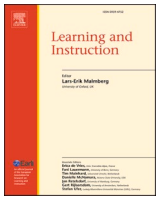
Learning and Instruction 97 (2025) 102104

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc



Teacher collaboration to elevate student achievement?

Andrea Wullschleger [a,1,*], Katharina Maag Merki [b,1], Urs Grob [b], Beat Rechsteiner [b], Miriam Compagnoni [b], András Vörös [c]

[a] Department of Research and Development, School of Education, University of Applied Sciences and Arts Northwestern Switzerland Bahnhofstrasse 6, 5210, Windisch, Switzerland
[b] Institute of Education, University of Zurich, Freiestrasse 36, 8032, Zurich, Switzerland
[c] School of Social Policy, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

ARTICLE INFO

Keywords:
Teacher collaboration
Social networks
Instructional quality
Mathematics achievement
Multilevel longitudinal mediation model

ABSTRACT

Background: Previous research on the relationship between teacher collaboration and student achievement yielded ambiguous results. From a theoretical perspective, an indirect association between high-quality teacher collaboration and student achievement is assumed, mediated by instructional quality. However, empirical evidence for this assumed theoretical mediation model is lacking.
Aim: This study analyses the relationship between the extent of high-quality teacher collaboration reported by teachers via social network ties and the development of fifth-grade students' mathematics achievement over one school year. It examines whether and how this relationship is mediated by instructional quality assessed as effective teaching from the perspective of the students.
Sample: The sample included 80 primary school teachers and 770 fifth-grade students from the German-speaking part of Switzerland.
Method: A two-level structural equation model was applied, comprising two latent change models, one for instructional quality and one for mathematics achievement, to examine student survey data on instructional quality, student mathematics achievement tests, and teacher social network survey data on perceived high-quality teacher collaboration.
Results: The results revealed a significant positive indirect effect between high-quality teacher collaboration and student achievement, mediated by instructional quality. However, the direct effect between high-quality teacher collaboration and mathematics achievement, although positive, was not significant. Both the total and indirect effects of the model were significant.
Conclusion: This study underpins and extends previous findings emphasizing the significance of perceived high-quality teacher collaboration by demonstrating that these collaborations are crucial for fostering instructional quality and students' mathematics achievement.

1. Introduction

Previous research findings on the effects of teacher collaboration on student achievement are far from conclusive. For one, only very few studies analyzed the effects in detail (e.g., Vangrieken et al., 2015). For another, teacher collaboration only seems to be consistently related to improvements in student achievement when the quality of teacher collaboration is considered (e.g., Antoniou & Kyriakides, 2011; Kılınç et al., 2023; Lomos et al., 2011; Ronfeldt et al., 2015). Furthermore, teacher collaboration appears to be indirectly rather than directly related to student achievement (Hochweber et al., 2012; Kılınç et al., 2023; Placklé et al., 2022).

However, the analysis of that mediation is seldom conducted. From a theoretical perspective, instructional quality is considered as a relevant mediator (Spillane & Louis, 2002), because it is a significant proximal predictor of student outcomes (e.g., Antoniou & Kyriakides, 2011). Respective analyses are lacking, however. Previous studies have only analyzed the mediation effect of teacher self-efficacy between teacher

* Corresponding author.
E-mail addresses: andrea.wullschleger@fhnw.ch (A. Wullschleger), kmaag@ife.uzh.ch (K. Maag Merki), grob@ife.uzh.ch (U. Grob), beat.rechsteiner@ife.uzh.ch (B. Rechsteiner), mcompagnoni@ife.uzh.ch (M. Compagnoni), a.voros@bham.ac.uk (A. Vörös).
[1] Andrea Wullschleger and Katharina Maag Merki share first authorship.

https://doi.org/10.1016/j.learninstruc.2025.102104
Received 23 July 2023; Received in revised form 14 February 2025; Accepted 15 February 2025

collaboration and student achievement, finding that teacher collaboration has a positive effect on teacher self-efficacy, which, in turn, improves student achievement (Kılınç et al., 2023; Moolenaar et al., 2012). Also, previous studies have investigated the effects of teacher collaboration on the development of teaching (e.g., Nguyen, Pietsch, & Gümüs, 2021) or student achievement (e.g., Goddard & Goddard, 2007) separately. Hence, a simultaneous longitudinal analysis is needed to test theoretical models to improve knowledge on how teacher collaboration fosters student achievement.

This study therefore examined the relationship between high-quality teacher collaboration among fifth-grade teachers in primary schools in the German-speaking part of Switzerland and the development of students' mathematics achievement, and whether this relationship is mediated by the development of instructional quality.

To analyze high-quality teacher collaboration, we relied on recent studies that point to the importance of teachers' experiences of teacher collaboration for supporting teachers' professional practice, such as teachers' perceived helpfulness of teacher collaboration (Ronfeldt et al., 2015) and perceived discussion utility of teacher collaboration (Sinnema et al., 2021). Hence, high-quality teacher collaboration in this study refers to the perceived usefulness of teacher collaboration for supporting school improvement. Our analysis of instructional quality was based on the concept of "opportunity and use" (Fend, 1981; Vieluf & Klieme, 2023). This concept is grounded in the idea that learners are active co-constructors of their own learning progress. Thus, the effectiveness of instruction depends not only on the learning opportunities provided in lessons but also on whether and how students engage with these opportunities. For this reason, our study focused on effective teaching from the perspective of the students: students' perception of the quality of classroom instruction.

2. The triad of teacher collaboration, student achievement, and instructional quality – theoretical framework

2.1. High-quality teacher collaboration on school improvement

The existence of various definitions makes it necessary to clarify the concept of high-quality teacher collaboration. Following Vangrieken et al.'s (2015) definition of collaboration as an "umbrella term" (p. 23), this study defined teacher collaboration as shared social activities among teachers to fulfil a common task or profession-related goal.

As our research aligns with theoretical approaches in school improvement research (e.g., Harris, 2002; Hopkins, 2001; Mitchell & Sackney, 2011) and social network research (e.g. Daly, 2010), we focused in this study on teacher collaboration as teachers' shared social activities to *develop teaching, improve working in teams, and optimize organizational procedures and structures* to establish their school as a learning organization with the ultimate aim of optimally promoting the learning of all pupils. Accordingly, this study included the following collaborative activities:

1. *Teaching improvement*: collaboration on teachers' primary tasks (James, Dunning, Connolly, & Elliott, 2007), which have the potential to increase teachers' professional capital (Hargreaves & Fullan, 2012). For example, teachers could jointly negotiate how students can be fostered or what methods could help improve students' learning
2. *Team improvement*: collaboration on how their working in teams can be improved (e.g., Decuyper et al., 2010)
3. *Organizational improvement*: collaboration on what norms and goals, procedures, and structures schools should pursue or how they could address significant social challenges (Mitchell & Sackney, 2011).

This emphasis, referred to as *teacher collaboration on school improvement*, broadens the focus of collaboration from solely classroom practice (Wullschleger et al., 2023) to individual, interpersonal, and organizational improvement in schools (Mitchell & Sackney, 2011) and considers that organizational and interpersonal improvement influence teachers' professional learning, which in turn affects student achievement (Kyriakides et al., 2015). Furthermore, interactions concerning work-related information or expertise form the core of collaboration, as they provide access to other actors' professional capital (e.g., Moolenaar et al., 2012). Hence, this approach is strongly oriented towards educational change (Daly, 2010).

