SCHOOL EFFECTIVENESS and SCHOOL IMPROVEMENT
An International Journal of Research, Policy and Practice

Routledge
Taylor & Francis Group

School Effectiveness and School Improvement
An International Journal of Research, Policy and Practice

ISSN: 0924-3453 (Print) 1744-5124 (Online) Journal homepage: www.tandfonline.com/journals/nses20

From school improvement to instructional practice: the role of school- versus team-level capacity

Andrea Wullschleger, Urs Grob, Katharina Maag Merki, Ladina Zumsteg & Beat Rechsteiner

To cite this article: Andrea Wullschleger, Urs Grob, Katharina Maag Merki, Ladina Zumsteg & Beat Rechsteiner (13 Oct 2026): From school improvement to instructional practice: the role of school- versus team-level capacity, School Effectiveness and School Improvement, DOI: 10.1080/09243453.2026.2744285

To link to this article: https://doi.org/10.1080/09243453.2026.2744285

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 13 Oct 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=nses20

SCHOOL EFFECTIVENESS AND SCHOOL IMPROVEMENT
https://doi.org/10.1080/09243453.2026.2744285

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

From school improvement to instructional practice: the role of school- versus team-level capacity

Andrea Wullschleger [a], Urs Grob [b], Katharina Maag Merki [b], Ladina Zumsteg [b] and Beat Rechsteiner [b]

[a]Institute of Research on Professional and Teaching Development, University of Teacher Education Lucerne, Lucerne, Switzerland; [b]Institute of Education, University of Zurich, Zurich, Switzerland

ABSTRACT

This study examines whether school improvement capacity (SIC), traditionally conceptualized as a school-wide property, benefits from incorporating a team-level perspective. Previous research has advanced our understanding of school-wide conditions that enable improvement but largely overlooked that improvement processes often unfold within teams. Drawing on data from 662 teachers at 32 primary schools organized into 104 teams, we investigated: (a) whether schools and teams differ in the degree of within-unit agreement in perceptions of SIC dimension, and (b) how average levels and within-unit variability in perceived SIC relate to instructional practice at both levels. We compared ICCs across levels and conducted regression analyses separately for schools and teams. Although differences in ICCs were not significant, substantial variance was located at the team level. Further, associations with implementation of new ideas in instructional practice were stronger at the team level. The findings underscore the relevance of extending SIC to the team level.

ARTICLE HISTORY
Received 20 March 2026
Accepted 30 September 2026

KEYWORDS
School improvement capacity; schools vs. teacher teams; within-unit variation; instructional practice

Introduction

In many educational systems, schools operate in increasingly complex and dynamic environments. Challenges such as ongoing reforms, demographic change, or growing student heterogeneity require schools to continually adapt and improve their instructional practices (OECD, 2023). In this context, school improvement capacity (SIC) has been identified as a crucial prerequisite for sustainable school improvement, as it captures a school's capacity to collectively learn, respond to challenges, and implement new ideas in instructional practice (Maag Merki, 2017; Sleegers et al., 2014; Thoonen et al., 2012).

Previous studies have predominantly examined SIC as a school-wide property, while paying limited attention to potential within-school variation in teachers' perceptions of SIC. Yet schools are internally differentiated organizations, and collective learning,

CONTACT Andrea Wullschleger andrea.wullschleger@phlu.ch

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

coordination, and instructional improvement often occur within teams (Drach-Zahavy & Somech, 2002). In these subgroups, distinct norms, values, and practices may emerge, with variation even within the same school (Meredith et al., 2017). Hence, to better understand the unfolding of school improvement, it is essential to consider SIC at both the school and team levels.

Moreover, even when previous studies applied multilevel designs, they focused on average levels of SIC and its relationship to instructional practice, while paying less attention to the extent to which perceptions of SIC are shared within schools and teams. Differences in the extent to which teachers share perceptions of SIC may be consequential for collective action and shape how effectively new ideas are implemented in instructional practice (Feldhoff & Radisch, 2021). But they remain underexplored.

Against this background, the present study examines the relevance of both the school and the team for school improvement. It compares SIC and its relation to instructional practice at the school and team levels; it analyses not only average levels of SIC but also variability within schools and teams. Addressing these issues is crucial for designing targeted professional development and intervention strategies.

Conceptualizing school improvement capacity

SIC is conceptualized within the broader perspective of organizational learning, which views schools as organizations that continuously adapt and learn in response to internal and external demands (Argyris & Schön, 2018; Mitchell & Sackney, 2011). From this perspective, SIC refers to a school's capacity for organizational learning, reflected in the organizational conditions, interaction processes, and collective practices that support reflection, coordination, adaptation, and improvement (Beckmann et al., 2022; Mitchell & Sackney, 2011; Sleegers et al., 2014; Thoonen et al., 2012). More specifically, SIC refers to a school's ability to respond effectively to demands and challenges and to develop its practices in a systematic and purposeful manner (Maag Merki, 2017; Sleegers et al., 2014; Thoonen et al., 2012). It encompasses an interplay of individual abilities, interpersonal relations, collective practice, and organizational structures and processes (Mitchell & Sackney, 2011). Through these interacting elements, schools work towards improving teaching, learning, and their conditions (Maag Merki, 2017; Sleegers et al., 2013; Stoll, 2009). Because such organizational capacities become relevant for collective action primarily through how they are perceived and interpreted by organizational members, the present study focuses on teachers' perceptions of SIC.

In the literature, SIC is characterized by a range of different dimensions reflecting organizational conditions, interaction processes, and collective practices that support organizational learning (Beckmann et al., 2022; Mitchell & Sackney, 2011). Although researchers differ in how they label or group these dimensions, they consistently emphasize the importance of leadership, collaborative structures, professional norms, and teacher learning (Beckmann et al., 2022; Maag Merki, 2017; Sleegers et al., 2014; Thoonen et al., 2012). Leadership is relevant for building trustful relationships, fostering shared responsibility, and motivating staff to engage in school improvement (Leithwood & Jantzi, 2006; Rechsteiner et al., 2022). Collaborative structures support shared goals, collective meaning-making, and coordinated action (Lomos et al., 2011; Stoll, 2009). Professional norms enable constructive responses to challenges – such as an open and

non-punitive approach to errors (Mitchell & Sackney, 2011; Uline et al., 2003). Finally, opportunities for teacher learning that foster the exchange of professional knowledge are viewed as essential for strengthening instructional practice (Newmann et al., 2000; Staples & Webster, 2008).

Building on this view of SIC, the question arises as to how schools in fact navigate demands and challenges in their daily work. Recent work (Maag Merki et al., 2023) suggests that understanding the regulation of routines offers an important process-oriented extension to existing SIC conceptualizations. Whereas established routines provide stability in everyday practice, regulation processes describe how schools collectively monitor, adapt, and further develop these routines in response to challenges, thereby constituting an important process dimension of SIC.

In everyday school work, challenging situations are typically addressed through established routines that follow an "if–then" structure (Winne & Hadwin, 2008). When a specific demand or challenge is perceived (*if*), routines provide clarity about the goals to be achieved, the actors responsible, and the strategies usually employed to reach these goals (*then*). However, if the enacted routine does not sufficiently resolve the situation, regulation processes are required. Drawing on Winne and Hadwin's (2008) framework of self-regulated learning, which was later extended to socially shared regulation within collaborative settings (Hadwin et al., 2018), this shift can be described as a move from *then* to *else*, meaning that the routine (*then*) must be adapted (*else*). The adaptation relies on cognitive, metacognitive, and motivational regulation strategies. Cognitive strategies support information processing and the development of new approaches; metacognitive strategies guide monitoring and adjustment when existing approaches prove inadequate; motivational strategies help sustain effort and focus throughout the adaptation process (Panadero, 2017). From this perspective, regulation processes represent mechanisms through which schools collectively adapt routines and further develop SIC.

