

### Effekte eines Unterrichtsfeedbacks von Schulleitungen auf die selbst eingeschätzte Entwicklung der Unterrichtskompetenz von Lehrkräften: Ergebnisse einer Interventionsstudie unter Berücksichtigung der Zielorientierungen von Lehrkräften

Kellermann, Christopher; Nachbauer, Max; Gärtner, Holger; Thiel, Felicitas

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) - Projektnummer 283241518 / Funded by the German Research Foundation (DFG) - Project number 283241518

#### Empfohlene Zitierung / Suggested Citation:

Kellermann, C., Nachbauer, M., Gärtner, H., & Thiel, F. (2022). Effekte eines Unterrichtsfeedbacks von Schulleitungen auf die selbst eingeschätzte Entwicklung der Unterrichtskompetenz von Lehrkräften: Ergebnisse einer Interventionsstudie unter Berücksichtigung der Zielorientierungen von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 25, 1267-1291. <https://doi.org/10.1007/s11618-022-01087-y>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

# Effekte eines Unterrichtsfeedbacks von Schulleitungen auf die selbst eingeschätzte Entwicklung der Unterrichtskompetenz von Lehrkräften – Ergebnisse einer Interventionsstudie unter Berücksichtigung der Zielorientierungen von Lehrkräften

Christopher Kellermann  · Max Nachbauer  · Holger Gärtner  · Felicitas Thiel 

Eingegangen: 24. November 2020 / Überarbeitet: 21. November 2021 / Angenommen: 25. Januar 2022 /  
Online publiziert: 25. April 2022  
© Der/die Autor(en) 2022

**Zusammenfassung** Unterrichtsfeedbacks von Schulleitungen sind häufig von geringer Qualität und daher nicht immer wirksam. Aus diesem Grund wurde eine Feedbackintervention entwickelt, die Schulleitungen dabei unterstützen soll, Lehrkräften ein qualitativ hochwertiges Feedback zu ihrem Unterricht zu geben, auf dessen Grundlage diese ihre Unterrichtskompetenz weiterentwickeln können. Die Wirksamkeit der Intervention wurde in einer Studie mit quasi-experimentellem Design überprüft. Im vorliegenden Beitrag wird der Frage nachgegangen, ob Lehrkräfte, die ein Unterrichtsfeedback von ihrer Schulleitung erhielten, ihre Kompetenzentwicklung positiver einschätzen als Lehrkräfte, die kein Feedback erhielten. Weil Lehrkräfte der Interventionsgruppe ihre Kompetenzentwicklung nach dem Unterrichtsfeedback im Mittel nicht positiver einschätzten als Lehrkräfte der Kontrollgruppe, wurde zusätzlich geprüft, ob der Effekt des Unterrichtsfeedbacks mit den Zielorientierungen der Lehrkräfte interagiert. Für die Lern- und Annäherungsleistungszielorientierung zeigten sich differenzielle Effekte. Lehrkräfte mit hoch ausgeprägter Lern- und Annäherungsleistungszielorientierung, die ein Feedback erhielten, schätzten ihre Kompetenzentwicklung positiver ein als Lehrkräfte der Kontrollgruppe.

---

Christopher Kellermann (✉) · Max Nachbauer · Prof. Dr. Felicitas Thiel  
Fachbereich Erziehungswissenschaft und Psychologie, Arbeitsbereich  
Schulpädagogik/Schulentwicklungsforschung, Freie Universität Berlin,  
Habelschwerdter Allee 45, 14195 Berlin, Deutschland  
E-Mail: christopher.kellermann@fu-berlin.de

Max Nachbauer  
E-Mail: max.nachbauer@fu-berlin.de

Prof. Dr. Felicitas Thiel  
E-Mail: felicitas.thiel@fu-berlin.de

Prof. Dr. Holger Gärtner  
Fachbereich Erziehungswissenschaft und Psychologie, Arbeitsbereich Unterrichts- und  
Schulevaluation, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Deutschland  
E-Mail: holger.gaertner@fu-berlin.de

**Schlüsselwörter** Unterrichtsbeobachtung · Feedback · Professionalisierung von Lehrkräften · Zielorientierungen

## **Effects of school leaders' feedback on teaching on teachers' self-rated instructional competence development—Results of an intervention study considering teachers' goal orientations**

**Abstract** Feedback by school leaders often is of poor quality and thus not always effective. Therefore, a feedback intervention was developed to assist school leaders in providing high-quality feedback to teachers based on which they can improve their instructional competence. The effectiveness of the intervention was tested within a quasi-experimental study. The present paper examines whether teachers who received feedback from their school leader rate their competence development more positively than teachers who did not receive feedback. After receiving feedback, teachers of the intervention group on average did not report a higher competence development than teachers of the control group. Therefore, interactions of school leaders' feedback and teachers' goal orientations were examined. With regard to teachers' learning goal orientation and performance approach goal orientation, an interaction was found. Teachers with a pronounced learning goal orientation and performance approach goal orientation who received feedback reported a higher competence development than teachers of the control group.

**Keywords** Classroom observation · Feedback · Goal orientation · Teachers' professional development

### **1 Einleitung**

Die Unterrichtskompetenz von Lehrkräften nimmt entscheidenden Einfluss auf die Qualität des in Schulen angebotenen Unterrichts und damit auf den Lernerfolg von Schülerinnen und Schülern (vgl. Lipowsky 2006; Muijs et al. 2014). Ein beobachtungs-basiertes Unterrichtsfeedback von Schulleitungen wird als erfolgversprechende Maßnahme diskutiert, die zu einer Weiterentwicklung der Unterrichtskompetenz beitragen und – vermittelt über eine verbesserte Unterrichtsqualität – einen positiven Effekt auf den Lernerfolg von Schülerinnen und Schülern haben kann. Studienergebnisse zeigen jedoch, dass Unterrichtsfeedbacks nicht immer wirksam sind (vgl. Kraft et al. 2018; Steinberg und Sartain 2015; Taylor und Tyler 2012). Eine mögliche Ursache für die teilweise ausbleibenden Feedbackeffekte ist die häufig geringe Qualität eines Feedbacks von Schulleitungen (vgl. Kimball und Milanowski 2009; Kraft und Gilmour 2016; Milanowski und Heneman 2001). Aus diesem Grund wurde eine Feedbackintervention entwickelt, die Schulleitungen dabei unterstützen soll, Lehrkräften ein qualitativ hochwertiges Feedback zu ihrem Unterricht zu geben, auf dessen Grundlage diese ihre Unterrichtskompetenz weiterentwickeln können. Die Wirksamkeit der Feedbackintervention wurde in einer Studie mit quasi-experimentellem Design überprüft. Im vorliegenden Beitrag wird der Frage nachgegangen, ob

Lehrkräfte, die ein Unterrichtsfeedback von ihrer Schulleitung erhielten, ihre Kompetenzentwicklung positiver einschätzen als Lehrkräfte, die kein Feedback erhielten.

Vor dem Hintergrund von Angebots-Nutzungs-Modellen, die die Wirkungsweise verschiedener Lerngelegenheiten auf die Kompetenzentwicklung von Lehrkräften beschreiben (vgl. Kunter et al. 2011; siehe auch Helmke 2014), ist anzunehmen, dass die Wirksamkeit von Unterrichtsfeedback nicht allein von der Feedbackqualität abhängt, sondern auch von der Feedbacknutzung, die ihrerseits von verschiedenen kognitiven und motivationalen Merkmalen der Lehrkraft beeinflusst wird. Hinsichtlich der motivationalen Merkmale haben sich Zielorientierungen als bedeutsam für die Nutzung verschiedener Lerngelegenheiten erwiesen (vgl. z. B. Butler 2007; Janke et al. 2019; Nitsche et al. 2013a, b). Daher wird zusätzlich geprüft, ob der Effekt des Unterrichtsfeedbacks auf die Kompetenzselbsteinschätzungen mit den Zielorientierungen der Lehrkräfte interagiert.

## 2 Theoretischer Rahmen und Forschungsbefunde

### 2.1 Beobachtungsbasiertes Unterrichtsfeedback und die *Feedback Intervention Theory*

Feedback kann als „bewusste Rückmeldung von Informationen an eine Person zu ihrem vorherigen Verhalten“ (Müller und Ditton 2014, S. 15) definiert werden. Ein beobachtungsbasiertes Unterrichtsfeedback stellt dementsprechend eine Rückmeldung von Informationen an eine Lehrkraft zu ihrem in einer beobachteten Unterrichtsstunde gezeigten Verhalten dar. Einem beobachtungsbasierten Unterrichtsfeedback liegt in der Regel ein theoretisches Unterrichtsmodell zugrunde, das lernwirksames Unterrichtsverhalten von Lehrkräften beschreibt. Das beschriebene Unterrichtsverhalten stellt den Soll-Standard dar, von dem ausgehend die Schulleitung das in der beobachteten Unterrichtsstunde gezeigte Verhalten einer Lehrkraft, den Ist-Stand, beurteilt. Das Unterrichtsfeedback enthält schließlich einen Vergleich des beobachteten Verhaltens mit dem im Unterrichtsmodell definierten Soll-Standard.

Die Wirkungsweise von Feedback lässt sich anhand der *Feedback Intervention Theory* (FIT) beschreiben (vgl. Kluger und DeNisi 1996). Nach der FIT ist Verhalten zielorientiert und wird über einen Soll-Ist-Vergleich reguliert. Die FIT unterscheidet vereinfacht dargestellt zwischen einer aufgaben- und einer selbstbezogenen Feedbackverarbeitung. Bei einer aufgabenbezogenen Feedbackverarbeitung zieht eine Soll-Ist-Diskrepanz eine erhöhte Anstrengungsbereitschaft nach sich (vgl. Kluger und DeNisi 1996): Macht ein kritisches Unterrichtsfeedback eine Lehrkraft auf eine Diskrepanz zwischen ihren aktuellen und den im Unterrichtsmodell beschriebenen Verhaltensweisen aufmerksam, sollte sie nach den Annahmen der FIT darum bemüht sein, diese Diskrepanz zu reduzieren. Bei einem bestätigenden Unterrichtsfeedback ist es denkbar, dass der Diskrepanzreduktion eine Diskrepanzproduktion vorausgeht (vgl. Bandura und Locke 2003), d. h. Lehrkräfte setzen sich nach einem bestätigenden Unterrichtsfeedback ggf. höhere Ziele und erzeugen so eine Diskrepanz zwischen ihrem aktuellen und einem erwünschten Unterrichtsverhalten. Führt ein Mehraufwand allein nicht zu einer Diskrepanzreduktion, werden nach den Annah-

men der FIT neue Problemlösestrategien entwickelt, die zu einer Reduktion der Soll-Ist-Diskrepanz führen können: „If working harder fails, people may try to work smarter by generating a hypothesis regarding means for improved performance“ (Kluger und DeNisi 1996, S. 263). Lehrkräfte sollten demnach neue Verhaltensstrategien für ihren Unterricht entwickeln, die – sofern sie sich bewähren – zu einer Reduktion der Diskrepanz zwischen ihren aktuellen und den erwünschten Verhaltensweisen führen. Neben einer aufgabenbezogenen Feedbackverarbeitung kann die Feedbackverarbeitung jedoch auch selbstbezogen erfolgen, was u. a. bei einem sehr kritischen Feedback zu erwarten ist (vgl. Ilgen und Davis 2000). Eine selbstbezogene Feedbackverarbeitung geht nach der FIT mit negativen Emotionen einher, hat einen Abzug von kognitiven Ressourcen von der Aufgabenbearbeitung zur Folge und kann zu einer Ablehnung des Feedbacks führen (vgl. Kluger und DeNisi 1996). Lehrkräfte könnten sich infolge eines sehr kritischen Unterrichtsfeedbacks z. B. fragen, ob sie für den Lehrerberuf ungeeignet sind und – um ihren Selbstwert zu schützen – das Unterrichtsfeedback zurückweisen.