In addition, this study was interested in analyzing the effect of *high-quality* teacher collaboration on the development of instructional quality and student achievement. Many studies have shown that collaboration can vary in terms of the depth of interactions (Decuyper et al., 2010; Havnes, 2009; Schippers, Den Hartog, & Koopman, 2007; Yang et al., 2018) and that collaboration that meets high-quality standards seems to have the potential to improve instructional quality and student achievement (e.g., Antoniou & Kyriakides, 2011; Moolenaar et al., 2012; Ronfeldt et al., 2015). In this study, we followed Sinnema et al. (2021) and Ronfeldt et al. (2015), who investigated the quality of interactions by focusing on perceived discussion utility (Sinnema et al., 2021) or the perceived helpfulness of teacher collaboration (Ronfeldt et al., 2015). Both argued that the perceived quality of interactions is particularly effective in terms of fostering professional practice and student achievement. Applied to collaboration, this signifies the extent to which teachers experience various forms of collaboration as beneficial for supporting their professional practice and student learning. When teachers experience several perceived useful collaborations in their social network, this can subsequently be understood as interaction routines on a deep level (Coburn & Russell, 2008).

Hence, we defined high-quality teacher collaboration as social interactions between teachers in a school's social network that are perceived as useful and have the goal of discussing new ideas for improving teaching, working in teams, and the school's organizational procedures and structures, all aimed at fostering students learning. As collaboration of this kind involves joint discussion and negotiation, it is close to the interdependence side of the continuum of teacher collaboration (Little, 1990) and co-constructive in nature (Gräsel et al., 2006).

2.2. Teacher collaboration and student achievement, an indirect relationship

Educational effectiveness theories (e.g., Creemers & Kyriakides, 2008; Spillane & Louis, 2002) emphasize that the preconditions of student outcomes must be modeled within a multilevel framework that distinguishes between distal factors, such as characteristics of the educational system (e.g., policy) and the school (e.g., school leadership), and proximal factors, such as teachers' competencies and teaching practice. Proximal factors are assumed to directly affect student outcomes. Distal factors influence student outcomes indirectly by improving instructional quality, which can be supported by two large-scale longitudinal studies (Heck & Hallinger, 2014; Kyriakides et al., 2015).

Instructional quality in general refers to what teachers do in their classroom that is positively related to student outcomes. However, researchers' conceptualization and operationalization of the construct vary (Nilsen et al., 2016; Senden et al., 2022; Praetorius and Charalambous, 2023). This study aligned with the concept of "opportunity and use" (Fend, 1981; Vieluf & Klieme, 2023) and focused on effective teaching from the students' perspective. The opportunity-use concept describes opportunities provided by teachers during teaching processes and how they are used by individual learners. 'Opportunity' refers to teaching processes orchestrated by the teacher, and 'use' refers to student's individual, cognitive, motivational and emotional learning processes. The same opportunities are not necessarily used by all students and not by all students in the same way. Therefore, "teaching has no universal quality, but needs to be adaptive to the particular needs of each individual student" (Vieluf & Klieme, 2023, p. 68). That is why

students' perceptions and experiences play a pivotal role in their learning outcomes. Integrating students' perspectives and perceptions of the quality of classroom instruction yields a deeper understanding of the effectiveness of teaching practices (Hattie, 2012).

The multilevel perspective, taking into account direct and indirect associations, also seems to be valid concerning the relationship between teacher collaboration and student achievement. Spillane and Louis (2002) argue that instruction is the proximal cause of students' opportunity to learn. Therefore, school improvement processes, in general, and the collaborative practice of teachers in particular, must focus on fostering instructional capacity because it enhances student achievement.

This assumption is also supported by Hargreaves and Fullan's (2012) theoretical framework on professional capital. Hargreaves and Fullan argued that social capital, referred to as the capital teachers share through their network of learning, strength of mutual support, shared professional development, and strong foundation of trust, adds value to individual human capital such as teachers' competencies (Hargreaves & Fullan, 2012). Accordingly, teacher collaboration as a social structure within the school and the transformation of social capital into individual human capital is vital for increasing the quality of teaching and, ultimately, student learning.

Ronfeldt et al. (2015) offered a significant differentiation of this perspective, arguing that teacher collaboration seems to be both a school and an individual phenomenon. Collaboration at the school level can be interpreted as an essential context for the teachers' individual collaborative practice. Accordingly, individual and collective mechanisms could explain how teacher collaboration is associated with student achievement: A collaborative culture in a school could enable teacher collaboration on a team level, which, in turn, could enable teachers to develop their teaching and improve student achievement (Ronfeldt et al., 2015).

Hence, robust theoretical evidence indicates that teacher collaboration is indirectly related to student achievement and mediated by enhanced classroom instruction quality. High-quality teacher collaboration should foster teachers' human capital (Hargreaves & Fullan, 2012), which improves teachers' instructional capabilities (Spillane & Louis, 2002). Increased instructional quality should lead to a higher level of student achievement.

3. Instructional quality as mediator between the relationship of teacher collaboration and student achievement – state of empirical results and research gaps

Empirical evidence for the assumed theoretical effect exists only partially. Studies have explored the relationship between teacher collaboration and student achievement (Akiba & Liang, 2016; Goddard & Goddard 2017; Kılınç et al., 2023; Lomos et al., 2011; Ronfeldt et al., 2015; Vescio et al., 2008) and between teacher collaboration and instructional quality (Holzberger & Schiepe-Tiska, 2021; Sleegers et al., 2014). These relationships have also been investigated using social network analyses (Daly et al., 2020; Daly et al., 2014; Horn et al., 2020; Pil & Leana, 2009; Sinnema et al., 2021). However, studies that consider all three elements together are rare. Furthermore, cross-sectional studies prevail (Weddle, 2022).

Hochweber et al. (2012) used multilevel regression analyses to examine whether teacher collaboration is related to instructional quality and student achievement in English as a foreign language. The results indicated that teacher collaboration is positively related to instructional quality (structure and cognitive challenge). Several dimensions revealed an effect on student achievement and motivation. However, Hochweber et al. failed to find a direct effect of teacher collaboration on student achievement. They also did not examine the mediation effects between teacher collaboration, instructional quality, and student achievement.

Placklé et al. (2022) related teaching qualifications, teacher collaboration, and effective teaching practices, individually and collectively, to student achievement in secondary vocational education. They found a strong positive correlation between teacher collaboration and almost all elements of effective teaching practice. The results revealed a significant positive correlation between effective teaching practices and mathematics achievement. However, no significant correlation was observed between teacher collaboration and student achievement or between teaching qualifications and student achievement. In addition, the study did not test a mediation model.

Given these findings and the theoretical arguments, it is notable that appropriate mediation analyses have not been conducted to empirically test the effect pathway from teacher collaboration to student achievement through instructional quality. Existing studies focus on either the effect of teacher collaboration on instruction or teacher collaboration on student achievement, without considering the two effects simultaneously. Studies that come close to addressing this gap have been conducted by Moolenaar et al. (2012), Kılınç et al. (2023), and Liu and Yin (2024). They examined the mediating effect of teachers' collective efficacy beliefs between teacher collaboration and student achievement and found no direct effect of teachers' social network structure or teacher collaboration on student achievement. However, their results indicated that the density of advice networks, the frequency of teacher collaboration, or professional learning communities affect teachers' efficacy, which, in turn, is linked to higher student achievement and, in the study by Liu and Yin (2024), also to cognitive activation in lessons. Those studies, however, were cross-sectional in nature and did not analyze instructional quality as a mediator. Nevertheless, their results provided further empirical support for the assumption of a mediation model.

In summary, the existing empirical literature on the relationship between teacher collaboration and student achievement is mixed. Most importantly, the results discussed in this section reveal that only a few studies examined the effects of collaboration, instructional quality, and student achievement simultaneously. However, these studies did not explore the mediating role of instructional quality between high-quality teacher collaboration and student achievement, as the relevant effects were analyzed separately.

