Check for updates

Factors influencing the implementation of a teacher professional development program to improve teaching quality

Lucas Silva Didier[1]*, Kim Schildkamp[2], Adrie J. Visscher[2] and Roel J. Bosker[1]

[1]GION Education/Research, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, Netherlands, [2]ELAN, Department of Teacher Development, University of Twente, Enschede, Netherlands

In this study, we examined why a Teacher Professional Development (TPD) program, designed to support teachers in using students' perceptions of teaching quality (SPTQ) data, faced significant implementation challenges in 17 secondary schools in Chile. Despite voluntary participation and initial interest, 15 of the 17 schools dropped out within 2−3 months of starting the program. Through 12 semi-structured interviews with professional learning community coordinators from nine schools, we investigated four key attributes of the TPD program to understand implementation challenges: its added value, compatibility, clarity, and tolerance. While coordinators valued several aspects of the program (including its structured manual, evidence-based teaching strategies, and integration of SPTQ data) significant implementation barriers emerged. Time constraints, lack of technological infrastructure, and insufficient organizational routines made the implementation of the TPD program too burdensome for most schools. We discuss how compatibility between TPD programs and schools' existing structures and routines acts as a critical bottleneck that can prevent successful implementation, even when participants see value in the program. This study provides important insights into the conditions necessary for successful TPD implementation in a global south country.

KEYWORDS

data-informed decision-making, teacher professional development, students' perceptions of teaching quality, professional learning communities, implementation challenges, educational design research

1 Introduction

Teaching quality is an important factor that supports student learning (Hattie, 2009; Muijs et al., 2014). However, teachers differ in the quality and effectiveness of their instructional practices (Escribano et al., 2020; Fernández-García et al., 2019). Therefore, it is important to support teachers in improving their teaching quality so they can offer quality learning experiences to as many students as possible. According to a recent meta-analysis, using data from students' perceptions of teaching quality (SPTQs) may help teachers to improve their teaching quality (Röhl et al., 2025). SPTQ data offers a unique perspective on teaching quality by capturing insights from students, who spend the most time experiencing teachers' practices (Gaertner, 2014; Kellermann et al., 2022). This method allows teachers to gather information

about aspects of their teaching quality that may be challenging to obtain through other approaches (e.g., classroom observation), such as clarity of explanations and emotional support (Helms-Lorenz and Visscher, 2021). Additionally, SPTQ data facilitates rapid feedback cycles, enabling timely adjustments to teaching practices (Bijlsma et al., 2019).

Teachers need to master diverse competencies to utilize SPTQ data to improve their practices, which as a whole are called data literacy (Mandinach and Gummer, 2016). Teachers usually find it difficult to use data to improve their practice (Cowie and Cooper, 2017; Sun et al., 2016). Therefore, we designed a Teacher Professional Development (TPD) program to support teachers in Chile in improving their teaching quality using SPTQ data by working in a Professional Learning Community (PLC) within the school. The design was informed by studies in several fields that provide insights into how to help teachers change their practices in the classroom (e.g., Anders et al., 2022; Kraft et al., 2018; van Merriënboer and Kirschner, 2018).

However, implementing TPD programs in real-world settings presents several challenges (McKenney and Reeves, 2019). For example, teachers can think that the TPD is too complicated, implies too much of a burden, is not aligned with teachers' values about education, or is useless in their context (McChesney and Aldridge, 2021; McKenney and Reeves, 2019). These challenges underscore the importance of considering both the program's design and the context in which it is implemented to facilitate its feasibility and impact.

In this study, we aim to investigate how PLC coordinators, who are teacher leaders and principals who led a PLC in a school, perceive four attributes of the TPD program that are key for implementing and scaling this program, according to educational design research (EDR). The attributes McKenney and Reeves (2019) specified are added value, compatibility, clarity, and tolerance. Given that 15 out of 17 schools dropped out, we use the four attributes by EDR to understand the causes of implementation challenges that led to a large dropout. Consequently, the research question that guides this study is: How do PLC coordinators perceive the added value, compatibility, clarity, and tolerance of a TPD program aimed to improve the teaching quality of secondary school teachers in Chile? Answering this question can help us to provide insights into important factors for the design and implementation of TPD programs to improve teaching quality, focused on using (SPTQ) data (in a global south country).

2 Theoretical framework

2.1 SPTQ data as a tool for improving teaching quality

Students' perceptions of teaching quality (SPTQ) surveys can help teachers improve their teaching quality and support students' learning (Röhl, 2021). In these surveys, students share their opinions about one teacher's instructional practices during an academic period (e.g., year, semester, lesson sequence, or a single lesson) (Helms-Lorenz and Visscher, 2021). Students can, for example, share their perceptions about topics such as classroom climate, feedback, instructional clarity, and lesson difficulty (e.g., Bijlsma et al., 2019; Herbert et al., 2022; van der Lans et al., 2015). Studies show that using SPTQ data can lead to identifying areas for improvement, stimulating improvement-oriented

undertakings, understanding students' social and emotional experiences in the classroom, and promoting democratic participation from students in their educational process (Finefter-Rosenbluh and Berry, 2024; Gaertner, 2014; Potvin, 2021).

2.2 Introduction to the TPD program

To gather student perspectives on their teaching practices, participating teachers employed the Impact! tool: a web platform with questionnaires developed and validated by researchers from the University of Twente in The Netherlands to measure teaching quality from students' perspective, teacher's perspective (self-report), and a classroom observation rubric. Besides, the Impact! tool has online dashboards to visualize the results. The following scales are included in all Impact! instruments: classroom climate, classroom management, instruction, and differentiation (Bijlsma et al., 2019, 2022).

Teachers learned how to use SPTQ data in PLCs led by PLC coordinators. PLC coordinators underwent three 1.5-h training sessions led by the first author. These sessions covered an introduction to the Impact! tool, the data-use cycle, and navigating the Impact! platform. Materials were also offered to teachers and PLC coordinators on a website, containing a detailed online manual outlining a specific agenda for each of the five planned PLC meetings comprising a data-use cycle, a directory of evidence-informed teaching strategies and activities that can be used to design an improvement plan for each of the items of Impact!, slides introducing the program and explaining each step of the data-use cycle. Table 1 presents how the data-use cycle was addressed in the PLC meetings. Teachers could conduct one or more data-use cycles throughout the program.

Regarding the content of the TPD, the learning activities were designed to learn a set of competencies helpful for conducting the eight steps of the data-use cycle mentioned previously (e.g., set SMART goals, describe the SPTQ data, design an improvement plan).