Taken together, these theoretical perspectives highlight that SIC is shaped not only by individual, cultural, and structural conditions but also by the processes through which staff perceive challenges, draw on routines, and regulate their actions.

Empirical evidence on school improvement capacity

Findings from previous research indicate that *schools differ systematically in their SIC*. In a recent contribution, Beckmann et al. (2022) operationalized SIC using six dimensions (school structure, commitment and collaboration, knowledge and skills, leadership, feedback and accountability, decision-making) and aggregated teacher ratings at the school level. Based on an average SIC index, schools were subsequently classified as low or high capacity. Comparisons between the two groups revealed significant differences across most SIC dimensions, with high-capacity schools scoring markedly higher. Beckmann et al. noted that conclusions must be drawn cautiously, however, as their analyses were based on a small sample with low survey response rates.

However, earlier studies support those findings. A strong line of research stems from studies conducted by Thoonen, Sleegers, and colleagues. The studies conceptualized SIC through a four-dimensional framework that encompasses leadership, organizational conditions, teachers' motivation, and opportunities for professional learning. In their cross-sectional analysis, Thoonen et al. (2011) found that the SIC dimensions vary

meaningfully across schools. Building on this, in a longitudinal study, Thoonen et al. (2012) demonstrated that schools differ not only in initial levels of SIC but also in how these levels develop over a 6-year period. Some schools strengthened their organizational and motivational conditions for improvement, whereas others showed stable or even declining trajectories.

A related series of studies by Hallinger and Heck (2010a, 2010b, 2011) further substantiates these findings. Drawing on large longitudinal samples of elementary schools, Hallinger and Heck conceptualized SIC as a set of organizational conditions that support teaching and learning. The conditions had four core dimensions: standards emphasis and implementation, focused and sustained action on improvement, quality of student support, and the school's professional capacity. Hallinger and Heck's analyses found that leadership influences student learning primarily through strengthening school-wide capacity, and that both SIC and leadership develop over time in ways that differ between schools. Their growth modeling revealed that schools follow distinct developmental trajectories, with some schools enhancing their SIC and others showing stagnation or decline.

Research on *the effects of SIC* indicates that it is meaningfully associated with both instructional practice and student learning outcomes. Cross-sectional evidence shows that schools with higher levels of SIC demonstrate more effective instructional practices (Thoonen et al., 2011). Longitudinal studies further support this link: schools showing increases in SIC over time are associated with improvements in teachers' classroom practices (Thoonen et al., 2012). Similarly, Sleegers et al. (2014) demonstrated that teacher motivation and professional learning play a crucial role in fostering changes in instructional practice across schools. Beyond effects on instruction, several studies by Hallinger and Heck (2010a, 2010b, 2011) found that SIC acts as a key organizational mechanism through which leadership influences student learning. Hallinger and Heck's longitudinal analyses indicated that schools with higher capacity for improvement exhibit more favorable trajectories in student achievement.

Taken together, these studies provide consistent evidence that SIC varies across schools and is an important predictor of instructional practice and student learning. Two important gaps remain, however: First, although several studies have employed multilevel models, SIC has typically been conceptualized as a school-wide property, with limited attention to the role of teams as potential sites of improvement processes. Second, previous studies have rarely reported or examined the extent of within-school homogeneity/heterogeneity in teachers' perceptions of SIC. The following two sections below address these gaps.

From an organizational to a team perspective on school improvement capacity

Increasingly, schools rely on team-based organizational structures to organize core work processes (Drach-Zahavy & Somech, 2002). This suggests that collective learning and adaptation are often situated within smaller work units. Team learning research conceptualizes teams as central sites in which collective reflection, knowledge sharing, coordination, and adaptation occur through interaction processes (Edmondson, 2002; Van Den Bossche et al., 2022). Teacher teams have been shown to play a crucial role in school improvement because meaningful organizational improvement depends on whether teams change the way they work (Bouwmans et al., 2017; Edmondson, 2002).

Accordingly, successful school improvement rests to a considerable extent on the effectiveness of teams as the basic building blocks of a school's organizational system (Park et al., 2005). Senge (1990) had suggested that teams constitute the fundamental learning unit of an organization. Edmondson (2002) built on this idea and argued that organizational learning is inherently "local, interpersonal, and variegated" (p. 128). Similarly, team learning has been described both as an interactional process and as the development of shared cognitions emerging through collaboration (Van Den Bossche et al., 2022). Edmondson (2002) made the perspective clear as follows:

> An organization 'learns' through actions and interactions that take place between people who are typically situated within smaller groups or teams. Through these subunits making appropriate changes in how they do their work—driven by both team-specific and organizational objectives—an organization maintains its effectiveness in a changing world. (p. 128)

Building on this perspective, school improvement can be understood as emerging within work groups, shaped by their social climate, and differing across organizational subunits. Because school members interact most frequently within these subgroups, distinct norms, values, and practices can develop within them – creating subcultures that vary significantly even within the same school (Meredith et al., 2017). At the same time, these interaction processes may foster shared understandings and collective perceptions of improvement-related capacities within teams. Given that schools have been described as loosely coupled systems (Weick, 1976), variation in SIC is likely to emerge within schools. Teams may therefore represent meaningful meso-level units for capturing that intra-school variation.

Shared versus divergent perceptions of school improvement capacity

Connected to the predominant school-level perspective, SIC has typically been conceptualized as a school-wide property. This may mask considerable variation within schools (Feldhoff & Radisch, 2021). It therefore remains an open question as to what extent teachers' perceptions of SIC are shared or differ within and between schools and teams, and whether these patterns of shared or divergent perceptions are related to instructional practice.

In organizational research, it is generally assumed that teams that are heterogeneous in expertise, knowledge, and information benefit from integrating diverse perspectives, thereby enhancing learning, performance, and innovation (Lewis, 2003; Rupert et al., 2016). In the school context, Drach-Zahavy and Somech (2002) distinguished between task-related heterogeneity (specific skills and abilities) and relations-oriented heterogeneity (personal characteristics). According to Drach-Zahavy and Somech, task-related heterogeneity may enhance team performance by expanding task-relevant resources, whereas relations-oriented heterogeneity may hinder performance by activating stereotypical perceptions. However, their findings revealed a less clear-cut pattern: Whereas significant relationships emerged between heterogeneity in functional roles and tenure, and team effectiveness, the proposed distinction between task-related and relations-oriented heterogeneity was not supported.

The present study, however, focuses on a different form of heterogeneity: variation in teachers' perceptions of SIC. Rather than reflecting differences in expertise or demographic composition, perceptual heterogeneity captures the extent to which

organizational members share – or differ in – their interpretations and evaluations of SIC. Perceptual heterogeneity may have ambivalent implications. On the one hand, divergent perceptions may stimulate reflection and adaptive learning by introducing multiple perspectives. On the other hand, low levels of shared understanding may hinder coordination and collective action, particularly when differences reflect conflicting goals, norms, or interpretations.

Examining shared versus divergent perceptions of SIC thus extends existing approaches by highlighting the relevance of within-school and within-team coherence for SIC and instructional practice. Given the central role attributed to teams in this study, the following section below clarifies its meaning in the school context.

What is a team in the school context?