A direct empirical test of the theoretical mediation model is necessary to better understand the relationship between high-quality collaboration, instruction, and student achievement. The findings could contribute to the development of effective professional learning tools to support teacher collaborations aimed at improving teaching practices.

4. Aims, research questions, and hypotheses

This study examined whether perceived high-quality teacher collaboration affects the development of instructional quality and student mathematics achievement in fifth graders. Teacher collaboration on school improvement was assessed at the beginning of the school year, capturing teachers' collaborative practice over the preceding six months. Teachers rated its perceived usefulness for school improvement, including improving teaching, teamwork, and the school's organizational procedures and structures. The study focused in particular on the mediating role of instructional quality in this process.

Given the temporal structure of the data, with two measurement points, and the possibility of a decrease in instructional quality and mathematics achievement over time, development was measured in terms of change and referred to as such. This approach is in line with statistical analyses based on latent change modelling (LCM; see section 5 below). This study investigated the following research questions:

1. To what extent does the perceived usefulness of teacher collaboration for school improvement have a positive effect on changes in student achievement?
2. Is this relationship mediated by changes in instructional quality?

Based on the theoretical framework of teachers' professional capital,

school improvement capacity, and collaboration routines (Coburn & Russell, 2008; Hargreaves & Fullan, 2012; Mitchell & Sackney, 2011; Ronfeldt et al., 2015; Spillane & Louis, 2002), we assumed that teachers' collaborative practice already in place at the beginning of the school year can be considered an organizational routine and a key component of school improvement capacity for academic learning. Research revealed that organizational routines inform teachers' discussion about instruction and influence subsequent professional activities as well as student learning (Hatch & Hill, 2016; Spillane et al., 2016). Furthermore, longitudinal studies revealed that school improvement capacity at t1 influences school improvement capacity and student achievement in the following year. (Hallinger & Heck, 2010).

Second, we assumed that the number of collaboration ties that teachers perceived as useful for school improvement has a positive indirect effect on changes in student achievement, mediated by changes in instructional quality. In the absence of studies analyzing the mechanism underlying how teacher collaboration affects student achievement mediated by instructional quality, it is unclear whether full or partial mediation of the relationship exists. Accordingly, the direct effects were also exploratively tested.

5. Method

5.1. Study context and sample

Data were drawn from a large-scale study investigating school improvement capacity in 59 public primary schools from the German-speaking part of Switzerland. In Switzerland, primary-level education comprises eight years, including two years of kindergarten, and all schools implement inclusive education. All schools took part in the study voluntarily, and all participants gave their informed consent to participate. All parents were required to sign a consent form for their child's participation in the study. The questionnaires and study procedures were approved by the Ethical Commission of the University of Zurich.

The sample schools differed in terms of regional context, size, and social context. The sample included slightly more urban schools than the average in the respective country. Large schools were slightly overrepresented in the sample compared to the population. Regarding social context, the sample was representative of the respective country. A total of 1630 teachers, specialist teachers, and principals (87% of whom were women) took part in the study. As in Switzerland where the study took place, all members of a school team are responsible for teaching, including the principal, we used the term 'teacher' for all participants. Their ages ranged from 21 to 67 years ($M = 42.81$, $SD = 11.51$), and they had worked in schools on average for 17 years ($SD = 11.01$, $min. = 0$, $max. = 45$). The sample of teachers was representative of employees in schools in the respective country. In addition, 1637 fifth graders (48% of whom were girls) from 128 classes participated in the study ($M = 11.79$ years, $SD = 0.48$). More than half of the children were taught in multigrade classes. In the multigrade classes, the number of fifth graders was lower than in the single-grade classes. On average, each class contained 13 fifth graders ($SD = 5.39$).

For the analyses of this study, a subsample of the fifth-grade teachers was used. This approach was chosen to enable the linking of student data on mathematics achievement and instructional quality with their respective teachers. However, the data of the entire sample of teachers was involved by utilizing their collaboration ties and their perception of discussion utility with the fifth-grade teachers.

The initial sample consisted of 147 fifth-grade teachers and their 1637 fifth graders from 59 schools. Three schools with their 12 teachers were excluded from the study because the network survey response rate was too low. We further excluded 19 schools with their 47 teachers respectively from very small or very large schools, to make the social network measure used in the analysis comparable between schools (see section 5.4 for details and justification). Due to missing data on teachers in two schools, the final sample for this study included 35 schools, 80 teachers, and 770 students.

5.2. Procedure

Data on perceived high-quality teacher collaboration ties were collected using an online questionnaire at the beginning of the 2019/20 school year. The survey response rate was 81.6%. The subsample of fifth-grade teachers yielded a response rate of 100%.

Student data were collected at the beginning (t1) and end (t2) of the school year. At t1, the response rate was 77.6% for the mathematics achievement assessment and 77.9% for the questionnaire on instructional quality. At t2, it was 71.2% for the mathematics achievement test and 72.8% for the instructional quality questionnaire. Instructional quality was assessed together with sociodemographic variables via a paper–pencil questionnaire, and mathematics achievement was assessed using a paper–pencil test. The questionnaire and tests were administered in the classroom by trained test administrators according to standardized instructions. See Fig. 1 for an overview. The duration of data collection was 2 lessons (school periods) in total.

5.3. Measures

5.3.1. Teacher collaboration

Perceived high-quality teacher collaboration was surveyed using social network questions in an online questionnaire distributed to all teachers in the sample. Three social network questions were connected to a rating of the perceived usefulness of the collaboration: With whom have you discussed improvement: (1) of your teaching, (2) of how you work in teams, and (3) of your school's organizational procedures and structures in the last six months? For each indicated collaboration, participants additionally rated the collaboration's usefulness for improving: (1) their teaching, (2) their work in teams, and (3) their school's organizational procedures and structures, using a 4-point scale ranging from 1 (not at all useful) to 4 (very useful). Participants were given a list of all teachers on their school team from which they could select interaction partners.

In this study, we focused on perceived high-quality of teacher collaboration and therefore selected only collaboration ties that teachers considered very useful for school improvement. In the teaching improvement network, 47.6% of the collaboration ties were perceived as very useful. This figure is 49.6% for the team improvement network and 40.9% for the organizational improvement network. The mean density of the networks, including all ties, was 0.46 for teaching improvement, 0.39 for team improvement, and 0.42 for organizational improvement. The mean density of the networks, including only the very useful ties, was 0.14 for teaching improvement, 0.11 for team improvement, and 0.11 for organizational improvement. Further, we used only the indegree of the actors, meaning the number of incoming relationships, which refers to the number of colleagues (teaching in any grade) who perceived the collaboration with the respective fifth-grade teacher as very useful for developing their teaching, their working in teams, and

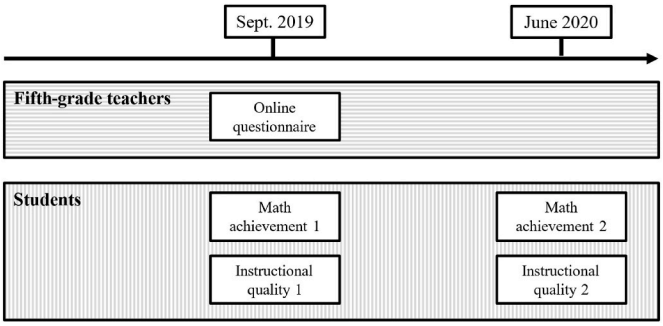


Fig. 1. Overview of data collection procedure.

their school's organizational procedures and structures.

The advantage of using the indegree was twofold: The perceived usefulness of teacher collaboration on school improvement was not assessed by the teachers themselves but by others. This means that information about each teacher's involvement in collaborations is based on reports from multiple respondents and is less affected by non-response. We also carried out the analyses with the outdegree (the number of ties sent by each teacher) and total degree (the sum of the teacher's in- and outdegree) for comparison, which led to very similar results.