The agenda of the PLC meetings (see Table 1) was inspired by the TPD "Embedding Formative Assessment" (Wiliam and Leahy, 2014), showing an effect size of 0.11 in a randomized controlled trial conducted by Anders et al. (2022). For example, similar to the TPD program mentioned above, we also created a program in which most of the work was done by the PLC coordinators and teachers, and we conducted a training at the beginning of the TPD program. Another source of inspiration for choosing the topics of the TPD program were other TPD activities for effective data-use TPD (e.g., Lai and McNaughton, 2016; Poortman and Schildkamp, 2016; van Geel et al., 2017). Many of the activities we proposed for teachers were based on previous data-use models such as the data team intervention (Schildkamp et al., 2018). Based on this all, we designed a TPD program with a clear structure. For each meeting, teachers would conduct activities for each step of the data use cycle (see Table 1).

2.3 Features of the TPD design

The TPD program included the five meetings in Table 1.

The design of the TPD program was inspired by the four pillars for effective teacher learning found in a recent meta-analysis (Visscher et al., 2023): performance standards, self-regulation by teachers, cooperation, and classroom coaching.

TABLE 1 Components of the data-use cycle covered in each of the five meetings.

Meeting	Components of the data-use cycle	Description
1	Prepare the data-use cycle	Organize the PLC to use SPTQ data (introduce the Impact! program, arrange when to meet, roles, etc.).
	Define an improvement goal – part 1	Define an area or topic the teacher wants to work on (e.g., classroom climate, feedback).
	Collect data with Impact!	Collect data on students' perceptions of teaching quality, classroom observation, and/or teachers' self-report.
2	Analyze Impact! data	Describe the SPTQ data.
	Interpret Impact! data	Make sense of the SPTQ data and explain possible causes of the results.
	Define an improvement goal – part 2	Define a SMART goal to improve teaching quality based on the SPTQ data.
3	Design an improvement plan(s)	Design an improvement plan to improve teaching quality. Ideally, teachers use evidence-informed teaching strategies.
4	Implement and monitor the teaching quality improvement plan(s)	Enact the improvement plan and adjustments in the classroom.
5	Evaluate the results of the teaching quality improvement plan(s)	Evaluate with Impact! to what extent the goal was achieved and evaluate the implementation process.

Performance standards refer to clearly communicating what constitutes effective teaching practices (Visscher et al., 2023). In the TPD program, the Impact! tool played a central role in defining quality teaching in five constructs: classroom climate, classroom management, differentiation, instruction, lesson result. These constructs are specified further in the items (e.g., for classroom climate "The teacher thinks I can be good at his subject"). Through SPTQ data, teachers gained a clear understanding of areas such as clarity of explanations or emotional support, enabling them to align their practices with these defined standards and improve their performance accordingly.

Self-regulation involves teachers monitoring their progress, reflecting on progress, making a planning on further improvement (Visscher et al., 2023). In the TPD program, self-regulation was taken into account through the activities embedded in the data-use cycle. Teachers were required to monitor their level of performance through SPTQ data, reflect on areas of strength and opportunities for growth, define specific improvement goals.

According to the Visscher et al. (2023) meta-analysis cooperation is fundamental for effective teacher learning, as collaboration fosters shared accountability, mutual support, exchange knowledge, and collective capacity building (Filderman et al., 2022; Kennedy, 2016; Wiliam and Leahy, 2014). In the TPD program, cooperation was encouraged through Professional Learning Communities (PLCs), where teachers worked together to discuss their SPTQ data, share strategies, and support each other in achieving their improvement goals.

The TPD program encouraged classroom coaching. Teachers had the PLC's colleagues to share experiences, exchange ideas, give feedback, and learn from one another. Additionally, the TPD program offered a directory of evidence-based strategies, such as retrieval practice and think-pair-share, to help teachers address areas identified in SPTQ feedback (e.g., Agarwal, 2019; Mundelsee and Jurkowski, 2021).

2.4 Attributes to measure TPD implementation and scale

Based on Educational Design Research, four intervention attributes can help implement and scale educational interventions, such as our TPD program: added value, compatibility, clarity, and tolerance (McKenney and Reeves, 2019).

2.4.1 Added value
This intervention attribute, in the case of our TPD program, refers to the extent to which the program offers something better than teachers' current alternatives to improve teaching quality (Rogers, 2003). Being involved in an intervention like a TPD to improve teaching quality requires effort, time, and resources that teachers need to believe are worthwhile to invest in, given the results they will obtain (Doyle and Ponder, 1977; Rogers, 2003). Added value is a delicate balance between advantages and acceptable costs (Doyle and Ponder, 1977). According to the Practicality theory (Doyle and Ponder, 1977), teachers are more inclined to modify their classroom practices when they perceive the new approach as useful. The greater the teachers' perceptions of the usefulness of an intervention, the more likely they will adopt changes in their classrooms (Janssen et al., 2013; McChesney and Aldridge, 2021). This is consistent with findings in the field of data use, where it has been found that teachers are more willing to use data when the perception of usefulness increases (Goffin et al., 2022). Teachers often value TPD more when they can apply new ideas in their own classrooms, collaborate during their learning processes (Datnow et al., 2013; OECD, 2019), and if they evaluate positively the feedback given by school leaders or colleagues about their performance (Yoon and Kim, 2023).

2.4.2 Compatibility
This attribute refers to the extent to which the TPD program is "congruent with existing values, cultures, practices, and beliefs" (McKenney and Reeves, 2019, p. 207). Zhao et al. (2002) found that innovations with technology that largely differ from school cultures were less successful. A TPD can bring fresh ideas to a school, but keeping the changes within reasonable limits is important. This is because if the changes are too drastic, they may be more difficult to implement (Chen, 2024; Guskey, 1988; Janssen et al., 2013; Zhao et al., 2002). Moreover, when teachers engage in practices that are not aligned with their beliefs, it may negatively influence their well-being and prevent teachers from implementing the desired change (Doyle and Ponder, 1977).

2.4.3 Clarity

According to McKenney and Reeves (2019), clarity refers to how clear the TPD program's goals, activities, and procedures are for its participants. They argue that it is desirable that teachers have a clear idea of the program's purpose and what is expected from them during the program (McKenney and Reeves, 2019). They also stated that an ill-structured TPD program would be difficult to implement. When implementing the desired change, unclear instructions, materials, and procedures cause friction. Similarly, McChesney and Aldridge (2021) emphasize developing clear instructions and materials and a detailed description of what the program's outcomes are supposed to be (McChesney and Aldridge, 2021).