Aus der FIT lässt sich schlussfolgern, dass Unterrichtsfeedback bei einer selbstbezogenen Feedbackverarbeitung weitgehend wirkungslos bleiben sollte. Erfolgt die Feedbackverarbeitung jedoch aufgabenbezogen, kann es zu einer Weiterentwicklung der Unterrichtskompetenz von Lehrkräften beitragen. In Anlehnung an das Kompetenzmodell der COACTIV-Studie (vgl. Baumert und Kunter 2006, 2011), wird unter Unterrichtskompetenz im vorliegenden Beitrag das für die Planung und Durchführung von Unterricht relevante professionelle Wissen und Können von Lehrkräften verstanden, das sich in ihrem Unterrichtsverhalten ausdrückt.

## 2.2 Forschungsstand zur Wirksamkeit von beobachtungsbasiertem Unterrichtsfeedback

Die Wirksamkeit eines beobachtungsbasierten Unterrichtsfeedbacks von Schulleitungen auf die Weiterentwicklung der Unterrichtskompetenz von Lehrkräften wurde nach unserem Kenntnisstand noch nicht erforscht. Es liegen aber Studien vor, die die Effekte von beobachtungsbasiertem Unterrichtsfeedback auf das Unterrichtsverhalten von Lehrkräften und den Lernerfolg von Schülerinnen und Schülern untersuchen. So zeigen Taylor und Tyler (2012) in einer quasi-experimentellen Studie, dass ein mehrmaliges Feedback von Schulleitungen und von vom Unterricht freigestellten Lehrkräften, das u. a. auf einer Unterrichtsbeobachtung basiert, einen positiven Effekt auf den Lernzuwachs von Schülerinnen und Schülern im Fach Mathematik hat. Auf den Lernzuwachs im Lesen hatte das Unterrichtsfeedback demgegenüber keinen Effekt (vgl. Taylor und Tyler 2012, S. 3648). Steinberg und Sartain (2015) berichten in einer experimentellen Studie von einem positiven Effekt eines beobachtungsbasierten Unterrichtsfeedbacks von Schulleitungen auf den Lernzuwachs im Lesen, jedoch nicht im Fach Mathematik. Weitere Hinweise auf die Wirksamkeit von Unterrichtsfeedbacks liefert eine aktuelle Meta-Analyse zu Coaching-Programmen, in der ausschließlich experimentelle Studien berücksichtigt wurden (vgl. Kraft et al. 2018). Coaching wird hier „broadly as all in-service PD [professional development] programs, where coaches or peers observe teachers’ instruction and provide feedback to help them improve“ (Kraft et al. 2018, S. 553) definiert. Im Mittel hatten

die Coaching-Programme einen großen positiven Effekt auf das Unterrichtsverhalten von Lehrkräften und einen kleinen positiven Effekt auf den Lernerfolg ihrer Schülerinnen und Schüler. Allerdings variieren die Effekte stark; einige Coaching-Programme haben keinen Effekt, andere sogar negative Effekte. Weil sich hinsichtlich des Umfangs und der Dauer des Coachings keine differenziellen Effekte zeigten, vermuten die Autoren, dass die Qualität und der Fokus des Coachings entscheidend seien (vgl. Kraft et al. 2018). Diese Schlussfolgerung lässt sich auf ein Unterrichtsfeedback von Schulleitungen übertragen; auch dessen Wirksamkeit hängt von seinem Fokus und seiner Qualität ab. Interviews mit Schulleitungen und Lehrkräften deuten jedoch darauf hin, dass Unterrichtsfeedbacks von Schulleitungen häufig eine geringe Qualität aufweisen (vgl. Kimball und Milanowski 2009; Kraft und Gilmour 2016; Milanowski und Heneman 2001), was die teilweise ausbleibenden Feedbackeffekte erklären kann.

### 2.3 Was zeichnet ein qualitativ hochwertiges Feedback aus?

Weil zu den Merkmalen eines qualitativ hochwertigen Unterrichtsfeedbacks vergleichsweise wenige Befunde vorliegen (für einen Überblick über den Forschungsstand siehe Scheeler et al. 2004; Thurlings et al. 2013), soll im Folgenden erneut Anschluss an die allgemeine Feedbackforschung gesucht werden, deren Theorien und Befunde der entwickelten Feedbackintervention zugrunde gelegt wurden. Nach der FIT lässt sich über die Feedbackgestaltung beeinflussen, ob die Feedbackverarbeitung aufgaben- oder selbstbezogen erfolgt (vgl. Kluger und DeNisi 1996): Während Hinweise auf die zu bearbeitende Aufgabe zu einer aufgabenbezogenen Feedbackverarbeitung führen sollten, macht ein hoher Personenbezug eine selbstbezogene Feedbackverarbeitung wahrscheinlich. Auch Hattie und Timperley (2007) postulieren, dass Feedback dann besonders wirksam ist, wenn es einen hohen Aufgabenbezug aufweist. Nach Hattie und Timperley (2007) sollte Feedback daher drei Arten von Informationen enthalten, die sie als FEED UP, FEED BACK und FEED FORWARD bezeichnen und die mit den Fragen *Where am I going?*, *How am I going?* und *Where to next?* korrespondieren. Feedback sollte Informationen zu den Zielen oder Standards enthalten (FEED UP), die die Beurteilung eines Lernertrags oder einer Leistung erst ermöglichen. Feedback sollte darüber hinaus aufzeigen, inwiefern der aktuelle Lern- oder Leistungsstand den Zielen oder Standards entspricht (FEED BACK) und es sollte um Informationen ergänzt werden, die den Lernenden oder Leistungserbringenden dabei unterstützen, diese Ziele oder Standards zukünftig zu erreichen (FEED FORWARD).

Kluger und DeNisi (1996) und Hattie und Timperley (2007) haben im Rahmen ihrer Meta-Analysen diverse Feedbackmerkmale identifiziert, die in einem positiven Zusammenhang mit Lernen und Leistung stehen und die Annahme stützen, nach der Feedback einen hohen Aufgabenbezug aufweisen sollte. Feedback sollte demnach auf vorab definierte Ziele oder Standards bezogen sein (vgl. Hattie und Timperley 2007) oder einen Vergleich zu vorherigen Lernerträgen und Leistungen beinhalten (vgl. Kluger und DeNisi 1996). Sowohl bestätigendes als auch kritisches Feedback kann einen positiven Effekt auf Lernen und Leistung haben (vgl. Hattie und Timperley 2007; Kluger und DeNisi 1996). Bestätigendes Feedback geht al-

lerdings nur mit positiven Effekten einher, wenn es auf konkrete Lernerträge oder Leistungen bezogen ist; einfaches Lob ist demgegenüber zu vermeiden (vgl. Hattie und Timperley 2007; Kluger und DeNisi 1996). Kritisches Feedback ist besonders wirksam, wenn es Hinweise dazu enthält, wie Lernende oder Leistungserbringende ihren Lernertrag steigern bzw. ihre Leistung verbessern können (vgl. Hattie und Timperley 2007; Kluger und DeNisi 1996). Darüber hinaus erhöhen Zielvereinbarungen, in denen konkret festgehalten wird, *welche* Lernerträge bzw. Leistungen in Zukunft *wie* realisiert werden sollen, die Wirksamkeit von Feedbackinterventionen (vgl. Kluger und DeNisi 1996). Ein weiterer Befund der Meta-Analysen von Hattie und Timperley (2007) und Kluger und DeNisi (1996) ist, dass computerbasiertes Feedback besonders wirksam ist, weil es vom Feedbackempfänger vermutlich als objektiver wahrgenommen wird.

## 2.4 Zielorientierungen von Lehrkräften und Feedbacknutzung

Die Wirksamkeit von Unterrichtsfeedback hängt jedoch nicht allein von der Feedbackqualität ab. Ein Unterrichtsfeedback stellt für Lehrkräfte eine Lerngelegenheit dar. Es führt nur dann zu einer Weiterentwicklung ihrer Unterrichtskompetenz, wenn Lehrkräfte die im Feedback enthaltenen Informationen auch nutzen: „No amount of feedback will result in professional growth if a teacher is unwilling or unable to co-construct and enact changes“ (Kraft und Gilmour 2016, S. 715). Forschungsbefunde deuten darauf hin, dass die Nutzung verschiedener Lerngelegenheiten in einem Zusammenhang mit den Zielorientierungen von Lehrkräften steht.

Bei Zielorientierungen handelt es sich um vergleichsweise stabile motivationale Orientierungen, die Einfluss auf das Erleben und die Verhaltensregulation in Lern- und Leistungssituationen nehmen (vgl. Nitsche et al. 2011). In der frühen Zielorientierungsforschung wurde zwischen einer Lern- (*LZO*) und einer Leistungszielorientierung unterschieden (vgl. Ames 1992; Dweck 1986; Nicholls 1984). Während Personen mit hoch ausgeprägter *LZO* in Lern- und Leistungssituationen danach streben, ihre Kompetenzen weiterzuentwickeln, geht es Personen mit hoch ausgeprägter Leistungszielorientierung darum, ihre Überlegenheit zu demonstrieren. Später wurde die Leistungszielorientierung in eine Annäherungs- (*ALZO*) und eine Vermeidungsleistungszielorientierung (*VLZO*) ausdifferenziert (vgl. Elliot und Harackiewicz 1996; Harackiewicz et al. 2002). Personen mit hoch ausgeprägter *ALZO* streben danach, ihre Überlegenheit zu demonstrieren, Personen mit hoch ausgeprägter *VLZO* wollen ihre Kompetenzdefizite verbergen.

Nach unserem Kenntnisstand wurde die Bedeutung der *LZO* für die Nutzung eines Unterrichtsfeedbacks bislang nicht untersucht. Befunde zu den lernrelevanten Einstellungen und Verhaltensweisen von Lehrkräften deuten aber darauf hin, dass eine hoch ausgeprägte *LZO* die Feedbacknutzung begünstigen könnte. So haben lernzielorientierte Lehrkräfte positive Einstellungen gegenüber externer Hilfe und nehmen diese häufig in Anspruch (vgl. Butler 2007; Janke et al. 2019; Nitsche et al. 2011). Außerdem besteht zwischen einer hoch ausgeprägten *LZO* und dem Einholen von Feedback ein positiver Zusammenhang (vgl. Chughtai und Buckley 2010; Runhaar et al. 2010). Lernzielorientierte Lehrkräfte lesen häufig fachbezogene



Literatur, sie haben positive Einstellungen gegenüber Fortbildungen und nehmen häufig an diesen teil (vgl. Nitsche et al. 2013a, b).

