

Institut Forschung und Entwicklung
Zentrum Lernen und Sozialisation

www.fhnw.ch/ph/zls

Schlussbericht: Evaluation der SCALA-Intervention

Markus P. Neuenschwander & Vanessa Rust

Eingereicht beim Volksschulamt, Bildungsdirektion Kanton Zürich

15. November 2024

Vorwort

Im Jahr 2015 wurde die Interventionsstudie SCALA lanciert mit dem Ziel, unfaire Erwartungen und Zuschreibungen von Lehrpersonen zu reduzieren, so dass die Schule chancengerechter die Schüler*innen fördert und beurteilt. Diese Intervention zeigte in einer randomisierten Kontrollgruppenstudie eine nachweisbare Wirkung (Neuenschwander et al., 2021). Es war möglich, dass durch eine geschickt durchgeführte Intervention Vorurteile von Lehrpersonen in der Regelschule reduziert und faire Beurteilungen gefördert werden können. Daher erteilte das Volksschulamt des Kantons Zürich dem ersten Autor der vorliegenden Studie den Auftrag, diese Intervention an QUIMS-Schulen des Kantons Zürich als schulinterne Weiterbildung durchzuführen und mit neuen Themen bzw. Modulen zu erweitern. Dieses Angebot erhielt grossen Zuspruch und wurde bis heute von über 50 Schulen gebucht und durchgeführt. Es wurde überdies auch in anderen Kantonen der Deutschschweiz durchgeführt, aber auch von Schulen aus dem benachbarten Ausland (Deutschland, Österreich) nachgefragt.

Allerdings kam die Frage auf, ob diese modifizierte Intervention die erhofften Wirkungen auch in der Umsetzungsphase zeigte. Daher gab das Volksschulamt Zürich eine kleine Evaluationsstudie in Auftrag, deren Schlussbericht hier vorliegt. Die Studie sollte prüfen, ob die Intervention in der modifizierten Variante und in der Phase der breiten Anwendung (scaling-up) die gleichen Wirkungen zeigt, die im ursprünglichen Projekt gefunden wurden. Der vorliegende Bericht beschreibt Idee, Methoden und Ergebnisse dieser Wirkungsstudie.

An dieser Stelle danken wir dem Volksschulamt des Kantons Zürich für das Vertrauen in unsere Arbeit und für den Auftrag. Daneben danken wir ganz herzlich allen Schulleitungen, Lehrpersonen und Eltern, welche an dieser Studie mitgewirkt haben. Obwohl der Schulalltag sehr viele Verpflichtungen mit sich bringt, haben sie sich die Zeit genommen, unsere Fragen zu beantworten. Nur dank diesen Daten können wir gesicherte Aussagen zur Wirksamkeit unserer Intervention machen und diese weiterentwickeln. Diese Rückmeldungen von den Lehrpersonen sind für den SCALA-Ansatz, aber auch für die Weiterentwicklung von QUIMS sehr wertvoll.

Zudem danken wir Dr. Sog Yee Mok für die konzeptuelle und operative Leitung der Studie in der Startphase der 1. Kohorte. Gobiga Ramachandran und Lea Ruf haben uns tatkräftig bei der Durchführung der Studie unterstützt, z.B. beim Verschicken von Briefen, Organisieren und Kodieren von Daten, u.a. Sie haben uns bei der operativen Durchführung des Projekts wesentlich geholfen.

Zudem danken wir herzlich der wissenschaftlichen Begleitgruppe dieses Projekts bestehend aus Dr. Sybille Bayard, Dr. Lena Hollenstein, Prof. Dr. Katharina Maag Merki und Jürg Schoch für die hilfreiche Beratung bei der Durchführung der Studie. Die Diskussion in dieser Gruppe trug bei, die Qualität der Studie zu verbessern.

Wir hoffen, dass wir mit diesem Bericht Einsichten in die Wirkungsweise der SCALA-Intervention geben können und werden versuchen, auf dieser Basis unsere Arbeit laufend zu verbessern.

Brugg, im November 2024 Markus Neuenschwander & Vanessa Rust

Zusammenfassung

Im Jahr 2015 wurde die Interventionsstudie SCALA lanciert mit dem Ziel, unfaire Erwartungen und Zuschreibungen von Lehrpersonen zu reduzieren, so dass die Schulen die Schüler*innen chancengerechter fördern und beurteilten. Diese Intervention zeigte in einer randomisierten Kontrollgruppenstudie eine nachweisbare Wirkung (Neuenschwander et al., 2021). Es war möglich, dass durch eine geschickt durchgeführte Intervention Vorurteile von Lehrpersonen in der Regelschule reduziert und faire Beurteilungen gefördert werden können. In Konsequenz sollte die Chancengerechtigkeit an Schulen zunehmen.

Daher erteilte das Volksschulamt des Kantons Zürich den Auftrag, diese Intervention an QUIMS-Schulen des Kantons Zürich als schulinterne Weiterbildung (Schilw) durchzuführen und mit neuen Themen bzw. Modulen zu erweitern. Dieses Angebot erhielt grossen Zuspruch und wurde bis heute in über 50 QUIMS-Schulen durchgeführt. Es wurde überdies auch in anderen Kantonen der Deutschschweiz realisiert, aber auch von Schulen aus dem benachbarten Ausland (Deutschland, Österreich) nachgefragt.

Ziele und Fragestellung

Immer wieder wird die Frage nach der Wirksamkeit der Massnahmen gestellt, welche an QUIMS-Schulen durchgeführt wird. Diese Frage soll anhand einer Überprüfung der Wirksamkeit des SCALA-Ansatzes, der an QUIMS-Schulen angeboten wird, beantwortet werden. Es wurde eine kleine Evaluationsstudie in Auftrag gegeben, deren Schlussbericht hier vorliegt. Der vorliegende Bericht beschreibt die Idee, die Methoden und die Ergebnisse dieser Wirkungsstudie. Der Schlussbericht soll Erkenntnisse bieten, inwiefern die schulinternen SCALA-Weiterbildungen dazu beitragen, Beurteilungsverzerrungen bei Lehrpersonen aufgrund von Stereotypen wirksam zu verringern.

SCALA-Ansatz

Zum SCALA-Ansatz werden fünf Module angeboten, die je in einem 1-jährigen Zyklus durchgeführt werden. Es können auch zwei oder mehr Module nacheinander belegt werden. Die Angebote richten sich grundsätzlich an alle drei Zyklen (Kindergarten bis Ende der obligatorischen Schule) und sind fächerübergreifend. Nur Modul E ist vorerst dem Zyklus 1 und 2 vorbehalten. In Modul A stehen Erwartungen und Beurteilungen von Lehrpersonen im Zentrum, die möglichst frei von Vorurteilen bzw. Stereotypen sein sollten. Modul B fokussiert auf Zuschreibungen (Erfolgsattributionen) und formative Rückmeldungen von Lehrpersonen, die förderlich und fair erfolgen sollten. Modul C thematisiert Erwartungen an Schulklassen und didaktische Möglichkeiten, Schüler*innen in kulturell und sozial heterogenen Schulklassen zu fördern. Modul D thematisiert den Zusammenhang von Elternarbeit und Beurteilung. Modul E geht auf Verhaltensprobleme und Beurteilung von Schüler*innen ein. Die Module D und E sind neu und waren beim Start dieses Projekts noch in Erprobung. Daher wurden nur die Module A bis C evaluiert. Konkret wurde angenommen, dass die Weiterbildung zu weniger verzerrten Erwartungen an die Schüler*innen und zu mehr internalen Erfolgsattributionen führt.

Vorgehen

Um die Wirkung der Schilw zu überprüfen, wurde ein Evaluationsdesign mit Vorher-Nachher-Messung und Kontrollgruppe verwendet. Eine Kontrollgruppe umfasst Lehrpersonen aus QUIMS-Schulen, die keine SCALA-Weiterbildung besucht haben. Es können damit die Einschätzungen von ähnlichen Lehrpersonen verglichen werden, wobei die erste Gruppe von Lehrpersonen (Interventionsgruppe) die SCALA-Weiterbildung besuchte und die zweite Gruppe von Lehrpersonen (Kontrollgruppe) keine Weiterbildung besuchte. Nur eine Kontrollgruppenstudie kann Wirkungen im strengen Sinn überprüfen.

Die Studie wurde in den Schuljahren 2022/23 (Kohorte 1) und 2023/24 (Kohorte 2) durchgeführt. Es wurden alle Schulen für die Interventionsgruppe angefragt, welche mindestens eines der

Module A, B oder C gebucht hatten. Die Kontrollgruppenschulen wurden aus dem Pool der QUIMS-Schulen zufällig ausgewählt, die keine SCALA-Weiterbildung besucht oder gebucht hatten. Insgesamt konnten Daten von 52 Lehrpersonen aus 12 Schulen zu beiden Messzeitpunkten erhoben werden. Diese Lehrpersonen füllten die Befragung für 470 Schüler*innen aus. In der Interventionsgruppe (IG) handelt es sich um 32 Lehrpersonen aus 5 Schulen mit 279 Schüler*innen. Bei der Kontrollgruppe (KG) sind es im gleichen Zeitraum 20 Lehrpersonen aus 7 Schulen mit 191 Schüler*innen. Die Zahl der teilnehmenden Lehrpersonen pro Schule variierte stark. Wie für QUIMS-Schulen typisch, war die Zahl von Schüler*innen mit Migrationshintergrund recht hoch (47%).

Die Lehrpersonen füllten einen standardisierten Online-Fragebogen zu Beginn und am Ende des Schuljahres aus. Sie gaben die Leistungserwartungen an die einzelnen Schüler*innen sowie die Erfolgsszuschreibung in verschiedenen Fächern an. Die Wirkungsprüfung wurde also einem strengen Test unterzogen. Zudem gaben sie die letzte und die vorletzte Note aus schriftlichen Prüfungen in diesen Fächern an. Zusätzlich wurden der Beruf der Eltern (und damit der sog. sozio-ökonomische Status (SES) der Familie) sowie das Geschlecht und der Migrationshintergrund der Schüler*innen erfasst. Die Eltern gaben ihr schriftliches Einverständnis zur Teilnahme an der Studie, die Datenschutzrichtlinien wurden eingehalten.

Ergebnisse

Modul A: Die Wirksamkeit der Intervention zu Modul A wurde mit komplexen statistischen Methoden überprüft. Es findet sich in Mathematik kein Interventionseffekt. Die Lehrpersonenerwartungen waren über das Schuljahr äusserst stabil. Sowohl in der Kontrollgruppe als auch in der Interventionsgruppe sagten die Erwartungen am Anfang des Schuljahres die Erwartungen am Ende des Schuljahres in hohem Mass vorher. D.h. der erste Eindruck vom Start des Schuljahres prägte die Erwartungen am Ende des Schuljahres entscheidend, sie konnten kaum verändert werden.

Die Lehrpersonenerwartungen in Deutsch unterschieden sich nicht signifikant zwischen den Gruppen. Allerdings gab es hier einen signifikanten Effekt des SES in der Kontrollgruppe, welcher in der Interventionsgruppe nicht zu finden war. D.h. die soziale Herkunft der Kinder beeinflusst die Lehrpersonenerwartungen nach der Intervention nicht mehr. In der Kontrollgruppe haben wir aber einen solchen unerwünschten Einfluss aber gefunden. Dieser Effekt zeigt in die erwartete Richtung, wonach die Intervention zum Abbau von verzerrten Lehrpersonenerwartungen in Deutsch beiträgt.

Die Schilw werden an die spezifischen Bedingungen und Bedürfnisse der Schule angepasst. Zudem unterscheiden sich die Voraussetzungen zwischen den Schulen stark. Die Schilw könnte in manchen Schulen wirksam sein, in anderen Schulen aber nicht. Die Interventionseffekte (d.h. die Wirkungen der Schilw) sollten daher in den einzelnen Schulen überprüft werden. Leider ist die Zahl der teilnehmenden Lehrpersonen an den meisten Schulen zu klein, dass aussagekräftige Aussagen über die einzelnen Schulen gemacht werden können. Nur in einer Schule ist die Zahl der Lehrpersonen, die an der Studie teilgenommen haben, für Auswertungen gross genug. Für Mathematik gibt es in dieser Schule keinen signifikanten Interventionseffekt. Für Deutsch gibt es aber an dieser Schule einen signifikanten Interventionseffekt: Der Effekt des SES auf die Erwartungen in Deutsch ist in der Kontrollgruppe signifikant grösser als in der Interventionsgruppe, d.h. die Intervention führte hier zu einer signifikanten Abnahme des SES-Einflusses auf die Lehrpersonenerwartungen in Deutsch. Dieser Effekt ist sehr erfreulich, weil er die Wirksamkeit der Weiterbildung bei den Lehrpersonen, die an der Studie mitgewirkt haben, belegt. Hier hat sich die SCALA Schilw also positiv ausgewirkt.

Modul B: Das Ziel von Modul B ist es, Lehrpersonen zu sensibilisieren, nicht die soziale oder kulturelle Herkunft für den Schulerfolg als Erklärung beizuziehen, sondern die individuellen Fähigkeiten und die Anstrengung der Schüler*innen. D.h. wir versuchten Lehrpersonen zu ermutigen, den schulischen Erfolg von Schüler*innen durch ihre Fähigkeiten und ihre Anstrengung (d.h.

Schlussbericht: Evaluation der SCALA-Intervention

internal) zu erklären. Die Intervention ist also dann wirksam, wenn Lehrpersonen den schulischen Erfolg ihrer Schülerinnen und Schüler eher internal und nicht external erklären. Die Ergebnisse aus der Anwendung von multivariaten statistischen Methoden zeigen erfreulicherweise einen signifikanten Interventionseffekt: In der Interventionsgruppe nimmt die durchschnittliche interne Attribution von schulischem Erfolg signifikant zu, während in der Kontrollgruppe die durchschnittliche interne Attribution von schulischem Erfolg abnimmt. Dieser Effekt zeigt, dass die Intervention in Modul B dazu beigetragen hat, die Erfolgsattribution von Lehrpersonen in die erwünschte Richtung zu beeinflussen.

Weil keine Schulen an dieser Studie mitgemacht haben, welche Modul C besucht haben, konnten keine Wirkungsanalysen zu Modul C durchgeführt werden.

Schlussfolgerungen

Die Ergebnisse zu Modul A zeigen, dass der Effekt des sozio-ökonomischen Status (SES) auf die Lehrpersonenerwartungen in Deutsch aufgrund der Intervention tendenziell abnimmt. In einer Einzelschule wurde dieser Interventionseffekt signifikant. Dies belegt, dass die SCALA-Schilw wirksam sein kann. Dieser Interventionseffekt für den SES ist erfreulich, weil er belegt, dass die Schulen nach dem Besuch der SCALA-Weiterbildung einen Vorteil haben.

Die Ergebnisse zu Modul B zeigen, dass Lehrpersonen den schulischen Erfolg ihrer Schüler*innen nach der SCALA Schilw verstärkt auf die Fähigkeiten und Motivation der Schüler*innen zurückführen und nicht auf ihre äusseren Lernbedingungen, die durch den Migrationshintergrund oder den sozio-ökonomischen Status bestimmt sind. Das Ergebnis ist erwartungskonform und erfreulich. D.h. der Besuch von Modul B führt zu günstigeren Erfolgsattributionen. Dies kann die Chancengerechtigkeit begünstigen.