2.4.4 Tolerance

Tolerance refers to how the program's key components must be enacted to achieve its goals (McKenney and Reeves, 2019). Because TPDs are commonly not implemented as intended (Cook et al., 2019; Doyle and Ponder, 1977), it is important to design TPD programs that allow for flexibility in adapting some components to make it effective in a variety of school sites, even under challenging circumstances (Chen, 2024; Clarke and Dede, 2009). Furthermore, schools vary tremendously in infrastructure, leadership, and school climate, among others, so it is important to create TPD programs with a core set of features that can be adapted to each of these varying contexts (Ryan et al., 2024). The idea is to copy the program in its essence but not to replicate it exactly (Brown and Flood, 2018). Tolerant TPD programs are those that have low levels of dependence on people and resources, which increases the chances of uptake by teachers (Zhao et al., 2002). Low-tolerant interventions are those created in a one-size-fits-all approach, where designers create a rigid solution that is expected to be implemented regardless of the differences in the school context (McKenney and Reeves, 2019).

3 Materials and methods

3.1 Research design

Schools were invited to participate in the Impact! program through social media, email campaigns by school districts, research centers, networks of schools, and snowball by participants enrolled in the TPD program. The criteria for selecting schools were that they be urban, have at least three participants for the PLC, and have teachers teaching classes from 7th grade onwards. The schools selected their PLC coordinators, ensuring they had the authority and capacity to oversee the implementation of the TPD program.

Teacher participation in the program was voluntary, as we emphasized to principals during recruitment. Principals and teachers were informed that participation required a commitment to collaborative meetings and data collection cycles. Principals were encouraged to ensure that the decision to join reflected a shared interest among their staff.

Because in Chile there are limited experiences of schools using SPTQ data, we sought feedback from different actors to adapt the research from North Global countries (e.g., Bijlsma et al., 2019; Röhl et al., 2025) to the Chilean context. Partnerships were established with local educational institutions, including school boards, faculties of education, and the Ministry of Education. These collaborations provided valuable insights into the cultural and contextual relevance of the TPD program and helped align the program's goals with national educational priorities.

We conducted online semi-structured interviews with PLC coordinators to investigate the PLC coordinators' opinions about the four attributes of the TPD program from educational design research, so we can understand the causes of implementation challenges. We invited by e-mail all 17 PLC coordinators from all schools that started in the TPD program. When they signed up for the study, they were aware that we would ask them later on to have an interview about their perceptions of the TPD program. Nine out of 17 PLC coordinators agreed to participate in the interviews, and they were working in public, charter, and private schools, reflecting the diversity of the overall sample. The interviews were conducted during the middle and the end of the TPD program so we could monitor PLCs' perceptions about the four intervention attributes during the implementation and if that changed afterward. In total, 12 interviews were conducted, with five taking place at the program's midpoint and seven at its end. Three PLC coordinators participated in both the mid- and end-point interviews. These interviews represent feedback from nine of the 17 schools that initially participated in the program. The remaining PLC coordinators did not respond to repeated invitations to participate in the interviews or inform us via email that they would not continue in the TPD program.

During the implementation of the TPD program, 15 out of 17 schools dropped out. All participating schools were urban. This large dropout will be revisited in the discussion section, where we explore the factors contributing to the implementation challenges.

3.2 Ethics statement

The University of Groningen granted ethics approvals for this study PED-2223-S-0057.

3.3 Educational design research attributes

The educational design research attributes informed the design and evaluation of the TPD program in several ways.

We considered the added value of the TPD program in several ways. First, we focused on practical activities for PLC meetings to help teachers see the program's utility (McChesney and Aldridge, 2021). The program also fostered collaborative learning, a highly valued form of professional development in Chile (OECD, 2019), enabling teachers to reflect with colleagues and build a support network for sustained improvement (Datnow et al., 2013). Additionally, the TPD program gave teachers autonomy to set goals and tailor improvement plans to their needs, making it personalized and relevant (Chen, 2024; Wiliam and Leahy, 2014). To minimize time and resource demands, PLC meetings were short, bi-weekly, and structured to complete tasks during sessions, reducing the need for extra work. Finally, the program introduced teachers to SPTQ data, a novel tool in the Chilean context (OECD, 2020).

To make the TPD program more compatible, the teaching dimensions in the Impact! questionnaire were explicitly related to Chile's national teaching standards, and participation in the TPD program was positioned as a strategic initiative to support schools' broader goals. Additionally, we communicated that PLC meetings can be integrated into existing schedules, such as weekly department meetings.

Clarity was addressed by providing clear instructions, scripts for PLC coordinators, templates for goal-setting and designing improvement plans, and straightforward explanations about what was measured in the Impact! questionnaire, avoiding jargon unfamiliar to teachers. These materials were discussed during PLC coordinators' training to prevent misunderstandings.

The TPD program was designed with flexibility in mind. Teachers had the autonomy to decide on what aspects of their teaching to improve (Chen, 2024; Thompson and Wiliam, 2007). They could administer the Impact! questionnaire either online or on paper, and PLC coordinators were allowed to schedule extra support meetings as needed. Low-tech tools like Word documents and PowerPoint ensured accessibility, and teachers were encouraged to enrich their SPTQ data with peer observations and self-assessments.

3.4 Participant characteristics

Table 2 shows that the interview participants were mostly academic coordinators who belong to the school's leadership team, and one was an instructional coach from the educational central office of the network of schools to which the schools belong.

3.5 Data collection

The first author collected the data through online semi-structured interviews conducted in Spanish. Each interview lasted 27–52 min and was conducted individually with each PLC coordinator. For the three schools with two PLC coordinators, both were interviewed together. Supplementary material 1 contains the interview protocol used. The interview questions were structured according to the four intervention attributes from Educational Design Research. Because the interview protocol was semi-structured, the interviewer did not always follow it in the same planned order. The interviewer used the participants' answers to the first two questions (which were about what aspects of the program they liked and disliked and why?) to direct the interview toward the relevant intervention attribute. For example, if participants expressed dissatisfaction at the beginning of the interview with the time required for the TPD program, the interviewer would

direct the conversation toward the compatibility attribute and thereafter explore the remaining attributes. The interviewer confirmed the interpretations with the interviewee during the interview and posed clarification questions if needed (Roulston, 2010).