Ob eine hoch ausgeprägte *ALZO* die Feedbacknutzung begünstigt, lässt sich nicht eindeutig vorhersagen. Theoretisch ist vorstellbar, dass annäherungsleistungszielorientierte Lehrkräfte ihre Unterrichtskompetenz auf der Grundlage eines Unterrichtsfeedbacks weiterentwickeln, um ihre Überlegenheit auch in Zukunft demonstrieren zu können. Die Befunde zu den lernrelevanten Einstellungen und Verhaltensweisen von annäherungsleistungszielorientierten Lehrkräften sind jedoch inkonsistent. Hinsichtlich der Einstellungen gegenüber und der Inanspruchnahme von Hilfe wurden in den zuvor zitierten Studien keine Zusammenhänge mit der *ALZO* gefunden (vgl. Butler 2007; Janke et al. 2019; Nitsche et al. 2011). Annäherungsleistungszielorientierte Lehrkräfte lesen aber häufig fachbezogene Literatur (vgl. Nitsche et al. 2013a). Zwischen einer *ALZO* und den Einstellungen gegenüber Fortbildungen besteht kein Zusammenhang, besucht werden sie von Lehrkräften mit hoch ausgeprägter *ALZO* dennoch häufig (vgl. Nitsche et al. 2013b).

Eine hoch ausgeprägte *VLZO* sollte demgegenüber in keinem Zusammenhang mit der Feedbacknutzung stehen. Vermeidungsleistungszielorientierte Lehrkräfte haben negative Einstellungen gegenüber externer Hilfe (vgl. Butler 2007; Janke et al. 2019; Nitsche et al. 2011) und Fortbildungen, die sie auch seltener besuchen (vgl. Nitsche et al. 2013b). Außerdem besteht zwischen dem Lesen fachbezogener Literatur und einer hoch ausgeprägten *VLZO* kein Zusammenhang (vgl. Nitsche et al. 2013b).

### 3 Fragestellungen und Hypothesen

Ein Unterrichtsfeedback von Schulleitungen stellt für Lehrkräfte eine Lerngelegenheit dar, die zu einer Weiterentwicklung ihrer Unterrichtskompetenz beitragen kann. Studienergebnisse zeigen, dass Unterrichtsfeedbacks jedoch nicht immer wirksam sind. Als ursächlich für die teilweise ausbleibenden Feedbackeffekte kann die häufig geringe Qualität von Unterrichtsfeedbacks von Schulleitungen gelten. Aus diesem Grund wurde eine Feedbackintervention entwickelt, die es Schulleitungen ermöglichen soll, Lehrkräften ein qualitativ hochwertiges Feedback zu ihrem Unterricht zu geben, auf dessen Grundlage sie ihre Unterrichtskompetenz weiterentwickeln können. Für die Unterrichtsbeobachtungen wurde ein standardisiertes Beobachtungsinstrument entwickelt. Für die Feedbackgespräche wurde – ausgehend von den Theorien und Befunden der allgemeinen Feedbackforschung – ein Feedbackverfahren entworfen, das zu einer aufgabenbezogenen Feedbackverarbeitung führen soll. Die Schulleitungen wurden in einem Beobachtungs- und Feedbacktraining in der Nutzung des Beobachtungsinstruments und der Anwendung des Feedbackverfahrens geschult. Die Effekte der Feedbackintervention auf die von Lehrkräften eingeschätzte Weiterentwicklung ihrer Unterrichtskompetenz wurde in einer Studie mit quasi-experimentellem Design überprüft.

Im vorliegenden Beitrag wird zunächst untersucht, wie Lehrkräfte die Qualität des Unterrichtsfeedbacks einschätzen (*Forschungsfrage 1*). Das Hauptaugenmerk gilt jedoch dem Effekt des Unterrichtsfeedbacks auf die selbst eingeschätzte Kompetenzentwicklung der Lehrkräfte. Es wird überprüft, ob Lehrkräfte, die ein Unter-



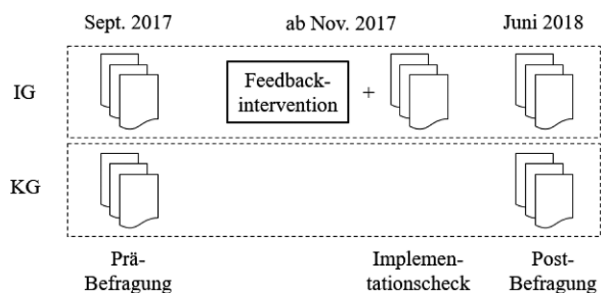
richtsfeedback von ihrer Schulleitung erhielten, ihre Kompetenzentwicklung positiver einschätzen als Lehrkräfte, die kein Feedback bekamen (*Forschungsfrage 2*). Die Wirksamkeit von Unterrichtsfeedback hängt jedoch nicht allein von der Feedbackqualität ab. Unter Rückgriff auf Angebots-Nutzungs-Modelle lässt sich die Kompetenzentwicklung als Interaktion des Feedbacks und verschiedener individueller Merkmale der Lehrkräfte verstehen, von denen die Feedbacknutzung abhängt (vgl. Kunter et al. 2011). Zielorientierungen spielen dabei vermutlich eine besondere Rolle. Daher wird zusätzlich untersucht, ob der Effekt des Unterrichtsfeedbacks auf die Kompetenzselbsteinschätzung mit den Zielorientierungen von Lehrkräften interagiert (*Forschungsfrage 3*). Vor dem Hintergrund des dargestellten Forschungsstands wird erwartet, dass Lehrkräfte mit hoch ausgeprägter *LZO* ihre Kompetenzentwicklung nach dem Unterrichtsfeedback positiver einschätzen als Lehrkräfte mit hoch ausgeprägter *LZO*, die kein Feedback erhielten. Hinsichtlich der *ALZO* wird aufgrund des inkonsistenten Forschungsstands keine gerichtete Hypothese formuliert. Darüber hinaus wird erwartet, dass der Effekt des Unterrichtsfeedbacks nicht mit der *VLZO* der Lehrkräfte interagiert. Auch im Hinblick auf die *VLZO* wird eine ungerichtete Hypothese getestet.

## 4 Methodik

### 4.1 Design

Alle teilnehmenden Lehrkräfte haben im September 2017 an einer Prä-Befragung teilgenommen (siehe Abb. 1). Das Beobachtungs- und Feedbacktraining fand im September und Oktober 2017 statt und wurde über einen Kurzfragebogen evaluiert. Während Lehrkräfte der Interventionsgruppe (IG) im Winter 2017/2018 ein Unterrichtsfeedback von ihrer Schulleitung erhielten, wurde den Lehrkräften der Kontrollgruppe (KG) kein Feedback gegeben. Die Post-Befragung der Lehrkräfte fand im Juni 2018 statt und wurde wie die Prä-Befragung von geschulten Testleiterinnen und Testleitern administriert. Darüber hinaus liegen die Ergebnisse eines Implementationschecks vor, den die Lehrkräfte der IG im Anschluss an den Erhalt des Unterrichtsfeedbacks ausfüllten und der Autorengruppe postalisch zuschickten. Die Anonymität der Lehrkräfte war jederzeit gewährleistet, sowohl gegenüber der Autorengruppe als auch gegenüber den Schulleitungen.

Abb. 1 Studiendesign



## 4.2 Stichprobe

Die Stichprobe umfasst  $N=6$  Berliner Oberstufenzentren (OSZ)<sup>1</sup> mit  $N=11$  Schulleitungen (inkl. stellvertretenden Schulleitungen, Abteilungs- und Fachleitungen) und  $N=81$  Lehrkräften.  $N=35$  Lehrkräfte haben ein Unterrichtsfeedback erhalten (IG),  $N=46$  Lehrkräfte haben kein Unterrichtsfeedback erhalten (KG). 55 % der teilnehmenden Lehrkräfte sind weiblich. 27,5 % der Lehrkräfte waren zum Zeitpunkt der Datenerhebung 30 bis 39 Jahre alt, 21,3 % waren 40 bis 49 Jahre alt, 45 % waren 50 bis 59 Jahre alt und 6,3 % waren 60 Jahre alt oder älter. Die Lehrkräfte der IG und KG unterscheiden sich hinsichtlich ihres Geschlechts und Alters nicht signifikant voneinander.

Die Schulen nahmen freiwillig an der Studie teil. Die Schulleitungen aller Berliner OSZ wurden auf einer turnusmäßigen Schulleitungskonferenz über die Studie informiert und konnten im Anschluss über eine Teilnahme entscheiden. Eine randomisierte Zuteilung der Lehrkräfte zur IG und KG war nicht möglich. Die Entscheidung, welche Lehrkräfte an der Studie teilnehmen und welcher Gruppe sie angehören würden, wurde schulintern getroffen. Es ist davon auszugehen, dass an einigen Schulen die Lehrkräfte über ihre Teilnahme und Gruppenzugehörigkeit entschieden, während diese Entscheidungen an anderen Schulen von der Schulleitung getroffen wurden. An jeder Schule gibt es Lehrkräfte der IG und KG, die sich weitgehend gleichmäßig auf die Gruppen verteilen.

Weil eine Randomisierung nicht möglich war, könnten sich die Lehrkräfte der IG und KG systematisch voneinander unterscheiden. In Fällen, in denen die Lehrkräfte über ihre Teilnahme und Gruppenzugehörigkeit entschieden, ist anzunehmen, dass sich vor allem Lehrkräfte für eine Teilnahme in der IG entschieden, die ohnehin eine hohe Bereitschaft zeigen, sich professionell weiterzuentwickeln. In Fällen, in denen diese Entscheidungen von der Schulleitung getroffen wurden, erscheint es denkbar, dass z. B. Lehrkräfte überproportional häufig für eine Teilnahme in der IG ausgewählt wurden, die kein Lehramtsstudium abgeschlossen haben und/oder über wenig Berufserfahrung verfügen. Möglichen Unterschieden zwischen der IG und KG wurde Rechnung getragen, indem weitere Lernaktivitäten und diverse Lehrkräftemerkmale erfasst wurden. Diese Variablen werden in den späteren Analysen als Kontrollvariablen berücksichtigt, um den Feedbackeffekt zu isolieren.

## 4.3 Beschreibung der Feedbackintervention

Bei der Feedbackintervention handelt es sich um ein einmaliges Unterrichtsfeedback von Schulleitungen an Lehrkräfte. Schulleitungen beurteilten den Unterricht der Lehrkräfte der IG mit einem standardisierten Beobachtungsinstrument und gaben ihnen ausgehend von der Unterrichtsbeobachtung ein Feedback zu ihrem Unterricht. Bei der Feedbackgabe orientierten sie sich an einem Feedbackverfahren, das auf Grundlage der Theorien und Befunde der allgemeinen Feedbackforschung

<sup>1</sup> An Berliner OSZ werden berufsvorbereitende und -qualifizierende sowie studienberechtigende Bildungsgänge angeboten, die zu verschiedenen Schulabschlüssen führen können (z. B. (erweiterte) Berufsbildungsreife, mittlerer Schulabschluss, Fach- sowie allgemeine Hochschulreife).

entworfen wurde und eine aufgabenbezogene Feedbackverarbeitung evozieren soll. Alle Schulleitungen wurden in einem Beobachtungs- und Feedbacktraining in der Nutzung des Beobachtungsinstruments und der Anwendung des Feedbackverfahrens geschult.

#### 4.3.1 Unterrichtsbeobachtung

Die Schulleitungen beobachteten je Lehrkraft zwei 90-minütige Unterrichtseinheiten und beurteilten das Unterrichtsverhalten der Lehrkräfte mit einem von der Autorengruppe entwickelten Beobachtungsinstrument (vgl. Gärtner et al. 2021). Das Beobachtungsinstrument soll vorrangig im Rahmen von Unterrichtsevaluationen durch Schulleitungen eingesetzt werden, die mit dem Ziel eines entwicklungsbezogenen Unterrichtsfeedbacks erfolgen. Daher liegt der Fokus des Beobachtungsinstruments auf dem Unterrichtsverhalten von Lehrkräften bzw. auf beeinflussbaren Verhaltensweisen von Schülerinnen und Schülern.