Wir vermuten, dass diese Wirkungen der SCALA-Schilw aufgrund der Einhaltung der folgenden Prinzipien aufgetreten sind, welche bei der Entwicklung des SCALA-Ansatzes leitend waren: (a) Vermittlung von einschlägigem Wissen und dieses mit vielen Beispielen illustrieren. (b) Betroffenheit schaffen und zu Einsicht/Reflexion motivieren, Aha-Erlebnis schaffen. (c) Bezug zur eigenen Schulklasse/Schüler*innengruppe herstellen (Praxistransfer). (d) Interventionstermine über ca. 6 Monate verteilen, immer wieder an Inhalte erinnern. (e) Hohes Engagement/Enthusiasmus der Schulleitung. Diese kann an pädagogischen Konferenzen die Lehrpersonen immer wieder an die Thematik erinnern und die Diskussion dazu schulintern weiterführen. Wir nehmen also an, dass die Einhaltung dieser Prinzipien in einer Schilw die Wirksamkeit erhöht.

Trotz einer strengen Evaluationsanlage konnten nachweisbare Wirkungen gefunden werden. Dies erhöht die Relevanz der vorliegenden Ergebnisse und des SCALA-Ansatzes. Der Besuch dieser Schilw trägt zu fairen Beurteilungen und damit zu Chancengerechtigkeit an QUIMS-Schulen bei. Auch wenn dieser Ansatz alleine nicht ausreicht, eine hohe Chancengerechtigkeit an QUIMS-Schulen zu schaffen, ermutigen die Ergebnisse zur Schlussfolgerung, dass sowohl die vorliegenden Module, aber auch die neuen SCALA-Module D und E an möglichst vielen Schulen durchgeführt werden. Die Schulen erhalten dadurch einen Vorteil. Zudem sollte das SCALA-Angebot mit neuen zusätzlichen Modulen erweitert werden.

Inhaltsverzeichnis

1	Ausgangslage, Auftrag	1
2	Fragestellung und Hypothesen	1
3	Theoretische Ausgangslage	2
3.1	Erwartungen und Attributionen	2
3.2	Prinzipien der Intervention	4
4	Beschreibung der Interventionen	4
4.1	Modul A: Chancengerechtigkeit bei der Beurteilung	5
4.2	Modul B: Zuschreibungen und Rückmeldungen	6
4.3	Modul C: Fördern in sozial heterogenen Klassen	7
4.4	Wie wird gearbeitet?	8
5	Methode	9
5.1	Durchführung der Studie	9
5.2	Stichprobe	10
5.2.1	Rekrutierung	10
5.2.2	Beschreibung der Stichprobe	11
5.3	Instrumente	15
5.4	Datenauswertung	18
6	Ergebnisse	18
6.1	Besuchte Weiterbildungen und weitere deskriptive Statistiken	18
6.2	Erwartungen (Modul A)	20
6.3	Attributionen (Modul B)	23
6.4	Unterrichtsformen in heterogenen Klassen (Modul C)	23
7	Diskussion	24
7.1	Wirkmechanismen	25
7.2	Herausforderungen	26
7.3	Ausblick	27
8	Empfehlungen	28
9	Literatur	29

1 Ausgangslage, Auftrag

Viele Studien belegen, dass Lehrpersonen Schüler*innen je nach sozio-ökonomischem Status (SES) und Migrationsuntergrund unfair formativ, summativ und prognostisch beurteilen (z.B. Niederbacher & Neuenschwander, 2021; Neuenschwander & Niederbacher, 2021; Neuenschwander et al., 2021). Diese negativ verzerrten Beurteilungen machen Lehrpersonen typischerweise unabsichtlich und beiläufig, werden aber von den Kindern wahrgenommen und beeinflussen deren Lernen und Handeln (Babad et al., 1989; Inan-Kaya & Rubie-Davies, 2021). Dadurch werden diese Gruppen von Kindern benachteiligt (Chancenungleichheit). Beurteilungen hängen eng mit Erwartungen und Attributionen von Lehrpersonen an die Schüler*innen zusammen. Hohe Leistungserwartungen der Lehrpersonen korrespondieren mit positiven Beurteilungen. Wenn Lehrpersonen den Schulerfolg der Schüler*innen internal und stabil erklären, resultieren ebenfalls günstigere Leistungen und Beurteilungen (Weiner, 1985). Diese verzerrten Beurteilungen können durch Stereotype erklärt werden, welche Lehrpersonen aktivieren und ihr Verhalten unabsichtlich beeinflussen (Frohlich, Martiny, Deaux, & Mok, 2015). Zahlreiche Medienberichte belegen eine wachsende Sensibilität für diese Problematik in der Öffentlichkeit und die Bereitschaft, Massnahmen zur Erhöhung der Chancengerechtigkeit zu ergreifen.

Zur Bearbeitung dieser Herausforderungen können Massnahmen auf der Ebene des Bildungssystems (z.B. der Sekundarstufe I in weniger Tracks, kooperative Sek-I-Modelle, Gestaltung der Selektionsverfahren in verschiedenen Etappen der Bildungslaufbahn, u.a.) und auf der Ebene des Unterrichts realisiert werden. Der SCALA-Ansatz fokussiert die Unterrichtsebene und thematisiert die Haltungen der Lehrpersonen und wie sie die Schüler*innen beurteilen. Der SCALA-Ansatz zeigt Wege, um Lehrpersonenerwartungen im Hinblick auf höhere Akkuratheit zu modifizieren (www.fhnw.ch/ph/scala). Die Ergebnisse einer Interventionsstudie belegen, dass dieses Ziel erreicht werden kann (Neuenschwander et al., 2021). Ausgehend davon wird seit 2019 die SCALA-Intervention an QUIMS-Schulen¹ des Kantons Zürich sowie an anderen Schulen durchgeführt (vgl. Modul A). Das Interesse am Angebot ist hoch, die Akzeptanz in der Lehrer*innen-schaft ist gemäss der vorliegenden Evaluationsergebnisse gegeben. Allerdings wissen wir nicht, ob die praktische Umsetzung dieses Angebots in den Schulen die Ziele erreicht, weil diese Umsetzung im Vergleich zur Intervention im ursprünglichen Projekt in verschiedener Hinsicht verändert worden ist (z.B. Schilw statt kursorischer Weiterbildung, inhaltliche Anpassungen der Intervention an die Bedürfnisse der QUIMS-Schulen, Verkürzung der Weiterbildung, Entwicklung neuer Module zu Attributionen und zu Heterogenität, Module B und C). Zudem sollten die ursprünglichen Studienergebnisse repliziert werden. Eine Rückmeldung zur Wirksamkeit dieser Intervention ist nicht nur für die Weiterentwicklung dieses Angebots wichtig, sondern auch für die Evaluation von QUIMS und für mehr Chancengerechtigkeit.

2 Fragestellung und Hypothesen

Der Schlussbericht soll Erkenntnisse liefern zu den Fragen, (a) wie sehr die schulinternen SCALA-Weiterbildungen dazu beitragen, Beurteilungsverzerrungen bei Lehrpersonen aufgrund von Stereotypen zu verringern und (b) ob Schilw wie diejenige von SCALA zur Verbesserung der Schulqualität und insbesondere der Chancengerechtigkeit beitragen kann, und zwar mit Blick auf die

¹ Das Programm „Qualität in multikulturellen Schulen (QUIMS)“ im Kanton Zürich unterstützt Schulen, die von überdurchschnittlich vielen Kindern aus eingewanderten, fremdsprachigen und sozial nicht privilegierten Familien besucht werden. Schulen, die beim Programm QUIMS beteiligt sind, betreiben langfristige Schul- und Unterrichtsentwicklung. Sie erarbeiten dazu Massnahmen, die allen Schülerinnen und Schülern zugutekommen und sie in ihrem Lernen fördern. Sie treffen ihre QUIMS-Massnahmen in drei Handlungsfeldern, in denen die Optimierungsmöglichkeiten besonders gross sind. Die Handlungsfelder betreffen die Förderung der Sprache, des Schulerfolgs und der Integration.

mögliche zukünftige Organisation der QUIMS-Schulen, für Schulen generell und für Schulen in sozial herausfordernden Lagen. Erreicht das SCALA-Angebot, so wie es an den QUIMS-Schulen umgesetzt wird, das Ziel, dass Lehrpersonen akkurate, nicht von SES und Migration verzerrten Erwartungen und Erfolgsattributionen an die Schüler*innen richten?

Diese Fragen besitzen vielfältige Relevanz.

- 1) Das Projekt trägt bei, Prozesse beim Entstehen von Chancengerechtigkeit zu verstehen und Grundlagen zu deren Reduktion zu bearbeiten.
- 2) Das Projekt dient der Qualitätssicherung bzw. Qualitätsentwicklung der SCALA-Intervention. Eine wirksame und nachhaltige Weiterbildung ist ein erklärtes Ziel von QUIMS-Schulen. Dies zu erreichen und zu überprüfen ist daher ein zentraler Auftrag des SCALA-Angebots.
- 3) Das Projekt liefert Erkenntnisse zur Wirksamkeit von Lehrpersonenweiterbildungen und Schulentwicklung.
- 4) Das Projekt bearbeitet eine innovative Forschungsfrage und generiert neue wissenschaftliche und praxisrelevante Ergebnisse, wie Lehrpersonen fair formativ, summativ und prognostische Urteile abgeben können.

3 Theoretische Ausgangslage

Der Effekt der selbsterfüllenden Prophezeiung von Erwartungen wurde bereits von Rosenthal und Jacobson (1974) untersucht. Im ursprünglichen Experiment wurde mit Schüler*innen ein Intelligenztest durchgeführt. Allerdings wurde den Lehrpersonen nicht das korrekte Ergebnis des Intelligenztestes ihrer Schüler*innen kommuniziert, sondern ein zufälliges. Das Ergebnis des Experiments zeigte nun, dass sich die Intelligenz nicht aufgrund des gemessenen Intelligenzwertes entwickelte, sondern aufgrund des Intelligenztestergebnisses, das der Lehrperson kommuniziert worden ist. Dieses Experiment ist ethisch problematisch und die Ergebnisse zur Entwicklung der Intelligenz konnten nicht repliziert werden. Allerdings belegen viele Folgestudien die selbsterfüllende Prophezeiung, welche mit Lehrpersonenerwartungen verbunden ist (Wang et al. 2018). Lehrpersonen richten Erwartungen an die Leistungen von Schüler*innen in einem Fach. Sie verhalten sich diesen Schüler*innen gegenüber entsprechend ihren Erwartungen, geben Rückmeldungen, Beurteilungen, Ermutigung oder Ermahnung. Diese Rückmeldungen werden von den Schüler*innen wahrgenommen und beeinflussen ihre Anstrengung, ihr Selbstkonzept, ihr Lernverhalten, was sich in ihren Leistungen abbildet. Entsprechend wurden in vielen Studien Effekte von Lehrpersonenerwartungen auf die fachlichen Leistungen von Schüler*innen in verschiedenen Schulstufen gezeigt.

3.1 Erwartungen und Attributionen

Erwartungen werden als Überzeugungen über die zukünftigen Leistungen von Kindern definiert. Wenn Lehrpersonen hohe Erwartungen haben und an das Entwicklungspotenzial der Kinder «glauben», nehmen die Leistungen der Schüler*innen zu. Solche Erwartungseffekte können auch langfristige Effekte auf den Bildungsverlauf (Neuenschwander et al., 2018), den Bildungsabschluss bzw. die intergenerationelle Bildungsmobilität haben (Neuenschwander, Ramseier, Garrote, eingereicht).

Allerdings zeigen viele Studien, dass Erwartungen von Lehrpersonen (sowie von anderen Menschen), nicht nur von den Leistungen der Schüler*innen beeinflusst sind, sondern auch vom sozioökonomischen Status, dem Migrationshintergrund und dem Geschlecht des Lernenden. In Übereinstimmung mit früherer Forschung zeigten Neuenschwander und Niederbacher (2021) gemäss Abbildung 1 am Beispiel von Schüler*innen des 4.-6. Schuljahres in der Deutschschweiz, dass die Leistungserwartungen in Mathematik zu Beginn eines Schuljahres nach Kontrolle der Leistungen

Schlussbericht: Evaluation der SCALA-Intervention

in Mathematik höher sind gegenüber Kindern mit höherem sozio-ökonomischem Status (SES), mit deutscher Familiensprache und gegenüber Mädchen. Gleichzeitig sind die Leistungserwartungen von Lehrpersonen selbsterfüllend, weil sie die zukünftigen Leistungen vorhersagen. Der Effekt von SES, Familiensprache und Geschlecht auf Leistungen am Ende des Schuljahres wird in dieser Analyse vollständig von den Leistungserwartungen von Lehrpersonen erklärt (mediert). Das heisst, die Leistungserwartungen von Lehrpersonen erklären vollständig die Effekte von SES, Migrationshintergrund und Geschlecht auf die Leistungen. Die Leistungserwartungen spielen daher eine zentrale Rolle beim Entstehen von Chancengerechtigkeit.

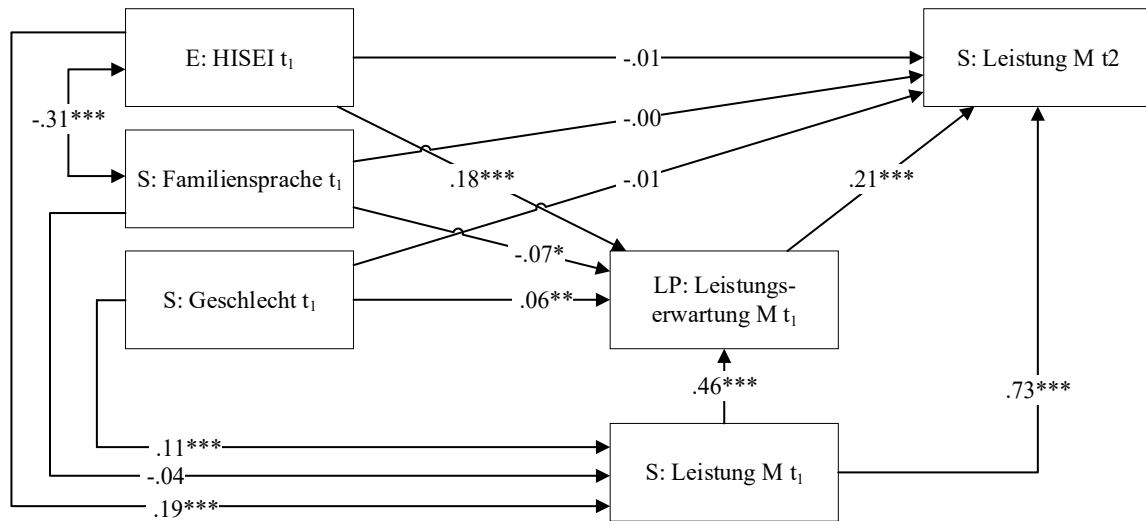


Abbildung 1: Pfadmodell für Mathematik (aus: Neuenschwander & Niederbacher, 2021)

Legende: S: Schüler*in, LP: Lehrperson; N=1150.