3.6 Coding framework

The interviews were transcribed verbatim into Spanish using Amberscript and then checked by the first author. Then, using Atlas.ti, 147 codes were inductively generated by the first author to diminish the data into a manageable undertaking. Next, codes were grouped into categories, which were grouped into four themes according to the four intervention attributes we studied. We chose an initial inductive coding approach to ensure that the richness of the data was captured before assigning it to Educational Design Research attributes. This approach also allowed us to explore whether the four TPD attributes from Educational Design Research were evident in the data in a more emergent manner. Finally, we used interrater reliability to enhance the consistency of the analytic process. The first and third authors met on two occasions to review a sample of codes (18%) and to ensure the shared interpretation of which theme each interview segment belongs (O'Connor and Joffe, 2020). In the first session agreement was 73%. After resolving the disagreement, we met again, and the agreement then was 100%. Inter-rater agreement among the two coders was determined using Cohen's Kappa coefficient (Cohen, 1960) using the irr package version 0.84.1 (Gamer et al., 2019) in R (R Core Team, 2024). Cohen's Kappa coefficient was estimated to account for agreement expected by chance (Cohen, 1960). The Kappa coefficient for all interview segments was 0.78 (p-value = <0.001), which is considered good (Altman, 1990).

In Supplementary material 2, we present the codebook used. Because the TPD program has several components, we organized the codes if they belong to the tools (Impact! platform, survey, materials) of the TPD program or the enactment by its users (meetings, interaction between participants, etc.).

4 Results

4.1 Added value

4.1.1 Tools

All PLC coordinators valued all the resources available to implement the program. For example, the meeting manual available for the PLC coordinator and teachers was considered useful to "guide the meetings" (PLC coordinator E) and to "do what you really needed to do" (PLC coordinator E) without focusing on other things that could deviate the community from the goal of the meeting.

One of the resources developed for participants was a directory with evidence-based teaching activities aligned with the dimensions measured in the Impact! questionnaire. This was considered useful by almost all PLC coordinators. PLC coordinator C noted, "We often say, 'You need to improve this' but lack the 'how'. This directory offers a toolbox of ideas, which is very helpful." Similarly, PLC coordinator H highlighted its benefit for novice teachers with limited teaching repertoires.

Despite its usefulness, the directory requires improvement. One PLC coordinator who used the tool extensively noted a lack of

TABLE 2 PLC coordinators' characteristics.

Who	Role	Type of school	When
PLC coordinator A	Instructional coach central office	Charter	End
PLC coordinator B	Academic Coordinator	Charter	End
PLC coordinator C	Academic Coordinator	Charter	End
PLC coordinator D	Academic Coordinator	Charter	Middle and end
PLC coordinator E	Principal	Charter	Middle and end
PLC coordinator F	Academic Coordinator	Private	Middle and end
PLC coordinator G	Head of the History department	Public	Middle
PLC coordinator H	Academic Coordinator	Charter	Middle
PLC coordinator I	Academic Coordinator	Charter	Middle

differentiation strategies, particularly for challenging high-performing students.

For all PLC coordinators, the fact that they could have students' perceptions about teaching quality as a source of data to improve was considered novel, useful, and insightful. None of them had had experience with using SPTQ data before.

PLC coordinator E highlighted how SPTQ data could be triangulated with classroom observations for a more complete perspective of teachers' instructional practices. The data from the Impact! questionnaire was useful for teachers to discover challenges in their daily practice that they had not noticed. One teacher said to his PLC coordinator when reflecting on the SPTQ data concerning student participation in the classroom: "I always say, 'Hey you, do you want to participate? But if I start to think about it, some kids have never participated in my classes.'"

4.1.2 Enactment of the TPD program

According to half of the PLC coordinators, using SPTQ data enhances student participation in education. One PLC coordinator mentioned that they usually ask students about their opinions on several topics, "but a survey is a more formal way of participation where we can make more visible the importance they have in the learning process" (PLC coordinator H). The same PLC coordinator emphasized the program's role in increasing student participation, which is typically limited in traditional educational settings.

Another PLC coordinator gave an example of how involving students in using SPTQ data was beneficial for them to reflect on their behavior in the classroom. When students were discussing the results of one of the teachers in front of the PLC coordinator, they mentioned: "Yes, but with Miss Sandra, we do not behave very well, so it is hard for her because we do not treat her well." Next, the other student said: "Yes, but we do not behave well because the teacher does not know how to set rules. This makes us behave poorly, which makes learning more difficult. So, if the teacher sets rules, we behave better and learn more."

Furthermore, students could give honest feedback when they felt secure answering the Impact! questionnaire. PLC coordinator E, who was a school principal, assured students that their responses would help teachers improve their lessons, which encouraged them to express their true thoughts about the learning experience.

When comparing the Impact! program with other TPD experiences, a few PLC coordinators highlighted that Impact! helps teachers to collaborate and use data, making the reflections more concrete and specific.

In addition, one PLC coordinator mentioned that the TPD program differs from the traditional mode of TPD, in which an external expert comes to the school to deliver content. The fact that the TPD program involved practical rather than theoretical learning activities was valued: "Teachers are very bored with professional development that involves a workshop where an expert comes in to present something. That kind of professional development at the school is no longer effective; it does not engage anyone" (PLC coordinator E).

Added value is not only about the benefits of the TPD program but also about how they are balanced with its costs. As a PLC coordinator eloquently said: "They were interested [in the TPD program], but having the disposition to say 'I will assign 2 h per week to this', they never have it. (…) They did not have the time to execute it" (PLC coordinator B).

The results show that the TPD program competed with other tasks in the school that were considered more of a priority for teachers. Schools needed to arrange the program amidst other priorities they had, such as administering a national standardized test, preparing the national teacher evaluation, assigning grades, preparing materials for their lessons, and interviewing parents, among others. "It is nothing against the platform [the TPD program], it is time" (PLC coordinator G). In other words, even though the participants saw the benefits of participating in the TPD program, the costs outweighed its benefits.

4.2 Compatibility

4.2.1 Tool

4.2.1.1 Alignment with national standards

All PLC coordinators considered the TPD program well-aligned with the national performance standards for teachers in Chile. For example, there were items in the Impact! questionnaire "that I could relate to one or more of the standards (…) There is alignment." (PLC coordinator D). Moreover, the program was also aligned with the strategic planning that schools had for the academic year and their existing instruments to measure teaching quality.