Das Beobachtungsinstrument basiert auf einem theoretischen Unterrichtsmodell, das an das Modell der drei Basisdimensionen anknüpft (vgl. Klieme et al. 2001; Praetorius et al. 2018). Weil Schulleitungen nicht in allen Fächern über fachliche und fachdidaktische Expertise verfügen, liegt der Schwerpunkt des Unterrichtsmodells auf der fachunabhängigen Gestaltung von Lehr-Lern-Prozessen. Auf der Grundlage von lehr-lernpsychologischen Theorien der Informationsverarbeitung, Motivations-theorien und Theorien der sozialen Interaktion im Klassenzimmer wurden drei übergeordnete Anforderungsbereiche des Unterrichts identifiziert: die Unterstützung von Schülerinnen und Schülern beim Wissenserwerb, die Motivierung von Schülerinnen und Schülern sowie das Klassenmanagement. Die theoretische Grundlage des Anforderungsbereichs *Unterstützung des Wissenserwerbs* bildet das Lehr-Lern-Prozessmodell von Klauer und Leutner (2012), das zwischen vier Komponenten eines kognitiven Lernprozesses unterscheidet, die in die Unterrichtsmerkmale (1) Erklärung und Präsentation, (2) Unterstützung von Verstehensprozessen, (3) Speicherung durch Üben sowie (4) Anwendung und Transfer überführt wurden. Die theoretische Grundlage des Anforderungsbereichs *Motivierung* bildet die Erwartungsmal-Wert-Theorie der Motivation nach Eccles und Wigfield (2002; siehe auch Rheinberg 2008), aus der die Unterrichtsmerkmale (5) Weckung von Interesse, (6) Förderung von Lernfreude, (7) Verdeutlichung des Nutzens sowie (8) Förderung der Selbstwirksamkeit hergeleitet wurden. Ausgehend von Theorien der sozialen Interaktion im Klassenzimmer (vgl. Doyle 2006; Ophardt und Thiel 2013) wurden hinsichtlich des Anforderungsbereichs *Klassenmanagement* die Unterrichtsmerkmale (9) Lerndienliches Unterrichtsklima, (10) Monitoring, (11) Steuerung des Unterrichtsflusses, (12) Gruppenaktivierung sowie (13) Umgang mit Störungen abgeleitet. Weil sich Schülerinnen und Schüler hinsichtlich ihrer kognitiven und motivationalen Lernvoraussetzungen unterscheiden, wird (14) Individualisierung als zu den Anforderungsbereichen *Unterstützung des Wissenserwerbs* und *Motivierung* querliegendes Unterrichtsmerkmal verstanden.

Die 14 Unterrichtsmerkmale wurden im Beobachtungsinstrument über 14 Skalen mit vier bis acht Items operationalisiert, die von den Schulleitungen auf einer

vierstufigen Likert-Skala eingeschätzt wurden.<sup>2</sup> Die theoretische Grundlage des Beobachtungsinstruments und Befunde zu seiner faktoriellen Validität und Reliabilität werden in einer weiteren Publikation der Autorengruppe ausführlicher beschrieben (vgl. Gärtner et al. 2021).

#### 4.3.2 Feedbackverfahren

Die Schulleitungen machten sich während der Unterrichtsbeobachtungen Notizen zum Unterrichtsverhalten der Lehrkräfte. Weil sich computerbasiertes Feedback als besonders wirksam erwiesen hat, füllten die Schulleitungen das Beobachtungsinstrument im Anschluss an die Beobachtungen in einem webbasierten Feedbackportal aus, das auf der Grundlage der Beobachtungsdaten lehrerspezifische Feedbackberichte generierte. Die Feedbackberichte enthielten eine Beschreibung des Unterrichtsmodells, ein überblicksartiges Feedback sowie bestätigendes und kritisches Feedback. Ein Algorithmus entschied in einem ersten Schritt über die Auswahl der Unterrichtsmerkmale: Die drei Unterrichtsmerkmale mit den höchsten Skalenmittelwerten wurden in den Feedbackberichten als bestätigendes Feedback aufgeführt, die drei Unterrichtsmerkmale mit den niedrigsten Mittelwerten als kritisches Feedback. In einem zweiten Schritt wurde das bestätigende Feedback um Beispiele gelungener Unterrichtspraxis ergänzt, die je Unterrichtsmerkmal aus den Items hergeleitet wurden, die die höchsten Ausprägungen innerhalb der entsprechenden Skalen aufwiesen. Das kritische Unterrichtsfeedback wurde um Hinweise auf alternative Verhaltensstrategien ergänzt, die je Unterrichtsmerkmal aus den Items hergeleitet wurden, die die geringsten Ausprägungen innerhalb der jeweiligen Skalen aufwiesen.

Die lehrerspezifischen Feedbackberichte bildeten die Grundlage für ein mündliches Unterrichtsfeedback. In den Feedbackgesprächen, die eine durchschnittliche Dauer von  $M=44$  min ( $SD=13$  min) hatten, wurde den Lehrkräften zu Beginn erläutert, dass das Ziel des Unterrichtsfeedbacks in der Weiterentwicklung ihres Unterrichts liegt und nicht der Kontrolle durch die Schulleitung dient. Anschließend erhielten die Lehrkräfte in Anlehnung an Hattie und Timperley (2007) drei Arten von Informationen: Zunächst wurde den Lehrkräften das dem Feedback zugrundeliegende Unterrichtsmodell vorgestellt, das den Soll-Standard der Unterrichtsbeurteilung definiert (FEED UP). Dadurch wurde den Lehrkräften verdeutlicht, dass dem Feedback eine sachliche Bezugsnorm zugrunde liegt, wodurch der „Fokus auf die Aufgabenstellung und die dafür nötige Anstrengung gelenkt [wird] und weniger auf das Selbst“ (Kopp und Mandl 2014, S. 151). Anschließend erhielten die Lehrkräfte Informationen zu ihrem beobachteten Verhalten (Ist-Stand) im Vergleich zum Soll-Standard (FEED BACK). Die Schulleitungen gaben den Lehrkräften bestätigendes und kritisches Unterrichtsfeedback zu den in den Feedbackberichten aufgeführten Unterrichtsmerkmalen, da beide Feedbackarten einen positiven Effekt auf Lernen und Leistung haben können. Das bestätigende Unterrichtsfeedback wurde anhand der Beispiele gelungener Unterrichtspraxis aus den Feedbackberichten erläutert, das kritische Unterrichtsfeedback um die Hinweise auf alternative Verhaltensstrategi-

<sup>2</sup> Eine tabellarische Übersicht über die 14 Unterrichtsmerkmale und Beispielitems finden sich im Online-Supplement 1.

en ergänzt (FEED FORWARD). Abschließend formulierten die Schulleitungen und die Lehrkräfte Entwicklungsziele im Hinblick auf die Umsetzung der alternativen Verhaltensstrategien. Weil ein schriftliches Feedback gegenüber einem mündlichen Feedback den Vorteil hat, dass es von der Lehrkraft mehrmals und zu einem späteren Zeitpunkt erneut gelesen werden kann (vgl. Kopp und Mandl 2014), händigten die Schulleitungen den Lehrkräften die Feedbackberichte im Anschluss an die Feedbackgespräche aus.<sup>3</sup>

#### 4.3.3 Beobachtungs- und Feedbacktraining

Alle Schulleitungen durchliefen ein Beobachtungs- und Feedbacktraining im Umfang von 18 h. Im Beobachtungstraining wurde den Schulleitungen ein grundlegendes Verständnis des dem Beobachtungsinstrument zugrundeliegenden Unterrichtsmodells vermittelt. Außerdem wurden häufig auftretende Beobachtungsfehler und Strategien zu ihrer Vermeidung vorgestellt. Die Nutzung des Beobachtungsinstruments wurde anhand von videobasierten Beobachtungsübungen trainiert. Im Feedbacktraining wurde den Schulleitungen erläutert, wie das Feedbackverfahren inklusive des webbasierten Feedbackportals und der lehrerspezifischen Feedbackberichte zu nutzen ist. Die Anwendung des Feedbackverfahrens wurde im Rahmen von Rollenspielen eingeübt.

### 4.4 Variablen

Die Feedbackqualität wurde mit einer an den Kontext eines Unterrichtsfeedbacks von Schulleitungen angepassten Skala der *Feedback Environment Scale* (vgl. Steelman et al. 2004) erfasst (5 Items; Beispielitem: „Meine Kollegin/mein Kollege hat mir nützliches Feedback gegeben“; Cronbachs  $\alpha = 0,82$ ). Eine Erfassung der Kompetenzentwicklung über einen Wissens- oder Performanztest konnte in der vorliegenden Studie nicht realisiert werden, weil aktuell kein geeignetes Testverfahren vorliegt und die zeitlichen Ressourcen eine Testentwicklung und -validierung nicht zuließen. Die Weiterentwicklung der Unterrichtskompetenz wurde stattdessen über eine retrospektive Selbsteinschätzung erfasst. Während eine retrospektive Selbsteinschätzung eine direkte Erfassung der Kompetenzentwicklung erlaubt, kann bei einer indirekten Messung der Kompetenzentwicklung, d. h. bei einer Selbsteinschätzung des Kompetenzstandes vor und nach der Feedbackintervention und einer anschließenden Differenzwertbildung, nicht ausgeschlossen werden, dass es zu einem *Response Shift* kommt, also zu einem neuen Bewertungsstandard bei der Beurteilung von Kompetenz oder einem veränderten Begriffsverständnis von Kompetenz (vgl. Piwowar und Thiel 2014; Schwartz und Sprangers 1999), was zu einer Über- oder Unterschätzung des Feedbackeffekts führen kann. Die Lehrkräfte wurden daher darum gebeten, ihre Kompetenzentwicklung für das Schuljahr 2017/2018 rückblickend zu beurteilen (4 Items; Beispielitem: „Ich habe viel dazugelernt, was die Verbesserung meines Unterrichts betrifft“; Cronbachs  $\alpha = 0,91$ ).

<sup>3</sup> Ein exemplarischer Ausschnitt eines Feedbackberichts findet sich im Online-Supplement 2.

Neben dem Unterrichtsfeedback der Schulleitung können weitere Faktoren die selbst eingeschätzte Kompetenzentwicklung der Lehrkräfte beeinflussen. Zuvorderst sind dies weitere Lernaktivitäten sowie verschiedene Merkmale der Lehrkräfte. Weil eine Randomisierung nicht möglich war und es hinsichtlich der Lernaktivitäten und Lehrkräftemerkmale systematische Unterschiede zwischen den Lehrkräften der IG und KG geben könnte, wurden diese als Kovariaten erfasst und in den späteren Analysen als Kontrollvariablen berücksichtigt. Im Hinblick auf die Lernaktivitäten wurde die Anzahl der im Schuljahr 2017/2018 besuchten Fort- und Weiterbildungen erfragt. Darüber hinaus wurde das Lesen fachbezogener Literatur mit einem in Anlehnung an Kwakman (2003) entwickelten Instrument erhoben (4 Items; Beispielitem: „Um auf dem aktuellen Stand zu bleiben, lese ich intensiv pädagogische Fachzeitschriften“; Cronbachs  $\alpha=0,84$ ). Außerdem wurde erfragt, ob Lehrkräfte im Rahmen einer kollegialen Hospitation ein Feedback erhalten oder sich ein Schülerfeedback eingeholt haben. Hinsichtlich der Lehrkräftemerkmale wurde der Abschluss eines Lehramtsstudiums und die Berufserfahrung erhoben.