Research on teams in organizations typically relies on definitions that specify strict criteria. Cohen and Bailey (1997), for instance, described teams as "a collection of individuals who are interdependent in their task, who share responsibility for outcomes, and who see themselves and are seen by others as an intact social entity embedded in a larger social system" (p. 241). However, as Vangrieken et al. (2017) noted, teacher teams do not always meet these criteria in a strict sense. Despite their growing importance in educational practice, teacher teams often fall somewhere between fully developed teams and more loosely organized collaborative groups.

To address this, recent work has argued for a more flexible understanding of what constitutes a team in the school context. Wageman et al. (2012) emphasized the value of extending team concepts to diverse forms of collaboration, proposing that the defining features of teams should be treated as dimensions that vary rather than as rigid boundaries. Building on this idea, Vangrieken et al. (2017) introduced the notion of team entitativity, which captures the extent to which a group exhibits "teamness" – that is, the degree to which it functions as a team.

In the present study, the focus is on formal teams. In schools, formal teams refer to organizationally defined subgroups – such as grade-level teams – whose membership and responsibilities are established by the school's formal structure. These teams are expected to coordinate their work within these assigned boundaries (Meredith et al., 2017). International evidence indicates that collaborative teaching arrangements are widespread across European countries, with classes often taught by more than one teacher, which requires regular coordination and joint responsibility (OECD, 2020). At the same time, formally established teams may differ substantially in the extent to which they function as cohesive teams and exhibit teamness.

Research questions and hypotheses

This study addresses the following research questions:

1. Do schools and teams differ in the degree of within-unit agreement in teachers' perceptions of SIC dimensions?

We expect teams to show higher within-unit agreement in perceptions of SIC dimensions than schools do, given their more similar work contexts and shared experiences.

2. To what extent are: (a) average levels, and (b) within-unit variability in teachers' perceptions of SIC dimensions associated with the implementation of new ideas in instructional practice at the school and team levels?

We hypothesize that these associations are more pronounced at the team level than at the school level, reflecting the teams' closer proximity to everyday instructional practice. Given the limited available evidence, we examined these associations exploratively.

Materials and methods

Research design and sample

Data for this study stem from a research project examining SIC in 59 public primary schools in the German-speaking part of Switzerland. Primary schooling in this context comprises 8 years, including 2 years of kindergarten. In Swiss primary schools, teachers teach multiple subjects rather than specializing in subject areas. Accordingly, the teams examined in this study consisted primarily of formally established grade-level or cycle-level teams.

Participation in the study was voluntary, and all respondents provided written informed consent prior to taking part. Compared to the national distribution, urban schools and larger schools were somewhat overrepresented, but the socioeconomic composition closely matched national patterns. Data used in this study was collected by an online questionnaire. The Ethical Commission of the University of Zurich approved the questionnaire and study procedures (approval no. 19.6.11).

A total of 1,627 teachers, specialist teachers, and principals were invited to participate in the study. After excluding cases with missing or incomplete data, 1,328 respondents remained in the analytic sample, corresponding to a response rate of 81.6%. As in Switzerland all members of a school team are responsible for teaching, we used the term "teacher" for all participants. Respondents ranged in age from 21 to 67 years (M = 42.81, SD = 11.51), reported an average of 17 years of professional experience (SD = 11.01, range: 0–45 years), and 87% were women.

Measures

Core dimensions of SIC

Building on previous research, this study operationalized core dimensions of SIC through four scales: leadership as transformational leadership, task cohesion as an aspect of collaborative structures, error culture as an aspect of professional norms, and knowledge sharing as an aspect of teacher learning. The teachers rated all items in terms of their school. All items on the four scales were rated on a 6-point Likert scale (1 = *strongly disagree*, 6 = *strongly agree*).

Leadership was conceptualized as a set of behaviors aimed at strengthening teachers' professional commitment and capacities. The construct was assessed using a reduced four-item version of the *intellectual stimulation subscale* developed by Sleegers et al. (2014). A sample item reads: "The school leader of our school encourages teachers to try new things in line with their own interests." A Cronbach's alpha value of .87 revealed a high reliability of the measurement instrument (N = 1,239; M = 4.73; SD = 0.83).

Task cohesion referred to the extent to which team members perceive a shared focus on and coordinated effort toward common developmental goals. The construct was assessed with a three-item scale adapted from the *group integration–task scale* from Carron et al. (1985). A sample item reads: "At our school, we work together to achieve our goals." A Cronbach's alpha value of .79 revealed acceptable reliability of the measurement instrument ($N = 1{,}307$; $M = 4.42$; $SD = 0.78$).

Error culture encompassed norms and practices for openly addressing mistakes and leveraging them as opportunities for collective learning. The construct was measured using a self-developed, previously piloted four-item scale. A sample item reads: "In our school, mistakes help us improve what we do." A Cronbach's alpha of .88 revealed a high reliability of the measurement instrument ($N = 1{,}305$; $M = 4.43$; $SD = 0.79$).

Knowledge sharing referred to the transfer and exchange of information, opinions, ideas, experiences, and materials to build a shared knowledge base (Staples & Webster, 2008; Widmann & Mulder, 2018). The construct was assessed with a four-item scale adapted from Staples and Webster (2008). A sample item reads: "People in this school are willing to share knowledge/ideas with others." A Cronbach's alpha value of .80 revealed a high reliability of the measurement instrument ($N = 1{,}312$; $M = 4.76$; $SD = 0.73$).

To address the second research question, we worked with a second-order construct of these core dimensions. Accordingly, we examined their factorial structure by conducting confirmatory factor analyses at the teacher level. A one-factor model (without taking into account the structure of the initial four core dimensions) showed poor fit to the data ($\chi^2(90) = 1430.26$, $p < .001$, CFI = .72, TLI = 0.68, RMSEA = .16, SRMR = .09). In contrast, a second-order model in which a higher-order latent construct captured the shared variance among the four scales showed good model fit ($\chi^2(86) = 293.01$, $p < .001$, CFI = .96, TLI == 0.95, RMSEA = .06, SRMR = .04). In addition, all first-order factors loaded strongly on the second-order factor, with standardized loadings ranging from $\lambda = .62$ to $\lambda = .93$. Reliability analyses indicated good to excellent internal consistency for all dimensions (Leadership: $\alpha = .86$, $\omega = .87$; Task cohesion: $\alpha = .80$, $\omega = .80$; Error culture: $\alpha = .88$, $\omega = .89$; Knowledge sharing: $\alpha = .81$, $\omega = .81$), with average variance extracted (AVE) ranging from .52 to .67. These findings justified the use of a second-order construct representing the core dimensions of SIC.

Additional dimensions of SIC

As argued in the section on conceptualizing SIC, routines and regulation processes were conceptualized as additional dimensions of SIC. The dimensions were assessed using four scales reflecting routines and different regulation strategies. The items were administered in parallel versions for the school and the team. All items were rated on a 6-point Likert scale (1 = *strongly disagree*, 6 = *strongly agree*).

Routines help staff to navigate the demands of daily school life effectively; they provide structure and support efficient task completion (Maag Merki et al., 2023). The construct was measured using a self-developed, previously piloted four-item scale. A sample item reads: "In our school (In our team), we have established strategies that help us effectively address challenges." For the school scale, Cronbach's alpha was .83, indicating high internal consistency ($N = 1{,}299$; $M = 4.68$; $SD = 0.62$), and a similarly high reliability was observed for the team scale ($\alpha = .84$; $N = 1{,}306$; $M = 4.78$; $SD = 0.61$).