4.2.1.2 Limited tech infrastructure

For a few PLCs, the technological part of the program worked well: students could complete the questionnaire online using their phones, tablets or by going to the computer classroom. But for most of the PLCs, the technological part was a hurdle:

> The use of technology is always a problem. Yes, because that involves making room for something that is kind of an extra, an extra effort or extra work. I mean, for us to be able to bring it [the questionnaire] to the kids, the class had to end at least 15 min earlier (…) moving to the computer classroom, students turning on the computer. (PLC coordinator C).

In the same school, teachers shared their phone internet with the computers because the school wi-fi was not working. At another school, they only used the computer classroom after 15:00 because that was when the wi-fi was fastest during the day.

4.2.2 Enactment of the TPD program

4.2.2.1 Attitude toward student voice

A prerequisite for using SPTQ data is to validate students as a data source to improve teaching quality. According to some PLC coordinators, this was difficult for some teachers. In a school a head of department had negative beliefs about students, saying that students were performing badly because "they are lazy and stupid" (PLC coordinator F). This type of attitude was sometimes accompanied by a general resistance from teachers to reflect on their own teaching practices data. For example, in two schools that dropped out, teachers were already reluctant to being observed by the teacher leaders and peers before implementing the Impact! program, so "imagine that you need to be open to what a kid will say about you" (PLC coordinator F). "Teachers have told me exactly with these words: I have done it

perfectly, everything I have been asked to do. If students do not learn, it is not my fault." (PLC coordinator F).

In contrast, teachers with a positive attitude toward student voice were more proactive, addressing internal factors that could explain the SPTQ results and persevering in case of challenges, such as using pen and paper when the internet was not working. These teachers also shared SPTQ results with students and two of them involved them in designing solutions to improve the learning environment in the classroom. In these cases, students became more engaged and honest in their subsequent feedback, leading to more reliable and valid data in later applications of the questionnaire (PLC coordinator E).

4.2.2.2 Attitude toward change and improvement

PLC coordinators mentioned that the Impact! program can effectively help teachers to improve their teaching quality because it comprises essential ingredients such as pedagogical reflection and collaboration. However, its success is conditional on the attitude of participating teachers. As one PLC coordinator who coordinated a school that did not drop out mentioned, "it really worked in our favor that teachers were super diligent. I'm not sure if this works as well with more resistant teachers." (PLC coordinator A). Her intuition was right: PLC coordinator F mentioned that many teachers were resistant to the TPD program and other improvement initiatives in the school, making it impossible to implement the TPD program.

Teacher attitudes toward improvement can also be influenced by their scores on the Impact! questionnaire. Some teachers were hesitant to reflect on their practices due to low scores, with one PLC coordinator explaining that insecurity often leads to resistance: "Teachers are very insecure, which is why they resist evaluations too much. And besides, imagine being open to what a kid will say about you" (PLC coordinator E).

Given the importance of teachers' attitudes, PLC coordinators emphasized that the leadership team in the school should carefully select teachers who are open to using SPTQ data: "If you do not address that, it will fail one way or another, even if your system and process are awesome" (PLC coordinator E). PLC coordinator D, advised selecting committed teachers for the program, emphasizing the need for those who are "always on time, plan their lessons, are committed, do not want to give the minimum" (PLC coordinator D).

4.2.2.3 Alignment with school routines and time management

The alignment of the Impact! program with existing school routines influenced its implementation according to several PLC coordinators. In some schools with established bi-weekly department meetings, the TPD program fits into the schedule seamlessly: "We have a culture of meeting regularly, so it is not that new for us" (PLC coordinator E). However, most schools lacked time to have recurrent meetings. One PLC coordinator shared that one of his teachers said, "This (TPD program) works; I'm not discussing that; I am saying I do not have time for this." (PLC coordinator G). In these circumstances, implementing a data-use cycle was too much of a burden and made the TPD program not realistic to their context:

> "The fact that having activities that constantly is like…is not really thought from teachers' reality (…) Because we have many activities and meetings, we are swamped with things. So, following such a rigorous and strict agenda is hard. That is what I see: there

is a perspective from the outside, but not from the teacher. I think there are really a few schools that really provide spaces to reflect about teaching practices (…) We should have more space to reflect about our own teaching, but they are not available" (PLC coordinator H).

Even though this PLC coordinator believes the program has been designed to help teachers, the TPD program is considered unrealistic given the current circumstances of these schools, where regular meetings to analyze data and address educational issues do not exist.

The lack of time, exacerbated by the program's timing near the end of the academic year, made participation difficult. In fact, all schools withdrew from the program within a range of 2–3 weeks between the last week of October and the first 2 weeks of November, just before or during the national standardized assessments. An important behavior from PLC coordinator A that, from her perspective, helped them not to drop out of the program was sending emails to teachers before the meetings, reinforcing the importance of the program and what they would do, and following the tasks after the meeting. For some PLCs, orchestrating the program properly was impossible, so they could not finish it.

4.2.2.4 School internal emergencies

PLC coordinators identified numerous unforeseen challenges that disrupted the TPD program's continuity. High-frequency events, such as medical leaves due to stress, burnout, accidents, urgent meetings with parents, and lawsuits among staff, often destabilized PLCs. In some cases, the absence of just one teacher dissolved a PLC group, reducing interactions to one-on-one sessions. In schools that completed the Impact! program, the principal and teacher leaders were able to protect teachers from unforeseeable events. Sometimes, these leaders took care of the issues by themselves, and on other occasions, they delegated them to someone else. This variety of challenges underscores the complexity of maintaining the continuity of PLCs under unpredictable circumstances.

4.3 Clarity

Clarity mostly relates to the communication between the first author, who acted as the point of contact with the PLC coordinators, the activities the coordinators needed to do in the PLC meetings, and the resources available for them to use (e.g., the Impact! platform with the online dashboards of SPTQ results).

4.3.1 Tool

All PLC coordinators consider the tools in the TPD program to be clear and easy to follow. For example, one PLC coordinator stated that the Impact! program was "structured and each stage has a proper reason to be there" (PLC coordinator E). As another PLC coordinator mentioned, "I liked the fact that everything was step by step. (…) When everything is ordered and structured for me it is very good and for them [teachers] it was also very comfortable." (PLC coordinator A).

The manual received positive feedback from PLC coordinators. One of the PLC coordinators said that it was so helpful that she was "carrying the manual in my hand" (PLC coordinator A), and it was so easy to follow that it was described as "for dummies" for PLC coordinator G. This facilitates clarity regarding the primary purpose

and activities for each meeting, allowing participants to envision their role in it.