We combined the number of ties each teacher received across the three collaboration items in a single measure. The internal consistency of the manifest scale was high ($M = 0.11$, $SD = 0.08$, $\alpha = 0.89$). Table 1 presents the factor loadings. The factor score represents a teacher's involvement in high-quality collaborations with their colleagues.

5.3.2. Instructional quality

Instructional quality was assessed by students' perception of teaching for effective learning on mathematics achievement using three items based on Jaekel and Göllner (2020). Children assessed the following three statements using a 4-point scale ranging from 1 (not at all true) to 4 (exactly true): (1) In math class taught by our math teacher, you can learn a lot, (2) I think our math teacher's lessons are good, and (3) I am satisfied with our mathematics teacher's lessons. The manifest scale's internal consistency was high at t1 ($M = 3.6$, $SD = 0.51$, $\alpha = 0.83$) and t2 ($M = 3.55$, $SD = 0.59$, $\alpha = 0.88$). Considering the intra-class correlations (Lüdtke et al., 2006), the variances between classes were in line with other studies in primary schools (e.g., Cohen et al., 2018), with an increase from t1 to t2 (t1: ICC 1 = 0.125; t2: ICC 1 = 0.189). The students' individual scores were latently aggregated to assess instructional quality at the class level (Lüdtke et al., 2008). The reliability was relatively low for t1 (ICC 2 = 0.59) and acceptable for t2 (ICC 2 = 0.69).

5.3.3. Mathematics achievement

We assessed mathematics achievement using items from the Check P5 item bank (Institut für Bildungsevaluation, 2020) to ensure strong curricular validity. These items align with the three mathematics competence areas outlined by the national curriculum for the respective country: numbers and variables; shape and space; and metrics, functions, data, and probability. The item pool consists of approximately 250 calibrated items for Grade 5, which have been completed by up to 40,000 students to date. A rotating test design was applied to avoid identical sets of items at t1 and t2. Four versions of the test booklet were created, containing overlapping item packages with matching difficulty levels at both time points.

A one-dimensional item response theory (IRT) model was employed (Van Linden & Hambleton, 1997) to avoid identical sets of items at t1 and t2. Specifically, a two-parameter logistic model (also known as the Birnbaum model; Birnbaum, 1968) was used to evaluate these dichotomous test items. The item parameters are based on t1. This model estimates both item difficulty (β) and item discriminability (α; De Ayala, 2009).

Data were scaled using the TAM software package by Kiefer et al. (2017) in the R development environment (http://www.rstudio.com). Weighted least square estimates (WLE) for mathematics competence were estimated and used as test scores (t1: $M = -0.42$, $SD = 1.22$; t2: $M = 0.27$, $SD = 1.26$). These estimates consider the item parameters and provide a weighted measure of mathematics competence based on the responses to the test items. Higher values indicate higher achievement. Differential item functioning (DIF) was verified by determining whether the link items functioned in the same way in our sample and the large Check P5 sample (Institut für Bildungsevaluation, 2020; Kiefer et al., 2017). The expected a posteriori reliability for the entire model was rEAP = 0.82, and the reliability of the WLE scores was rWLE = 0.81.

Considering the intra-class correlations (Lüdtke et al., 2006), the variances between classes remained the same at t1 and t2 (t1: ICC 1 = 0.143; t2: ICC 1 = 0.140). The students' individual scores were latently aggregated to assess mathematics achievement at the class level (Lüdtke et al., 2008). The reliability was acceptable for t1 (ICC 2 = 0.61) and relatively low for t2 (ICC 2 = 0.59).

Table 1
Factor loading for very useful teacher collaboration.

Item		Loadings
1	With whom have you discussed the improvement of your teaching in the last 6 months?	0.89
2	With whom have you discussed the improvement of how you work in the team in the last 6 months?	0.77
3	With whom have you discussed the improvement of your school's organizational procedures and structures in the last 6 months?	0.78

5.3.4. Control variables

Student age was collected by asking for date of birth. Age was then calculated in years ($N = 1{,}183$, $M = 11.81$, $SD = 0.48$, $Min = 10.28$, $Max = 14.59$). *Student gender* was controlled for ($f = 48\%$).

Socioeconomic context was assessed following Bos et al. (2016), using five indicators of wealth, which were answered 'yes' or 'no' (e.g., Do you have a room at home just for yourself?). A sum score was derived from these five indicators, with values ranging from 0 to 5. Children from less affluent backgrounds were defined as those with scores <4. They constituted 14.2% of the sample.

Migration background was assessed using an item from the KESS study (Bos et al., 2016). Students were asked if their father and mother were born in the country of current residence. This item was measured in binary form: both parents were born abroad (1 = migration background) vs. both parents or one parent was born in the country of residence (0 = no migration background). For 27% of the students, both parents were born abroad.

5.4. Statistical analysis

It should be noted that social network data consist of ties between actors and are, accordingly, not independent (Snijders, 2011). Schools must be of similar size to ensure the comparability of the number of ties received in different schools: The possible number of colleagues who perceive the collaboration of a given teacher as useful is much larger in a school with many teachers than in a school with few teachers. To be able to consider the largest possible number of schools for the study, the social network data were examined to determine the relationship between school size and the average indegree of teachers in the schools. Linear and curvilinear correlations were assessed. The sample was gradually reduced by dropping the largest and smallest schools until a significant relationship between school size and average indegree values was no longer present. The sample was narrowed to schools with no fewer than 15 teachers and no more than 50 teachers.

Next, a two-level structural equation model (SEM) was applied, comprising two latent change models (LCM)—one for instructional quality and one for mathematics achievement (McArdle, 2009)—and using MPlus statistical software (Version 8.8; Muthén & Muthén, 2017). LCMs represent the intraindividual change between two measurement points as the value of a latent variable. In our analysis, instructional quality assessed by students and students' estimated mathematics ability scores were represented at the between level (teacher level) in the format of a latently aggregated manifest score (Lüdtke et al., 2008).

The model was controlled at level 1 for the students' gender, socioeconomic context, migration background, and age. At level 2, classroom composition was controlled for in terms of socioeconomic context and migration background. The latter two control variables, due to being binary at level 1, were manifestly aggregated at level 2 and group-mean centered at level 1.

Given the limited number of cases and the model's complexity at the

between level, we refrained from applying a doubly latent approach for the LCM on instructional quality (Marsh et al., 2009). We isolated the effects of collaboration on student achievement mediated by changes in instructional quality by controlling for student age, socioeconomic context, and migration background at the within level and socioeconomic context and migration background at the between level. We tested theoretically driven one-sided hypotheses to determine the significance of the effects.

6. Results

6.1. Correlations

Table 2 shows the correlation coefficients for the within level (lower triangular) and the between level (upper triangular) between instructional quality, mathematics achievement, the control variables, and useful collaboration (only between-level).

6.2. Measurement invariance

Scalar measurement invariance is required for the comparison of mean values over time (Millsap, 2011). This also applies to latent change models based on manifest scale variables, as used in the present analyses. We conducted a two-level confirmatory factor analysis and tested the doubly-latent measurement model, based on three indicators at both levels, behind the scale variable instructional quality, for scalar measurement invariance over time (pairwise identical loadings and identical intercepts at level 1 and level 2). Given the sample size of n = 770 cases, a delta CFI criterion of 0.010 and a delta RMSEA criterion of 0.015 were considered adequate (Putnick & Bornstein, 2016).

The analyses revealed a significant Chi-Square value ($p < .001$) for the model with scalar restrictions applied to both levels, but otherwise a good fit (CFI = 0.979, RMSEA = 0.047). For level 1, with the intercepts necessarily set to zero, and therefore identical, already in the baseline model, delta CFI was 0.001, and delta RMSEA was 0.001. For level 2, delta CFI for establishing scalar measurement invariance was 0.010, and delta RMSEA was 0.011. These results indicated that scalar measurement invariance was confirmed at both levels.