Hinweis zur Modellgüte: $\chi^2=2.16$, $df=2$, $\chi^2/df=1.1$, CFI=1.0, RMSEA=.01, SRMR=.01

Mit Erwartungen sind Erfolgs- bzw. Misserfolgsattributionen verwandt (Weiner, 2010). Erfolgsattributionen bezeichnen die subjektive Zuschreibung eines Grundes für einen Erfolg. Weiner (2010) unterschied drei Dimensionen: (a) Lokalität: der Erfolg kann bei der Person selber (internal) oder in der Umwelt der Person (external) liegen. (b) Temporalität: der Erfolg kann einem stabilen Merkmal (z.B. Intelligenz) oder einem variablen Merkmal (z.B. Anstrengung, Glück/Pech) zugeschrieben werden. (c) Kontrollierbarkeit: Der Mensch kann über diesen Grund die Kontrolle ausüben oder nicht. Im Folgenden gehen wir auf die Lokalität und Temporalität ein, welche sich als besonders wichtig herausgestellt haben. Studien zeigen, dass die Zuschreibung von Bildungserfolg (z.B. gute Leistungen) mit internal – stabilen Merkmalen zu einer Anstrengungs- und Leistungssteigerung führt (Wang & Hall, 2018; Reyna, 2000; 2008). Das heisst die Art, wie Lehrpersonen den Erfolg bzw. Misserfolg eines Schülers bzw. einer Schülerin erklären, beeinflusst ihre formativen Beurteilungen, Rückmeldungen, welche von den Schüler*innen wahrgenommen werden und sich in deren Lernverhalten abbilden.

Reyna (2000; 2008) zeigte Effekte von askriptiven Merkmalen wie SES, Migration und Geschlecht auch auf Attributionsmuster: Lehrpersonen attribuieren Erfolg von Kindern mit höherem SES und/oder mit Schweizer Pass eher internal und stabil, während sie Misserfolg eher external und variabel attribuieren. Sie kommunizieren dadurch dieser Gruppe von Kindern eine höhere Anstrengung und Zuversicht, welche die Leistungszunahme beeinflusst.

Die Überprüfung der eigenen Leistungserwartungen und Erfolgsattributionen gegenüber Schüler*innen bildet ein Anlass, die Rückmeldungen und Verhaltensweisen gegenüber Gruppen von Schüler*innen zu reflektieren und wenn nötig anzupassen. Diese Überprüfung fokussiert auf die

Haltung der Lehrpersonen, die sich in vielen verschiedenen Verhaltensweisen äussern, die oft implizit gezeigt werden. So zeigte Babad et al. (1989), dass verzerrte Lehrpersonenerwartungen ihre nonverbale Kommunikation gegenüber den Schüler*innen beeinflussen. Diese Kommunikation wird von den Schüler*innen wahrgenommen und beeinflusst deren Lernmotivation und Leistungen.

3.2 Prinzipien der Intervention

Die Interventionsforschung zu Lehrpersonenerwartungen im Nachgang zu sozialpsychologischen Erkenntnissen zur Modifikation von Überzeugungen (Rubie-Davies, Peterson, Sibley & Rosenthal, 2015; Yeager & Walton, 2011) zeigen, dass sich Überzeugungen und stereotype Erwartungen durch Sensibilisierung, Reflexion und markante Ereignisse in Kombination mit konkreten Übungen modifizieren lassen. Die Veränderung von Einstellungen braucht nicht viel Zeit, wenn damit nicht neue, zusätzliche Kompetenzen erforderlich sind, sondern basiert oft auf einem Aha-Erlebnis (Yeager & Walton, 2011). Einstellungsänderungen können nach Yeager & Walton (2011) ausgelöst werden, wenn die Intervention (a) situativ und personenbezogen ist, (b) glaubwürdig und überzeugend ist, (c) die Emotionen anspricht bzw. Betroffenheit auslöst. Solche Einstellungsänderungen können nachhaltig sein (Cohen & Garcia, 2008).

Die Interventionen setzen an Überlegungen zur Pädagogik der Vielfalt an, wonach Vielfalt und Individualität als Norm und Ressource gedacht werden. Sie thematisieren die Konzepte Erwartungen, Attribution, Feedback, aber auch Formen von enrichment, kooperativem Lernen, um kulturell- und leistungsheterogene Gruppen optimal zu fördern. Beispielsweise zeigte Slavin (1996) sehr effektive Formen von kooperativem Lernen für leistungs- und kulturell heterogene Gruppen, die in der Schweiz weniger bekannt sind. Schliesslich erhält die Kooperation von Lehrpersonen mit Eltern mit verschiedenem kulturellen Hintergrund Aufmerksamkeit. Studien zeigten, dass Eltern je nach sozialem und kulturellem Hintergrund die Erwartungen und Beurteilungen, aber auch die Förderstrategien von Lehrpersonen wesentlich beeinflussen (z.B. Niederbacher & Neuenchwander, 2020; Niederbacher & Neuenchwander, 2021). Zudem werden die Konzepte sozioökonomischer Status und Migration erläutert und ihre Bedeutung für die Leistungsbeurteilung gezeigt. Zusammenfassend orientiert sich die Intervention an den folgenden Prinzipien (Neuenchwander et al., 2021):

a) *Information*: Die Lehrpersonen werden über die einschlägigen Theorien und ausgewählte empirische Befunde informiert bzw. das Wissen zu diesen Themen wird in Erinnerung gerufen.

b) *Sensibilisierung*: Die Lehrpersonen werden für die Bedeutung dieser Effekte im eigenen Schulalltag anhand von vielen Beispielen sensibilisiert. Zudem wird den Lehrpersonen die hohe Bedeutung dieser Prozesse gezeigt, um sie zu motivieren, diese zu reflektieren und gegebenenfalls ihr Verhalten anzupassen. Wichtig ist, dabei keine Schuldzuweisungen vorzunehmen, um eine Abwehrreaktion zu vermeiden. Vielmehr ist eine ehrliche Auseinandersetzung in einem freundlichen Umfeld erforderlich.

c) *Anwendung auf eigene Schulklasse*: Die Konzepte werden auf die Kinder der eigenen Schulklasse angewendet. Es werden Situationen im eigenen Schulalltag betrachtet, in denen sich diese Phänomene konkret zeigen. Es werden Wege gesucht, alternative Verhaltensweisen in diesen Situationen zu entwickeln.

4 Beschreibung der Interventionen

Im Lehrplan 21 gewinnen neben der summativen Beurteilung auch die formative und die prognostische Beurteilung an Bedeutung. Die Beurteilung soll einerseits das Lernen und die Entwicklung der Kinder fördern, andererseits soll sie aber auch gerecht sein. Gestützt auf frühere Forschung und Erfahrung wurde der SCALA-Ansatz entwickelt, in der Zusammenarbeit mit Lehrpersonen

erprobt und seine Wirksamkeit überprüft und nachgewiesen. Der SCALA-Ansatz arbeitet mit den Lehrpersonen (Unterricht) und der Schulleitung (Schule) und zeigt konkrete Wege, wie lernförderlich und sozial gerecht beurteilt werden kann.

Es werden fünf Module angeboten, die je in einem 1-jährigen Zyklus durchgeführt werden. Es können auch zwei oder mehr Module nacheinander belegt werden. Die Angebote richten sich grundsätzlich an alle drei Zyklen (Kindergarten bis Ende der obligatorischen Schule) und sind fächerübergreifend. Nur Modul E ist vorerst Zyklus 1 und 2 vorbehalten. Jedes Modul beinhaltet ein Orientierungs- und ein Anwendungselement (Abbildung 2).

	Orientierungselemente		Anwendungselemente
Modul A Chancengerechtigkeit bei der Beurteilung	A1 Wie Herkunft den Schulerfolg beeinflusst		A2 Erwartungen und Beurteilungen beim Übertritt
Modul B Zuschreibungen und Rückmeldungen	B1 Was Zuschreibungen von Lehrpersonen bewirken		B2 Rückmeldung und formative Beurteilung
Modul C Fördern in sozial heterogenen Klassen	C1 Strategien zur Steigerung positiver Erwartungen		C2 Enrichment in sozial heterogenen Klassen
Modul D Elternarbeit und Beurteilung	D1 Eltern und Beurteilung	D2 Herkunft und Kommunikation	D3 Leitideen Elternarbeit
Modul E: Unterrichtsstörung und Beurteilung	E1 Herausforderndes Verhalten und Beurteilung		E2 Strategien im Umgang mit herausforderndem Verhalten

Abbildung 2: Module des SCALA-Ansatzes

Zudem wird ein Fachcoaching angeboten, welches die Umsetzung der Inhalte in den Unterricht unterstützen soll. Damit soll die Intervention umfassender in den Unterricht implementiert werden.

4.1 Modul A: Chancengerechtigkeit bei der Beurteilung

Im Modul A wird die Bedeutung der Herkunft für den Schulerfolg von Kindern anhand empirischer Befunde dargestellt. Es wird aufgezeigt, wie die Herkunftseffekte mit Erwartungen von Lehrpersonen an die Leistung der Lernenden zusammenhängen. Anhand von praktischen Situationen werden Lehrpersonen für Beurteilungstendenzen im Unterricht und in Übertrittsverfahren sensibilisiert. Es werden konkrete Massnahmen vermittelt, wie Lehrpersonen regelmässig prüfen können, ob sie alle Kinder unabhängig von ihrer familiären Herkunft lernförderlich und fair wahrnehmen und beurteilen.

Im Modul A verfolgen die Schulleitungen und Lehrpersonen folgende Ziele:

- (a) Ursachen für Chancenungerechtigkeit im Bildungssystem, für lernungünstige und ungerechte Beurteilungen im Unterricht und bei Übertrittsentscheiden erkennen und verstehen.

(b) Lehrpersonen reflektieren die eigene Beurteilungspraxis und werden für die eigenen Beurteilungstendenzen sensibilisiert. Sie richten ihre Beurteilung konsequent auf eine lernförderliche und faire (formative, summative, prognostische) Beurteilung aus.

(c) Die Übertrittspraxis in die Primarstufe, die Sekundarstufe I und Sekundarstufe II im eigenen Schulhaus wird geplant und umgesetzt.

Praxisbeispiel aus der Mittelstufe

In der 6. Klasse von Frau M. steht die Zuteilung der Schülerinnen und Schüler in eine Klasse der Oberstufe kurz bevor. Wie jedes Mal hat sie auch in dieser Klasse Kinder, welche «auf der Kippe» stehen. Dazu gehören auch Mohammed und Lisa. Die beiden haben einen guten 4,5-er im Schnitt, Mohammed ist in Mathe etwas besser, Lisa in Sprache. Mohammed kann gut abstrahieren und logisch schlussfolgern. Lisa liest gerne und viel und schreibt wunderbare Erzählungen. Beide sind fleissig und möchten gerne in die Sek A übertreten. Im Elterngespräch merkt Frau M. einmal mehr, dass Lisa bei schulischen Problemen grosse Unterstützung von ihren Eltern erhält. Mohammeds Eltern hingegen können ihn kaum unterstützen, sie sprechen beide schlecht Deutsch, da sie erst seit 6 Jahren in der Schweiz sind. Beide Eltern wünschen sich, dass ihr Kind in die Sek A übertritt. Frau M. entscheidet sich schliesslich dafür, Lisa in die Sek A einzuteilen und Mohammed in die Sek B. Zum Glück, so denkt sie, ist eine Durchlässigkeit in der Oberstufe gewährleistet.

Ausgehend von solchen Beispielen diskutieren die Lehrpersonen darüber, aufgrund welcher Überzeugungen die Entscheidung von Frau M. zustande gekommen ist. In einer interaktiven Arbeitssequenz werden mögliche Entwicklungsszenarien von Mohammed und Lisa aufgezeigt und die förderlichen und hinderlichen Faktoren derselben beschrieben. In einem nächsten Schritt werden die Argumente der Lehrpersonen mit der Forderung des Bildungsrates konfrontiert, allen Kindern unabhängig von deren sozialen, sprachlichen, kulturellen oder religiösen Hintergrund die gleichen, guten Bildungschancen zu geben. Im letzten Schritt werden gemeinsam konkrete Massnahmen und Richtlinien erarbeitet, welche im Schulhaus und in Zusammenarbeit mit der Oberstufe der Gemeinde diesen bildungsrätlichen Forderungen gerecht werden. Ziel ist es, die Umsetzung der gemeinsamen Massnahmen im Schulprogramm zu verankern sowie die thematische Zusammenarbeit mit der Sekundarstufe zu intensivieren.

4.2 Modul B: Zuschreibungen und Rückmeldungen

Das Modul B stellt Zuschreibungen und lernförderliche Rückmeldungen von Lehrpersonen an die Kinder ins Zentrum. Lehrpersonen erfahren, dass sich ihre Zuschreibungen, im Sinne von Erklärungen für schulische Erfolge und Misserfolge, in ihren Rückmeldungen widerspiegeln. Die Lehrpersonen reflektieren, dass Rückmeldungen – auch beiläufige und nonverbale – die Motivation, den Selbstwert und die Leistungen des Kindes stark beeinflussen. Sie prüfen ihre eigenen Zuschreibungstendenzen und lernen Strategien, um zu prüfen, wie sie lernförderliche und gerechte Rückmeldungen geben können.

Im Modul B verfolgen die Schulleitungen und Lehrpersonen folgende Ziele:

(a) Zuschreibungen gegenüber Schüler*innen und deren Wirkungen erkennen.

(b) Lehrpersonen reflektieren eigene Zuschreibungstendenzen.

(c) Lehrpersonen wenden lernförderliche Zuschreibungen und Rückmeldungen für die formative Beurteilung an.

(d) Die Schule entwickelt eine Kultur der lernförderlichen und gerechten Zuschreibung und Rückmeldung.

Ein Praxisbeispiel aus der Oberstufe

Die Sekundarklasse von Herrn P. hat in den vergangenen Wochen an einem Geschichtsvortrag gearbeitet. In den Themenbereichen Absolutismus und Französische Revolution wurde den Schülerinnen und Schülern ein Vortragsthema zugestellt. Lisa und Samira sind zwei Lernende, welche Herr P. zu den mittelmässigen Schülerinnen zählt. Beide Elternteile von Lisa sind Ärzte. Samira betreut zu Hause oft ihre fünf kleineren Geschwister. Ihre Eltern arbeiten beide für ein Reinigungsinstitut, ihre Mutter häufig auch abends. Die Familie ist erst seit fünf Jahren in der Schweiz. Lisa bekam das Thema «Das Leben am Hofe von Louis XIV» und Samira «Merkantilismus» zugestellt. Beide Vorträge waren gut aufgebaut und sorgfältig recherchiert. Herr P. gibt den Schülerinnen und Schülern immer gleich nach dem Vortrag eine kurze mündliche Rückmeldung. Zu Lisas Vortrag meinte er: «Du hast gut recherchiert, sachlogisch aufgebaut und souverän vorgetragen. Insgesamt eine sehr gute Leistung». Samira hörte: «Na, passabel. Du hast ja erstaunlich gute Quellen gefunden. Manchmal hilft einem ja der Zufall. Du hast dir wirklich Mühe gegeben.»

Ausgehend von dieser Fallbeschreibung werden Zusammenhänge von Zuschreibungen bei guten und schlechten Leistungen analysiert, und zwar hinsichtlich ihrer Wirkung auf die formative Beurteilung. In einem nächsten Schritt werden in Gruppen lernförderliche Rückmeldungen formuliert, welche den individuellen Voraussetzungen von verschiedenen Schülerinnen und Schülern und deren Kompetenzniveau entsprechen. Der Praxisauftrag beinhaltet das Einüben von lernförderlichen Zuschreibungen. In der abschliessenden Veranstaltung mit Erfahrungsaustausch werden die Resultate und Erfahrungen zusammengetragen. Die Schule erarbeitet sich einen Leitfaden zur Implementierung von lernförderlichen Zuschreibungen. Die Möglichkeit, die Erkenntnisse aus der Weiterbildung auch als Grundhaltung in ein Leitbild zu integrieren, wird diskutiert.