PLC coordinators also considered the survey "super friendly to answer, it is fast (…) it does not require much work" (PLC coordinator E). They also stated that during the questionnaire's application, students did not have questions about the items, "so the items were really well written" (PLC coordinator G).

Other aspects of the TPD were not clear to some PLC coordinators. For example, creating accounts for teachers in the Impact! platform was difficult for many teachers. Some PLC coordinators mentioned that teachers did not receive an e-mail with the credentials to log in to the platform.

4.3.2 Enactment of the TPD program

Initially, at PLC coordinator E's school, students viewed the Impact! questionnaire as an opportunity to express dissatisfaction with their teachers, using it as a form of protest because they felt unheard. This attitude changed dramatically after the school principal reiterated the purpose of the data collection before the second measurement, emphasizing that their feedback was essential for improving teaching practices. As a result, students began to see themselves as active contributors to the improvement process. One student remarked to PLC coordinator E, "It is so good you are doing this because I feel that we are never taken into account. Teachers come and do their lessons. Some of them do it very well, others not really (…) Maybe the teacher does not know that I did not learn, or she does not know that that way of teaching is not effective for us."

This shift highlights the importance of clear communication with students about the purpose of feedback. When students felt their input was valued and could lead to tangible improvements, they responded more positively, providing more honest and constructive feedback. This transformation underscores the role of clarity in fostering trust and engagement in the feedback process, leading to more reliable and useful data for improving teaching quality.

4.4 Tolerance

All participating schools made changes to the TPD program. The Impact! program was considered flexible to adapt for almost all PLC coordinators: "There is a structure, but it is not rigid" (PLC coordinator D), and "It is super adaptable," said another (PLC coordinator A). Below, we describe the adaptations users made to the tools and during the enactment.

4.4.1 Tools

Users made one adaptation to the tools regarding the administration of the Impact! questionnaire for students. Instead of using the online Impact! platform, students answered the questionnaire on paper, and their responses were put into Excel (PLC coordinator E and PLC coordinator A). This change was made because smartphones are banned in these schools and the lack of a reliable technological infrastructure.

A few users skipped some activities in the manual because they did not have much time to meet. It is unclear what type of adaptation teachers made, and whether these changes could have negatively affected the impact of the TPD program.

4.4.2 Enactment of the TPD program

PLC coordinators said the Impact! program allowed them to choose their classes and lessons to use in the program (PLC coordinator E), choose which activities to use in their lessons to improve their score in the Impact! questionnaire (PLC coordinator C), and adjust the duration of their meeting as needed (PLC coordinator E and PLC coordinator H).

A few PLC coordinators included individual meetings with teachers to discuss the low scores they obtained in the Impact! questionnaire before having the PLC meeting. This was done to take care of the psychological aspects involved in the teacher's improvement process:

> I had a 1:1 meeting with Ana because her results were really bad. So, I first met with her and asked if she wanted to share her results with Susan. She said "yes," so we met again in the formal [PLC] meeting.

Individual meetings also took place when PLC coordinators could not find time to meet all together, so they decided to meet individually to continue making progress in the program (PLC coordinator G). However, these individual meetings had a downside: there were no opportunities for "collective feedback" (PLC coordinator G), which meant teachers missed the chance to share common challenges and collaborate to find solutions to their needs.

Due to the short meetings one school was having (20–30 min), PLC coordinator H opted for creating a WhatsApp group and coordinating the program through this platform so she could keep the meeting about more substantive reflections and discussions. The WhatsApp group was used to remind teachers about administering the survey to students and submit the URLs with the link to the resources of the program (manual, teaching strategies, etc.). The PLC coordinators preferred WhatsApp to email because email "is too formal and takes more time to write" (PLC coordinator H).

5 Discussion

We designed a program to improve teaching quality using students' perceptions of teaching quality in professional learning communities. However, due to several implementation challenges, 15 out of 17 schools dropped out from the TPD program within the first 2–3 months. We explored how PLC coordinators perceived the added value, compatibility, clarity, and tolerance of the designed TPD program. This can help us understand which factors can better explain the large dropout.

Based on the 12 interviews conducted with PLC coordinators from nine out of 17 schools, we found the PLC coordinators hold positive perceptions about the Impact! program. In terms of *added value*, it was highlighted how novel and insightful it was to collect students' perceptions about teaching quality, to use the tools, to collaborate in the PLC, and to engage students' participation in the improvement process. Regarding *compatibility*, the Impact! program proved to be aligned with Chilean national standards of teaching. Also, the TPD program was congruent with teachers who were motivated to improve their practices and those who had positive attitudes toward students as a source of feedback. However, most schools faced challenges due to limited technology infrastructure and

a lack of time to meet and conduct structured improvement cycles, making it difficult to implement the Impact! program. In terms of clarity, the tools were considered self-explanatory, and PLC coordinators and teachers could envision their role in the TPD program. Finally, the Impact! program was also perceived as *tolerant*, given that it allowed adaptation to their context.

5.1 Using SPTQ data in collaboration with students

Using SPTQ data becomes significantly more fruitful when it is actively discussed with students, as suggested in the literature (Bijlsma and Röhl, 2023; Gaertner, 2014), and corroborated by PLC coordinators in this study. Rather than merely being data sources, students play a crucial role in interpreting the data and contributing to developing improvement plans (Finefter-Rosenbluh and Berry, 2024; Geurts et al., 2023). PLC coordinators observed that actively involving students in using SPTQ data can help teachers adjust their practices and encourage students to reflect on their behaviors and learning. To facilitate students' involvement in the use of SPTQ data, it helps teacher leaders and teachers communicate clearly with students about the purpose of collecting SPTQ data and its intended use.

5.2 Compatibility as a bottleneck for TPD implementation

Although participants found the intervention beneficial, most did not implement it as expected due to a lack of time and routines for regular meetings to enhance teaching quality using SPTQ data. This positions compatibility as a critical bottleneck for successful TPD implementation.