Cognitive regulation referred to strategies that support information processing and the development of new approaches (Panadero, 2017). The focus was on how teachers engage with different forms of feedback and professional learning activities to advance instructional practice. The construct was measured using a self-developed, previously piloted five-item scale. A sample item reads: "In our school (In our team), we draw connections between new information and our educational practice." For the school scale, Cronbach's alpha was .86, revealing high reliability ($N = 1{,}304$; $M = 4.51$; $SD = 0.67$), and a similarly high reliability was observed for the team scale ($\alpha = .85$; $N = 1{,}308$; $M = 4.64$; $SD = 0.63$).

Metacognitive regulation referred to strategies that guide monitoring and adjustment when existing approaches prove inadequate (Panadero, 2017). Metacognitive regulation was measured using a self-developed, previously piloted six-item scale. A sample item reads: "In our school (In our team), we often think about what works in our own practice and what needs to be changed." For the school scale, Cronbach's alpha was .88, revealing high reliability ($N = 1{,}296$; $M = 4.70$; $SD = 0.66$), and a similarly high reliability was observed for the team scale ($\alpha = .88$; $N = 1{,}302$; $M = 4.89$; $SD = 0.62$).

Emotional and motivational regulation referred to strategies that help sustain effort and focus throughout the adaptation process (Panadero, 2017). It was assessed using a six-item scale adapted from Schwinger et al. (2007). A sample item reads: "In our school (In our team), we have suitable strategies for regulating negative emotions at work so that we can continue working effectively." For the school scale, Cronbach's alpha was .87, revealing high reliability ($N = 1{,}292$; $M = 4.61$; $SD = 0.64$), and a similarly high reliability was observed for the team scale ($\alpha = .87$; $N = 1{,}296$; $M = 4.70$; $SD = 0.62$).

To address the second research question, we specified a second-order construct for the three regulation strategies. Confirmatory factor analyses were conducted separately for the school and the team. A one-factor model (without taking into account the structure of the initial three regulation strategies) showed poor fit to the data (school: $\chi^2(119) = 1{,}388.63$, $p < .001$, CFI = .76, TLI = 0.72, RMSEA = .13, SRMR = .08; team: $\chi^2(119) = 1{,}419.64$, $p < .001$, CFI = .75, TLI = 0.72, RMSEA = .14, SRMR = .08). In contrast, a second-order model specifying regulation as a higher-order construct provided good fit (school: $\chi^2(116) = 492.44$, $p < .001$, CFI = .93, TLI = 0.92, RMSEA = .07, SRMR = .04; team: $\chi^2(116) = 424.33$, $p < .001$, CFI = .94, TLI = 0.93, RMSEA = .07, SRMR = .04). All first-order factors loaded strongly on the second-order regulation factor. Standardized loadings ranged from $\lambda = .74$ to $\lambda = .88$ for the school and from $\lambda = .75$ to $\lambda = .91$ for the team. Reliability analyses indicated high internal consistency for all three dimensions (school: $\alpha = .84$–$.88$, $\omega = .84$–$.88$; team: $\alpha = .84$–$.88$, $\omega = .84$–$.88$). Average variance extracted (AVE) values ranged from .51 to .56 for the school and from .52 to .56 for the team, indicating adequate convergent validity. Together, these findings supported the use of a second-order regulation construct in subsequent analyses.

Implementation

Building on the evidence that SIC is closely linked to instructional practice, we conceptualized the implementation of new ideas in instructional practice as a proximal behavioral indicator of instructional practice. The measure reflected the extent to which schools and teams translate new knowledge, feedback, and improvement impulses into concrete adjustments to instruction.

The construct was measured using a self-developed five-item scale for the school and the team. A sample item reads: "In our school (In our team), when we try out something new, we implement the insights gained in our practice." For the school scale, Cronbach's alpha was .78, indicating acceptable internal consistency ($N = 1296$; $M = 4.38$; $SD = 0.63$), and a similarly acceptable internal consistency was observed for the team scale ($\alpha = .77$; $N = 1305$; $M = 4.55$; $SD = 0.61$).

Control variables

As *control variables*, we included school and team size, defined as the number of teachers per unit. The average school size was 20.7 teachers ($SD = 12.1$), with school sizes ranging from 8 to 49 teachers. The average team size was 6.37 teachers ($SD = 2.92$), with team sizes ranging from 3 to 16 teachers. Because the variability of within-unit standard deviations depended on the number of respondents, we additionally included the inverse of unit size as robustness controls.

Analytical strategy

Preliminary analyses

Before addressing the research questions, we conducted preliminary analyses to determine the analytical unit of teams within schools. The online questionnaire asked participants to indicate the formal team most relevant to their work and to identify the individuals belonging to that team.

Team-level analyses were restricted to teachers who could be assigned to clearly identifiable formal teams. To determine analytically distinct teams, a two-step procedure was applied. First, teams were identified using the Louvain community detection algorithm (Blondel et al., 2008). This algorithm detects clusters by maximizing modularity within the reported network of team memberships, thereby providing an initial structure of densely connected groups. In a second step, because this procedure may assign individuals to overlapping clusters, further criteria for team membership were applied. Teams were required to exhibit a high proportion of reciprocal nominations and to comprise at least three individuals. Based on the structure derived from the Louvain algorithm and these criteria, final team assignments were established through independent review by two researchers and subsequently reconciled through discussion to ensure consensus.

This procedure resulted in a refined sample of 662 teachers at 32 primary schools, organized into 104 teams. To assess potential selection bias, teachers included and excluded in the final sample were compared on all focal study variables (Appendix). Although some differences emerged for core SIC dimensions, effect sizes were small. Moreover, no significant differences in variability were observed between the two groups across any of the focal variables. Overall, the findings suggest limited risk of substantial selection bias.

These teams were assigned to different formal team types. As in Switzerland the municipalities are responsible for school organization, structural arrangements vary (Figure 1). In some municipalities, teams are organized by grade level; in others, schools are divided into lower elementary (Grades 1–3) and upper elementary (Grades 4–6). Further, some municipalities structure teams by cycle, with Cycle 1 covering kindergarten through Grade 2 and Cycle 2 covering Grades 3 through 6. Special needs teachers typically form a separate team, providing targeted support to individual students or small groups and advising classroom teachers.

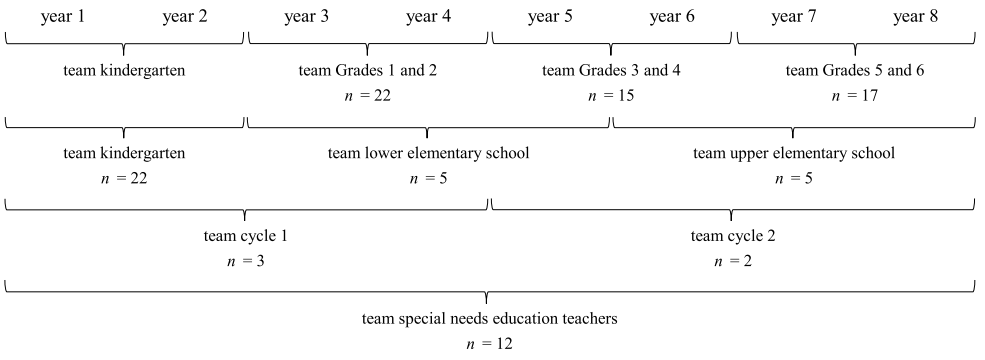


Figure 1. Overview of teacher allocation to formal teams types ($n = 104$).