4.3 Modul C: Fördern in sozial heterogenen Klassen

Im Modul C wird die Bedeutung von Erwartungen von Lehrpersonen an die Kinder und Jugendlichen thematisiert und aufgezeigt, wie sie sich bei der formativen, summativen und prognostischen Beurteilung auswirken bzw. wie Unterricht besonders lernförderlich wird. Lehrpersonen lernen konkrete Strategien kennen, um hohe und zugleich auch realistische Erwartungen an die ganze Klasse und an die einzelnen Schülerinnen und Schüler aufzubauen. Sie erproben konkrete Hilfsmittel, wie Kinder und Jugendliche, welche zuhause wenig schulische Unterstützung erhalten, gefördert werden können und wenden diese in ihrem Unterricht an. Im Modul C verfolgen die Schulleitungen und Lehrpersonen folgende Ziele:

- (a) Lehrpersonen erkennen die lernförderliche Wirkung von hohen Erwartungen an die ganze Klasse und an die einzelnen Schüler*innen.
- (b) Lehrpersonen wenden im eigenen Unterricht konkrete Strategien an, um hohe Erwartungen sowohl an die Klasse als auch an die einzelnen Schüler*innen zu richten.
- (c) Sie wenden konkrete Massnahmen für die Arbeit mit sozial heterogenen Klassen und zur Förderung von Lernenden mit geringer familiärer Unterstützung an.
- (d) Die Schule entwickelt eine Kultur, in welcher Lehrpersonen die sozial heterogenen Klassen als Ressource verstehen und nutzen.

Ein Praxisbeispiel aus dem Kindergarten

Die Kindergartenschülerinnen und -schüler im zweiten Jahr von Frau A. haben verschiedene Regelspiele gelernt. Regelspiele sind anspruchsvoll: Man muss nicht nur eine ganze Zeit lang ruhig sitzen können, man muss sich auch Regeln merken und diese einhalten. Nun ist das Rushhour-Spiel zum absoluten Lieblingsspiel der grossen Kiga-Kinder geworden. Manchmal dürfen sie dieses Spiel nun auch in der Auffangzeit des Kindergartens spielen. Das gilt aber nur für die älteren Kinder. Eines Morgens bricht ein Streit aus. Der sehr lernbegierige und etwas wilde Anton besucht die erste Kindergartenklasse und will partout nicht verstehen, dass das Rushhour-Spiel nicht für

ihn gedacht ist. Die grösseren Kinder versuchen es ihm zu erklären, aber er wird nur wütend und wirft alle Autos auf den Boden. Frau A. verbietet ab nun, dass dieses Spiel in der Auffangzeit gespielt wird. Die grösseren Kinder machen Anton und sein Verhalten für diese Regeländerung verantwortlich und in den Pausen gibt es deshalb immer wieder Streit.

Die Anforderungen, welche an Lehrpersonen aller Stufen durch die grosse Heterogenität in Schulklassen gestellt werden, sollen in diesem Modul thematisiert werden. Ausgehend von diesem Praxisbeispiel wird diskutiert, welche Überlegungen die Kiga-Lehrerin dazu bewegen haben, das Rushhour- Spiel nur den grösseren Kindern zu erlauben. Sowohl die Erwartungen der Lehrpersonen an die Kompetenzen der Kinder, wie auch die Zuschreibungen, welche zu solchen Regeln führen, werden diskutiert. Im Anwendungselement wird aufgezeigt, dass die Heterogenität in einer Klasse als soziale Ressource genutzt werden kann, indem etwa die Kompetenzen der Kinder mit Verantwortungsübernahme für die Gruppe verknüpft werden. Die Umsetzung in die Unterrichtspraxis wird geplant und im Praxisauftrag umgesetzt.

4.4 Wie wird gearbeitet?

Nachdem in einem Gespräch mit der Schulleitung und meist der QUIMS-Beauftragten der Umfang und die Themenschwerpunkt des Angebots besprochen wurden, konstituiert sich die schulinterne Begleitgruppe (Abbildung 3). Diese besteht aus Lehrpersonen, welche in dieser Schule besondere Verantwortung übernehmen, wie etwa Stufenleitung, QUIMS-Beauftragte etc. sowie mindestens eine Vertretung aus der Schulleitung. Der Begleitgruppe kommt die wichtige Aufgabe zu, den Kontakt zwischen der Schule und der Leitung der externen Weiterbildung herzustellen. Gemeinsam mit der Begleitgruppe werden die Teamweiterbildungen vorbesprochen und die Umsetzung in den Unterrichtsalltag nachbereitet. Die Begleitgruppe stellt sicher, dass sowohl Reflexionsaufträge als auch Praxisaufträge von den Teammitgliedern umgesetzt werden. Bei Bedarf moderiert sie zusätzliche Teamsitzungen, welche mit der Weiterbildung im Zusammenhang stehen bzw. diese ergänzen. Das folgende Arbeitsphasenmodell zeigt die Arbeitsweise in den verschiedenen Modulen. Die gelben Bereiche stehen für die Arbeit mit der Begleitgruppe, die blauen Felder sind Arbeitssequenzen mit dem ganzen Schulhausteam und die grünen Felder werden an Teamsitzungen ohne das Kursleitungsteam durchgeführt.

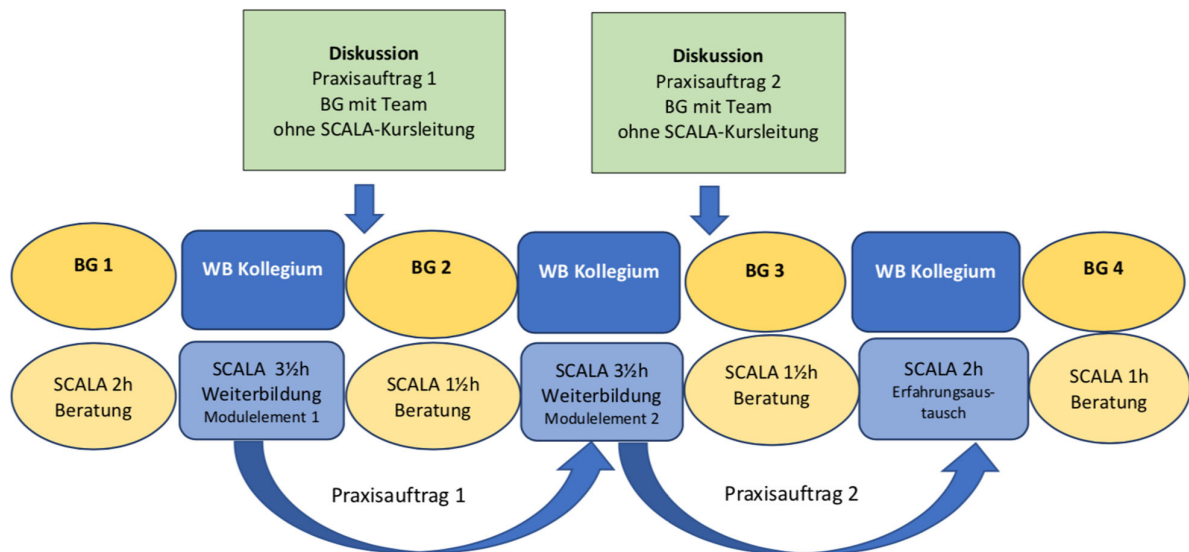


Abbildung 3: Die Arbeitsphasen in den Modulen A, B, C, E
 Legende: BG – Begleitgruppe, WB - Weiterbildung

Nachdem in der letzten gemeinsamen Veranstaltung des Weiterbildungszyklus ein Erfahrungsaustausch mit den Lehrpersonen stattgefunden hat, ist die Begleitgruppe dafür verantwortlich, dass passende Inhalte der Weiterbildung ins Schulprogramm einfließen und nachhaltig umgesetzt werden. In Modul D wird die Organisation der Weiterbildung leicht angepasst.

5 Methode

Um die Wirkung der SCALA-Weiterbildung zu prüfen, wurde ein Untersuchungsdesign entwickelt, welches aufzeigen soll, inwiefern sich die Lehrpersonenerwartungen und die Erfolgsattribution in Bezug auf einzelne Schüler*innen innerhalb eines Schuljahres durch den Besuch der Weiterbildung veränderten. Für die Untersuchung wurde deshalb ein Prä-Post-Design mit Kontrollgruppe verwendet, damit die Wirksamkeit der Weiterbildung überprüft werden kann. Es wurden Veränderungen durch die Intervention systematisch erfasst und mit den Veränderungen einer zufällig gewählten zweiten Gruppe ohne Intervention verglichen. Nur aufgrund dieses Vergleichs können schlüssige Aussagen zur Wirksamkeit der Intervention vorgenommen werden. Zudem wurde ein harter Test verwendet zur Messung der Beurteilungspraxis. Es wurden nämlich die effektiven Einschätzungen der Lehrpersonen über die Schüler*innen ihrer Klasse vor und nach der Weiterbildung erfasst. Dadurch können Veränderungen dieser Einschätzungen konkret analysiert werden. Die Wirksamkeitseinschätzung sowie die Akzeptanz der Lehrpersonen gegenüber der Weiterbildung wurden nicht erfragt, weil sie keine Aussagen über die Wirkung der Intervention im strengen Sinn zulassen.

5.1 Durchführung der Studie

Die Studie wurde mittels Prä-Post-Design, also mit zwei standardisierten Befragungen, durchgeführt. Die Befragungen wurden in der Interventionsgruppe vor und nach dem letzten Weiterbildungstermin durchgeführt. Das entspricht grob dem Anfang und dem Ende des Schuljahres (Abbildung 4). Die Befragungen der Kontrollgruppe fanden in einem vergleichbaren Zeitraum statt (Abbildung 5).

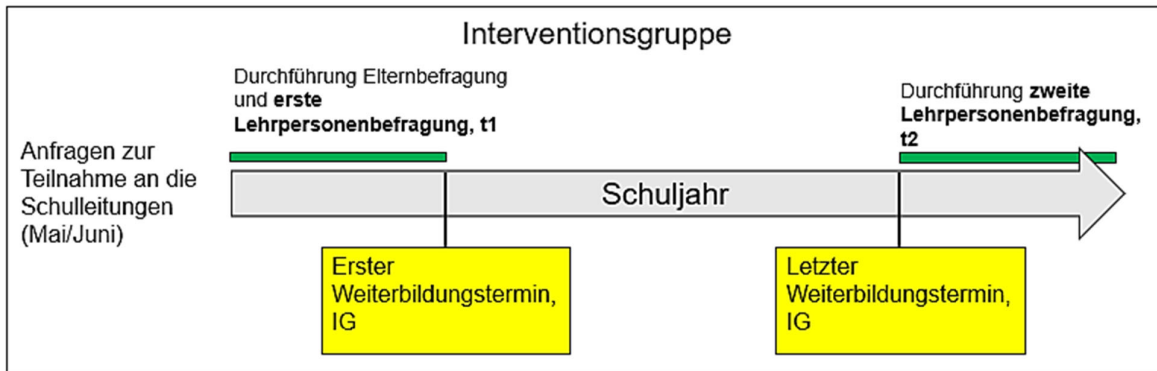


Abbildung 4: Vorgehen Interventionsgruppe

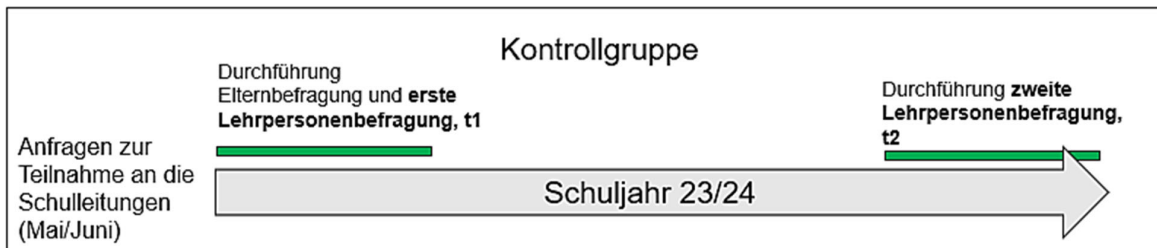


Abbildung 5: Vorgehen Kontrollgruppe

5.2 Stichprobe

5.2.1 Rekrutierung

Die Untersuchung wurde an QUIMS-Schulen im Kanton Zürich durchgeführt. Für die Interventionsgruppe waren von den rund 150 QUIMS-Schulen (154 Schulen im Schuljahr 22/23 und 153 Schulen im Schuljahr 23/24) diejenigen von Interesse, welche in den Schuljahren 2022/23 und 2023/24 jeweils entweder Modul A, B oder C der SCALA-Weiterbildung besucht haben. Die Module D und E waren zum Untersuchungszeitpunkt noch in Erprobung und wurden deshalb nicht untersucht. Im ursprünglichen Untersuchungsdesign war lediglich eine einmalige Durchführung der Befragung vorgesehen (eine Kohorte). Weil in diesem Schuljahr zu wenig Daten gesammelt werden konnten, wurde im folgenden Schuljahr die Datenerhebung wiederholt (2. Kohorte).

Für die Kontrollgruppe wurden ebenfalls QUIMS-Schulen im Kanton Zürich gesucht, welche keine SCALA-Weiterbildung besucht hatten. Um den Effekt von Störvariablen zu reduzieren bzw. auszuschliessen, wurden Kontrollgruppenschulen gesucht, die in Bezug auf Grösse und Schulstufe mit den Interventionsgruppen vergleichbar waren. So sollte sichergestellt werden, dass sich die Schulen nicht grundlegend in ihren Voraussetzungen unterschieden. In der ersten Kohorte wurden 20 QUIMS-Schulen angefragt, von welchen zunächst fünf zugesagt hatten, jedoch nur drei sowohl am Prä- als auch am Posttest teilgenommen hatten. Für die zweite Kohorte wurden 29 Schulen angefragt, von welchen vier zugesagt und an der Befragung teilgenommen hatten. Die Diskrepanz zwischen Anfrage und Zusage lässt sich damit erklären, dass eine Teilnahme in der Kontrollgruppe (ohne Weiterbildungsteilnahme) für die Schulen weniger attraktiv zu sein schien.

Die Rekrutierung der teilnehmenden Lehrpersonen fand in den beiden Kohorten teilweise unterschiedlich statt. In der ersten Kohorte wurden die Schulen, beziehungsweise die Schulleitungen vom Volksschulamt Zürich angefragt. Insgesamt drei Schulen nahmen am Prätest und Posttest teil. Eine vierte Schule nahm nur am Prätest teil. In der zweiten Kohorte wurden hingegen die Schulleitungen vom Projektteam kontaktiert, wobei der Anfrage ein Empfehlungsschreiben des Volksschulamtes Zürich beigelegt wurde. Es wurden sieben Schulen, welche eine SCALA-Weiterbildung vereinbart hatten, vom Projektteam angefragt, von welchen drei einer Teilnahme zugestimmt und teilgenommen hatten. Das weitere Vorgehen war in beiden Kohorten gleich. Die Schulleitungen informierten das Kollegium bezüglich der Teilnahme an der Untersuchung. Die Lehrpersonen bestätigten die Teilnahme an der Untersuchung, indem sie dem Projektteam die Anzahl Schüler*innen mitteilten. Dadurch konnte den Schüler*innen ein Code zugeordnet werden, welcher wiederum zur Identifikation der Elternbefragung verwendet wurde.