For some schools, the program was perceived as an additional burden rather than an integrated part of the school's routines. In addition to the TPD program-specific factors, such as the need for technological infrastructure, broader contextual challenges likely played a significant role in the implementation difficulties and dropouts observed in this study. Teacher workload emerged as a critical issue, as many participants struggled to balance the demands of the TPD program with their existing responsibilities. Secondary teachers in Chile have more compulsory instruction time than the average of OECD countries (OECD, 2023), and 52% of teachers reported feeling stressed by their work (Agencia de Calidad de la Educación, 2022). Teachers' stress can be partly associated with Chile's high-stakes accountability measures, including the national teacher evaluation system and standardized assessments such as SIMCE, which place additional pressure on educators (Contreras et al., 2024; Falabella, 2020). Participants in our study reported that teachers prioritized tasks related to these mandatory programs over the Impact! program. Besides, the school system faces a shortage of 13,630 teachers in 2020, and a projected shortage of 32,000 teachers by 2030 (Elige Educar, 2021). Creating new routines and reallocating tasks to accommodate a TPD program like the one we studied requires significant effort and energy, making it costly to implement (Guskey, 2002; Janssen et al., 2013; Rogers, 2003; van den Boom-Muilenburg et al., 2023). This finding aligns with other educational interventions that struggle with practicality, meaning that they are difficult to implement within participants' contexts due to similar constraints in the Chilean educational system, such as lack of time, limited resources and support available, and teachers' openness to improve their practices (Bowen et al., 2009; Finefter-Rosenbluh et al., 2021; McChesney and Aldridge, 2021).

5.3 Leaders can influence how SPTQ data is used

Despite constraints on time and routines, two schools remained actively engaged in the TPD program, mainly due to strong leadership. Aligned with findings from previous studies (e.g., Hawlitschek et al., 2024; van den Boom-Muilenburg et al., 2023), teachers leaders played a crucial role by providing individualized support (Schildkamp et al., 2019), attending PLC meetings, and showing interest in the data-use process (van den Boom-Muilenburg et al., 2023). Additionally, teacher leaders and principals safeguarded teachers' time for joining PLC's meetings and created a safe environment for engaging with SPTQ data (Schildkamp et al., 2019).

In one school, the principal communicated the importance of SPTQ data to students, previous to a teacher's second measurement, helping them understand the impact of their feedback on teachers. This exemplifies how teacher leaders can create a supportive climate for data use where students feel secure and valued. In sum, teacher leaders can facilitate SPTQ data use by influencing both teachers and students.

5.4 Practical implications

The results of this study suggest that PLC coordinators valued different aspects of the TPD program (see a complete list in Supplementary material 3). There are also several improvements to enhance the compatibility of the TPD program with schools in Chile and in other countries in the global south that share similar contextual characteristics. Implementing the program during the first semester of the academic year is recommended to avoid conflicts with mandatory end-of-year tasks. A mixed approach of Professional Learning Communities (PLCs) and one-on-one meetings between PLC coordinators and teachers could address scheduling difficulties and help teachers make sense of the data in a more secure environment. In addition, future iterations of the TPD program could incorporate a pre-implementation phase (Alley et al., 2023) to assess compatibility with schools' existing routines and develop tailored adaptation plans or adaptation protocols, (Holtrop et al., 2022) to facilitate compatibility with the school's context. Lastly, to make the TPD program more feasible for Chilean schools, it would be beneficial to keep the number of months but reduce the number of meetings required to complete one data-use cycle. This would enable teachers to derive benefits even with limited time and quickly recognize the TPD program's value.

A long-term approach to diminishing the dropout of STPQ-related TPD programs is to provide training opportunities before implementing similar TPD programs. Using SPTQ data was new for all participating teachers. Incorporating SPTQ-related training into

teacher education programs could assist future educators in recognizing its value, promoting greater openness and commitment to these initiatives once they enter the profession. Early exposure through practicum experiences or coursework may better equip teachers to utilize SPTQ data. Furthermore, collaborations with teacher induction programs could offer structured support and mentorship, strengthening the application of SPTQ insights and enhancing long-term engagement with TPD initiatives focused on using SPTQ data.

5.5 Limitations and future studies

This study relied on interviews with PLC coordinators from nine out of 17 schools, so we cannot generalize the results to the total sample of schools participating in the TPD program. Even when the insights of the PLC coordinators were valuable, we did not capture the full range of perspectives, particularly those of teachers and students directly involved in the TPD program. Additionally, due to various challenges, none of the participating schools fully completed the TPD program as designed. This incomplete implementation restricts our ability to compare the attributes of the TPD program between schools that completed it and those that did not.

Future studies should consider a longer-term approach (e.g., at least one academic year) to examine how schools adapt to and integrate the Impact! program over time, including multiple cycles of implementation. Furthermore, it would be valuable to evaluate the effects of the TPD program on teaching quality and students' achievement. In addition, it is essential to include the perceptions of teachers and students to gain a more holistic view of the TPD program's attributes. Based on the findings of this study, future research could focus on developing and testing an improved version of the TPD program that addresses the challenges mentioned by PLC coordinators. Additionally, contrasting cases with high and low improvement in teaching quality after the TPD program may provide deeper insights.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by University of Groningen–Faculty of Behavioural and Social Sciences. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LS: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. KS: Conceptualization, Methodology, Supervision, Writing – review & editing. AV: Conceptualization, Methodology, Supervision, Writing – review & editing. RB: Supervision, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1546448/full#supplementary-material

References

Agarwal, P. K. (2019). Retrieval practice & Bloom's taxonomy: do students need fact knowledge before higher order learning? *J. Educ. Psychol.* 111, 189–209. doi: 10.1037/edu0000282

Agencia de Calidad de la Educación. (2022). *Estudio de bienestar socioemocional docente* (p. 60). Agencia de Calidad de la Educación. Available at: https://archivos. agenciaeducacion.cl/Estudios%20Bienestar%20Socioemocional%20Docente.pdf (Accessed November 15, 2024).

Alley, Z. M., Chapman, J. E., Schaper, H., and Saldana, L. (2023). The relative value of pre-implementation stages for successful implementation of evidence-informed programs. *Implement. Sci.* 18:30. doi: 10.1186/s13012-023-01285-0

Altman, D. G. (1990). Practical statistics for medical research. New York: Chapman and Hall/CRC.

Anders, J., Foliano, F., Bursnall, M., Dorsett, R., Hudson, N., Runge, J., et al. (2022). The effect of embedding formative assessment on pupil attainment. *J. Res. Educ. Eff.* 15, 748–779. doi: 10.1080/19345747.2021.2018746

Bijlsma, H. J. E., Glas, C. A. W., and Visscher, A. J. (2022). Factors related to differences in digitally measured student perceptions of teaching quality. *Sch. Eff. Sch. Improv.* 33, 360–380. doi: 10.1080/09243453.2021.2023584

Bijlsma, H. J. E., and Röhl, S. (2023). "Fostering effective teaching at schools through measurements of student perceptions: processes, risks and chances" in Effective teaching around the world: Theoretical, empirical, methodological and practical insights. eds. R. Maulana, M. Helms-Lorenz and R. M. Klassen (Cham, Switzerland: Springer International Publishing), 119–135.

