26.06.2014. Verfügbar unter <http://www.lehrerfortbildung.schulministerium.nrw.de/vielfaltfoerdern/regionen.asp>
- Müncher, A., & Sturm-Schubert, T. (2015). *Leitfaden zu den Inhalten der Fortbildungsmodule*. Verfügbar unter http://www.lehrerfortbildung.schulministerium.nrw.de/Fortbildung/Vielfalt-f%C3%B6rdern-NRW/pdf/Brosch%C3%BCre_Inhalte_Vf_150707.pdf
- Neuenhaus, N. (2011). *Metakognition und Leistung. Eine Längsschnittuntersuchung in den Bereichen Lesen und Englisch bei Schülerinnen und Schülern der fünften und sechsten Jahrgangsstufe*. Dissertation, Otto-Friedrich-Universität. Bamberg.
- Nitsche, S., Praetorius, A.-K., Janke, S., Drexler, K., Fasching, M. S., Dresel, M. et al. (2017). Berufliche Zielorientierungen von Lehrkräften. In C. Gräsel & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals* (S. 17–35). Wiesbaden: Springer.
- Pröbstel, C. H. (2008). *Lehrerkooperation und die Umsetzung von Innovationen. Eine Analyse der Zusammenarbeit von Lehrkräften aus Perspektive der Bildungsforschung und der Arbeits- und Organisationspsychologie*. Berlin: Logos.
- Richter, A., Diehl, T., Krüger, J., & Vigerske, S. (2011). Einflussfaktoren auf die Fort- und Weiterbildung von Lehrkräften – Erste Ergebnisse eines Forschungsprojekts. *bwp@ Berufs- und Wirtschaftspädagogik – online*, 1–21. Verfügbar unter www.bwpat.de/ausgabe19/diehl_etal_bwpat19.pdf

- Richter, D., Engelbert, M., Weirich, S., & Pant, H. A. (2013). Differentielle Teilnahme an Lehrerfortbildungen und deren Zusammenhang mit professionsbezogenen Merkmalen von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 27(3), 193–207.
- Richter, D., Kuhl, P., Reimers, H., & Pant, H. A. (2012). Aspekte der Aus- und Fortbildung von Lehrkräften in der Primarstufe. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 237–250). Stuttgart: Kohlhammer.
- Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). Rochester: The University of Rochester Press.
- Rzejak, D., Künsting, J., Lipowsky, F., Fischer, E., Dezhgahi, U., & Reichardt, A. (2014). Facetten der Lehrerfortbildungsmotivation – eine faktorenanalytische Betrachtung. *Journal for Educational Research Online*, 6(1), 139–159.
- Rzejak, D., Lipowsky, F., & Künsting, J. (2013). Lehrerinnen und Lehrer als Lernende – Welche Merkmale beeinflussen den selbstberichteten Lernertrag von Lehrpersonen in Fortbildungsmaßnahmen? *Erziehung und Unterricht*, 163(1–2), 90–99.
- Schaarschmidt, U. (2001). Lehrerbelastung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (2., überarb. und erw. Aufl., S. 373–381). Weinheim: Beltz.
- Schulministerium Nordrhein-Westfalen. (2016). Schulgesetz für das Land Nordrhein-Westfalen. Schulgesetz NRW – SchulG. Verfügbar unter <https://www.schulministerium.nrw.de/docs/Recht/Schulrecht/Schulgesetz/Schulgesetz.pdf>
- Schwarzer, R., & Warner, L. M. (2014). Forschung zur Selbstwirksamkeit bei Lehrerinnen und Lehrern. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. Aufl., S. 662–678). Münster: Waxmann.
- Timperley, H. S., & Parr, J. M. (2009). Chain of influence from policy to practice in the New Zealand literacy strategy. *Research Papers in Education*, 24(2), 135–154.
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration (BES)*. Wellington.
- Wischer, B. (2007). Wie sollen LehrerInnen mit Heterogenität umgehen? Über „programmatische Fallen“ im aktuellen Reformdiskurs. *Die deutsche Schule*, 99(4), 422–433.
- Zurbriggen, C. (2016). *Schulklasseneffekte. Schülerinnen und Schüler zwischen komparativen und normativen Einflüssen*. Wiesbaden: Springer.