Die Eltern wurden schriftlich über die Studie informiert und angefragt, ob ihr Kind teilnehmen durfte. Die Lehrpersonen gaben die entsprechenden Unterlagen ihren Schüler*innen für die Eltern nach Hause, sammelten diese wieder ein und schickten sie gesammelt dem Projektteam zurück. Auf dieser Basis füllten die Lehrpersonen die Fragebogen nur zu denjenigen Kindern ihrer Klasse aus, deren Eltern ihr Einverständnis gegeben hatten.

Im Prätest der Kohorte 1 zeigte sich, dass die Lehrpersonen den Beruf der meisten Eltern nicht kannten. Daher wurde im Posttest der Kohorte 1 bzw. in Kohorte 2 bereits im Prätest zusätzlich eine Elternbefragung durchgeführt, um den sozio-ökonomischen Status der Eltern zu erfassen. Wie beim Einholen des Elterneinverständnisses wurde die Befragung über die Lehrpersonen durchgeführt. Die Eltern füllten auf Papier zwei Fragen aus, gaben das Blatt in einem verschlossenen Couvert über das Kind der Lehrperson, welche das Couvert dem Forschungsteam weiterleitete. Mit diesem Verfahren wurde ein höherer Rücklauf erzielt, als wenn die Eltern die Unterlagen direkt dem Projektteam geschickt hätten.

5.2.2 Beschreibung der Stichprobe

Insgesamt konnten in den zwei Kohorten (Schuljahr 2022/23 und Schuljahr 2023/24) Daten von 52 Lehrpersonen aus 12 Schulen erhoben werden. Diese Lehrpersonen füllten die Befragung für 470 Schüler*innen aus. In der Interventionsgruppe (IG) handelt es sich um 32 Lehrpersonen aus 5 Schulen mit 279 Schüler*innen. Bei der Kontrollgruppe (KG) sind es im gleichen Zeitraum 20 Lehrpersonen aus 7 Schulen mit 191 Schüler*innen.

Der Rücklauf aus den beiden Kohorten ist den Tabelle 1 und 2 zu entnehmen. Die Zahlen stellen die Anzahl der angefragten Lehrpersonen dar. In den Tabelle 3 ist die Stichprobe nach Kohorte und nach Gruppe aufgeteilt dargestellt.

Tabelle 1: Rücklauf Kohorte 1: Anzahl Lehrpersonen

Rücklauf Kohorte 1					
	Anfragen t1	Rücklauf t1 absolut / relativ	Anfragen t2	Rücklauf t2 absolut / relativ	Rücklauf gesamt (relativ)
IG	87	33 / 38%	23	9 / 39%	10%
KG	141	31 / 22%	20	7 / 35%	5%

Tabelle 2: Rücklauf Kohorte 2: Anzahl Lehrpersonen

Rücklauf Kohorte 2					
	Anfragen t1	Rücklauf t1 absolut / relativ	Anfragen t2	Rücklauf t2 absolut / relativ	Rücklauf gesamt (relativ)
IG	177	69 / 39%	69	23 / 33%	13%
KG	77	30 / 39%	30	13 / 43%	17%

Legende: IG – Interventionsgruppe, KG - Kontrollgruppe

Tabelle 3: Zusammensetzung der Stichprobe – Kohorte 1 - IG

Interventionsgruppe – Kohorte 1					
Lehrpersonen				Fehlend	Gesamt
Stufe	Kindergarten	Primarstufe	Sekundarstufe I		
	0	2	7	0	9
Funktion	KLP	SHP	Andere		
	8	1	0	0	9
SCALA Modul	Modul A	Modul B	Modul C		
	2	7	0	0	9
Schüler*innen				Fehlend	Gesamt
Geschlecht	Weiblich	Männlich	Divers		
	36	32	0	0	68
Migrationshintergrund	Ja	Nein			
	26	42		0	68
Alter	Zwischen 8 und 16 Jahren, Mittelwert 13.2 Jahre				

Tabelle 4: Zusammensetzung der Stichprobe – Kohorte 2 - IG

Interventionsgruppe – Kohorte 2					
Lehrpersonen				Fehlend	Gesamt
Stufe	Kindergarten	Primarstufe	Sekundarstufe I		
	3	12	8	0	23
Funktion	KLP	SHP	Andere		
	17	3	3	0	23
SCALA Modul	Modul A	Modul B	Modul C		
	18	5	0	0	23
Schüler*innen				Fehlend	Gesamt
Geschlecht	Weiblich	Männlich	Divers		
	84	80	0	47	211
Migrationshintergrund	Ja	Nein			
	115	49		47	211
Alter	Zwischen 6 und 16 Jahren, Mittelwert 10.8 Jahre				

Tabelle 5: Zusammensetzung der Stichprobe – Kohorte 1 - KG

Kontrollgruppe – Kohorte 1					
Lehrpersonen				Fehlend	Gesamt
Stufe	Kindergarten	Primarstufe	Sekundarstufe I		
	1	6	0	0	7
Funktion	KLP	SHP	Andere		
	7	0		0	7
Schüler*innen				Fehlend	Gesamt
Geschlecht	Weiblich	Männlich	Divers		
	37	31	0	2	70
Migrationshintergrund	Ja	Nein			
	31	38		1	70
Alter	Zwischen 4 und 12 Jahren, Mittelwert 8.4 Jahre				

Tabelle 6: Zusammensetzung der Stichprobe – Kohorte 2 - KG

Kontrollgruppe – Kohorte 2					
Lehrpersonen				Fehlend	Gesamt
Stufe	Kindergarten	Primarstufe	Sekundarstufe I		
	0	5	8	0	13
Funktion	KLP	SHP	Andere		
	9	1	3	0	13
Schüler*innen				Fehlend	Gesamt
Geschlecht	Weiblich	Männlich	Divers		
	65	55	0	1	121
Migrationshintergrund	Ja	Nein			
	79	38		4	121
Alter	Zwischen 7 und 16 Jahren, Mittelwert 11.6 Jahre				

Legende: KLP – Klassenlehrperson, SHP – Schulische*r Heilpädagog*in

Ein Vergleich zwischen der Interventions- und der Kontrollgruppe zu t1 zeigt, dass sich die beiden Gruppen in den untersuchten Variablen nicht oder nur gering unterscheiden. In Bezug auf die

Unterrichtsstufe und die Funktion der teilnehmenden Lehrpersonen unterscheiden sich die Interventions- und die Kontrollgruppe nicht. Auf der Schüler*innenebene gibt es aber Unterschiede in Bezug auf Alter, Migrationshintergrund und SES. Die Schüler*innen in der Interventionsgruppe ($M = 11.51$, $SD = 2.71$) sind signifikant älter als diejenigen in der Kontrollgruppe ($M = 10.44$, $SD = 2.93$, $t(418) = 3.87$, $p < .001$), wobei es sich um einen kleinen Effekt handelt (Cohen's $d = .38$). In Bezug auf den Migrationshintergrund kann ebenfalls ein signifikanter Unterschied zwischen den beiden Gruppen festgestellt werden ($\chi^2(1) = 23.154$, $p < .001$). Auch hier handelt es sich um einen kleinen Effekt (Cramer's $V = .24$). Schüler*innen aus der Kontrollgruppe hatten einen höheren SES ($M = 60.08$, $SD = 18.66$) im Vergleich zur Interventionsgruppe ($M = 52.54$, $SD = 21.16$, $t(409) = -3.78$, $p < .001$). Die Effektstärke ist auch hier relativ klein (Cohen's $d = .38$). Die Schüler*innenleistungen unterschieden sich nicht zwischen den beiden Gruppen (letzte und vorletzte Mathematiknote, letzte und vorletzte Deutschnote, letzte und vorletzte Hauptfachnote (wenn nicht DE oder MA unterrichtet wird)). Ebenso unterschieden sich die Lehrpersonenerwartungen an die Leistungen der Schüler*innen nicht zwischen den beiden Gruppen. Für die Leistungsattribution (internal vs. external) konnte in den Fächern Deutsch und Mathematik ein Unterschied zwischen den beiden Gruppen gefunden werden. Für das Fach Deutsch attribuierten die Lehrpersonen der Interventionsgruppe die Leistungen der Schüler*innen stärker internal ($M = 36.52$, $SD = 19.53$) als die Lehrpersonen der Kontrollgruppe ($M = 43.38$, $SD = 21.75$, $t(309) = -2.93$, $p = .004$). Es handelt sich um eine kleine Effektstärke (Cohen's $d = -.33$). Ein ähnlicher Effekt zeigt sich für Mathematik, wo Lehrpersonen der Kontrollgruppe die Leistungen stärker external attribuierten ($M = 41.84$, $SD = 21.82$) als Lehrpersonen der Interventionsgruppe ($M = 35.43$, $SD = 20.13$, $t(268) = -2.50$, $p = .013$). Auch hier findet sich ein kleiner Effekt (Cohen's $d = .31$). Wir erachten daher die Ausgangswerte der Interventionsgruppe und der Kontrollgruppe als ähnlich und vergleichbar.

Zudem wurden die Werte der Variablen zwischen den Kohorten unterschieden. Die erste und zweite Kohorte unterscheiden sich, wie nachfolgend dargestellt, kaum in den untersuchten Variablen. Die Vergleiche ergaben, dass sich die beiden Kohorten in Bezug auf das von den Lehrpersonen besuchte Weiterbildungsmodul ($\chi^2(1) = 8.67$, $p = .003$) mit einem mittleren Effekt (Cramer $V = .52$) unterschieden. Für die unterrichtete Schulstufe sowie Funktion der Lehrpersonen wurde kein Unterschied gefunden. Auf der Schüler*innenebene unterschieden sich die beiden Kohorten signifikant in ihrem Migrationshintergrund ($\chi^2(1) = 6.08$, $p = .014$), die Effektstärke ist jedoch klein (Cramer $V = .12$). Ein t-Test für unabhängige Stichproben zeigt, dass der durchschnittliche SES der Kohorte 1 ($M = 60.11$, $SD = 19.98$) signifikant höher war als derjenige der Kohorte 2 ($M = 53.78$, $SD = 20.35$, $t(409) = 2.97$, $p = .002$). Die Effektstärke war jedoch ebenfalls klein (Cohen's $d = .31$). Bei Alter und Geschlecht der Schüler*innen gibt es keinen Unterschied zwischen den Kohorten. In Bezug auf die Schüler*innenleistung gibt es jedoch Aspekte, in denen sich die beiden Kohorten signifikant unterscheiden. Die *letzte vergebene Mathematiknote* ist in der ersten Kohorte höher ($M = 4.31$, $SD = .86$) als in der zweiten Kohorte ($M = 3.72$, $SD = .93$, $t(286) = 5.19$, $p < .001$). Dieser Unterschied hat einen mittleren Effekt (Cohen's $d = .64$). Ein ähnlicher Unterschied kann bei der *letzten vergebenen Deutschnote* gefunden werden. Diese Note in der ersten Kohorte ($M = 4.21$, $SD = .91$) ist signifikant höher als die der zweiten Kohorte ($M = 3.94$, $SD = .99$) ($t(331) = 2.30$, $p = .022$). Hierbei handelt es sich allerdings nur um einen geringen Effekt (Cohen's $d = .28$). Die übrigen erhobenen Leistungsvariablen (vorletzte Note Mathematik, vorletzte Note Deutsch, letzte Note und vorletzte Note Hauptfach) unterschieden sich nicht zwischen den beiden Kohorten. In Bezug auf die Leistungsattribution (internal vs. external) und die erwartete Schulleistung (*Welche Leistungen erwarten Sie von Kind in den nächsten vier Schulwochen in den folgenden Fächern?*) konnte zwischen den Kohorten ebenfalls kein Unterschied festgestellt werden. Zusammenfassend sind die Unterschiede zwischen den Kohorten daher klein. Deswegen und aufgrund der kleinen Stichproben werden die Auswertungen für beide Kohorten gemeinsam vorgenommen.

5.3 Instrumente

Zur Erhebung der Lehrpersonen-Daten wurden zwei Online-Befragungen verwendet. Diese enthielten Fragen zur Lehrperson, sowie Fragen zu den Erfolgsfaktoren der Weiterbildung wie die Unterrichtsgestaltung und Einschätzungen (d.h. Erwartungen und Attributionen) zu den einzelnen Schüler*innen. Die Lehrpersonen gaben die eigene Schulstufe an (Itemformulierung: *«Auf welcher Stufe unterrichten Sie?»*), Item aus Eigenentwicklung), wobei 7.7% der Befragten im Kindergarten unterrichten, 44.2% in der Primarstufe und 44.2 % in Sekundarstufe I. Zusätzlich wurde die Funktion der Lehrperson erfragt (Itemformulierung: *«Welche Funktion haben Sie in der Klasse?»*), Item aus Eigenentwicklung). 80.8% der Befragten gaben an, Klassenlehrperson zu sein, 7.7 % Schulische Heilpädagog*in und 7.7% Fachlehrperson. In der Interventionsgruppe wurde zudem erhoben, welches Weiterbildungsmodul die Befragten besuchten (Itemformulierung: *«Welche SCALA-Weiterbildung zur Beurteilung von Schüler:innenleistungen haben Sie im Schuljahr 2022/2023 / 2023/2024 besucht und abgeschlossen?»*), Item aus Eigenentwicklung). 62.5% der Lehrpersonen aus der Interventionsgruppe besuchten das Modul A, 37.5 % besuchten Modul B. Leider konnte keine Schule mit dem Modul C für eine Teilnahme gewonnen werden.

In Bezug auf die einzelnen Schüler*innen waren zunächst die Erwartungen von Interesse (Itemformulierung: *«Welche Leistungen erwarten Sie von Kind X in den nächsten vier Schulwochen in den folgenden Fächern?»*), Neuenschwander et al., 2021). Die Erwartungen wurden für die Fächer Mathematik und Deutsch, sowie zusätzlich für ein selbstgewähltes Hauptfach (sofern weder Mathematik noch Deutsch unterrichtet wurde) angegeben. Kindergartenlehrpersonen gaben ihre Erwartungen in Bezug auf den von ihnen gewählten fachlichen Schwerpunkt an. Die Einschätzungen konnten auf einer Skala von «überhaupt nicht gute Leistungen» (1) bis «sehr gute Leistungen» (6) gegeben werden.

Weiter wurde die Erfolgsattribution (external vs. internal) in den gleichen Fächern erhoben (Itemformulierung: *«Erinnern Sie sich an die letzte Situation, als Kind X gute Leistungen in Deutsch/ Mathematik/ Hauptfach/ fachlichem Schwerpunkt gezeigt hat. Worauf führen Sie diese guten Leistungen primär zurück?»*), in Anlehnung an Weiner, 1985). Mit einem Schieberegler konnten Werte zwischen 1 und 100 angegeben werden, wobei 1 «grösstenteils vom Kind abhängig» und 100 «grösstenteils von äusseren Faktoren abhängig» entspricht.

Es wurden zusätzlich die Leistungen der Schüler*innen mit Noten erfasst (Itemformulierung: *«Welche Noten haben Sie Kind X in den letzten beiden schriftlichen Prüfungen in Mathematik gegeben?»*), Item aus Eigenentwicklung). Hierbei wurden die Noten in Halbnotenschritten erhoben.