6.3. Structural equation models

Descriptive statistics of the change scores were estimated separately based on unconditional change models without any directional effects. As a result of the latent aggregation of instructional quality at t1 and t2 as well as math achievement at t1 and t2, which implies centering around the group mean at level 1, all four scores show mean values of zero at level 1. Therefore, also the two estimated change scores for instructional quality and math achievement show mean values of zero. The respective variances were $s^2 = 0.324$ ($SE = 0.025$, $p = .000$) for instructional quality and $s^2 = 0.769$ ($SE = 0.045$, $p = .000$) for math achievement.

At level 2, again estimated in separate unconditional models, the latent change score for instructional quality showed a mean value that was negative but different from zero only by chance ($M = -0.147$, $SD = 0.152$, $p = .333$). Moreover, its variance of $s^2 = 0.023$ was very small and not significantly different from zero ($SE = 0.017$, $p = .171$). For mathematics achievement at level 2, the latent change score showed a mean value that was significantly different from zero ($M = 0.717$, $SD = 0.326$, $p = .014$) and a variance of $s^2 = 0.106$ ($SE = 0.033$, $p = .001$).

The estimated parameters of the structural equation model indicated a good model fit ($\chi^2 = 52.52$, $df = 23$, $p = .000$, $\chi^2/df = 2.28$, CFI = 0.972, TLI = 0.922; RMSEA = 0.041, $SRMR_{within} = 0.035$, $SRMR_{between} = 0.051$). The results revealed no significant direct relationship between the number of colleagues who perceived the collaboration with the respective teacher as useful for school improvement and the change in mathematics achievement ($\beta = 0.114$, $p = .294$, one-sided). However, significant positive relationships were evident between the extent of the perceived usefulness of teacher collaboration and the change in their instructional quality as perceived by their students ($\beta = 0.362$, $p = .001$, one-sided). A significant positive relationship was also found between instructional quality change and mathematics achievement change ($\beta = 0.723$, $p = .015$, one-sided). The total effect of the analyzed model was significant ($\beta = 0.376$, $p = .005$, one-sided), as was also the indirect effect between the perceived usefulness of teacher collaboration and change in student achievement, mediated by the change in instructional quality ($\beta = 0.262$, $p = .045$, one-sided). Although the correlation between the change in math achievement and math achievement at t1 at the between level was not significant ($\beta = -0.135$, p = .525), the respective correlation for instructional quality was positive and significant ($\beta = 0.799$, p = .009). This indicates that the instructional quality of teachers with higher values at t1 had a more positive development from t1 to t2.

Fig. 2 provides an overview of the results. For improved readability, the effects of the control variables are not depicted. At the within level (students), significant positive effects were observed between all control variables and mathematics achievement at the first measurement time point. Also, a significant positive effect was observed between socioeconomic context and instructional quality at the first measurement time point. At the between level (teachers), significant positive effects were observed between migration background and change in mathematics achievement as well as between socioeconomic context and change in instructional quality.

7. Discussion

Although theoretical models assume indirect effects between high-quality teacher collaboration and the development of student achievement, expecting instructional quality to act as a mediator, studies analyzing this mediation in a multilevel framework are lacking. For this reason, this study examined longitudinally the effects of the number of collaboration ties that teachers perceived as useful for school improvement on changes in instructional quality and mathematics achievement among fifth-graders. Significantly, we focused on the mediating role of the change in instructional quality.

Table 2
Within-level correlation (lower triangular) and between-level correlations (upper triangular).

Variable	1	2	3	4	5	6	7	8	9
1. Instructional quality t1	–	**.63**	−0.08	0.12	−0.47	0.29	−0.12	−0.20	−0.18
2. Instructional quality t2	**.42**	–	0.22	**.54**	−0.74	0.14	0.52	0.11	0.18
3. Math achievement t1	0.02	−0.01	–	**.76**	−0.04	0.21	0.43	**.50**	−0.17
4. Math achievement t2	0.00	0.05	**.73**	–	−0.22	0.49	0.34	**.37**	0.13
5. Student age	0.03	0.01	**−.22**	**−.24**	–	−0.17	−0.25	−0.41	−0.42
6. Gender	−0.08	**−.14**	**.08**	0.06	**.12**	–	−0.49	0.34	0.42
7. Socioeconomic context	0.02	0.03	0.08	**.09**	−0.05	−0.04	–	0.01	−0.01
8. Migration background	−0.03	0.03	**.18**	**.20**	**−.15**	0.01	**.13**	–	**.26**
9. Useful collaboration									–

Note. Statistically significant coefficients at α = 0.05 are printed in boldface. $n_{within} = 559$, $n_{between} = 76$.

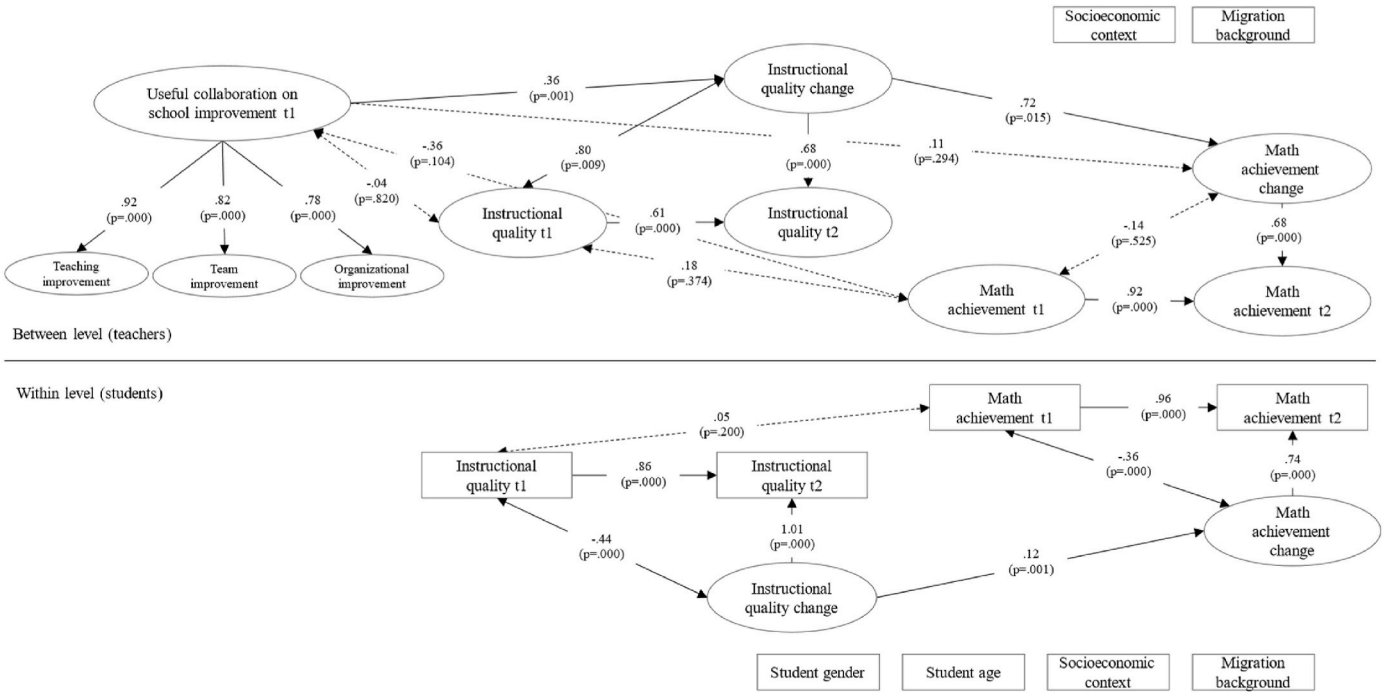


Fig. 2. Results of the Multilevel Structural Equation Model
Note. Continuous lines represent significant effects, dashed lines non-significant effects.