Abbildungsverzeichnis

Abbildung 1. Vereinfachtes Angebots-Nutzungs-Modell, angelehnt an Lipowsky (2014)	10
Abbildung 2. Vier-Ebenen-Evaluationsmodell nach Kirkpatrick (1994)	11
Abbildung 3. Modell zum Fortbildungsprozess von Lehrpersonen nach Guskey (1985)	12
Abbildung 4. Inhalte der Module 1–4 (Müncher & Sturm-Schubert, 2015).....	16
Abbildung 5. Prä-Prozess-Posterhebungsstruktur am Beispiel von Kohorte 1.	21
Abbildung 6. Erhebungsdesign bei Kohorte 1 (K1) und Kohorte 2 (K2).....	25
Abbildung 7. Anordnung von Items und Antwortkategorien am Beispiel der Berufszufriedenheit von Lehrpersonen.....	26
Abbildung 8. Beispielabbildung für eine einfache lineare Regression.....	31
Abbildung 9. Beispielabbildung für eine multiple lineare Regression.....	31
Abbildung 10. Beispielabbildung für ein Mediatormodell.....	32
Abbildung 11. Beispiel einer hierarchischen Datenstruktur.....	33
Abbildung 12. Beispiel für ein Mehrebenenmodell mit Prädiktoren auf Ebene 1 und Ebene 2	33
Abbildung 13. Anzahl der Schule nach Schulformen am Anfang der Projektlaufzeit von Kohorte 1	37
Abbildung 14. Anzahl der Schule nach Schulformen am Anfang der Projektlaufzeit von Kohorte 2.....	39
Abbildung 15. Prozentualer Anteil der Schulformen in der Haupt- und Vertiefungsstudie bei Kohorte 2	44
Abbildung 16. Grafische Darstellung der Lösung mit vier Voraussetzungsprofilen.....	56
Abbildung 17. Mittelwerte und Standardabweichungen für die Intensität der Nutzung individueller Lernaktivitäten bei den vier Profilen	60
Abbildung 18. Mittelwerte und Standardabweichungen für die Intensität der kooperativen Umsetzung bei den vier Profilen.....	61
Abbildung 19. Mittelwerte und Standardabweichungen für die Umsetzung der Fortbildungsinhalte bei den vier Profilen	61
Abbildung 20. Mittelwerte und Standardabweichungen für die zeitliche Belastung bei den vier Profilen	62

Abbildung 21. Mittelwerte und Standardabweichungen für die Zufriedenheit bei den vier Profilen	63
Abbildung 22. Mittelwerte und Standardabweichungen für den selbstberichteten Lernertrag bei den vier Profilen.....	63
Abbildung 23. Vorhersage des Fortbildungsklimas durch die Anpassungsquote.	66
Abbildung 24. Vorhersage der kooperativen Erprobung zu t_3 durch die Anpassungsquote...67	
Abbildung 25. Vorhersage der kooperativen Erprobung zu t_6 durch die Anpassungsquote...68	
Abbildung 26. Vorhersage der Zufriedenheit zu t_4 durch die Anpassungsquote.....	68
Abbildung 27. Vorhersage der Zufriedenheit zu t_7 durch die Anpassungsquote.....	69
Abbildung 28. Vorhersage des selbstberichteten Lernertrags zu t_4 durch die Anpassungsquote.....	69
Abbildung 29. Vorhersage des selbstberichteten Lernertrags zu t_7 durch die Anpassungsquote.....	70
Abbildung 30. Mittelwerte und Standardabweichungen für die Skala Fokus auf das Lernen der Schülerinnen und Schüler zu t_2 (Ausgangswerte).....	75
Abbildung 31. Mittelwerte und Standardabweichungen für das Fortbildungsklima zwischen den Lehrpersonen in Modul 2 (t_3).	76
Abbildung 32. Mittelwerte und Standardabweichungen für die kooperative Umsetzung während Modul 2 (t_3).	76
Abbildung 33: Mittelwerte und Standardabweichungen für die wahrgenommene Moderatorenkompetenz in Modul 2 (t_3).	77
Abbildung 34. Mittelwerte und Standardabweichungen für die Deprivatisierung zu t_2 (Ausgangswerte).	78
Abbildung 35. Mittelwerte und Standardabweichungen für das Fortbildungsklima zwischen den Lehrpersonen in Modul 2 (t_3).	78
Abbildung 36. Mittelwerte und Standardabweichungen für die kooperative Umsetzung während Modul 2 (t_3).	79
Abbildung 37. Mittelwerte und Standardabweichungen für die wahrgenommene Moderatorenkompetenz in Modul 2 (t_3).	80
Abbildung 38. Mittelwerte und Standardabweichungen für die kollektive Selbstwirksamkeitserwartung zu t_3	81