Analyses RQ 1

To examine whether teams and schools differ in their degree of within-unit agreement on the SIC dimensions, intraclass correlations coefficients (ICCs) were estimated separately for each dimension at the team and school levels. Three-level Bayesian variance-components models were specified to decompose variance into within-team, between-team, and between-school components. Because ICCs are bounded at zero, inference regarding whether variance components differed from zero was based on 90% credible intervals, corresponding to a directional hypothesis. In contrast, differences between team- and school-level ICCs could occur in either direction and were therefore evaluated using 95% credible intervals.

Models were estimated in Mplus using Bayesian MCMC procedures. Bayesian estimation was chosen to ensure non-negative variance component estimates and to avoid inadmissible negative variances commonly observed under maximum likelihood estimation. Four chains with 20,000 iterations each were run. Convergence was achieved rapidly for all models, as indicated by the potential scale reduction factor (PSR) and inspection of trace plots. All effective sample sizes (ESS) exceeded 400, indicating adequate sampling efficiency.

Differences between team- and school-level ICCs were defined within Mplus's model constraint framework (ICCteam minus ICCschool). Inference was based on the posterior distribution of this difference; effects were considered credible when the 95% credible interval excluded zero. Credible intervals (CrIs) indicate the range within which the parameter lies with a given posterior probability, conditional on the model and the observed data.

Analysis RQ 2

To examine the extent to which the average level and the degree of within-unit agreement in SIC were associated with the implementation of new ideas in instructional practice, a set of regression analyses was conducted separately for the team and school levels. Specifically, we examined whether schools and teams with higher average levels of perceived SIC, and with more or less shared perceptions of SIC within units, differed in their reported implementation of new ideas in instructional practice.

To operationalize SIC in the regression models, aggregated indicators of average levels and within-unit variability were constructed. We relied on meta scales for the core SIC dimensions and for regulation. The use of meta scales was justified by the strong intercorrelations among the respective subdimensions. For routines, which were conceptualized as a distinct construct, the scale was retained. For each school and team, we computed: (a) mean levels of the core SIC meta scale, the regulation meta scale, and the routine scale, and (b) within-unit variability, operationalized as within-team and within-school standard deviations (*SDs*). Higher *SD* values indicate greater divergence in teachers' perceptions within a school or team, whereas lower *SD* values indicate a higher degree of shared perceptions. For the core SIC dimensions, which were rated in terms of the school, the same teacher ratings were aggregated to both the school and team levels. In contrast, routines and regulation were rated in terms of the school and the team and aggregated accordingly. All predictors were z-standardized prior to analysis.

Because our hypothesis stated that SIC would be associated with implementation but did not specify a direction, the models were specified exploratively and included terms beyond main effects. Interaction terms between mean levels and degree of within-unit agreement ($M \times SD$) were included to test whether the association depended on the construct's average level. Quadratic terms (SD^2) were added to capture potential curvilinear relationships. Because the distributional assumptions for *SD* and SD^2 terms were not fully met and heteroskedasticity could not be ruled out, all regression models were estimated with heteroskedasticity-robust standard errors (HC3 sandwich estimator).

Three regression models predicting implementation of new ideas in instructional practice were estimated for the school and for the team:

- Model 1 (level-only): Included mean levels of the SIC meta scale, the regulation meta scale, and routines, as well as the control variables school/team size and its inverse (1/n).
- Model 2 (level & heterogeneity): Added heterogeneity indicators (*SDs*) and $M \times SD$ interaction terms.
- Model 3 (level, heterogeneity & squared heterogeneity): Added quadratic terms (SD^2) for each construct to examine potential curvilinear associations between heterogeneity and implementation.

Results

Differences in homogeneity of SIC for school and team (RQ1)

Table 1 and Figure 2 present ICCs for each SIC dimension. Across most SIC dimensions, ICC point estimates were higher at the team level than at the school level; this suggested a greater degree of within-team agreement than within-school agreement. Even for SIC dimensions referring to the school, teachers' perceptions were more homogeneous within teams than within schools. However, comparisons of the 90% credible intervals derived from the three-level Bayesian variance-components models showed substantial overlap between team- and school-level ICCs for all dimensions. Accordingly, no statistically distinguishable differences in homogeneity between the two levels could be established.

Table 1. ICCs for SIC dimensions at the team and school levels.

	School			Team			Difference		
Variable	ICC	CrI[1] low	CrI[1] high	ICC	CrI[1] low	CrI[1] high	Δ ICC	CrI[2] low	CrI[2] high
Leadership (sl)	.081	.025	.169	.109	.052	.184	.026	−.116	.163
Task cohesion (sl)	.104	.039	.201	.085	.031	.161	−.022	−.171	.129
Error culture (sl)	.088	.028	.180	.081	.029	.154	−.009	−.151	.130
Knowledge sharing (sl)	.059	.013	.138	.089	.033	.163	.028	−.103	.153
Routine (tl)	.030	.005	.080	.054	.012	.121	.022	−.060	.117
Routine (sl)	.036	.006	.100	.088	.032	.162	.049	−.058	.158
Cognitive regulation (tl)	.025	.005	.074	.053	.012	.116	.025	−.055	.114
Cognitive regulation (sl)	.022	.004	.069	.083	.033	.149	.058	−.028	.148
Metacognitive regulation (tl)	.034	.006	.091	.039	.007	.098	.004	−.081	.091
Metacognitive regulation (sl)	.036	.006	.101	.075	.026	.142	.036	−.069	.136
Motivational regulation (tl)	.013	.003	.046	.055	.013	.119	.039	−.024	.122
Motivational regulation (sl)	.014	.003	.048	.056	.014	.118	.039	−.023	.119

Notes: Bayesian estimation of a three-level model with teachers at level 1, teams at level 2, and schools at level 3. sl = school level, tl = team level. CrI[1] credible interval of 90% (one-tailed test against 0 with $p < .05$). CrI[2] credible interval of 95% (two-tailed test against 0 with $p < .05$). Δ ICC = ICC team − ICC school.

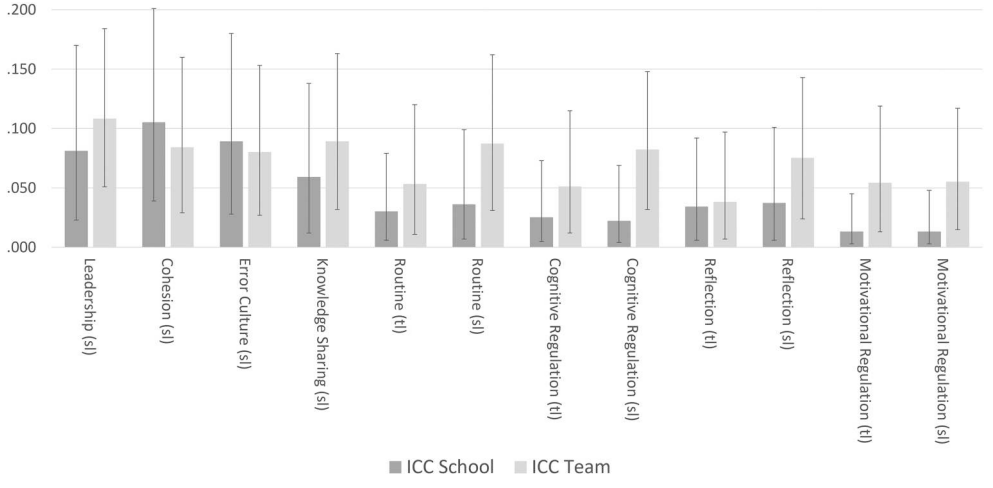


Figure 2. ICCs for SIC dimensions at the team and school levels with 90% credible intervals.