Die Zielorientierungen wurden mit einem von Nitsche et al. (2011) entwickelten Instrument in der Prä-Befragung erfasst, das aus erhebungsökonomischen Gründen gekürzt wurde. Während die *LZO* mit der originalen Skala erfasst wurde, wurde auf eine nach Adressaten differenzierte Erfassung der *ALZO* und *VLZO* verzichtet. Die Kürzungen stehen im Einklang mit einer Kurzversion des Instruments von Nitsche et al. (2011), die von Janke et al. (2019) entwickelt und validiert wurde. Ein Beispielitem der *LZO* (9 Items; Cronbachs  $\alpha=0,90$ ) lautet: „In meinem Unterricht strebe ich danach, mit kritischen Unterrichtssituationen besser umgehen zu können“. Ein Beispielitem der *ALZO* (3 Items; Cronbachs  $\alpha=0,91$ ) lautet: „In meinem Unterricht strebe ich danach, zu zeigen, dass ich kritische Unterrichtssituationen besser bewältigen kann als andere Lehrkräfte“. Ein Beispielitem der *VLZO* (3 Items; Cronbachs  $\alpha=0,83$ ) lautet: „In meinem Unterricht strebe ich danach, nicht zu zeigen, wenn mir die Unterrichtsanforderungen schwerer fallen als anderen Lehrkräften“.

## 4.5 Analysen

Die Qualitätseinschätzungen der Lehrkräfte werden deskriptiv ausgewertet. Ein Einstichproben-T-Test wird genutzt, um zu prüfen, ob der Mittelwert der Skala *Qualität des Feedbacks* signifikant vom theoretischen Skalennittel abweicht. Die Effekte des Unterrichtsfeedbacks auf die selbst eingeschätzte Kompetenzentwicklung werden mithilfe von Regressionsanalysen überprüft. Da von jeder Schule mehrere Lehrkräfte an der Studie teilgenommen haben, liegt eine hierarchische Datenstruktur vor (Lehrkräfte geclustert in Schulen), der mit einer Mehrebenenregressionsanalyse Rechnung getragen werden könnte (vgl. Raudenbush und Bryk 2002). Die im vorliegenden Fall relevanten Random-Intercept-Modelle setzen jedoch mindestens  $N=30$  Cluster voraus, um zu nicht verzerrten Parameterschätzungen zu kommen (vgl. McNeish und Stapleton 2016). Weil diese Voraussetzung bei  $N=6$  Schulen nicht erfüllt ist, werden die Effekte des Unterrichtsfeedbacks auf die Kompetenzselbsteinschätzung der Lehrkräfte mithilfe multipler Regressionsanalysen überprüft. Der Modellaufbau erfolgt schrittweise; mit jedem Schritt wird die Modellierung restriktiver. Zunächst wird das Unterrichtsfeedback als einziger Prädiktor model-

liert. Im Anschluss werden weitere Lernaktivitäten eingeführt und abschließend die Lehrkräftemerkmale ergänzt. Nicht signifikante Prädiktoren ( $p \geq 0,10$ ) werden aus den Analysen ausgeschlossen. Weil sich standardisierte Regressionskoeffizienten für dichotome Variablen inhaltlich nicht sinnvoll interpretieren lassen und sie bei einer 0/1-Kodierung von der Größe der Gruppe abhängen, die mit 1 kodiert ist, wurde eine Teilstandardisierung vorgenommen: Während die metrischen Variablen z-standardisiert wurden, ist von einer z-Standardisierung der dichotomen Variablen abgesehen worden (vgl. Hayes 2013; Urban und Mayerl 2018). Die Interaktionsterme zur Berechnung der differenziellen Effekte der Zielorientierungen wurden aus den z-standardisierten Moderatorvariablen und der dichotomen Feedbackvariable gebildet. Da signifikante Moderationseffekte nur Auskunft darüber geben, ob die Zielorientierungen mit der Feedbackvariable interagieren, sich anhand der Moderationseffekte aber nicht beurteilen lässt, ob der Feedbackeffekt bei einer bestimmten Ausprägung der Zielorientierungen das Niveau statistischer Signifikanz erreicht, werden alle signifikanten Moderationseffekte mit sogenannten *Simple-Slopes*-Tests überprüft, mit denen sich der Feedbackeffekt bei verschiedenen Ausprägungen der Zielorientierungen ( $-1 SD$ ,  $M$  und  $+1 SD$ ) schätzen lässt (vgl. Hayes 2013). Die Regressionsanalysen wurden mit Mplus 8 durchgeführt (vgl. Muthén und Muthén 2017). Weil nicht alle Variablen eine Normalverteilung aufweisen, wurde ein robustes *Maximum Likelihood Schätzverfahren* genutzt. Fehlende Werte wurden mit der *Full Information Likelihood Methode* behandelt (vgl. Lüdtke et al. 2007).

## 5 Ergebnisse

### 5.1 Evaluation des Beobachtungs- und Feedbacktrainings und Implementationscheck

Bevor die Ergebnisse zu den Forschungsfragen präsentiert werden, sollen die Evaluationsergebnisse zum Beobachtungs- und Feedbacktraining der Schulleitungen sowie die Ergebnisse des Implementationschecks vorgestellt werden. Das Beobachtungs- und Feedbacktraining wurde über selbst entwickelte Einzelitems evaluiert, denen die Schulleitungen auf einer 6-stufigen Likert-Skala zustimmen konnten. Hinsichtlich des Beobachtungstrainings gaben die Schulleitungen überwiegend an, sie hätten ihr Wissen über Unterrichtsbeobachtungen erweitert ( $M = 5,08$ ,  $SD = 0,95$ ) und ihre Beobachtungskompetenz weiterentwickelt ( $M = 5,08$ ,  $SD = 0,86$ ). Auch das Feedbacktraining hat in der Wahrnehmung der Schulleitungen zu einer Erweiterung ihres Wissens über wirksames Feedback ( $M = 5,29$ ,  $SD = 1,07$ ) und einer Weiterentwicklung ihrer Feedbackkompetenz geführt ( $M = 5,07$ ,  $SD = 0,62$ ).

Anhand des Implementationschecks lässt sich prüfen, ob die Schulleitungen das Unterrichtsfeedback in Übereinstimmung mit dem entwickelten Feedbackverfahren gestaltet haben. Der Implementationscheck liegt für  $N = 29$  Lehrkräfte vor (82,9% aller Lehrkräfte der IG). Auch die Umsetzung des Feedbackverfahrens wurde über selbst entwickelte Einzelitems mit einer 6-stufigen Likert-Skala erfasst. Die Lehrkräfte der IG gaben überwiegend an, das Unterrichtsfeedback habe aus ihrer Sicht der Weiterentwicklung ihres Unterrichts ( $M = 5,17$ ,  $SD = 1,14$ ) und nicht der Kontrolle



durch die Schulleitung ( $M=1,52$ ,  $SD=0,91$ ) gegolten. Neben einer hohen Zielklarheit berichten die Lehrkräfte davon, dass ihnen das theoretische Unterrichtsmodell vorgestellt wurde ( $M=5,55$ ,  $SD=0,78$ ). Das bestätigende Unterrichtsfeedback wurde anhand von Beispielen gelungener Unterrichtspraxis erläutert ( $M=5,55$ ,  $SD=0,63$ ). Auch der Aussage, das kritische Feedback habe konkrete Hinweise auf alternative Verhaltensstrategien für die Weiterentwicklung ihres Unterrichts enthalten, stimmten die Lehrkräfte überwiegend zu ( $M=4,97$ ,  $SD=1,18$ ). Die Zielvereinbarungen wurden akzeptiert, als realistisch eingeschätzt und als konkret wahrgenommen ( $M=5,58$ ,  $SD=0,55$ ).

Zusammenfassend lässt die Kurzevaluation des Beobachtungs- und Feedbacktrainings darauf schließen, dass die Schulleitungen ihr Wissen und ihre Kompetenzen über Unterrichtsbeobachtungen und -feedbacks erweitert haben. Die Daten des Implementationschecks zeigen, dass die Schulleitungen das Unterrichtsfeedback in Übereinstimmung mit dem entwickelten Feedbackverfahren gestaltet haben.

## 5.2 Qualität des Unterrichtsfeedbacks

Die erste Forschungsfrage bezog sich auf die Feedbackqualität. Der Mittelwert der Skala *Qualität des Feedbacks* fällt mit  $M=5,33$  ( $SD=0,71$ ) deutlich höher aus als das theoretische Mittel von  $\mu=3,50$ . Der Mittelwertunterschied ist statistisch signifikant ( $t(27)=14,39$ ,  $p<0,01$ ) und praktisch bedeutsam ( $d=2,6$ ) (vgl. Cohen 1988). Weiterführende Analysen zeigen zudem, dass die Qualitätseinschätzungen der Lehrkräfte mit den meisten der Gestaltungsmerkmale des Feedbacks korrespondieren: Je stärker die Lehrkräfte mit dem Unterrichtsfeedback das Ziel der Weiterentwicklung ihres Unterrichts assoziierten ( $r_s=0,51$ ,  $p<0,01$ ), je besser das bestätigende Unterrichtsfeedback anhand von Beispielen gelungener Unterrichtspraxis erläutert wurde ( $r_s=0,44$ ,  $p<0,05$ ), je konkreter die alternativen Verhaltensstrategien im Hinblick auf das kritische Feedback waren ( $r_s=0,71$ ,  $p<0,01$ ) und je mehr Akzeptanz die Zielvereinbarungen erfuhren ( $r_s=0,62$ ,  $p<0,01$ ), desto höher bewerteten die Lehrkräfte die Feedbackqualität.

## 5.3 Effekte des Unterrichtsfeedbacks auf die selbst eingeschätzte Kompetenzentwicklung

### 5.3.1 Deskriptive Analysen

In der zweiten Forschungsfrage wurde nach dem Haupteffekt des Unterrichtsfeedbacks auf die selbst eingeschätzte Kompetenzentwicklung der Lehrkräfte gefragt. Weil sich die Lehrkräfte der IG und KG systematisch voneinander unterscheiden könnten, werden zunächst die Mittelwerte der Kovariaten zwischen den Gruppen verglichen (siehe Tab. 1). Das Signifikanzniveau der Gruppenunterschiede wurde bei metrischen Variablen mithilfe von T-Tests für unabhängige Stichproben oder dem Mann-Whitney-U-Test überprüft und bei dichotomen Variablen mithilfe des  $\chi^2$ -4-Felder-Tests. Um die Größe der Gruppenunterschiede zu bestimmen, wurde für metrische Variablen Cohens  $d$  und für dichotome Variablen Cohens  $h$  berechnet (vgl. Cohen 1988).