The findings support our hypotheses. In line with the reviewed theoretical framework (see section 2 above), we found a significant positive indirect relationship between perceived high-quality teacher collaboration and change in student achievement, mediated by change in instructional quality. Both the model's indirect and total effect were significant.

This study thus confirms empirical studies that found significant effects of teacher collaboration on instructional quality (Hochweber et al., 2012; Kılınç et al., 2023; Moolenaar et al., 2012; Placklé et al., 2022; Samaranayake et al., 2018) and no significant bivariate correlation between teacher collaboration and student achievement (Hochweber et al., 2012, 2018; Placklé et al., 2022). Further, our study is also in line with Placklé et al.'s (2022) finding of significant positive correlations between effective teaching practices and mathematics achievement.

We argued that students' achievement gains are the result of individual collaborative practices among teachers in a school network and their effects on instructional quality. Hence, in line with Ronfeldt et al. (2015) and Hargreaves and Fullan (2012), the interaction between school-wide collaborative structures and teachers' individual collaborative practices helps improve teachers' competence to provide high-quality instruction, which, in turn, promotes student learning. Accordingly, the empirical results support the theoretical assumptions (e.g., Creemers & Kyriakides, 2008; Spillane & Louis, 2002) that teacher collaboration can be interpreted as a distal factor in relation to student achievement, whereas teaching quality has a closer relationship with student achievement and can therefore be described as proximal.

For further research, it could be important to notice that educational effectiveness theories (e.g., Creemers & Kyriakides, 2008; Spillane & Louis, 2002) are even more complex; the tested model may include a 'black box' between high-quality teacher collaboration and its effect on the change in instructional quality. Studies on teachers' professional competencies (e.g., Blömeke et al., 2022) for instance support the assumption that teachers' motivations, knowledge, and skills are essential preconditions of instructional quality. Accordingly, it could be valuable to extend the mediation model by including these factors as mediators between high-quality teacher collaboration and the development of instructional quality.

Beyond analyzing the mediation model, the study results reveal one important additional aspect that extends previous research. As the analyses included teacher collaboration perceived as useful for school improvement, the results provide evidence of the importance of involvement in several collaborative relationships perceived as useful. In other words, the higher the social capital of teachers (Hargreaves & Fullan, 2012), quantified by the extent of teacher collaboration perceived as useful for school improvement, the stronger the gain in instructional quality and student achievement. This result extends previous findings emphasizing the significance of high-quality teacher collaboration (e.g., Decuyper et al., 2010; Gräsel et al., 2006) by demonstrating that it is crucial for fostering instructional quality and student achievement. Thus, from a practical point of view, it would be valuable to not only foster high-quality teacher collaboration within schools but to extend this individual perspective by considering the social network throughout the school to ensure that teachers are involved in several interactions with others that are perceived as useful.

Further research could investigate the question of whether a linear or curvilinear relationship exists between the extent of high-quality teacher collaboration and the development of both instructional quality and student achievement. Due to their workloads and the importance of perceived autonomy in their work (Vangrieken et al., 2017), teachers could benefit from involvement in an optimal number of high-quality teacher collaborations. If this theory holds true, it would be valuable to identify the optimal balance between individual autonomy and collaborative practice (Vangrieken et al., 2017).

8. Limitations

Several limitations must be considered.

1. Information from social network measures was used to analyze collaboration practice. Collaboration ties related to teaching and team and organization improvement were each measured using a single item only. Although this approach is common practice in social network research, as a multi-item approach to collecting

network data has a significant participant burden, it must be critically discussed in terms of validity (Vörös & Snijders, 2017).

2. We only measured high-quality teacher collaboration once, at the beginning of the school year, capturing teachers' collaborative practice from the previous six months rather than during the school year when changes in instructional quality and student achievement were analyzed. We argued that this approach allows us to identify teachers' interpersonal routines within their schools (Hatch & Hill, 2016; Spillane et al., 2016) as an important aspect of school improvement capacity for academic learning (Mitchell & Sackney, 2011). Both routines and school improvement capacity lend structure to teachers' discussion about instruction in the subsequent school year. Accordingly, our indicator is understood as a predictor of the change of instruction, measured prior to assessing instructional quality during the school year. While this design enables us to assess the potential time-lagged effects of collaboration on the development of teaching quality over the course of the year, it provides limited insight into the dynamic nature of teacher collaboration and its ongoing role in enhancing teaching quality. As a result, reciprocal effects between collaborative practices and improvements in teaching quality could not be explored. To address this limitation, future research could incorporate multiple time points for measuring teacher collaboration, allowing for the analysis of reciprocal effects between collaboration practice on school improvement and teaching improvement on the change of student achievement. Daily analysis of collaboration practice in a longitudinal study using experience-sampling methods (Maag Merki et al., 2022) has the potential to investigate the relationship between collaboration practice and the development of instructional quality and student achievement in a more differentiated way.
3. Other school-level variables, such as school policy or school leadership, could be important factors that affect student achievement indirectly (Heck & Hallinger, 2014; Kyriakides et al., 2015). This indirect effect is very likely mediated through teacher collaboration as well (Dorukbaşi & Cansoy, 2024; Woodland & Mazur, 2019). Therefore, an even more complex model might be promising that analyzes the relationship between school policy and school leadership, with an effect on teacher collaboration, instructional quality, and, ultimately, student achievement.
4. Instructional quality in mathematics was only measured using a global scale, without considering central sub-dimensions, as has been suggested by research on the basic dimensions of instructional quality (Praetorius et al., 2018). Differentiated recording of instructional quality could produce different results. Some aspects could be easier to change (e.g., classroom management), whereas others (e.g., cognitive activation) are more demanding.
5. Although scalar measurement invariance over time was tested and confirmed in a separate CFA, the instructional quality scale showed some anomalies. First, the variance of the change score was, in comparison, very small and was considered not significant based on a Wald test. Although it is acknowledged that the Wald test may not be entirely appropriate for testing variances, a likelihood ratio (LR) test could not be applied in this case. However, the manifest level 2 difference scores demonstrated a significant degree of variation ($p < .001$). Second, the correlation between the initial level of instructional quality and the change term, again at level two, was positive, significant, and unexpectedly high. This is in contrast to the level one correlation, which was negative and significant. Future research may focus on whether this is a substantial finding or an indication of estimation issues resulting from latent variance decomposition and the reliance on level two variance components for modeling changes at that level, given the sample size. In addition to this methodological attempt at explanation, it could also be an indication of the limitations in the operationalization of instructional quality. In both cases, caution should be taken when evaluating the results.
6. We only measured mathematics achievement at two time points. A study by Kyriakides et al. (2015) demonstrated that a longer-term assessment of students' competence development provides a more valid assessment of the combined effects of school factors. Therefore, further studies should investigate the mediating effect of instructional quality over a longer period.
7. We conducted the study in primary schools. It might be important to consider other school levels. At the secondary level for instance, teaching is organized more subject-specifically and teachers teach individual subjects, which very likely affects collaboration among teachers.
8. As a substantial percentage of the classes are multigrade, the mean number of fifth-grade students per class was only 13. In combination with the substantial within-class variation of perceived instructional quality, the reliability of the scale at the class level (ICC 2) was relatively low at t1 (ICC 2 = 0.59) and only acceptable at t2 (ICC 2 = 0.69; Lüdtke et al., 2006). Consequently, even the effects of substantial sizes were not statistically significant. Hence, in further studies, more students per class would be needed.