Associations between average levels and degree of within-unit agreement of SIC dimensions and implementation (RQ2)

To examine the extent to which average levels and degree of within-unit agreement of SIC dimensions within schools and teams are associated with the implementation of new ideas in instructional practice, a sequence of three regression models was estimated separately for schools and teams.

At the *school level* (Table 2), Model 1 explained 70% of the variance in implementation ($R^2_{adj} = .64$, $F(5, 26) = 11.91$, $p < .001$, $AIC = -26.64$). Among the predictors, only the average level of regulation was positively associated with implementation, indicating that schools reporting higher average levels of regulation also reported higher implementation.

Table 2. Regression models predicting implementation, school level.

Effect	Model 1			Model 2			Model 3		
	Estimate	*SE*	*p*	Estimate	*SE*	*p*	Estimate	*SE*	*p*
Intercept	4.38	0.03	< .001	4.40	0.03	< .001	4.40	0.08	< .001
M core SIC	.05	0.06	.339	.04	0.07	.551	.04	0.08	.594
SD core SIC				−.01	0.07	.902	−.02	0.07	.767
M x *SD* core SIC				.02	0.09	.856	−.01	0.08	.917
SD^2 core SIC							−.03	0.07	.718
M Routine	.04	0.04	.283	.04	0.05	.392	.03	0.07	.630
SD Routine				.01	0.07	.871	.02	0.09	.831
M x *SD* Routine				−.00	0.03	.980	−.02	0.08	.767
SD^2 Routine							−.05	0.08	.507
M Regulation	**.13**	**0.04**	**.002**	**.14**	**0.06**	**.032**	**.13**	**0.06**	**.067**
SD Regulation				.06	0.10	.535	.05	0.10	.580
M x *SD* Regulation				.04	0.09	.709	.03	0.10	.783
SD^2 Regulation							.05	0.09	.578
School Size	.02	0.05	.634	.05	0.07	.470	.04	0.07	.516
School Size 1/*n*	.05	0.06	.420	.11	0.09	.267	.08	0.11	.446
R^2	.70			.78			.82		
ΔR^2				.08			.04		

Notes: Values are standardized regression coefficients (β), standard errors (*SE*), and *p*-values for predicting implementation at the school level ($N = 32$).

Model 2 showed descriptively higher explained variance overall ($R^2 = .78$, $R^2_{adj} = .66$, $F(11, 20) = 6.49$, $p < .001$, $AIC = -25.15$). However, a robust Wald test indicated that the increase in explained variance compared to Model 1 was not significant, $F(6, 20) = 0.79$, $p = .589$. The mean level of regulation remained the only significant predictor.

In Model 3 the explained variance showed a small descriptive increase ($R^2 = .82$, $R^2_{adj} = .67$, $F(14, 17) = 5.54$, $p = .001$, $AIC = -25.45$). However, the additional quadratic terms did not lead to a significant improvement over Model 2 according to the robust Wald test, $F(3, 17) = 0.17$, $p = .917$. The average level of regulation reached only marginal significance.

At the *team level* (Table 3), Model 1 explained 61% of the variance in implementation ($R^2_{adj} = .59$, $F(5, 98) = 30.66$, $p < .001$, $AIC = -34.91$). Among the predictors, the average level of core SIC and the average level of regulation were positively associated with implementation. This suggested that teams reporting higher average levels of core SIC and regulation also report higher implementation.

Model 2 showed descriptively higher explained variance overall ($R^2 = .66$, $R^2_{adj} = .62$, $F(11, 91) = 16.08$, $p < .001$, $AIC = -35.97$). However, a robust Wald test indicated that the increase in explained variance compared to Model 1 was not significant, $F(6, 91) = 0.95$, $p = .465$. The positive associations of regulation and core SIC average levels with implementation remained significant. In addition, regulation heterogeneity emerged as a significant positive predictor, indicating that teams with greater variability in perceived regulation reported higher implementation.

Model 3 showed a small descriptive increase in explained variance relative to the previous models ($R^2 = .68$, $R^2_{adj} = .63$, $F(14, 88) = 13.16$, $p < .001$, $AIC = -35.09$). However, the additional quadratic terms did not significantly improve model fit over Model 2, $F(3, 88) = 1.14$, $p = .336$. Consistent with the previous models, the average levels of core SIC and regulation were positively associated with implementation. In addition, routine heterogeneity was negatively associated with implementation, indicating that teams with more homogeneous perceptions of routines reported higher levels of implementation. The effect of regulation heterogeneity reached only marginal significance.

Table 3. Regression models predicting implementation, team level.

Effect	Model 1			Model 2			Model 3		
	Estimate	*SE*	*p*	Estimate	*SE*	*p*	Estimate	*SE*	*p*
Intercept	4.57	0.02	< .001	4.56	0.02	< .001	4.54	0.03	< .001
M core SIC	**.07**	**0.02**	**< .001**	**.07**	**0.03**	**.005**	**.06**	**0.02**	**.010**
SD core SIC				−.03	0.03	.225	−.04	0.03	.134
M x *SD* core SIC				−.02	0.02	.356	−.02	0.03	.498
SD^2 core SIC							.01	0.02	.654
M Routine	.04	0.03	.233	.04	0.03	.267	.03	0.03	.388
SD Routine				−.04	0.03	.147	**−.05**	**0.03**	**.042**
M x *SD* Routine				−.01	0.03	.713	−.02	0.03	.503
SD^2 Routine							−.01	0.02	.480
M Regulation	**.17**	**0.03**	**< .001**	**.18**	**0.03**	**< .001**	**.18**	**0.03**	**< .001**
SD Regulation				**.07**	**0.03**	**.034**	**.05**	**0.03**	**.085**
M x *SD* Regulation				.00	0.03	.867	.04	0.04	.250
SD^2 Regulation							.03	0.02	.116
Team Size	−.03	0.04	.505	−.04	0.04	.363	−.05	0.04	.291
Team Size 1/*n*	−.03	0.04	.467	−.03	0.05	.559	−.04	0.05	.474
R^2	.61			.66			.68		
ΔR^2				.05			.02		

Notes: Values are standardized regression coefficients (β), standard errors (*SE*), and *p*-values for predicting implementation at the team level ($N = 104$).

Across all models, the discrepancy between overall explained variance and individual effect sizes was consistent with the highly interrelated nature of all SIC dimensions and indicated that implementation is best understood as a function of their combined configuration.

Discussion

This study examined whether SIC, traditionally conceptualized as a school-wide property, benefits from incorporating a team-level perspective. To this end, we investigated whether schools and teams differ in the degree of homogeneity in perceptions of SIC dimensions and to what extent these dimensions – both in their average levels and homogeneity – are associated with instructional practice.

As to within-unit agreement in 8 SIC dimension, although differences in ICCs between the school and team levels were not significant, a substantial proportion of variance across most dimensions was located at the team level. The fact that with few exceptions, the proportion of variance attributable to the team level was descriptively larger than the proportion attributable to the school level suggests a somewhat stronger agreement in teachers' perceptions of SIC within teams than within schools. Notably, this pattern was also evident for dimensions that teachers were asked to rate in terms of their school. Thus, teachers within the same team tended to perceive even school-level dimensions more similarly than teachers from different teams within the same school. This pattern aligns with conceptualizations of schools as loosely coupled but internally differentiated systems (Weick, 1976), in which coordination and shared meaning-making are more likely to emerge in smaller units (Meredith et al., 2017).