Der sozio-ökonomische Status (SES) der Schüler*innen wurde mittels Elternbefragung erhoben. Diese enthielt Fragen zum Beruf (Itemformulierung: *«Welchen Beruf üben Sie im Moment aus? Wenn Sie aktuell nicht erwerbstätig sind, welchen Beruf haben Sie zuletzt ausgeführt oder gelernt?»*), Neuenschwander et al., 2021) und Arbeitspensum (Itemformulierung: *«Welches Arbeitspensum haben Sie im Moment?»*), Neuenschwander et al., 2021). Dies wurde jeweils für Elternteil 1 und Elternteil 2 angegeben (sofern zwei Elternteile oder Erziehungsberechtigte vorhanden waren). Die Berufe der Eltern wurden zunächst nach ISCO (International Standard Classification of Occupations) codiert und anschliessend in den ISEI (International Socio-Economic Index of Occupational Status nach Ganzeboom et al. 1992) umgewandelt. Der ISEI dient als Indikator für den sozio-ökonomischen Status der Familien. Es wurde jeweils der höhere ISEI (sog. HISEI) der beiden Elternteile verwendet. Im Fragebogen für Lehrpersonen wurde zudem die Familiensprache der Schüler*innen basierend auf dem Wissen der Lehrkraft erfasst (Itemformulierung: *«Welche Sprache spricht Kind X in der Familie am häufigsten?»*), Neuenschwander et al., 2021). Auf Grundlage dieser Information wurde die Variable Migrationshintergrund erstellt. Schüler*innen, die zu Hause eine andere Sprache als Deutsch/ Schweizerdeutsch sprachen, wurden als «mit Migrationshintergrund» kategorisiert.

Schlussbericht: Evaluation der SCALA-Intervention

Tabelle 7: Deskriptive Statistiken der verwendeten Items

	Interventions- gruppe t1			Interventions- gruppe t2			Kontrollgruppe t1			Kontrollgruppe t2		
	N	M	S	N	M	S	N	M	S	N	M	S
Erwartungen Ma	144	4.48	1.28	105	4.38	1.30	148	4.30	1.44	138	4.11	1.43
Erwartungen De	189	4.15	1.12	136	4.10	1.30	155	4.32	1.34	146	4.25	1.37
Erwartungen HF	10	4.70	1.05	12	4.75	1.14	29	4.52	1.18	10	3.70	1.42
Erwartungen SP KiGa	9	4.89	.93	10	5.0	1.05	8	5.13	1.13	8	5.00	1.20
Erfolgsattributionen Ma	100	43.28	21.35	99	37.12	21.32	77	40.30	24.13	137	44.22	22.17
Erfolgsattributionen De	136	43.07	19.35	130	38.64	19.98	86	44.81	24.11	147	43.74	23.58
Erfolgsattributionen HF	49	42.02	15.83	12	44.17	24.85	50	42.40	23.37	10	46.50	22.64
Erfolgsattributionen SP KiGa	31	37.74	20.76	10	42.60	21.05	32	33.41	18.61	8	39.75	16.04
Leistungsbeurtei- lung KiGa SP	18	3.72	.96	10	3.90	.74	36	4.03	.91	8	4.00	1.20
Letzte Note Ma	140	4.09	1.03	125	3.63	1.02	107	4.17	1.16	119	3.66	1.10
Vorletzte Note Ma	138	4.46	1.18	126	3.65	1.17	106	4.25	1.18	122	3.60	1.13

Schlussbericht: Evaluation der SCALA-Intervention

	Interventions- gruppe t1			Interventions- gruppe t2			Kontrollgruppe t1			Kontrollgruppe t2		
	N	M	S	N	M	S	N	M	S	N	M	S
Letzte Note De	173	4.21	1.20	160	3.63	.84	117	4.18	1.08	131	3.88	1.04
Vorletzte Note De	156	4.75	1.07	134	3.74	1.01	117	4.16	1.06	131	3.81	1.03
Letzte Note HF	10	5.50	.00	12	3.92	.67	27	4.30	1.00	10	3.30	.75
Vorletzte Note HF	10	5.50	.00	11	3.96	.79	27	4.11	.84	10	3.40	.78
SES	231	52.54	21.16				180	60.08	18.66			

Legende: N: Anzahl Personen, M: Mittelwert, S: Standardabweichung; Ma: Mathematik, De: Deutsch, HF: Hauptfach, SP: Schwerpunkt, KiGa: Kindergarten

Tabelle 8: Migrationshintergrund Schüler*innen Interventions- und Kontrollgruppe

		Anzahl	Migrationshintergrund		Gesamt
			Kein Migrationshinter- grund	Migrationshintergrund	
Interventionsgruppe oder Kontrollgruppe	Interventionsgruppe		91	141	232
	Kontrollgruppe		117	69	186
Gesamt			208	210	418

Die Tabelle 7 enthält eine Übersicht von allen in den Berechnungen verwendeten Items. Diese sind jeweils mit Stichprobe, Mittelwert und Standardabweichung zu Zeitpunkt 1 (t1) und Zeitpunkt (2) angegeben. Zudem wird zwischen Interventions- und Kontrollgruppe unterschieden. Tabelle 8 enthält die deskriptiven Analysen zum Migrationshintergrund.

5.4 Datenauswertung

Die Stichproben zu den einzelnen Modulen sind leider zu klein, dass systematisch schulspezifische oder modulspezifische Wirkungsanalysen durchgeführt werden konnten. Einzig zu einer Schule konnte aufgrund der Stichprobengrösse eine schulspezifische Analyse durchgeführt werden.

Die Wirksamkeit des Moduls A der SCALA-Weiterbildung wurde mit Regressionsanalysen im Gruppenvergleich unter Berücksichtigung der Mehrebenenstruktur der Daten mit MPlus 8 (Type=complex) ausgewertet.

Die Wirksamkeit der Weiterbildung zu Modul B wurde mit Ko-Varianzanalysen mit Messwiederholung überprüft. Weil die Zahl der Lehrpersonen aus Schulen, welche Modul B besucht hatten, klein ist, sind die Angaben nur teilweise generalisierbar. Leider hat keine Schule, die Modul B besucht hat, mit einer genügend grossen Anzahl Lehrpersonen an dieser Studie teilgenommen, weshalb hier keine schulspezifischen Analysen durchgeführt werden können.

Leider haben keine Schulen, welche Modul C besucht haben, an der Studie teilgenommen, obwohl solche Schulen angefragt worden sind. Daher können keine spezifischen Auswertungen zur Wirksamkeit von Modul C durchgeführt werden. Zur groben Orientierung werden mit allen Lehrpersonen der Interventionsgruppe im Vergleich zur Kontrollgruppe die Ergebnisse von Varianzanalysen dargestellt im Hinblick auf den Einsatz von kooperativen Lernformen und digitalen Lernformen aus Lehrpersonensicht.

6 Ergebnisse

Nachfolgend werden die Ergebnisse der Datenauswertung präsentiert.

6.1 Besuchte Weiterbildungen und weitere deskriptive Statistiken

Die Teilnehmenden der Kontrollgruppe haben keine der SCALA-Weiterbildungen besucht. Zudem wurden im Erhebungszeitraum keine vergleichbaren Weiterbildungen besucht, welche mögliche Einflüsse auf die Beurteilungspraktiken der Lehrpersonen hätten nehmen können.

Die Tabelle 9 enthält die Pearson-Korrelationen zwischen den einbezogenen Variablen. Sie zeigt unter anderem, dass der Migrationshintergrund und der SES mit den erwarteten Leistungen, den externalen Attributionen zu t1 sowie mit den durchschnittlichen Noten in Mathematik und Deutsch korrelieren. Die erwarteten Leistungen sowie die externalen Attributionen korrelieren mit den durchschnittlichen Noten in Mathematik und Deutsch, auch zwischen den Fächern.

Schlussbericht: Evaluation der SCALA-Intervention

Tabelle 9: Korrelationstabelle

	2	3	4	5	6	7	8	9	10
1 Erwartete Leistung MA t ₁	.73***	.82***	.64***	-.39***	-.46***	.52***	.46***	.20***	.24***
2 Erwartete Leistung DE t ₁		.66***	.80***	-.44***	-.39***	.51***	.53***	.33***	.32***
3 Erwartete Leistung MA t ₂			.77***	-.38***	-.45***	.45***	.41***	.16*	.21**
4 Erwartete Leistung DE t ₂				-.29***	-.34***	.41***	.43***	.35***	.34***
5 externale Attribution (Mittel der Fächer) t ₁					.47***	-.31***	-.38***	-.11*	-.11*
6 externale Attribution (Mittel der Fächer) t ₂						-.34***	-.25***	.01	-.07
7 Durchschnittliche Noten MA t ₁							.64***	.15*	.18**
8 Durchschnittliche Noten DE t ₁								.07	.19**
9 CH (vs. Migration)									.35***
10 SES									

Legende. * $p < .05$, ** $p < .01$, *** $p < .001$. 173 < N < 391

MA = Mathematik, DE = Deutsch, CH = Schweizer*in

6.2 Erwartungen (Modul A)

Vor der Prüfung der Wirksamkeit von Modul A wurde analysiert, wie sehr die Lehrpersonenerwartungen in Mathematik von Migrationshintergrund und SES der Schüler*innen abhängen. Im Rahmen der Studie konnten keine Leistungstests eingesetzt werden, weil alle Klassenstufen vom Kindergarten bis zur 9. Klasse einbezogen worden sind – wobei keine Leistungstests für alle Stufen vorliegen – und weil der Aufwand der Lehrpersonen möglichst geringgehalten werden sollte. Als Ersatz für die Leistungstests wurden die letzte und vorletzte Note des Kindes erhoben. Um die Reliabilität der Noten zu steigern, wurde der Mittelwert dieser Noten verwendet. Gemäss Tabelle 9 korrelieren die Noten und die Lehrpersonen-Leistungseinschätzungen hoch (im Bereich $r = 0.5$).

In Weiterführung zu den Korrelationen in Tabelle 99 wurden Regressionsanalysen unter Kontrolle der Mehrebenenstruktur der Daten (Schüler*innen sind in Klassen verschachtelt) mit MPlus 8 gerechnet. Es wurden alle Lehrpersonen der Interventionsgruppe und der Kontrollgruppe zum ersten Messzeitpunkt in die Analyse einbezogen. Gemäss Tabelle 10 hängen die Lehrpersonenerwartungen in Mathematik am Anfang des Schuljahres vom SES, aber nicht vom Migrationshintergrund des Kindes ab. Wenn die Noten zusätzlich in die Gleichung einbezogen werden, werden die Lehrpersonenerwartungen ausschliesslich von den Noten erklärt. Der Effekt des SES ist offenbar in den Noten eingebunden.

Die Lehrpersonenerwartungen in Deutsch werden von dem Migrationshintergrund der Schüler*innen, aber nicht vom SES beeinflusst. Dieser Effekt des Migrationshintergrunds bleibt erhalten, wenn die Noten in die Regressionsgleichung zusätzlich einbezogen werden.

Tabelle 10: Effekte von SES und Migration auf Erwartungen in Dt und M (beta mit Standardmessfehler, Mehrebenenanalysen)

	Lp- Erw M t1	Lp- Erw M t1	Lp-Erw D t1	Lp-Erw D t1
Noten M t1		.50 (.16)***		
Noten D t1				.48 (.06)***
CH (vs NCH)	.12 (.09)	.02 (.09)	.26 (.06)***	.23 (.05)***
SES	.25 (.06)***	.13 (.08)	.29 (.07)	.15 (.07)
R ²	0.10****	.30***	.18***	.38***
N	245	200	298	237

Legende: *: $p < .05$, **: $p < .01$, ***, $p < .001$, zweiseitig, Lp-Erw M: Lehrpersonenerwartung in Mathematik, LP-Erw D: Lehrpersonenerwartung in Deutsch, Durchschnitt der letzten und der vorletzten Note. N: Anzahl Personen, R²: erklärte Varianz

In Abweichung zu früheren Studien findet sich in dieser Stichprobe nur ein SES-Effekt auf die Lehrpersonenerwartungen in Mathematik und ein Effekt des Migrationshintergrunds auf die Lehrpersonenerwartungen in Deutsch. Der fehlende Effekt des Migrationshintergrunds auf die Leistungserwartungen in Mathematik könnte möglicherweise damit zusammenhängen, dass Lehrpersonen an QUIMS-Schulen für Einflüsse der Kultur auf die Beurteilung bereits gut sensibilisiert sind. Der fehlende SES-Effekt auf die Lehrpersonenerwartungen in Deutsch könnte mit einer geringen Varianz des SES in dieser Stichprobe zusammenhängen. In den ausgewählten QUIMS-Schulen gehört die grosse Mehrheit der Schüler*innen unteren sozialen Schichten an. Es könnte aber auch mit

der Verwendung von Noten statt der Leistungstestdaten zusammenhängen. Lehrpersonenerwartungen und Noten sind hoch korreliert.

Die Wirksamkeit der Intervention zu Modul A wurde anhand von Regressionsanalysen im Gruppenvergleich unter Kontrolle der Mehrebenenstruktur der Daten mit Mplus 8 überprüft. Es wurde geprüft, ob sich die konkreten, kindbezogenen Einschätzungen der Lehrpersonen aufgrund der Intervention und im Vergleich zur Kontrollgruppe veränderten. Die Ergebnisse in Tabelle 11 zeigen keine signifikanten Gruppenunterschiede. Weder in der Kontrollgruppe noch in der Interventionsgruppe erklärten der Migrationshintergrund und SES die Lehrpersonenerwartungen in Mathematik. Die Lehrpersonenerwartungen waren über das Schuljahr sehr stabil. Sowohl in der Kontrollgruppe als auch in der Interventionsgruppe sagten die Erwartungen am Anfang des Schuljahres die Erwartungen am Ende des Schuljahres in hohem Mass vorher. D.h. der erste Eindruck vom Start des Schuljahres prägte die Erwartungen bis am Ende des Schuljahres.

Die Lehrpersonenerwartungen in Deutsch unterschieden sich nicht signifikant zwischen den Gruppen. Allerdings gibt es hier einen signifikanten Effekt des SES in der Kontrollgruppe, welcher in der Interventionsgruppe nicht zu finden ist. D.h. im Unterschied zur Kontrollgruppe ist der Herkunftseffekt auf Lehrpersonenerwartungen in Deutsch nach der Intervention nicht mehr signifikant. Dieser Unterscheid zeigt in die erwartete Richtung. Allerdings sind die Effekte des Migrationshintergrunds auf die Lehrpersonenerwartungen in Deutsch in der Interventionsgruppe und in der Kontrollgruppe signifikant. In dieser Stichprobe konnte dieser Effekt durch die Intervention nicht beseitigt bzw. reduziert werden.