Abbildung 39. Mittelwerte und Standardabweichungen für das Fortbildungsklima zwischen den Lehrpersonen in Modul 2(t ₃).	81
Abbildung 40. Mittelwerte und Standardabweichungen für die kooperative Umsetzung während Modul 2 (t ₃).	82
Abbildung 41. Mittelwerte und Standardabweichungen für die wahrgenommene Moderatorenkompetenz in Modul 2 (t ₃).	83
Abbildung 42. Vorhersage der Häufigkeit von Gesprächen über konkrete Lernschwierigkeiten einzelner Schülerinnen und Schüler durch die kollektive Innovationsbereitschaft des Kollegiums.	88
Abbildung 43. Vorhersage der Häufigkeit einer gemeinsamen Durchführung einer Unterrichtseinheit durch die kollektive Innovationsbereitschaft des Kollegiums.	88
Abbildung 44. Vorhersage der Häufigkeit von Unterrichtsbesuchen von einer Kollegin/einem Kollegen im eigenen Unterricht durch die kollektive Innovationsbereitschaft des Kollegiums.	89
Abbildung 45. Vorhersage der Häufigkeit von Unterrichtsbesuchen bei einer Kollegin/einem Kollegen durch die kollektive Innovationsbereitschaft des Kollegiums.....	89
Abbildung 46. Vorhersage der Häufigkeit von Gesprächen zur Verbesserung von Unterricht durch das Klima während der Fortbildung.....	90
Abbildung 47. Vorhersage der Häufigkeit einer gemeinsamen Durchführung einer Unterrichtseinheit durch das Klima während der Fortbildung.....	91
Abbildung 48. Vorhersage der Häufigkeit von Unterrichtsbesuchen von einer Kollegin/einem Kollegen im eigenen Unterricht durch das Klima während der Fortbildung.....	91
Abbildung 49. Vorhersage der Häufigkeit von Unterrichtsbesuchen bei einer Kollegin/einem Kollegen durch das Klima während der Fortbildung.....	92
Abbildung 50. Einrichtung fester Arbeits- oder Steuergruppen	93
Abbildung 51. Einrichtung einer gemeinsamen Klassenführung durch zwei Lehrpersonen, denen feste Teamzeiten zur Verfügung stehen.....	94
Abbildung 52. Einrichtung von Jahrgangsteams	94
Abbildung 53. Einrichtung einer kontinuierlichen Doppelbesetzung in einzelnen wöchentlich festgelegten Unterrichtsstunden.....	95
Abbildung 54. Einrichtung von Teams, die gegenseitig Unterrichtshospitation praktizieren ..	96

Abbildung 55. Anzahl an eingerichteten Teamstrukturen pro Schule	96
Abbildung 56. Vorhersage der Entwicklung der Motivation zur Diagnostik durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung.	103
Abbildung 57. Vorhersage der Entwicklung der Motivation zur Diagnostik durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung.....	104
Abbildung 58. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich der lernbegleitenden Diagnostik durch Appraisals, die zeitliche Belastung und die subjektive Verarbeitung.	105
Abbildung 59. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich der lernbegleitenden Diagnostik durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung.....	106
Abbildung 60. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich Klassenführung durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung.	107
Abbildung 61. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich Klassenführung durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung.	108
Abbildung 62. Vorhersage der Selbstwirksamkeitserwartung im Bereich adaptive Unterrichtsgestaltung durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung.	109
Abbildung 63. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich adaptive Unterrichtsgestaltung durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung. ...	110
Abbildung 64. Vorhersage der Transferabsichten durch Appraisals, die erlebte zeitliche Belastung und die subjektive Verarbeitung.	112
Abbildung 65. Vorhersage der Transferabsichten durch die individuelle Nutzung von Lernaktivitäten, die Bewertung als Bedrohung und die zeitliche Belastung. ...	113
Abbildung 66. Vorhersage der Entwicklung der Motivation zur Diagnostik durch die Intensität der Thematisierung von Lerngelegenheiten aus den Modul Diagnostik.....	120
Abbildung 67. Vorhersage der Entwicklung der Selbstwirksamkeitserwartung im Bereich der adaptiven Unterrichtsgestaltung durch die Intensität der Thematisierung von Lerngelegenheiten aus den Modulen zum Lernen und Lehren.	120