Regarding the second research question, more SIC dimensions (meta scales and routine) were significantly associated with implementation at the team level than at the school level. At the team level, higher perceived average levels of core SIC dimensions were consistently associated with higher levels of assessed implementation. Teams that

perceive themselves as strong in these capacities tended to report higher levels of implementation. At the school level, this association was also positive but not statistically significant. These findings align with previous research demonstrating that core SIC dimensions at the school level provide an important foundation for instructional practice (Sleegers et al., 2014; Thoonen et al., 2012).

Furthermore, at both the school and team levels, higher average perceived regulation was positively and significantly associated with implementation. In addition, teams with greater variability in perceived regulation reported higher levels of implementation. One possible explanation is that regulation reflects a set of cognitive, metacognitive, and motivational skills (Panadero, 2017) and that differences in how these skills are perceived and enacted within a team may be associated with reflection and adaptive responses to challenges. Some degree of perceptual heterogeneity in regulation may be productive, as it can foster learning-oriented dialogue and problem solving. Together, these findings suggest that regulation emerges as the most robust predictor of implementation across levels and indicates that regulation processes are not only theoretically meaningful (Maag Merki et al., 2023) but also empirically relevant. This highlights that improvement capacity is reflected not only in shared norms and structures but also in the processes through which teachers collectively navigate challenges in daily practice.

For teams, routine heterogeneity was negatively associated with implementation, suggesting that teams with more homogeneous perceptions of routines reported higher implementation. Unlike regulation, routines primarily refer to established patterns of action. Shared understandings of such routines may enable smooth coordination and reduce ambiguity in daily work. When perceptions of routines diverge, coordinated action may be hindered, potentially limiting the effective enactment of (new) practices. Together, these findings underscore that homogeneity and heterogeneity may play different roles depending on the nature of the SIC dimension under consideration (Drach-Zahavy & Somech, 2002).

Taken together, the findings support the conceptual extension of SIC to the team level. They are consistent with organizational and team learning perspectives that view collective learning and improvement as emerging within smaller work units rather than exclusively at the organizational level (Edmondson, 2002; Van Den Bossche et al., 2022). SIC has traditionally been conceptualized as a school-level construct, but the present findings suggest that important variation in SIC may be located within teams.

Importantly, the teams analyzed were formally established units rather than self-selected subgroups, so that the findings cannot be attributed to teachers clustering together based on shared views. The results suggest that even within formal teams, shared meaning-making and joint action are more likely than within schools (Meredith et al., 2017), and that SIC showed more consistent associations with implementation at the team level than at the school level. At the same time, formal teams differ in the extent to which they function as cohesive units (Vangrieken et al., 2017), a pattern reflected in the somewhat higher team-level ICCs.

Several limitations should be considered. First, the cross-sectional design does not allow causal conclusions. Longitudinal designs would be particularly valuable to examine how team-level capacity and homogeneity evolve over time and how they relate to instructional practice. Second, as all variables were assessed through self-reports, common method bias may have contributed to the observed associations and

the comparatively high explained variance in some models. Future research should therefore incorporate multiple data sources measurements. Third, although robust standard errors were applied, the relatively small number of schools and the aggregation of data at the school and team levels may have reduced statistical power, obscured relationships at the individual level, and increased the risk of overfitting, particularly given the substantial shared variance among predictors. Fourth, the study focused on formal teams; future research could examine whether informal teams further explain within-school variation. Finally, future research should more systematically report variance components to make within- and between-unit variation transparent.

Acknowledgements

The author acknowledges the use of ChatGPT (version 5.3) for language improvement. Final editing was performed by a professional native-English academic proofreader.

CRediT roles

Andrea Wullschleger: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, visualization, writing – original draft. Urs Grob: Conceptualization, formal analysis, methodology, visualization, writing – review & edition. Katharina Maag Merki: Conceptualization, funding acquisition, investigation, project administration, writing – review & edition. Ladina Zumsteg: Data curation, writing – review & edition. Beat Rechsteiner: Data curation, investigation, writing – review & edition. All authors approved the final version to be published and agree to be accountable for all aspects of the work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by the Swiss National Science Foundation (SNSF) under Grant number 00019_175872/1.

ORCID

Andrea Wullschleger http://orcid.org/0000-0003-4390-3835
Urs Grob http://orcid.org/0000-0002-6671-7085
Katharina Maag Merki http://orcid.org/0000-0002-0215-1684
Beat Rechsteiner http://orcid.org/0000-0003-2555-431X

References

Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method, and practice*. Addison-Wesley.

Beckmann, L., Kötter-Mathes, S., Klein, E. D., Bremm, N., & Van Ackeren, I. (2022). School's improvement capacity and responses to the COVID-19 pandemic: Evidence from schools serving

disadvantaged communities. *Frontiers in Education*, *7*, 1008813. https://doi.org/10.3389/feduc.2022.1008813

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bouwmans, M., Runhaar, P., Wesselink, R., & Mulder, M. (2017). Fostering teachers' team learning: An interplay between transformational leadership and participative decision-making? *Teaching and Teacher Education*, *65*, 71–80. https://doi.org/10.1016/j.tate.2017.03.010

Carron, A. V., Widmeyer, W. N., & Brawley, L. R. (1985). The development of an instrument to assess cohesion in sport teams: The group environment questionnaire. *Journal of Sport Psychology*, *7*(3), 244–266. https://doi.org/10.1123/jsp.7.3.244

Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, *23*(3), 239–290. https://doi.org/10.1177/014920639702300303

Drach-Zahavy, A., & Somech, A. (2002). Team heterogeneity and its relationship with team support and team effectiveness. *Journal of Educational Administration*, *40*(1), 44–66. https://doi.org/10.1108/09578230210415643

Edmondson, A. C. (2002). The local and variegated nature of learning in organizations: A group-level perspective. *Organization Science*, *13*(2), 128–146. https://doi.org/10.1287/orsc.13.2.128.530

Feldhoff, T., & Radisch, F. (2021). Why must everything be so complicated? Demands and challenges on methods for analyzing school improvement processes. In A. Oude Groote Beverborg, T. Feldhoff, K. Maag Merki, & F. Radisch (Eds.), *Concept and design developments in school improvement research* (pp. 9–25). Springer International Publishing.

Hadwin, A. F., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 83–106). Routledge.

Hallinger, P., & Heck, R. H. (2010a). 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

Hallinger, P., & Heck, R. H. (2010b). Leadership for learning: Does collaborative leadership make a difference in school improvement? *Educational Management Administration & Leadership*, *38*(6), 654–678. https://doi.org/10.1177/1741143210379060

Hallinger, P., & Heck, R. H. (2011). Exploring the journey of school improvement: Classifying and analyzing patterns of change in school improvement processes and learning outcomes. *School Effectiveness and School Improvement*, *22*(1), 1–27. https://doi.org/10.1080/09243453.2010.536322

Leithwood, K., & Jantzi, D. (2006). Transformational school leadership for large-scale reform: Effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement*, *17*(2), 201–227. https://doi.org/10.1080/09243450600565829

Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology*, *88*(4), 587–604. https://doi.org/10.1037/0021-9010.88.4.587

Lomos, C., Hofman, R. H., & Bosker, R. J. (2011). Professional communities and student achievement – a meta-analysis. *School Effectiveness and School Improvement*, *22*(2), 121–148. https://doi.org/10.1080/09243453.2010.550467

Maag Merki, K. (2017). School Improvement Capacity als ein Forschungsfeld der Schulentwicklungs- und Schuleffektivitätsforschung. Theoretische und methodische Herausforderungen. In U. Steffens, K. Maag Merki, & H. Fend (Eds.), *Schulgestaltung. Aktuelle Befunde und Perspektiven der Schulqualitäts- und Schulentwicklungsforschung* (pp. 225–242). Springer VS.