**Tab. 1** Gruppenvergleich IG und KG (Mittelwerte (*M*), Standardabweichungen (*SD*) sowie Gruppenunterschiede (*p* und *d/h*))

	Gesamt		IG		KG		Gruppenunterschiede	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>	<i>d/h</i>
<b>Lernaktivitäten</b>								
Besuchte Fort- und Weiterbildungen	3,36	1,88	3,33	1,56	3,39	2,17	0,80	-0,03
Lesen fachbezogener Literatur	3,06	1,03	3,10	1,06	3,03	1,02	0,80	0,07
Kollegiales Feedback erhalten <sup>a</sup>	0,26	–	0,15	–	0,35	–	0,06	-0,47
Schülerfeedback erhalten <sup>a</sup>	0,30	–	0,24	–	0,35	–	0,32	-0,24
<b>Lehrkraftmerkmale</b>								
Lehramtsstudium abgeschlossen <sup>a</sup>	0,75	–	0,74	–	0,76	–	0,84	-0,02
Berufserfahrung (in Jahren)	13,87	9,47	13,91	9,13	13,84	9,81	0,97	0,01
<i>LZO</i>	4,93	0,88	5,16	0,66	4,76	0,99	0,05	0,48
<i>ALZO</i>	1,75	0,94	1,86	1,02	1,67	0,87	0,39	0,20
<i>VLZO</i>	1,90	0,99	1,88	1,02	1,92	0,98	0,83	-0,04

*LZO* Lernzielorientierung, *ALZO* Annäherungsleistungszielorientierung, *VLZO* Vermeidungsleistungszielorientierung

<sup>a</sup>Dichotome Variable (0 „nein“, 1 „ja“)

Im Hinblick auf die Lernaktivitäten gibt es Unterschiede zwischen den Lehrkräften der IG und KG. Während sich die Gruppen hinsichtlich der besuchten Fort- und Weiterbildungen und dem Lesen fachbezogener Literatur kaum voneinander unterscheiden, erhielten Lehrkräfte der KG signifikant häufiger ein kollegiales Feedback ( $h = -0,47, p < 0,10$ ) und häufiger ein Schülerfeedback ( $h = -0,24, p > 0,10$ ) als Lehrkräfte der IG. Auch im Hinblick auf die Lehrkräftemerkmale gibt es Unterschiede zwischen der IG und KG. Während der Anteil an Lehrkräften mit abgeschlossenem Lehramtsstudium und die Berufserfahrung in den Gruppen vergleichbar sind, haben Lehrkräfte der IG im Mittel eine signifikant höher ausgeprägte *LZO* ( $d = 0,48, p < 0,05$ ) und eine höher ausgeprägte *ALZO* ( $d = 0,20, p > 0,10$ ) als Lehrkräfte der KG. Hinsichtlich der *VLZO* unterscheiden sich die Gruppen kaum voneinander.

### 5.3.2 Haupteffekte auf die Kompetenzentwicklung

Die Haupteffekte des Unterrichtsfeedbacks können den ersten drei Spalten von Tab. 2 entnommen werden.<sup>4</sup> Für dichotome Prädiktoren werden teilstandardisierte und für metrische Prädiktoren standardisierte Regressionskoeffizienten angegeben. Zunächst wurde der Effekt des Unterrichtsfeedbacks in einer einfachen Regressionsanalyse überprüft (Modell 1). Im bivariaten Modell hat das Unterrichtsfeedback keinen Ef-

<sup>4</sup> Eine Korrelationstabelle der bivariaten Zusammenhänge zwischen der Kompetenzselbsteinschätzung und den Kovariaten findet sich im Online-Supplement 3.

**Tab. 2** Ergebnisse der Regressionsanalysen (teilstandardisierte und standardisierte Regressionskoeffizienten ( $\beta$ ) und Standardfehler ( $SE$ ))

	Modell 1 $\beta$ ( $SE$ )	Modell 2 $\beta$ ( $SE$ )	Modell 3 $\beta$ ( $SE$ )	Modell 4 $\beta$ ( $SE$ )	Modell 5 $\beta$ ( $SE$ )	Modell 6 $\beta$ ( $SE$ )
<b>Feedbackintervention</b>						
Feedbackintervention erhalten <sup>a</sup>	0,32 (0,24)	0,30 (0,22)	0,09 (0,19)	0,11 (0,22)	0,15 (0,21)	0,09 (0,19)
<b>Lernaktivitäten</b>						
Lesen fachbezogener Literatur	–	0,36** (0,13)	0,23* (0,11)	–	0,24+ (0,13)	0,21+ (0,11)
<b>Lehrkraftmerkmale</b>						
Lehramtsstudium <sup>a</sup>	–	–	–0,54* (0,26)	–	–	–0,50* (0,25)
Berufserfahrung (in Jahren)	–	–	–0,30** (0,11)	–	–	–0,28* (0,11)
<i>LZO</i>	–	–	0,18* (0,09)	0,13 (0,11)	0,09 (0,10)	0,11 (0,10)
<i>ALZO</i>	–	–	0,21* (0,09)	–0,16 (0,23)	–0,17 (0,22)	0,02 (0,22)
<i>VLZO</i>	–	–	–	0,12 (0,18)	0,17 (0,17)	0,06 (0,18)
<b>Interaktionsterme</b>						
<i>LZO</i> * Unterrichtsfeedback	–	–	–	0,40+ (0,21)	0,36+ (0,20)	0,26 (0,17)
<i>ALZO</i> * Unterrichtsfeedback	–	–	–	0,62* (0,26)	0,56* (0,27)	0,24 (0,25)
<i>VLZO</i> * Unterrichtsfeedback	–	–	–	–0,31 (0,22)	–0,28 (0,22)	–0,03 (0,22)
<b>Erklärte Varianz</b>						
$R^2$	0,03	0,15+	0,43**	0,28+	0,29*	0,39**

*LZO* Lernzielorientierung, *ALZO* Annäherungsleistungszielorientierung, *VLZO* Vermeidungsleistungszielorientierung

+ $p < 0,10$ , \* $p < 0,05$ , \*\* $p < 0,01$

<sup>a</sup>Dichotome Variable (0 „nein“, 1 „ja“)

fekt auf die selbst eingeschätzte Kompetenzentwicklung der Lehrkräfte ( $\beta = 0,32$ ,  $p > 0,10$ ).

Im Modell 2 wurden weitere Lernaktivitäten bei der Modellbildung berücksichtigt. Von den erhobenen Lernaktivitäten hat nur das Lesen fachbezogener Literatur einen signifikanten Effekt auf die Kompetenzselbsteinschätzung. Die besuchten Fort- und Weiterbildungen, das kollegiale Feedback und das Schülerfeedback wurden bei der weiteren Modellbildung daher nicht berücksichtigt. Auch unter Kontrolle des Lesens fachbezogener Literatur hat das Unterrichtsfeedback keinen Effekt auf die selbst eingeschätzte Kompetenzentwicklung der Lehrkräfte ( $\beta = 0,30$ ,  $p > 0,10$ ).

Im Modell 3 wurden abschließend die Lehrkräftemerkmale als Kontrollvariablen in die Analysen aufgenommen. Die *VLZO* wurde nicht modelliert, weil sie keinen Effekt auf die Kompetenzselbsteinschätzung hat. Auch im Modell 3 erreicht der Feedbackeffekt nicht das Niveau statistischer Signifikanz ( $\beta = 0,09$ ,  $p > 0,10$ ). Zu-

sammenfassend lässt sich festhalten, dass das Unterrichtsfeedback in keinem der Regressionsmodelle einen Effekt auf die selbst eingeschätzte Kompetenzentwicklung hat. Lehrkräfte, die ein Feedback zu ihrem Unterricht erhielten, schätzten ihre Kompetenzentwicklung nicht positiver ein als Lehrkräfte, die kein Feedback bekamen.

### 5.3.3 Differenzielle Effekte der Zielorientierungen

In der dritten Forschungsfrage wurde gefragt, ob der Feedbackeffekt mit den Zielorientierungen der Lehrkräfte interagiert. Die differenziellen Effekte der Zielorientierungen können den letzten drei Spalten von Tab. 2 entnommen werden. Für dichotome Prädiktoren und die Interaktionsterme werden teilstandardisierte und für metrische Prädiktoren standardisierte Regressionskoeffizienten angegeben. Modell 4 umfasst das Unterrichtsfeedback, die Moderatorvariablen sowie die Interaktionsterme. Für die *LZO* ( $\beta=0,40$ ,  $p<0,10$ ) und die *ALZO* ( $\beta=0,62$ ,  $p<0,05$ ) konnte ein signifikant positiver Moderationseffekt und für die *VLZO* kein Moderationseffekt gefunden werden ( $\beta=-0,31$ ,  $p>0,10$ ).

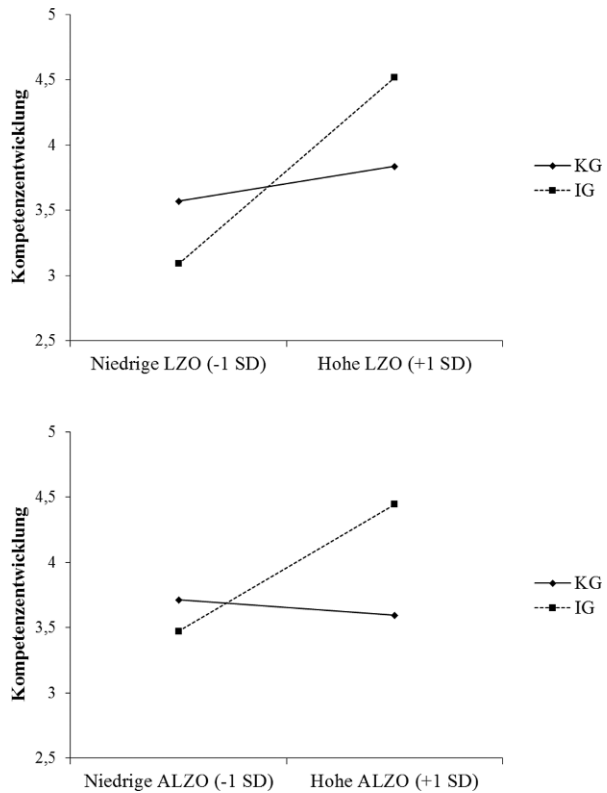
Im Modell 5 wurde das Lesen fachbezogener Literatur ergänzt. Auch unter Kontrolle des Lesens fachbezogener Literatur moderieren die *LZO* ( $\beta=0,36$ ,  $p<0,10$ ) und die *ALZO* ( $\beta=0,56$ ,  $p<0,05$ ) den Effekt des Unterrichtsfeedbacks. Der Moderationseffekt der *VLZO* erreicht weiterhin nicht das Niveau statistischer Signifikanz ( $\beta=-0,28$ ,  $p>0,10$ ). Modell 6 ist am restriktivsten und enthält zusätzlich die Lehrkräftemerkmale abgeschlossenes Lehramtsstudium und Berufserfahrung. Im Modell 6 konnte weder für die *LZO* ( $\beta=0,26$ ,  $p>0,10$ ) noch die *ALZO* ( $\beta=0,24$ ,  $p>0,10$ ) oder die *VLZO* ( $\beta=-0,03$ ,  $p>0,10$ ) ein Moderationseffekt gefunden werden.

Werden die Effekte des Lehramtsstudiums und der Berufserfahrung nicht kontrolliert, interagiert sowohl die *LZO* als auch die *ALZO* signifikant positiv mit dem Feedbackeffekt. Um zu überprüfen, bei welchen Ausprägungen der *LZO* und *ALZO* der Feedbackeffekt das Niveau statistischer Signifikanz erreicht, wurden *Simple-Slopes*-Tests durchgeführt. Für die *LZO* zeigte sich bei einer um  $+1$  *SD* über dem Mittelwert liegenden *LZO* ein signifikant positiver Feedbackeffekt ( $\beta=0,68$ ,  $p<0,05$ ). Auch für die *ALZO* konnte bei einer um  $+1$  *SD* über dem Mittelwert liegenden *ALZO* ein signifikant positiver Feedbackeffekt gefunden werden ( $\beta=0,81$ ,  $p<0,05$ ). Die differenziellen Effekte der *LZO* und *ALZO* werden in Abb. 2 veranschaulicht.<sup>5</sup>

Zusammenfassend lässt sich festhalten, dass der Effekt des Unterrichtsfeedbacks auf die Kompetenzselbsteinschätzungen mit der *LZO* und der *ALZO* der Lehrkräfte interagiert. Lehrkräfte mit hoch ausausgeprägter *LZO* und hoch ausgeprägter *ALZO*, die ein Unterrichtsfeedback erhielten, schätzten ihre Kompetenzentwicklung positiver ein als lern- bzw. annäherungsleistungszielorientierte Lehrkräfte, die kein Feedback bekamen.