9. Conclusion

This study underpins and extends previous international research by demonstrating that the extent of perceived high-quality teacher collaboration is crucial for fostering instructional quality and student achievement. Most importantly for school improvement practice, the findings demonstrate that teacher collaboration should emphasize not only the development of teaching practice but also the development of team work and the school as an organization. All three areas of school improvement are relevant preconditions for increasing the quality of teaching and, ultimately, student learning. However, we encourage researchers conducting future studies to analyze the identified relationships in a larger sample and in other student age groups and to investigate additional mediating variables between high-quality teacher collaboration and student achievement.

CRediT authorship contribution statement

Andrea Wullschleger: Writing – original draft, Project administration, Investigation, Formal analysis, Conceptualization. **Katharina Maag Merki:** Writing – original draft, Project administration, Investigation, Funding acquisition, Conceptualization. **Urs Grob:** Writing – review & editing, Formal analysis. **Beat Rechsteiner:** Writing – review & editing, Investigation, Data curation. **Miriam Compagnoni:** Writing – review & editing, Data curation. **András Vörös:** Writing – review & editing, Data curation.

Data availability

Data will be made available on request.

References

Akiba, M., & Liang, G. (2016). Effects of teacher professional learning activities on student achievement growth. *Journal of Educational Research, 109*(1), 99–110. https://doi.org/10.1080/00220671.2014.924470

Antoniou, P., & Kyriakides, L. (2011). The impact of a dynamic approach to professional development on teacher instruction and student learning: Results from an experimental study. *School Effectiveness and School Improvement, 22*(3), 291–311. https://doi.org/10.1080/09243453.2011.577078

Birnbaum, A. (1968). Test scores, sufficient statistics, and the information structures of tests. In F. M. Lord, & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 425–435). Addison-Wesley.

Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction, 79*, Article 101600. https://doi.org/10.1016/j.learninstruc.2022.101600

Bos, W., Pietsch, M., List, M. K., Guill, K., Gröhlich, C., Scharenberg, K., et al. (2016). KESS 4: Skalenhandbuch zur Dokumentation der Erhebungsinstrumente. Retrieved from https://doi.org/10.25656/01:12711.

Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis, 30*(3), 203–235. https://doi.org/10.3102/0162373708321829

Cohen, J., Ruzek, E., & Sandilos, L. (2018). Does teaching quality cross subjects? Exploring consistency in elementary teacher practice across subjects. *AERA Open, 4*(3). https://doi.org/10.1177/2332858418794492, 2332858418794492.

Creemers, B. P. M., & Kyriakides, L. (2008). A theoretical based approach to educational improvement. Establishing links between educational effectiveness research and school improvement. In W. Bos, H. G. Holtappels, H. Pfeiffer, H.-G. Rolff, & R. Schulz-Zander (Eds.), *Jahrbuch der Schulentwicklung. Band 15. Daten, Beispiele und Perspektiven* (pp. 41–61). Juventa.

Daly, A. J. (Ed.). (2010). *Social network theory and educational change*. Harvard University Press.

Daly, A. J., Liou, Y.-H., & Der-Martirosian, C. (2020). A capital idea: Exploring the relationship between human and social capital and student achievement in schools. *Journal of Professional Capital and Community, 6*(1), 7–28. https://doi.org/10.1108/JPCC-10-2020-0082

Daly, A. J., Moolenaar, N. M., Der-Martirosian, C., & Liou, Y.-H. (2014). Accessing capital resources: Investigating the effects of teacher human and social capital on student achievement. *Teachers College Record: The Voice of Scholarship in Education, 116*(7), 1–42. https://doi.org/10.1177/016146811411600702

De Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.

Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning: An integrative model for effective team learning in organisations. *Educational Research Review, 5*, 111–133. https://doi.org/10.1016/j.edurev.2010.02.002

Dorukbaşi, E., & Cansoy, R. (2024). Examining the mediating role of teacher professional learning between perceived instructional leadership and teacher instructional practices. *European Journal of Education, 59*(3), Article e12672. https://doi.org/10.1111/ejed.12672

Fend, H. (1981). *Theorie der Schule*. Urban & Schwarzenberg.

Goddard, Y. L., & Goddard, R. D. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record, 109*(4), 877–896. https://doi.org/10.1177/016146810710900401

Gräsel, C., Fussangel, K., & Parchmann, I. (2006). Lerngemeinschaft in der Lehrerfortbildung. Kooperationserfahrungen und -überzeugungen von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*(4), 545–561. https://doi.org/10.25656/01:4453

Hallinger, P., & Heck, R. H. (2010). Collaborative leadership and school improvement: Understanding the impact on school capacity and student learning. *School Leadership & Management, 30*(2), 95–110. https://doi.org/10.1080/13632431003663214

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.

Harris, A. (2002). *School improvement. What's in it for schools?* Routledge.

Hatch, T., & Hill, K. (2016). Investigating the role of instructional rounds in the development of social networks and district-wide improvement. *American Educational Research Journal, 53*(4), 1022–1053.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Havnes, A. (2009). Talk, planning and decision-making in interdisciplinary teacher teams: A case study. *Teachers and Teaching: Theory and Practice, 15*(1), 155–176. https://doi.org/10.1080/13540600802661360

Heck, R. H., & Hallinger, P. (2014). Modeling the longitudinal effects of school leadership on teaching and learning. *Journal of Educational Administration, 52*(5), 663–681. https://doi.org/10.1108/JEA-08-2013-0097

Hochweber, J., Steinert, B., & Klieme, E. (2012). Lehrerkooperation, Unterrichtsqualität und Lernergebnisse im Fach Englisch. *40*(4), 351–370.

Holzberger, D., & Schiepe-Tiska, A. (2021). Is the school context associated with instructional quality? The effects of social composition, leadership, teacher collaboration, and school climate. *School Effectiveness and School Improvement, 32*(3), 465–485. https://doi.org/10.1080/09243453.2021.1913190

Hopkins, D. (2001). *School improvement for real*. Routledge.

Horn, I., Garner, B., Chen, I.-C., & Frank, K. A. (2020). Seeing colleagues as learning resources: The influence of mathematics teacher meetings on advice-seeking social networks. *AERA Open, 6*(2), Article 233285842091489. https://doi.org/10.1177/2332858420914898

Institut für Bildungsevaluation. (2020). *Check P5 2020: Ergebnisbericht für den Bildungsraum Nordwestschweiz*. Zürich, Switzerland: Institut für Bildungsevaluation.

Jaekel, A., & Göllner, R. (2020). *Unterricht aus Schülersicht (UNITAS) – Dokumentation der Erhebungsinstrumente*. Tübingen, Germany: Hector-Institut für Empirische Bildungsforschung.

James, C. R., Dunning, G., Connolly, M., & Elliott, T. (2007). Collaborative practice: A model of successful working in schools. *Journal of Educational Administration, 45*(5), 541–555. https://doi.org/10.1108/09578230710778187

Kiefer, T., Robitzsch, A., & Wu, M. (2017). *Tam: Test analysis modules* (Version 1.99999 – 31) [Software].