Tabelle 11: Wirksamkeit auf Erwartungen
(beta mit Standardmessfehler, Mehrebenenanalysen mit Mplus)

	Interventions- gruppe Lp-Erw M t ₂	Kontrollgruppe Lp-Erw M t ₂	Interventions- gruppe Lp-Erw D t ₂	Kontrollgruppe Lp-Erw D t ₂
Lp-Erw M t ₁	.60 (.11)***	.81 (.04)***		
Lp-Erw D t ₁			.67 (.07)***	.78 (.08)***
Noten M t ₁	.15 (.08)	.13 (.06)*		
Noten D t ₁			.09 (.06)	-.06 (.13)
CH (vs NCH)	.03 (.05)	.05 (.04)	.12 (.05)*	.10 (.04)*
SES	.16 (.11)	.00 (.04)	.10 (.09)	.11 (.04)**
R ²	.53***	.78***	.61***	.69***
N	70	95	108	103

Legende: *: p<.05, **: p<.01, ***. p<.001, zweiseitig

Lp-Erw M: Lehrpersonenerwartung in Mathematik, LP-Erw D: Lehrpersonenerwartung in Deutsch, Durchschnitt der letzten und der vorletzten Note. N: Anzahl Personen, R²: erklärte Varianz

Die Schulen sowohl der Interventionsgruppe als auch der Kontrollgruppe haben vor und während dieser Studie unterschiedliche Weiterbildungen besucht. Die Situation an Schulen unterscheidet sich daher in verschiedener Hinsicht. Um die schulspezifischen Situationen zu berücksichtigen, sollten daher die Effekte in den einzelnen Schulen der Interventionsgruppe mit der Kontrollgruppe verglichen werden. Leider ist die Zahl der teilnehmenden Lehrpersonen an den meisten Schulen zu klein für solche Analysen. Nur in einer Schule ist die Zahl der Lehrpersonen, die an der Studie teilgenommen haben, so gross, dass statistische Auswertungen vorgenommen werden können. Für die anderen Schulen konnten diese Detailanalysen nicht durchgeführt werden. Die Analysen wurden deshalb noch einmal spezifisch für die Lehrpersonen dieser Schule im Vergleich zu den Lehrpersonen der Kontrollgruppe wiederholt.

Die Ergebnisse in Tabelle 12 zeigen für die Lehrpersonenerwartungen in Mathematik ein ähnliches Bild wie für die Gesamtstichprobe. Es gibt keine Effekte von Migrationshintergrund und SES auf die Lehrpersonenerwartungen in Mathematik. Die Lehrpersonenerwartungen in Deutsch wurden aber in den beiden Gruppen unterschiedlich erklärt. Es gibt einen signifikanten Gruppenunterschied des Effekts des SES (Wald $\chi^2(1) = 5.6, p < .05$). Der Effekt des SES ist in der Kontrollgruppe signifikant stärker als in der Interventionsgruppe, d.h. die Intervention führte zu einer signifikanten Abnahme des SES-Einflusses auf die Lehrpersonenerwartungen in Deutsch. Dieser Effekt ist sehr erfreulich, weil er die Wirksamkeit der Weiterbildung bei den Lehrpersonen, die in der Interventionsgruppe an der Studie mitgewirkt haben, zeigt. Für den Effekt des Migrationshintergrunds wurde aber kein signifikanter Gruppenunterschied gefunden (Wald $\chi^2(1) = .44, ns$).

Tabelle 12: Wirksamkeit der Intervention auf Erwartungen
(beta mit Standardmessfehler, Mehrebenenanalysen mit Mplus)

	Interventions- gruppe Lp-Erw M t ₂	Kontrollgruppe Lp-Erw M t ₂	Interventions- gruppe Lp-Erw D. t ₂	Kontrollgruppe Lp-Erw D t ₂
Lp-Erw M t ₁	.62 (.13)***	.81 (.04)***		
Lp-Erw D t ₁			.77 (.05)***	.78(.08)***
Noten M t ₁	.11 (.13)	.13(.06*		
Noten D t ₁			-.07 (.11)	-.06 (.13)
CH (vs NCH)	.06 (.05)	.04 (.04)	.16 (.06)**	.10 (05)*
SES	.04 (.13)	.00 (.04)	-.01 (.04)	.11 (.04)**
R ²	50%	78%	70%***	69%
N	49	95	60	103

Legende: *: $p < .05$, **: $p < .01$, ***: $p < .001$, zweiseitig

Lp-Erw M: Lehrpersonenerwartung in Mathematik, LP-Erw D: Lehrpersonenerwartung in Deutsch, Durchschnitt der letzten und der vorletzten Note im jeweiligen Fach.

6.3 Attributionen (Modul B)

Das Modul B thematisiert Erfolgsattributionen von Lehrpersonen gegenüber ihren Schüler*innen. Das Ziel der Intervention ist, Lehrpersonen zu sensibilisieren, nicht die soziale oder kulturelle Herkunft für die Erklärung des Schulerfolgs beizuziehen, sondern die individuellen Fähigkeiten und die Anstrengung der Schüler*innen. Die Intervention ist also dann wirksam, wenn Lehrpersonen den schulischen Erfolg ihrer Schülerinnen und Schüler eher internal und nicht external erklären. Dabei sollte der Effekt von SES, Migrationshintergrund und der Noten in Deutsch und Mathematik im Pretest kontrolliert werden, um Störeffekte auszuschliessen. Die Schüler*innen sind in Schulklassen eingebettet. Die Kontrolle der Klassenzugehörigkeit ist aber in dieser Kovarianzanalyse mit Messwiederholung aufgrund der zu kleinen Zellbesetzungen nicht möglich. Zudem ist die Klassenzugehörigkeit abhängig von der Zugehörigkeit zur Experimental- und Kontrollgruppe.

In Tabelle 13 sind die Ergebnisse der Kovarianzanalyse dargestellt. Der Haupteffekt Gruppe ist nicht signifikant ($F(1, 117) = .16$, ns), der Zeiteffekt ist auch nicht signifikant ($F(1, 117) = .6$, ns), der Interaktionseffekt ist aber signifikant ($F(1, 117) = 4.8$, $p < .05$). In der Interventionsgruppe nimmt die durchschnittliche externale Attribution von schulischem Erfolg ab, während in der Kontrollgruppe die durchschnittliche externale Attribution von schulischem Erfolg zunimmt. Dieser Effekt zeigt, dass die Intervention in Modul B dazu beigetragen hat, die Erfolgsattribution von Lehrpersonen in die erwünschte Richtung zu korrigieren.

Tabelle 13: Veränderung der Attribution durch Intervention (Mittelwerte mit Standardmessfehlern)

	Interventions- gruppe	Kontrollgruppe
externale Attribution t ₁	41.01 (2.9)	37.6 (2.4)
externale Attribution t ₂	37.9 (3.1)	44.2 (2.5)

Legende: Skala: 0-100, Kontrollvariablen: SES, Migrationshintergrund Noten in Deutsch und Mathematik t₁.

6.4 Unterrichtsformen in heterogenen Klassen (Modul C)

Die Wirksamkeit der Intervention mit Modul C sollte anhand von zwei Items überprüft werden, nämlich ob die Lehrpersonen berichteten, nach der Intervention häufiger kooperative Lehrformen einzusetzen und ob sich der Einsatz digitaler Medien zur Förderung von Kindern mit wenig familiärer Unterstützung veränderte. Leider nahm keine Schule an dieser Evaluation teil, welche das Modul C besuchte, so dass die Wirksamkeit von Modul C leider nicht überprüft werden kann.

Die Module A und B thematisieren die Förderung von Kindern mit Benachteiligungen und sprechen am Rande auch didaktische Formen wie den Einsatz von kooperativen Lehrformen und den Einsatz digitaler Medien an. Explorativ wird daher untersucht, ob der Besuch der Module A oder B auch diese didaktischen Formen beeinflusste. Die Stichprobe umfasst 28 Lehrpersonen in der Interventionsgruppe und 18 Lehrpersonen in der Kontrollgruppe und ist daher sehr klein. Gemäss Tabelle 144 nimmt der regelmässige Einsatz kooperativer Lernformen in kulturell heterogenen Klassen marginal zu, in der Kontrollgruppe nimmt er ab. Dieser erwartete Trend ist aber nicht signifikant

Schlussbericht: Evaluation der SCALA-Intervention

(Haupteffekt: $F(1, 44) = .86$, ns; Zeiteffekt: $F(1, 44) = .40$, ns; Interaktionseffekt: $F(1, 44) = .76$, ns). D.h. es gibt keinen Interventionseffekt von Modul A bzw. B auf diese didaktischen Unterrichtsformen.

Tabelle 14: Regelmässiger Einsatz kooperative Lernformen in kulturell heterogenen Klassen, um den sozialen Zusammenhalt zu stärken (least square means)

	Interventions- gruppe	Kontrollgruppe
Pretest	4.61 (.17)	4.78 (.21)
Posttest	4.64 (.18)	4.56 (.23)

Der Einsatz digitaler Medien in heterogenen Klassen durch die Lehrpersonen in der Interventionsgruppe und Kontrollgruppe unterscheidet sich nicht signifikant (Haupteffekt: $F(1, 44) = 0.02$, ns; Zeiteffekt: $F(1, 44) = 4.01$, ns; Interaktionseffekt: $F(1, 44) = .54$, ns). In der Interventionsgruppe nimmt die Nutzung digitaler Medien während des Schuljahres leicht ab, in der Kontrollgruppe zu (Tabelle 15). Möglicherweise besuchten Lehrpersonen der Kontrollgruppe Interventionen, die sich mit dem Einsatz digitaler Medien beschäftigen ausserhalb von SCALA, weil diese Thematik in diesen Schulen eine hohe Aktualität besitzt. Grundsätzlich gibt es keinen wesentlichen Unterschied zwischen den Gruppen.

Tabelle 15: Einsatz digitaler Medien in heterogenen Klassen (least square means)

	Pretest	Posttest
Interventionsgruppe	3.93 (.24)	3.83 (.29)
Kontrollgruppe	4.14 (.29)	4.33 (.28)

Zusammenfassend führt der Besuch der Module A und B nicht zu einer Änderung im didaktischen Verhalten bezüglich kooperativen Lernens und dem Einsatz von digitalen Medien. Dieser Effekt wurde auch nicht erwartet, weil die Modul A und B diese didaktischen Formen nicht explizit thematisierten.

7 Diskussion

In diesem Bericht sollte dargestellt werden, wie sehr die schulinterne SCALA-Weiterbildung dazu beiträgt, Beurteilungsverzerrungen von Lehrpersonen aufgrund von Stereotypen zu reduzieren. Leider sind aufgrund der geringen Beteiligungsbereitschaft der Lehrpersonen die Stichproben klein. Daher müssen die Ergebnisse zurückhaltend interpretiert werden.

Für Modul A weist das Ergebnis, dass die soziale Herkunft (SES) die Lehrpersonenerwartungen in Deutsch in der Kontrollgruppe aber nicht in der Interventionsgruppe beeinflusst, in die erwartete Richtung. Für eine Schule, in welcher eine hinreichend grosse Gruppe von Lehrpersonen an der Studie mitgemacht haben, wird dieser Effekt statistisch signifikant. Offenbar gelingt es in der Umsetzung ansatzweise den Effekt der sozialen Herkunft auf die Lehrpersonenerwartungen in Deutsch zu reduzieren. Dies belegt, dass die Umsetzung der Intervention in der Schule wirksam sein kann.

Für Mathematik wurde dieser Effekt nicht gefunden, weil es in der Kontrollgruppe wider Erwarten für Mathematik keinen Herkunftseffekt auf die Lehrpersonenerwartungen gibt. Daher konnte kein Interventionseffekt auftreten. Ebenfalls finden sich keine Interventionseffekte der kulturellen Herkunft auf Lehrpersonenerwartungen in Deutsch und Mathematik.

Der Interventionseffekt für den SES in dieser Schule ist deshalb erfreulich, weil der SES-Effekt in der ursprünglichen Studie von Neuenschwander et al. (2021) nicht gefunden worden ist. Dem SES wurde in der adaptierten Intervention mehr Gewicht beigemessen. Weiter ist dieser Befund erwartungskonform, weil der Effekt des SES auf die Erwartungen in verschiedenen Studien besonders ausgeprägt ist (Neuenschwander & Niederbacher, 2021). Der verstärkte Einbezug der Diskussion zum SES in die Intervention im Unterschied zur ursprünglichen Intervention hat sich offenbar gelohnt.

Ein anderes Ergebnis zeigt, dass Lehrpersonen nach Besuch von Modul B schulischen Erfolg ihrer Schüler*innen eher internal erklären und dabei den Effekt der familiären, kulturellen und sozialen Herkunft reduzieren. Das ist ein wichtiges Ziel von Modul B. Das Ergebnis zeigt, dass der Besuch von Modul B diese erwarteten, günstigen Effekte haben kann.

Die Wirksamkeit von Modul C konnte leider nicht überprüft werden, weil keine Schule, welche das Modul C besucht hat, an dieser Wirkungsstudie teilgenommen hat.

Obwohl die Umsetzung der Intervention ins Schulfeld sehr anspruchsvoll ist und obwohl sehr strenge Erfolgsmasse (effektive Beurteilungen von Schüler*innen statt Lehrpersonen-Einschätzungen, Kontrollgruppendesign) verwendet wurden, wurden Interventionseffekte nachgewiesen. Die erfolgreiche Umsetzung in die QUIMS-Schulen kann also gelingen. Ergebnisse zeigen, dass auch bei einer breit angelegten Umsetzung einer Intervention die erwünschten Effekte auftreten können.

7.1 Wirkmechanismen

Eine wichtige Frage zur Interpretation der Ergebnisse lautet, welche Wirkmechanismen diese Ergebnisse erklären. Das vorliegende kleine Projekt erlaubt keine detaillierte Analyse der Wirkmechanismen. Doch erlaubt der Vergleich der ursprünglichen Intervention (Neuenschwander et al., 2021) mit der hier realisierten Intervention gewisse Spekulationen, wie die Effekte entstanden sein könnten. Die folgenden Überlegungen können aber zur Beantwortung dieser Frage beitragen.

Prinzipien der Intervention sind die Folgenden: (a) Vermittlung von einschlägigem Wissen und dieses mit vielen Beispielen illustrieren (b) Betroffenheit schaffen und zu Einsicht/Reflexion motivieren, Aha-Erlebnis schaffen, (c) Bezug zu eigenen Schulklassen herstellen (Praxistransfer). (d) Interventionstermine über ca. 6 Monate verteilen, immer wieder an Inhalte erinnern. Das erhöht die Chance auf Umsetzung in den Unterricht. (e) Hohes Engagement/Enthusiasmus der Schulleitung, diese kann an pädagogischen Konferenzen die Lehrpersonen immer wieder an die Thematik erinnern und die Diskussion dazu schulintern weiterführen. Diese Prinzipien waren sowohl im ursprünglichen Projekt als auch in der Implementation der Intervention in die Breite leitend und könnten möglicherweise zu einer Erklärung der Wirkmechanismen beitragen.

Die ursprüngliche Intervention umfasste 4 Halbtage und war als kursorisches Angebot angelegt. D.h. motivierte Lehrpersonen aus ganz unterschiedlichen Schulen, welche an den Klassenstufen 4-6 unterrichteten, konnten an der Intervention teilnehmen. Ein grosser Teil der Interventionsinhalte bezog sich auf die interkulturelle Kommunikation, auf die Reflexion über Migration und wie sich diese auf die Schüler*innen im Unterricht auswirken könnte.

Im Unterschied dazu wurde die hier vorliegende Intervention mit einem ganzen Lehrkollegium in 2.5 Halbtagen durchgeführt. D.h. die Lehrpersonen aller Stufen an der jeweiligen Schule nahmen teil. Die Plenumsteile wurden eingebettet in Begleitgruppensitzungen mit der erweiterten Schulleitung durchgeführt, zudem wurden Praxisaufträge zur Anwendung der Inhalte auf die eigene Klasse

durchgeführt. Das Konzept des sozio-ökonomischen Status erhielt im Vergleich zur Migrationsdebatte mehr Raum.