<sup>5</sup> Für die grafische Veranschaulichung wurde ein von Dawson (2014) entwickeltes Online-Tool genutzt (<http://www.jeremydawson.co.uk/slopes.htm>).

**Abb. 2** Differenzielle Effekte der Lern- (*LZO*) und Annäherungsleistungszielorientierung (*ALZO*) auf die selbst eingeschätzte Weiterentwicklung der Unterrichtskompetenz



## 6 Diskussion

Im vorliegenden Beitrag wurde eine Feedbackintervention vorgestellt, die Schulleitungen dabei unterstützen soll, Lehrkräften ein qualitativ hochwertiges Feedback zu ihrem Unterricht zu geben, auf dessen Grundlage diese ihre Unterrichtskompetenz weiterentwickeln können. Die Qualität des Unterrichtsfeedbacks wurde von den Lehrkräften als hoch eingeschätzt. Trotz der hohen Qualitätseinschätzungen hatte das Feedback jedoch keinen Haupteffekt auf die Kompetenzselbsteinschätzungen der Lehrkräfte. Dieser Befund steht im Widerspruch zu den Ergebnissen einer weiteren Studie der Autorengruppe, in der der Effekt der Feedbackintervention auf die von Schülerinnen und Schülern wahrgenommene Unterrichtsqualität untersucht wurde (vgl. Kellermann et al. [im Druck](#)): Schülerinnen und Schüler, die von Lehrkräften der IG unterrichtet wurden, schätzten die Unterrichtsqualität nach dem Unterrichtsfeedback bezogen auf 3 der 14 Unterrichtsmerkmale des theoretischen Unterrichtsmodells signifikant besser ein als Schülerinnen und Schüler von Lehrkräften der KG. Die widersprüchlichen Befunde der beiden Studien sollen später erneut aufgegriffen werden.

Auch wenn das Unterrichtsfeedback keinen Haupteffekt auf die selbst eingeschätzte Kompetenzentwicklung der Lehrkräfte hatte, zeigten sich für die *LZO* und die *ALZO* differenzielle Effekte, für die *VLZO* nicht. Lehrkräfte mit hoch ausge-

prägter *LZO* und hoch ausgeprägter *ALZO*, die ein Unterrichtsfeedback erhielten, schätzten ihre Kompetenzentwicklung positiver ein als lern- oder annäherungsleistungszielorientierte Lehrkräfte, die kein Feedback bekamen. Der positive Moderationseffekt der *LZO* und der ausbleibende Moderationseffekt der *VLZO* sind hypothesenkonform. Im Hinblick auf die *ALZO* wurde aufgrund der inkonsistenten Befunde zu den lernrelevanten Einstellungen und Verhaltensweisen von annäherungsleistungszielorientierten Lehrkräften keine Hypothese formuliert. Eine Erklärung für den positiven Moderationseffekt bietet die Operationalisierung der *ALZO*. Die inkonsistenten Befunde, die auch aus anderen Bereichen bekannt sind, können häufig auf die Konzeptualisierung der *ALZO* zurückgeführt werden; teilweise liegt der Fokus auf dem Streben, *Kompetenz* zu demonstrieren, teilweise auf dem Streben, *Überlegenheit* zu demonstrieren (vgl. Senko et al. 2011). Letztgenannte Konzeptualisierung steht häufig in einem positiven Zusammenhang mit Lernen und Leistung (vgl. Hulleman et al. 2010; Senko und Dawson 2017) und liegt der Operationalisierung im vorliegenden Beitrag zugrunde. Es wird vermutet, dass das Ziel, andere zu übertreffen, mit einem hohen Leistungsstandard und einer hohen Leistungsmotivation einhergeht (vgl. Senko et al. 2011). Insofern erscheint es plausibel, dass Lehrkräfte mit hoch ausgeprägter *ALZO* ihre Unterrichtskompetenz auf Grundlage des Unterrichtsfeedbacks weiterentwickeln, um auch in Zukunft ihre Überlegenheit demonstrieren zu können. Vor dem Hintergrund der Zusammenhänge von Zielorientierungen mit den lernrelevanten Einstellungen und Verhaltensweisen von Lehrkräften ist vorstellbar, dass annäherungsleistungs- ebenso wie lernzielorientierte Lehrkräfte das Unterrichtsfeedback mit weiteren Lernaktivitäten verknüpft haben; so könnten sie z. B. gezielt fachbezogene Literatur gelesen oder Fort- und Weiterbildungen besucht haben (vgl. Nitsche et al. 2013a, b), um neue Verhaltensstrategien für ihren Unterricht kennenzulernen.

Einschränkend ist zu erwähnen, dass die Moderationseffekte der *LZO* und der *ALZO* nur dann das Niveau statistischer Signifikanz erreichen, wenn die Variablen abgeschlossenes Lehramtsstudium und Berufserfahrung nicht kontrolliert werden. Dies kann vorrangig auf die hohe geteilte Varianz der beiden Variablen mit der Kompetenzselbsteinschätzung zurückgeführt werden. Es ist plausibel, dass Lehrkräfte, die über wenig Berufserfahrung verfügen und kein Lehramtsstudium abgeschlossen haben, ihre Kompetenzentwicklung positiver einschätzen als Lehrkräfte mit viel Berufserfahrung sowie Lehrkräfte mit einem abgeschlossenen Lehramtsstudium. Offen bleibt zudem, worauf die Feedbackeffekte bei den Lehrkräften mit hoch ausgeprägter *LZO* und *ALZO* zurückzuführen sind. Welchen Anteil das standardisierte Beobachtungsinstrument, das Feedbackverfahren und die lehrerspezifischen Feedbackberichte an den Feedbackeffekten haben, lässt sich anhand der vorliegenden Daten nicht beurteilen. Auch über die Wirksamkeit des Beobachtungs- und Feedbacktrainings liegen nur wenige Informationen vor. Die Kurzevaluation deutet zwar darauf hin, dass die Schulleitungen ihr Wissen und ihre Kompetenz über Unterrichtsbeobachtungen und -feedbacks erweitert haben, inwiefern die Feedbackeffekte bei lern- und annäherungsleistungszielorientierten Lehrkräften auf das Training der Schulleitungen zurückzuführen sind, lässt sich jedoch nicht beurteilen.

Darüber hinaus weist die vorliegende Studie weitere Limitationen auf. Eine Limitation betrifft die nicht randomisierte Zuordnung der Lehrkräfte zur IG und KG,

die eine kausale Interpretation erschwert. Durch die Berücksichtigung diverser Kontrollvariablen konnte der Feedbackeffekt nach unserer Einschätzung jedoch weitgehend isoliert werden. Eine weitere Limitation betrifft die Operationalisierung der Weiterentwicklung der Unterrichtskompetenz. Durch eine retrospektive Selbsteinschätzung der Kompetenzentwicklung kann es zwar nicht zu einem *Response Shift* kommen (vgl. Piowar und Thiel 2014; Schwartz und Sprangers 1999), im Vergleich zu einem Wissens- oder Performanztest kann eine retrospektive Kompetenzselbsteinschätzung jedoch mit Validitätsproblemen wie Verzerrungen einhergehen. Beispielsweise ist vorstellbar, dass Lehrkräfte der IG ihre Kompetenzentwicklung rückblickend allein schon deshalb positiver einschätzten, weil sie ein Unterrichtsfeedback erhalten haben, was in besonderer Weise auf Lehrkräfte mit hoch ausgeprägter *LZO* und *ALZO* zutreffen könnte. Inwiefern retrospektive Selbsteinschätzungen der Kompetenzentwicklung tatsächliche Veränderungen im professionellen Wissen und Können von Lehrkräften abbilden, wurde nach unserem Kenntnisstand noch nicht erforscht. Es liegen aber Studien vor, die nahelegen, dass Selbsteinschätzungen von Lehrkräften im Allgemeinen und Kompetenzselbsteinschätzungen im Besonderen positive Zusammenhänge mit extern erhobenen Kriterien aufweisen. So zeigt die Unterrichtsforschung, dass die von Lehrkräften selbsteingeschätzte Unterrichtsqualität überwiegend moderate Zusammenhänge mit den Qualitätseinschätzungen von Schülerinnen und Schülern aufweist (vgl. Fauth et al. 2014; Kunter und Baumert 2006; Kunter et al. 2008; Wagner et al. 2016). Darüber hinaus liegen Befunde vor, die zeigen, dass die Kompetenzselbsteinschätzung von angehenden Lehrkräften zwar schwach mit ihrem Abschneiden in Wissens- bzw. Performanztests zusammenhängt (vgl. König et al. 2012; Seifert et al. 2018), dass dieser Zusammenhang aber größer ausfällt, nachdem Lehramtsstudierende eigene Unterrichtserfahrungen sammeln konnten (vgl. Mertens und Gräsel 2018). Auch wenn es vereinzelt zu Verzerrungen bei der selbsteingeschätzten Kompetenzentwicklung kommen kann, gehen wir vor dem Hintergrund der zuvor zitierten Befunde davon aus, dass die in der vorliegenden Studie befragten Lehrkräfte, die überwiegend über eine langjährige Unterrichtserfahrung verfügen, die Weiterentwicklung ihrer Unterrichtskompetenz auch rückblickend differenziert beurteilen können.

Eine weitere Limitation betrifft die Passung der Feedbackintervention und der selbst eingeschätzten Kompetenzentwicklung. Während sich das Unterrichtsfeedback auf einzelne Unterrichtsmerkmale des theoretischen Unterrichtsmodells bezog, wurde die selbst eingeschätzte Kompetenzentwicklung global erfasst. Dass Schülerinnen und Schüler die Unterrichtsqualität von Lehrkräften der IG im Hinblick auf einzelne Merkmale des theoretischen Unterrichtsmodells signifikant besser einschätzten als Schülerinnen und Schüler von Lehrkräften der KG (vgl. Kellermann et al. im Druck), kann als Hinweis darauf interpretiert werden, dass eine nach den 14 Unterrichtsmerkmalen differenzierte Erfassung der Kompetenzentwicklung die Wahrscheinlichkeit, einen Haupteffekt zu finden, erhöhen könnte. Denkbar ist auch, dass erst ein mehrmaliges und ggf. regelmäßiges Unterrichtsfeedback der Schulleitung, das in Übereinstimmung mit der entwickelten Feedbackintervention gegeben wird, zu einer Weiterentwicklung der Unterrichtskompetenz von Lehrkräften führt. Der ausbleibende Haupteffekt könnte darüber hinaus auf die kleine Stichprobe von



$N=81$  Lehrkräften und eine damit einhergehende geringe Teststärke zurückzuführen sein.