Kılınç, A.Ç., Polatcan, M., & Çepni, O. (2023). Exploring the association between distributed leadership and student achievement: The mediation role of teacher professional practices and teacher self-efficacy. *Journal of Curriculum Studies*, 1–17. https://doi.org/10.1080/00220272.2023.2216770

Kyriakides, L., Creemers, B. P. M., Antoniou, P., Demetriou, D., & Charalambous, C. Y. (2015). The impact of school policy and stakeholders' actions on student learning: A longitudinal study. *Learning and Instruction, 36*, 113–124. https://doi.org/10.1016/j.learninstruc.2015.01.004

Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teacher's professional relations. *Teachers College Record, 91*, 509–536. https://doi.org/10.1177/016146819009100403

Liu, S., & Yin, H. (2024). Opening the black box: How professional learning communities, collective teacher efficacy, and cognitive activation affect students' mathematics achievement in schools. *Teaching and Teacher Education, 139*, Article 104443. https://doi.org/10.1016/j.tate.2023.104443

Lomos, C., Hofman, R. H., & Bosker, R. J. (2011). The relationship between departments as professional communities and student achievement in secondary schools. *Teaching and Teacher Education, 27*(4), 722–731. https://doi.org/10.1016/j.tate.2010.12.003

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*(3), 203–229. https://doi.org/10.1037/a0012869

Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2006). Reliability and agreement of student ratings of the classroom environment: A reanalysis of TIMSS data. *Learning Environments Research, 9*(3), 215–230. https://doi.org/10.1007/s10984-006-9014-8

Maag Merki, K., Wullschleger, A., & Rechsteiner, B. (2022). Adapting routines in schools when facing challenging situations: Extending previous theories on routines by considering theories on self-regulated and collectively regulated learning. *Journal of Educational Change, 24*, 583–604. https://doi.org/10.1007/s10833-022-09459-1

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., et al. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*(6), 764–802. https://doi.org/10.1080/00273170903333665

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. Routledge.

Mitchell, C., & Sackney, L. (2011). *Profound improvement: Building learning-community capacity on living-system principles*. Routledge.

Moolenaar, N. M., Sleegers, P. J. C., & Daly, A. J. (2012). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education, 28*(2), 251–262. https://doi.org/10.1016/j.tate.2011.10.001

Muthén, L. K., & Muthén, B. (2017). *Mplus user's guide: Statistical analysis with latent variables*. Muthén & Muthén.

Nguyen, D., Pietsch, M., & Gümüs, S. (2021). Collective teacher innovativeness in 48 countries: Effects of teacher autonomy, collaborative culture and professional learning. *Teaching and Teacher Education, 106*, Article 103463. https://doi.org/10.1016/j.tate.2021.103463

Nilsen, T., Gustafsson, J.-E., & Blömeke, S. (2016). Conceptual framework and methodology of this report. In T. Nilsen, & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (pp. 1–19). Springer.

Pil, F. K., & Leana, C. (2009). Applying organizational research to public school reform: The effects of teacher human and social capital on student performance. *Academy of Management Journal, 52*(6), 1101–1124. https://doi.org/10.5465/amj.2009.47084647

Placklé, I., Könings, K. D., Jacquet, W., Libotton, A., Merriënboer, J. J. G., & Engels, N. (2022). Improving student achievement through professional cultures of teaching in Flanders. *European Journal of Education, 57*(2), 325–341. https://doi.org/10.1111/ejed.12504

Praetorius, A.-K., & Charalambous, C. Y. (Eds.). (2023). *Theorizing teaching: Current status and open issues*. Springer. https://doi.org/10.1007/978-3-031-25613-4.

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM - Mathematics Education, 50*, 407–426. https://doi.org/10.1007/s11858-018-0918-4

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal, 52*(3), 475–514. https://doi.org/10.3102/0002831215585562

Samaranayake, G., Premadasa, K., Amarasinghe, R., & Paneru, K. (2018). Teacher change through Lesson Study collaboration. *International Journal for Lesson and Learning Studies, 7*(4), 263–276. https://doi.org/10.1108/IJLLS-12-2017-0055.

Schippers, M. C., Den Hartog, D. N., & Koopman, P. L. (2007). Reflexivity in teams: A measure and correlates. *Applied Psychology: International Review, 56*(2), 189–201. https://doi.org/10.1111/j.1464-0597.2006.00250.x

Senden, B., Nilsen, T., & Blömeke, S. (2022). Instructional quality: A review of conceptualizations, measurement approaches, and research findings. In M. Blikstad-Balas, K. Klette, & M. Tengberg (Eds.), *Ways of analyzing teaching quality* (pp. 140–172). Scandinavian University Press. https://doi.org/10.18261/9788215045054-2021-05.

Sinnema, C., Liou, Y.-H., Daly, A., Cann, R., & Rodway, J. (2021). When seekers reap rewards and providers pay a price: The role of relationships and discussion in improving practice in a community of learning. *Teaching and Teacher Education, 107*, Article 103474. https://doi.org/10.1016/j.tate.2021.103474

Sleegers, P. J. C., Thoonen, E. E. J., Oort, F. J., & Peetsma, T. T. D. (2014). Changing classroom practices: The role of school-wide capacity for sustainable improvement. *Journal of Educational Administration, 52*(5), 617–652. https://doi.org/10.1108/JEA-11-2013-0126

Snijders, T. A. (2011). Statistical models for social networks. *Annual Review of Sociology, 37*, 131–153. https://doi.org/10.1146/annurev.soc.012809.102709

Spillane, J. P., & Louis, K. S. (2002). School improvement processes and practices: Professional learning for building instructional capacity. *The Yearbook of the National Society for the Study of Education, 101*(1), 83–104. https://doi.org/10.1111/j.1744-7984.2002.tb00005.x

Spillane, J. P., Shirrell, M., & Hopkins, M. (2016). Designing and deploying a professional learning community (PLC) organizational routine: Bureaucratic and collegial arrangements in tandem. *Les Dossiers Des Sciences De L'éducation, 35*, 97–122. https://doi.org/10.4000/dse.1283

Van Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. Springer.

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review, 15*, 17–40. https://doi.org/10.1016/j.edurev.2015.04.002

Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teacher and Teacher Education, 61*, 47–59. https://doi.org/10.1016/j.tate.2016.10.001

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education, 24*(1), 80–91. https://doi.org/10.1016/j.tate.2007.01.004

Vieluf, S., & Klieme, E. (2023). Teaching effectiveness revisited through the lens of practice theories. In A.-K. Praetorius, & C. Y. Charalambous (Eds.), *Theorizing teaching: Current status and open issues* (pp. 57–95). Springer. https://doi.org/10.1007/978-3-031-25613-4.

Vörös, A., & Snijders, T. A. B. (2017). Cluster analysis of multiplex networks: Defining composite network measures. *Social Networks, 49*, 93–112. https://doi.org/10.1016/j.socnet.2017.01.002

Weddle, H. (2022). Approaches to studying teacher collaboration for instructional improvement: A review of literature. *Educational Research Review, 35*, Article 100415. https://doi.org/10.1016/j.edurev.2021.100415

Woodland, R. H., & Mazur, R. (2019). Of teams and ties: Examining the relationship between formal and informal instructional support networks. *Educational Administration Quarterly, 55*(1), 42–72. https://doi.org/10.1177/0013161X18785868

Wullschleger, A., Vörös, A., Rechsteiner, B., Rickenbacher, A., & Maag Merki, K. (2023). Improving teaching, teamwork, and school organization: Collaboration networks in school teams. *Teaching and Teacher Education, 121*. https://doi.org/10.1016/j.tate.2022.103909

Yang, R., Zhang, Z., & Yu, S. (2018). Understanding teacher collaboration processes from a complexity theory perspective: A case study of a Chinese secondary school. *Teachers and Teaching, 24*(5), 520–537. https://doi.org/10.1080/13540602.2018.1447458

Andrea Wullschleger is a professor and co-leads the Center for Research on Teachers' Professional Development. Katharina Maag Merki is Professor of Educational Science with focus on Theory and Empirical Research in Educational Processes in Schools. Urs Grob is an academic associate; a main focus of his work is on quantitative data analysis. Beat Rechsteiner and Miriam Compagnoni are senior research assistants with a main focus on school improvement. András Vörös is Associate Professor Quantitative Methods.