Maag Merki, K., Wullschleger, A., & Rechsteiner, B. (2023). 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*. https://doi.org/10.1007/s10833-022-09459-1

Meredith, C., Moolenaar, N. M., Struyve, C., Vandecandelaere, M., Gielen, S., & Kyndt, E. (2017). The measurement of collaborative culture in secondary schools: An informal subgroup approach. *Front Learning Research*, *5*(2), 24–35. https://doi.org/10.14786/flr.v5i2.283

Mitchell, C., & Sackney, L. (2011). *Profound improvement: Building capacity for a learning community* (2nd ed). Routledge.

Newmann, F. M., King, M. B., & Youngs, P. (2000). Professional development that addresses school capacity: Lessons from urban elementary schools. *American Journal of Education*, *108*(4), 259–299. https://doi.org/10.1086/444249

OECD. (2020). TALIS 2018 Results (Volume II): Teachers and school leaders as valued professionals. *TALIS*. https://doi.org/10.1787/19cf08df-en

OECD. (2023). *Equity and inclusion in education: Finding strength through diversity*. OECD Publishing.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*(422), 1–28. https://doi.org/10.3389/fpsyg.2017.00422

Park, S., Henkin, A. B., & Egley, R. (2005). Teacher team commitment, teamwork and trust: Exploring associations. *Journal of Educational Administration*, *43*(5), 462–479. https://doi.org/10.1108/09578230510615233

Rechsteiner, B., Compagnoni, M., Wullschleger, A., Schäfer, L. M., Rickenbacher, A., & Maag Merki, K. (2022). Teachers involved in school improvement: Analyzing mediating mechanisms of teachers' boundary-crossing activities between leadership perception and teacher involvement. *Teaching and Teacher Education*, *116*, 103774. https://doi.org/10.1016/j.tate.2022.103774

Rupert, J., Blomme, R. J., Dragt, M. J., & Jehn, K. (2016). Being different, but close: How and when faultlines enhance team learning. *European Management Review*, *13*(4), 275–290. https://doi.org/10.1111/emre.12083

Schwinger, M., von der Laden, T., & Spinath, B. (2007). Strategien zur Motivationsregulation und ihre Erfassung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *39*(2), 57–69. https://doi.org/10.1026/0049-8637.39.2.57

Senge, P. M. (1990). *The fifth discipline: The art & practice of the learning organization*. Currency Doubleday.

Sleegers, P. J. C., Den Brok, P., Verbiest, E., Moolenaar, N. M., & Daly, A. J. (2013). Toward conceptual clarity: A multidimensional, multilevel model of professional learning communities in Dutch elementary schools. *The Elementary School Journal*, *114*(1), 118–137. https://doi.org/10.1086/671063

Sleegers, P. J. C., Thoonen, E., Oort, F., & Peetsma, T. (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

Staples, D. S., & Webster, J. (2008). Exploring the effects of trust, task interdependence and virtualness on knowledge sharing in teams. *Information Systems Journal*, *18*(6), 617–640. https://doi.org/10.1111/j.1365-2575.2007.00244.x

Stoll, L. (2009). Capacity building for school improvement or creating capacity for learning? A changing landscape. *Journal of Educational Change*, *10*(2-3), 115–127. https://doi.org/10.1007/s10833-009-9104-3

Thoonen, E. E. J., Sleegers, P. J. C., Oort, F. J., & Peetsma, T. T. D. (2012). Building school-wide capacity for improvement: The role of leadership, school organizational conditions, and teacher factors. *School Effectiveness and School Improvement*, *23*(4), 441–460. https://doi.org/10.1080/09243453.2012.678867

Thoonen, E. E. J., Sleegers, P. J. C., Oort, F. J., Peetsma, T. T. D., & Geijsel, F. P. (2011). How to improve teaching practices: The role of teacher motivation, organizational factors, and leadership practices. *Educational Administration Quarterly*, *47*(3), 496–536. https://doi.org/10.1177/0013161X11400185

Uline, C. L., Tschannen-Moran, M., & Perez, L. (2003). Constructive conflict: How controversy can contribute to school improvement. *Teachers College Record: The Voice of Scholarship in Education*, *105*(5), 782–816. https://doi.org/10.1177/016146810310500505

Van Den Bossche, P., Gabelica, C., & Koeslag-Kreunen, M. (2022). Team learning. In C. Harteis, D. Gijbels, & E. Kyndt (Eds), *Research approaches on workplace learning* (Vol. 31, pp. 201–218). Springer International Publishing.

Vangrieken, K., Boon, A., Dochy, F., & Kyndt, E. (2017). Group, team, or something in between? Conceptualising and measuring team entitativity. *Frontline Learning Research*, *5*(4), 1–41. https://doi.org/10.14786/flr.v5i4.297

Wageman, R., Gardner, H., & Mortensen, M. (2012). The changing ecology of teams: New directions for teams research. *Journal of Organizational Behavior*, *33*(3), 301–315. https://doi.org/10.1002/job.1775

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21*(1), 1–19. https://doi.org/10.2307/2391875

Widmann, A., & Mulder, R. H. (2018). Team learning behaviours and innovative work behaviour in work teams. *European Journal of Innovation Management*, *21*(3), 501–520. https://doi.org/10.1108/EJIM-12-2017-0194

Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Erlbaum.

Appendix. Comparison of teachers included and excluded from the final sample.

Variable	Included			Excluded			*p*	Cohen's *d*	BF *F*	BF *p*
	n	*M*	*SD*	*n*	*M*	*SD*				
Leadership (sl)	590	4.71	0.82	649	4.75	0.85	.364	0.05	0.40	.527
Task cohesion (sl)	593	4.23	0.74	712	4.45	0.68	< .001	0.31	1.28	.259
Error culture (sl)	590	4.32	0.79	715	4.53	0.77	< .001	0.27	0.03	.860
Knowledge sharing (sl)	595	4.65	0.73	717	4.85	0.72	< .001	0.28	0.28	.599
Routine (sl)	591	4.67	0.63	708	4.69	0.62	.532	0.03	0.02	.896
Routine (tl)	594	4.78	0.62	712	4.78	0.61	.795	−0.01	0.05	.824
Cognitive regulation (sl)	587	4.49	0.65	717	4.53	0.69	.320	0.05	0.18	.668
Cognitive regulation (tl)	589	4.65	0.62	719	4.64	0.65	.850	−0.01	0.42	.519
Metacognitive regulation (sl)	589	4.67	0.67	707	4.72	0.65	.216	0.07	1.91	.167
Metacognitive regulation (tl)	590	4.89	0.64	712	4.89	0.61	.924	0.01	2.16	.142
Motivational regulation (sl)	586	4.6	0.65	706	4.62	0.62	.559	0.03	0.31	.575
Motivational regulation (tl)	586	4.68	0.64	710	4.7	0.6	.537	0.03	1.31	.252
Implementation (sl)	585	4.34	0.63	711	4.41	0.64	.057	0.11	0.51	.476
Implementation (tl)	591	4.54	0.6	714	4.55	0.62	.968	0.00	0.13	.722

Notes: Included teachers were assigned to clearly identifiable formal teams and constituted the final sample. Group differences were assessed using independent-samples t-tests. Cohen's *d* values indicate effect sizes. BF *F* and BF *p* refer to Brown–Forsythe tests of homogeneity of variance.