Abbildung 68. Durchschnittlich wahrgenommene Moderatorenkompetenz eines Moderatorenteams an unterschiedlichen Schulen	122
Abbildung 69. Vorhersage der systematischen Verarbeitung der Fortbildungsinhalte durch den Enthusiasmus des Moderatorenteams für das Fortbilden aus Moderatorenperspektive.....	123
Abbildung 70. Vorhersage der systematischen Verarbeitung der Fortbildungsinhalte durch den Enthusiasmus des Moderatorenteams für das Thema Diagnostik aus Moderatorenperspektive.....	123
Abbildung 71. Vorhersage der systematischen Verarbeitung der Fortbildungsinhalte durch den Enthusiasmus des Moderatorenteams aus Lehrerperspektive.	124
Abbildung 72. Vorhersage der kooperativen Umsetzung von Fortbildungsinhalten zwischen den Bausteinen durch den Enthusiasmus des Moderatorenteams für das Fortbilden aus Moderatorenperspektive.....	124
Abbildung 73. Vorhersage der kooperativen Umsetzung von Fortbildungsinhalten zwischen den Bausteinen durch den Enthusiasmus des Moderatorenteams für das Thema Diagnostik aus Moderatorenperspektive.	125
Abbildung 74. Vorhersage der kooperativen Umsetzung von Fortbildungsinhalten zwischen den Bausteinen durch den Enthusiasmus des Moderatorenteams aus Lehrerperspektive.....	125
Abbildung 75. Vorhersage des selbstberichteten Lernertrags durch den Enthusiasmus des Moderatorenteams für das Fortbilden aus Moderatorenperspektive.....	125
Abbildung 76. Vorhersage des selbstberichteten Lernertrags durch den Enthusiasmus des Moderatorenteams für das Thema Diagnostik aus Moderatorenperspektive..	126
Abbildung 77. Vorhersage des selbstberichteten Lernertrags durch den Enthusiasmus des Moderatorenteams aus Lehrer- und Moderatorenperspektive.	126
Abbildung 78. Mediatormodell mit der Teilnahmemotivation externale Erwartungsanpassung als Prädiktor, dem Enthusiasmus des Moderatorenteams aus Lehrerperspektive als Mediator und der Nutzungsmotivation externale Erwartungsanpassung als Kriterium.	129
Abbildung 79. Mediatormodell mit der Teilnahmemotivation Schulentwicklungsorientierung als Prädiktor, dem Enthusiasmus des Moderatorenteams aus Lehrerperspektive als Mediator und der Nutzungsmotivation Schulentwicklungsorientierung als Kriterium.	130

Abbildung 80. Mediatormodell mit der Teilnahmemotivation soziale Interaktion als Prädiktor, dem Enthusiasmus des Moderatorenteams aus Lehrerperspektive als Mediator und der Nutzungsmotivation soziale Interaktion als Kriterium.....	131
Abbildung 81. Vorhersage der Fortbildungsnutzungsmotivationsfacette externale Erwartungsanpassung durch den Einbezug von Teilnehmererfahrungen aus Moderatorenperspektive.....	133
Abbildung 82. Vorhersage der Fortbildungsnutzungsmotivationsfacette externale Erwartungsanpassung durch den Einbezug von Teilnehmererfahrungen aus Lehrerperspektive.....	134
Abbildung 83. Vorhersage der Fortbildungsnutzungsmotivationsfacette Schulentwicklungsorientierung durch den Einbezug von Teilnehmererfahrungen aus Moderatorenperspektive.....	135
Abbildung 84 Vorhersage der Fortbildungsnutzungsmotivationsfacette Schulentwicklungsorientierung durch den Einbezug von Teilnehmererfahrungen aus Lehrerperspektive.....	135
Abbildung 85. Vorhersage der Fortbildungsnutzungsmotivationsfacette soziale Interaktion durch den Einbezug von Teilnehmererfahrungen aus Moderatorenperspektive.	136
Abbildung 86. Vorhersage der Fortbildungsnutzungsmotivationsfacette soziale Interaktion durch den Einbezug von Teilnehmererfahrungen aus Lehrerperspektive.....	136

Tabellenverzeichnis

Tabelle 1. Ausgewählte Schritte im Rahmen der qualitativ-strukturierten Inhaltsanalyse	35
Tabelle 2. Stichprobenumfang bei Kohorte 1	38
Tabelle 3. Stichprobenumfang zu den verschiedenen Erhebungszeitpunkten bei Kohorte 2	40
Tabelle 4: Stichprobe der Unterrichtsstudie bei Kohorte 2	42
Tabelle 5. Korrelation zwischen den Voraussetzungsmerkmalen für die Profilanalyse	55
Tabelle 6. Kennwerte der Profillösungen	55
Tabelle 7. Skalenkennwerte und Beispielitem der abhängigen Variablen	59
Tabelle 8. Korrelation zwischen der Anpassungsquote und anderen Merkmalen der schulischen Ausgangslage.....	65
Tabelle 9. Schulquote für die Beteiligung aufgrund des Mehrheitsentscheids.....	65
Tabelle 10. Schulische Differenzwerte für den Fokus auf das Lernen der Schülerinnen und Schüler, die Deprivatisierung und die kollektive Selbstwirksamkeitserwartung des Kollegiums	73
Tabelle 11. Mittelwert, Standardabweichung und ICC-Werte für Kooperationsaktivitäten	85
Tabelle 12. Korrelation der Prädiktorvariablen auf Individualebene	102
Tabelle 13. Mittelwerte, Standardabweichungen, Differenzen und Korrelationen für die Lerngelegenheiten aus Sicht der Lehrpersonen und Moderatorinnen und Moderatoren	117
Tabelle 14. Die drei relevantesten Themen/Inhalte von <i>Vielfalt fördern</i> aus Sicht der Vertiefungsschulen	154
Tabelle 15. Lernförderliche Eigenschaften einer Moderatorin/eines Moderators aus Sicht der Vertiefungsschulen	160