Zusammenfassend kann festgehalten werden, dass die Feedbackintervention Schulleitungen dabei unterstützt, ein qualitativ hochwertiges Unterrichtsfeedback zu gestalten. Dass das Unterrichtsfeedback nur bei Lehrkräften mit hoch ausgeprägter *LZO* und *ALZO* mit einer Weiterentwicklung ihrer selbsteingeschätzten Unterrichtskompetenz einhergeht, zeigt aber auch, dass individuelle Merkmale von Lehrkräften für die Wirksamkeit von Unterrichtsfeedbacks von Relevanz sind. Dies sollte bei der Entwicklung von Feedbackinterventionen in Zukunft berücksichtigt werden. Insbesondere gilt es zu klären, wie ein Unterrichtsfeedback zu gestalten ist, von dem auch Lehrkräfte mit niedrig ausgeprägter *LZO* und *ALZO* profitieren. Hierbei könnte eine Verzahnung des Unterrichtsfeedbacks mit weiteren Formaten der professionellen Weiterentwicklung eine entscheidende Rolle spielen.

**Zusatzmaterial online** Zusätzliche Informationen sind in der Online-Version dieses Artikels (<https://doi.org/10.1007/s11618-022-01087-y>) enthalten.

**Förderung** Die vorliegende Studie ist im Rahmen des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekts „Unterrichtsbeurteilung und -feedback durch Schulleitungen“ (Projektnummer 283241518) entstanden.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

**Interessenkonflikt** C. Kellermann, M. Nachbauer, H. Gärtner und F. Thiel geben an, dass kein Interessenkonflikt besteht.

## Literatur

- Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3), 261–271. <https://doi.org/10.1037/0022-0663.84.3.261>.
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88(1), 87–99. <https://doi.org/10.1037/0021-9010.88.1.87>.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>.
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert & W. Blum (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.

- Butler, R. (2007). Teachers' achievement goal orientations and associations with teachers' help seeking. Examination of a novel approach to teacher motivation. *Journal of Educational Psychology*, 99(2), 241–252. <https://doi.org/10.1037/0022-0663.99.2.241>.
- Chughtai, A. A., & Buckley, F. (2010). Assessing the effects of organizational identification on inrole job performance and learning behaviour. *Personnel Review*, 39(2), 242–258. <https://doi.org/10.1108/00483481011017444>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsday: Lawrence Erlbaum.
- Dawson, J. F. (2014). Moderation in management research: what, why, when, and how. *Journal of Business and Psychology*, 29(1), 1–19. <https://doi.org/10.1007/s10869-013-9308-7>.
- Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson & C. S. Weinstein (Hrsg.), *Handbook of classroom management* (S. 97–126). Mahwah: Lawrence Erlbaum.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10), 1040–1048. <https://doi.org/10.1037//0003-066X.41.10.1040>.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology*, 53, 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>.
- Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: a mediational analysis. *Journal of Personality and Social Psychology*, 70(3), 461–475. <https://doi.org/10.1037/0022-3514.70.3.461>.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive: Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28(3), 127–137. <https://doi.org/10.1024/1010-0652/a000129>.
- Gärtner, H., Thiel, F., Nachbauer, M., & Kellermann, C. (2021). Unterrichtsentwicklung durch Unterrichtsfeedback – Erste Entwicklungsschritte eines Beobachtungsbogens für Schulleitungen. *Zeitschrift für Bildungsforschung*, 89(11), 529–547. <https://doi.org/10.1007/s35834-021-00321-7>.
- Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: necessary and illuminating. *Journal of Educational Psychology*, 94(3), 638–645. <https://doi.org/10.1037//0022-0663.94.3.638>.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. New York: Guilford.
- Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität – Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Friedrich.
- Hulleman, C. S., Schragger, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, 136(3), 422–449. <https://doi.org/10.1037/a0018947>.
- Ilgén, D. R., & Davis, C. A. (2000). Bearing bad news: reactions to negative performance feedback. *Applied Psychology: An International Review*, 49(3), 550–565. <https://doi.org/10.1111/1464-0597.00031>.
- Janke, S., Bardach, L., Oczlon, S., & Lüftenegger, M. (2019). Enhancing feasibility when measuring teachers' motivation: a brief scale for teachers' achievement goal orientations. *Teaching and Teacher Education*, 83, 1–11. <https://doi.org/10.1016/j.tate.2019.04.003>.
- Kellermann, C., Nachbauer, M., Gärtner, H., & Thiel, F. (under review). The effect of school leaders' feedback on teaching quality perceived by students—Results of an intervention study.
- Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, 45(1), 34–70. <https://doi.org/10.1177/0013161X08327549>.
- Klauer, K. J., & Leutner, D. (2012). *Lehren und Lernen. Einführung in die Instruktionspsychologie* (2. Aufl.). Weinheim: Beltz.
- Klieme, E., Schümer, G., & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In BMBF (Hrsg.), *TIMSS – Impulse für Schule und Unterricht, Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–58). Bonn: BMBF.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>.
- König, J., Kaiser, G., & Felbrich, A. (2012). Spiegelt sich pädagogisches Wissen in den Kompetenzselbst-einschätzungen angehender Lehrkräfte? Zum Zusammenhang von Wissen und Überzeugungen am Ende der Lehrerausbildung. *Zeitschrift für Pädagogik*, 58(4), 476–491.

- Kopp, B., & Mandl, H. (2014). Aspekte der Feedbacknachricht. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 151–162). Münster: Waxmann.
- Kraft, M. A., & Gilmour, A. (2016). Can principals promote teacher development as evaluators? A case study of principals' views and experiences. *Educational Administration Quarterly*, 52(5), 711–753. <https://doi.org/10.1177/0013161X16653445>.
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: a meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251. <https://doi.org/10.1007/s10984-006-9015-7>.
- Kunter, M., Kleickmann, T., Klusmann, U., & Richter, D. (2011). Die Entwicklung professioneller Kompetenz von Lehrkräften. In M. Kunter, J. Baumert & W. Blum (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 55–69). Münster: Waxmann.
- Kunter, M., Tsai, Y.-M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, 18(5), 468–482. <https://doi.org/10.1016/j.learninstruc.2008.06.008>.
- Kwakman, K. (2003). Factors affecting teachers' participation in professional learning activities. *Teaching and Teacher Education*, 19(2), 149–170. [https://doi.org/10.1016/S0742-051X\(02\)00101-4](https://doi.org/10.1016/S0742-051X(02)00101-4).
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern* (Zeitschrift für Pädagogik: Beiheft 51, S. 47–70). Weinheim: Beltz.
- Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 58(2), 103–117. <https://doi.org/10.1026/0033-3042.58.2.103>.
- McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: a review and illustration. *Educational Psychology Review*, 28(2), 295–314. <https://doi.org/10.1007/s10648-014-9287-x>.
- Mertens, S., & Gräsel, C. (2018). Entwicklungsbereiche bildungswissenschaftlicher Kompetenzen von Lehramtsstudierenden im Praxissemester. *Zeitschrift für Erziehungswissenschaft*, 21(6), 1109–1133. <https://doi.org/10.1007/s11618-018-0825-z>.
- Milanowski, A. T., & Heneman, H. G. (2001). Assessment of teacher reactions to a standards-based teacher evaluation system: a pilot study. *Journal of Personnel Evaluation in Education*, 15(3), 193–212. <https://doi.org/10.1023/A:1012752725765>.
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art—teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>.
- Müller, A., & Ditton, H. (2014). Feedback: Begriffe, Formen und Funktionen. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 11–28). Münster: Waxmann.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus. Statistical analysis with latent variables. User's guide*. Los Angeles: Muthén & Muthén.
- Nicholls, J. G. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91(3), 328–346. <https://doi.org/10.1037/0033-295X.91.3.328>.
- Nitsche, S., Dickhäuser, O., Dresel, M., & Fasching, M. S. (2013a). Zielorientierungen von Lehrkräften als Prädiktoren lernrelevanten Verhaltens. *Zeitschrift für Pädagogische Psychologie*, 27(1–2), 95–103. <https://doi.org/10.1024/1010-0652/a000092>.
- Nitsche, S., Dickhäuser, O., Fasching, M. S., & Dresel, M. (2011). Rethinking teachers' goal orientations. Conceptual and methodological enhancements. *Learning and Instruction*, 21(4), 574–586. <https://doi.org/10.1016/j.learninstruc.2010.12.001>.
- Nitsche, S., Dickhäuser, O., Fasching, M. S., & Dresel, M. (2013b). Teachers' professional goal orientations: Importance for further training and sick leave. *Learning and Individual Differences*, 23, 272–278. <https://doi.org/10.1016/j.lindif.2012.07.017>.
- Ophardt, D., & Thiel, F. (2013). *Klassenmanagement. Ein Handbuch für Studium und Praxis*. Stuttgart: Kohlhammer.

- Piwowar, V., & Thiel, F. (2014). Evaluating response shift in training evaluation: comparing the retrospective pretest with an adapted measurement invariance approach in a classroom management training program. *Evaluation Review*, 38(5), 420–448. <https://doi.org/10.1177/0193841X14546932>.
- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of three basic dimensions. *ZDM*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods*. Thousand Oaks: SAGE.
- Rheinberg, F. (2008). *Motivation*. Stuttgart: Kohlhammer.
- Runhaar, P., Sanders, K., & Yang, H. (2010). Stimulating teachers' reflection and feedback asking. An interplay of self-efficacy, learning goal orientation, and transformational leadership. *Teaching and Teacher Education*, 26(5), 1154–1161. <https://doi.org/10.1016/j.tate.2010.02.011>.
- Scheeler, M. C., Ruhl, K. L., & McAfee, J. K. (2004). Providing performance feedback to teachers: a review. *Teacher Education and Special Education*, 27(4), 396–407. <https://doi.org/10.1177/088840640402700407>.
- Schwartz, C. E., & Sprangers, M. A. G. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48, 1531–1548. [https://doi.org/10.1016/S0277-9536\(99\)00047-7](https://doi.org/10.1016/S0277-9536(99)00047-7).
- Seifert, A., Schaper, N., & König, J. (2018). Bildungswissenschaftliches Wissen und Kompetenzeinschätzungen von Studierenden im Praxissemester: Veränderungen und Zusammenhänge. In J. König, M. Rothland & N. Schaper (Hrsg.), *Learning to Practice, Learning to Reflect? Ergebnisse aus der Längsschnittstudie LtP zur Nutzung und Wirkung des Praxissemesters in der Lehrerbildung* (S. 325–347). Wiesbaden: Springer VS.
- Senko, C., & Dawson, B. (2017). Performance-approach goal effects depend on how they are defined: meta-analytic evidence from multiple educational outcomes. *Journal of Educational Psychology*, 109(4), 574–598. <https://doi.org/10.1037/edu0000160>.
- Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: old controversies, current challenges, and new directions. *Educational Psychologist*, 46(1), 26–47. <https://doi.org/10.1080/00461520.2011.538646>.
- Steelman, L. A., Levy, P. E., & Snell, A. F. (2004). The feedback environment scale. Construct definition, measurement, and validation. *Educational and Psychological Measurement*, 64(1), 165–184. <https://doi.org/10.1177/0013164403258440>.
- Steinberg, M. P., & Sartain, L. (2015). Does teacher evaluation improve school performance? Experimental evidence from Chicago's Excellence in Teaching Project. *Education Finance and Policy*, 10(4), 535–572. [https://doi.org/10.1162/EDFP\\_a\\_00173](https://doi.org/10.1162/EDFP_a_00173).
- Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, 102(7), 3628–3651. <https://doi.org/10.1257/aer.102.7.3628>.
- Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: a learning theory perspective. *Educational Research Review*, 9, 1–15. <https://doi.org/10.1016/j.edurev.2012.11.004>.
- Urban, D., & Mayerl, J. (2018). *Angewandte Regressionsanalyse: Theorie, Technik und Praxis*. Wiesbaden: Springer. <https://doi.org/10.1007/978-3-658-01915-0>.
- Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, 108(5), 705–721. <https://doi.org/10.1037/edu0000075>.