Auch wenn nicht beurteilt werden kann, welche Massnahme sich wie ausgewirkt hat, scheint die zeitliche Kürzung der Intervention die Wirkung nicht reduziert zu haben. Umgekehrt waren die Effekte der schulinternen Intervention im Vergleich zur kursorischen Intervention nicht stärker. Möglicherweise ist es wichtig, wie ehrlich und engagiert die einzelnen Lehrpersonen bereit sind, sich mit der Thematik auseinanderzusetzen und auf den eigenen Unterricht zu beziehen. Beispielsweise schafft ein Praxisauftrag die Möglichkeit des Transfers von der Intervention in den eigenen Unterricht. Ob dieser Transfer aber gelingt, hängt davon ab, ob sich die Lehrpersonen mit der Thematik ganz persönlich auseinandersetzen. Es stellt sich daher die Frage, wie jede einzelne Lehrperson für die Sensibilisierung der Thematik angesprochen werden kann.

7.2 Herausforderungen

Die Ergebnisse zeigen eine Reihe von Herausforderungen.

Zuerst zeigen die Daten die Schwierigkeiten, eine genügend grosse Anzahl von Lehrpersonen zu finden, welche bereit sind, an einer solchen Studie mitzuwirken. Aufgrund der eher kleinen Stichprobe sind die Effekte eher klein, aber doch immer wieder signifikant. Die Schulen konnten nicht zufällig der Interventionsgruppe bzw. der Kontrollgruppe zugewiesen werden, wie ein Experimentaldesign dies forderte. Die vorliegende Studie wollte untersuchen, wie die Umsetzung in die Breite gelang, weshalb eine randomisierte Zuteilung zu den Bedingungen nicht möglich und sinnvoll war. An einigen Schulen ist die Zahl der Lehrpersonen, die teilgenommen haben, zu klein, um schulspezifische aussagekräftige Ergebnisse zu ermöglichen.

In Bezug auf die erfolgreiche Durchführung von Weiterbildungen zeigen sich folgende Herausforderungen.

- Die schulischen Voraussetzungen beim Start einer schulinternen Weiterbildung sind sehr heterogen. Es gibt Schulen mit einem sehr harmonischen Umgang untereinander, während in anderen Schulen heimliche oder offene Konflikte im Kollegium oder/und gegenüber der Schulleitung vorliegen. Während es in manchen Schulen eine hohe Fluktuation der Lehrpersonen gibt, sich die Lehrpersonen überlastet fühlen oder die Schüler*innenzusammensetzung sehr heterogen ist, ist die Personalsituation in anderen Schulen stabil und wertschätzend. Konflikte im Kollegium können die offene und ehrliche Auseinandersetzung mit der Thematik erschweren, weil sie mit einer Angst vor Blossstellung und Kritik verbunden sein kann. Die Einstellung der Lehrpersonen gegenüber der Weiterbildung variiert, je nachdem wie die Lehrpersonen in die Entscheidung einbezogen worden sind, ob diese Weiterbildung an der Schule durchgeführt werden sollte.
- Die Bereitschaft der Lehrpersonen, sich mit Vorurteilen und Stereotypen auseinanderzusetzen, variiert sehr stark. Während manche Lehrpersonen sagen, dass sie bereits ausreichend für die Thematik sensibilisiert worden sind, setzen sich andere Lehrpersonen engagiert und offen mit dem Thema (erneut) auseinander. Die Arbeit an Haltungen/Überzeugungen/Erwartungen/Attributionen erfordert eine grosse Ehrlichkeit und Offenheit von Lehrpersonen, sich engagiert auseinanderzusetzen.
- Unklar ist, ob die Effekte der Intervention nachhaltig sind, d.h. es ist unklar, ob Lehrpersonen, die ihre Einstellungen während der Intervention im Hinblick auf mehr Fairness korrigiert haben, diese unverzerrten Einstellungen auch auf die Schüler*innen einer neuen Klasse anwenden können. Es ist zu prüfen, ob diese selbstkritische Auseinandersetzung bei jeder Klasse neu erfolgen oder ob sie auf neue Schüler*innen angewendet werden kann. Die Prüfung dieser Transferleistung auf eine neue Klasse erfordert ein aufwändiges Forschungsdesign, das mit den vorliegenden Mitteln nicht realisiert werden konnte.

7.3 Ausblick

Trotz dieser Limitationen und trotz des anspruchsvollen Versuchsdesigns konnten gewisse wichtige Interventionseffekte gefunden werden. Dabei soll nochmals betont werden, dass es sich aufgrund der Kontrollgruppenstudie mit Wiederholungsmessung um effektive Wirkungen handelt und nicht um Wirkungs- oder Akzeptanzeinschätzungen. Das erhöht die Relevanz der Ergebnisse. Aus diesen Gründen lohnt es sich, das SCALA-Angebot weiterzuführen. Die oben formulierten Prinzipien für die Ausgestaltung des Angebots waren für die Entwicklung der weiteren Module (Modul D (Elternarbeit), E (Unterrichtsstörungen und Beurteilung), F (diskriminierungskritische Schulen) leitend.

Die Ergebnisse zeigen, dass mit der Durchführung des SCALA-Ansatzes alleine die Bildungsgerechtigkeit in der Schweiz nicht ausreichend erhöht werden kann. Es wird angenommen, dass diese Intervention dazu zwar einen Beitrag leisten kann, aber dass dieses Ziel ohne weitere Massnahmen nicht erreicht wird. D.h. neben der Thematisierung von Vorurteilen braucht es weitere Massnahmen zur Erhöhung der Chancengerechtigkeit im Bildungskontext.

Aus unserer Sicht sind Massnahmen auf institutioneller Ebene, auf Schulebene und Unterrichtsebene denkbar. Auf institutioneller Ebene stehen die äussere Gliederung des Schulsystems, vor allem in der Sekundarstufe I und II, sowie die Verfahren der Selektion, die der Zuweisung in diese verschiedenen Formen zugrunde liegen, im Vordergrund. Wenn die soziale und die kulturelle Herkunft die Zuweisung von Schüler*innen in ein Schulniveau beeinflusst, verstärken sich mit der äusseren Gliederung des Bildungssystem Herkunft- und Kultureffekte und damit die Chancenungerechtigkeit. Dies wurde in vielen Studien und Berichten dokumentiert (z.B. Becker & Schoch, 2018). Die Aufhebung oder Reduktion der äusseren Gliederung könnte sekundäre Herkunftseffekte nach Boudon (1974) reduzieren. Neuenchwander (2017) zeigte überdies, dass die Leistungszunahme der Mehrheit der Schüler*innen grösser ist, wenn der Übertritt in die gegliederte Sekundarstufe I nach dem 6. Schuljahr statt nach dem 5. Schuljahr erfolgt.

Weitere Ergebnisse zeigen, dass die Schule (bzw. Lehrpersonen) die Schüler*innen in ein Niveau der Sekundarstufe I zuweisen sollten und dass die Eltern darüber nur informiert werden (Neuenchwander & Schneitter, 2014). Je mehr Mitbestimmung Eltern bei der Schulniveauzuweisung haben, desto stärkere Herkunftseffekte treten auf. D.h. aus der Perspektive der Chancengerechtigkeit sollten die Lehrpersonen die Selektionsmacht bekommen, weil sie am ehesten unparteiische Entscheidungen treffen können. Darum ist es wichtig, dass Lehrpersonen faire Urteile abgeben. Es könnte zudem sinnvoll sein, Lehrpersonen vermehrt über die rechtlichen Grundlagen der Selektion aufzuklären, so dass souveräne Antworten auf juristische Fragen bei Bedarf gegeben und Unsicherheiten bei juristischen Anfragen reduziert werden können.

Auf Schulebene sind Konzepte nötig, wie Lehrpersonen zu einer von den Elternerwartungen unabhängigen Förderung und Beurteilung der Schüler*innen gelangen. Die Förderung sollte sich am kognitiven Potenzial der Schüler*innen orientieren und nicht an den Erwartungen der Eltern oder am Verhalten der Schüler*innen (vgl. SCALA Modul E). Ein Bildungssystem wird chancengerechter, wenn das Potenzial bzw. die Leistungen der Schüler*innen im Zentrum stehen. Dazu gehört, dass Lehrpersonen Kinder und Jugendliche mit hohem Potenzial identifizieren und zusätzlich fördern. Insbesondere Schüler*innen mit tieferem sozio-ökonomischem Status, deren Potenzial nicht ausgeschöpft ist, könnten dadurch profitieren. Das ist einerseits eine liberale Forderung, wonach nicht die Herkunft, sondern die Leistungen der Kinder und Jugendlichen über die Chancen bestimmen sollen. Andererseits wird damit auch das soziale Programm realisiert, dass das Potenzial von allen Kindern gefördert werden soll. An Schulen sollte sich daher eine Kultur etablieren, wonach auch versteckte Diskriminierung verhindert und die Förderung der einzelnen Kinder bzw. deren faire Beurteilung aktiv gefördert wird. Auch wenn kaum eine Schule diesen Zielen widersprechen würde, können sie im intensiven und vielschichtigen Unterrichtsalltag leicht in den Hintergrund geraten.

8 Empfehlungen

1. Die Prinzipien der SCALA Weiterbildung (Information, Sensibilisierung, Anwendung auf eigene Klasse) sollten bei zukünftigen Schilw beachtet werden.
2. Der SCALA-Ansatz bewährt sich grundsätzlich, ist wirksam und wird nachgefragt. Daher sollte die neuen Module (D, E, F) nun durchgeführt werden.
3. Diskriminierung und Chancenungleichheit kann nicht alleine auf der Unterrichtsebene beseitigt werden, sondern braucht komplementäre Massnahmen auf der Ebenen der Schule und des Bildungssystems (Gliederung Sek I, Gestaltung Selektionsverfahren, Rückhalt von Lehrpersonen, die sich für gerechte Schullaufbahnentscheidungen einsetzen).
4. Migrationsthemen wurden an Quims-Schulen breit diskutiert. Es braucht aber mehr Auseinandersetzung mit Wirkungen der sozialen Herkunft und der Hautfarbe auf Förderung und Beurteilung.
5. Mehr wissenschaftliche Begleitforschung ist nötig, um die Qualität und Innovation zu sichern.

9 Literatur

- Babad, E., Bernieri, F., & Rosenthal, R. (1989). Nonverbal Communication and Leakage in the Behavior of Biased and Unbiased Teachers. *Journal of Personality and Social Psychology*, 56(1), 89-94.
- Becker, R., & Schoch, J. (2018). *Soziale Selektivität*. Schweizer Wissenschaftsrat.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. Wiley.
- Cohen, Geoffrey L., & Garcia, Julio. (2008). Identity, Belonging, and Achievement. *Current Directions in Psychological Science*, 17(6), 365–369.
- Froehlich, L., Martiny, S. E., Deaux, K., & Mok, S. Y. (2015). “It’s Their Responsibility, Not Ours”. *Social Psychology*, 47(2), 74–86. <https://doi.org/10.1027/1864-9335/a000260>.
- Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1-56. [https://doi.org/10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B).
- Inan-Kaya, G., & Rubie-Davies, C. M. (2021). Teacher classroom interactions and behaviours: Indications of bias. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2021.101516>
- Neuenschwander, M. P., & Niederbacher, E. (2021). Disparitäten in Anstrengungsbereitschaft und Leistung nach SES, Familiensprache und Geschlecht: Folgen von Sozialisation oder von Diskriminierung durch verzerrte Lehrpersonenerwartungen. *Zeitschrift für Soziologie der Sozialisation und Erziehung*, 41(4), 449–466. <https://doi.org/10.3262/ZSE2104449>
- Neuenschwander, M. P., Fräulin, J., Schumann, S., & Jüttler, A. (2018). Gymnasium oder Berufsausbildung? Zur Vorhersage der Bildungsentscheidung in der Sekundarstufe II durch Schülerleistungen sowie Leistungserwartungen von Lehrpersonen und Eltern in der Primarschulzeit. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 38(4), 364–382.
- Neuenschwander, M. P. (2017). Anpassungsprozesse beim Übergang in die Sekundarstufe I. In M. P. Neuenschwander & C. Nägele (Hrsg.), *Bildungsverläufe von der Einschulung bis in den ersten Arbeitsmarkt: Theoretische Ansätze—Empirische Befunde—Beispiele* (S. 143–161). VS Verlag für Sozialwissenschaften.
- Neuenschwander, M. P., & Schneitter, J. (2014). *Analyse des Selektionsverfahrens in die Sekundarstufe I im Kanton Bern*. Zentrum Lernen und Sozialisation der PH FHNW.
- Neuenschwander, M. P., Mayland, C., Niederbacher, E., & Garrote, A. (2021). Modifying biased teacher expectations in mathematics and German: A teacher intervention study. *Learning and Individual Differences*, 87, 101995. <https://doi.org/10.1016/j.lindif.2021.101995>
- Niederbacher, E., & Neuenschwander, M. P. (2020). Herkunftsbedingte Leistungsdisparitäten: Die Rolle von Selbstwirksamkeitsüberzeugungen und Unterstützungshandlungen von Eltern und Leistungserwartungen von Lehrpersonen Generalisierbarkeit eines Mediationsmodells für einsprachige und fremd- bzw. Mehrsprachige Schülerinnen und Schüler. *Zeitschrift für Erziehungswissenschaft*. <https://doi.org/10.1007/s11618-020-00955-9>
- Niederbacher, E., & Neuenschwander, M. P. (2021). Erklärung von Schulniveauerwartungen von Eltern und Lehrpersonen durch Herkunftsmerkmale und Überzeugungen und Folgen von Erwartungsdiscrepanzen für Leistungen. *Empirische Pädagogik*, 12(1), 164-182.
- Reyna, C. (2000). Lazy, Dumb, or Industrious: When Stereotypes Convey Attribution Information in the Classroom. *Educational Psychology Review*, 12(1), 85–110.
- Reyna, C. (2008). Ian is intelligent but Leshawn is lazy: Antecedents and consequences of attributional stereotypes in the classroom. *European Journal of Psychology of Education*, 13(439–458).

Rosenthal, R., & Jacobson, L. (1974). *Pygmalion im Unterricht*. Belz.

Rubie-Davies, C. M., Peterson, E. R., Sibley, C. G., & Rosenthal, R. (2015). A teacher expectation intervention: Modelling the practices of high expectation teachers. *Contemporary Educational Psychology, 40*, 72–85. <https://doi.org/10.1016/j.cedpsych.2014.03.003>

Slavin, R. E. (1990). Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis. *American Educational Research Journal, 60*(3), 471–499.

Wang, H., & Hall, N. C. (2018). A Systematic Review of Teachers' Causal Attribution: Prevalance, Correlates, and Consequences. *Frontiers in Psychology, 9*, 1–22.

Wang, S., Rubie-Davies, C. M., & Meissel, K. (2018). A systematic review of the teacher expectation literature over the past 30 years. *Educational Research and Evaluation, 24*(3–5), 124.179. <https://doi.org/10.1080/13803611.2018.1548798>

Weiner, B. (1985). An Attributional Theory of Achievement Motivation and Emotion. *Psychological Review, 92*(4), 548-573.

Weiner, B. (2010). The Development of an Attribution-Based Theory of Motivation: A History of Ideas. *Educational Psychologist, 45*(1), 28–36.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267–301.