Bijlsma, H. J. E., Visscher, A. J., Dobbelaer, M. J., and Veldkamp, B. P. (2019). Does smartphone-assisted student feedback affect teachers' teaching quality? *Technol. Pedagog. Educ.* 28, 217–236. doi: 10.1080/1475939X.2019.1572534

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., et al. (2009). How we design feasibility studies. *Am. J. Prev. Med.* 36, 452–457. doi: 10.1016/j.amepre.2009.02.002

Brown, C., and Flood, J. (2018). Lost in translation? Can the use of theories of action be effective in helping teachers develop and scale up research-informed practices? *Teach. Teach. Educ.* 72, 144–154. doi: 10.1016/j.tate.2018.03.007

Chen, B. (2024). A framework for infrastructuring sustainable innovations in education. *J. Learn. Sci.* 33, 583–612. doi: 10.1080/10508406.2024.2320159

Clarke, J., and Dede, C. (2009). Design for Scalability: a case study of the River City curriculum. *J. Sci. Educ. Technol.* 18, 353–365. doi: 10.1007/s10956-009-9156-4

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46. doi: 10.1177/001316446002000104

Contreras, J., Martín, C. S., and Muñoz, S. (2024). Uses and consequences in the theory of action of high-stakes teacher evaluation in Chile. *Educ. Assess. Eval. Account.*:1–23. doi: 10.1007/s11092-024-09448-3

Cook, C. R., Lyon, A. R., Locke, J., Waltz, T., and Powell, B. J. (2019). Adapting a compilation of implementation strategies to advance school-based implementation research and practice. *Prev. Sci.* 20, 914–935. doi: 10.1007/s11121-019-01017-1

Cowie, B., and Cooper, B. (2017). Exploring the challenge of developing student teacher data literacy. *Assessment Educ.* 24, 147–163. doi: 10.1080/0969594X.2016.1225668

Datnow, A., Park, V., and Kennedy-Lewis, B. (2013). Affordances and constraints in the context of teacher collaboration for the purpose of data use. *J. Educ. Adm.* 51, 341–362. doi: 10.1108/09578231311311500

Doyle, W., and Ponder, G. A. (1977). The practicality ethic in teacher decision-making. *Interchange* 8, 1–12. doi: 10.1007/BF01189290

Elige Educar. (2021). *Análisis y proyección de la dotación docente en Chile*. Elige Educar. Available at: https://eligeeducar.cl/content/uploads/2021/01/deficitactualizado2021.pdf (Accessed November 15, 2024).

Escribano, R., Treviño, E., Nussbaum, M., Torres Irribarra, D., and Carrasco, D. (2020). How much does the quality of teaching vary at under-performing schools? Evidence from classroom observations in Chile. *Int. J. Educ. Dev.* 72:102125. doi: 10.1016/j.ijedudev.2019.102125

Falabella, A. (2020). The ethics of competition: accountability policy enactment in Chilean schools' everyday life. *J. Educ. Policy* 35, 23–45. doi: 10.1080/02680939.2019.1635272

Fernández-García, C.-M., Maulana, R., Inda-Caro, M., Helms-Lorenz, M., and García-Pérez, O. (2019). Student perceptions of secondary education teaching effectiveness: general profile, the role of personal factors, and educational level. *Front. Psychol.* 10:533. doi: 10.3389/fpsyg.2019.00533

Filderman, M. J., Toste, J. R., Didion, L., and Peng, P. (2022). Data literacy training for K–12 teachers: a Meta-analysis of the effects on teacher outcomes. *Remedial Spec. Educ.* 43, 328–343. doi: 10.1177/07419325211054208

Finefter-Rosenbluh, I., and Berry, A. (2024). With or without you? Exploring student perception survey assessment data as a teacher-student (un)shared process. *Teach. Teach.* 30, 227–240. doi: 10.1080/13540602.2023.2236955

Finefter-Rosenbluh, I., Ryan, T., and Barnes, M. (2021). The impact of student perception surveys on teachers' practice: teacher resistance and struggle in student voice-based assessment initiatives of effective teaching. *Teach. Teach. Educ.* 106:103436. doi: 10.1016/j.tate.2021.103436

Gaertner, H. (2014). Effects of student feedback as a method of self-evaluating the quality of teaching. *Stud. Educ. Eval.* 42, 91–99. doi: 10.1016/j.stueduc.2014.04.003

Gamer, M., Lemon, J., Fellows, I., and Singh, P. (2019). *Irr: various coefficients of interrater reliability and agreement.* Available at: https://cran.r-project.org/package=irr (Accessed November 07, 2024).

Geurts, E. M. A., Reijs, R. P., Leenders, H. H. M., Jansen, M. W. J., and Hoebe, C. J. P. A. (2023). Co-creation and decision-making with students about teaching and learning: a systematic literature review. *J. Educ. Chang.* 25, 103–125. doi: 10.1007/s10833-023-09481-x

Goffin, E., Janssen, R., and Vanhoof, J. (2022). The interplay of user beliefs and situated characteristics in explaining school performance feedback use. *Sch. Eff. Sch. Improv.* 33, 456–478. doi: 10.1080/09243453.2022.2041048

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teach. Teach. Educ.* 4, 63–69. doi: 10.1016/0742-051X(88)90025-X

Guskey, T. R. (2002). Professional Development and Teacher Change. *Teachers and Teaching*, 8, 381–391. doi: 10.1080/135406002100000512

Hattie, J. (2009). Visible learning. A synthesis of over 800 Meta-analyses relating to achievement. London, UK: Routledge.

Hawlitschek, P., Henschel, S., Richter, D., and Stanat, P. (2024). The relationship between teachers' and principals' use of results from nationwide achievement tests: the mediating role of teacher attitudes and data use culture. *Stud. Educ. Eval.* 80:101317. doi: 10.1016/j.stueduc.2023.101317

Helms-Lorenz, M., and Visscher, A. J. (2021). Unravelling the challenges of the data-based approach to teaching improvement. School Effectiveness and School Improvement

Herbert, B., Fischer, J., and Klieme, E. (2022). How valid are student perceptions of teaching quality across education systems? *Learn. Instr.* 82:101652. doi: 10.1016/j.learninstruc.2022.101652

Holtrop, J. S., Gurfinkel, D., Nederveld, A., Phimphasone-Brady, P., Hosokawa, P., Rubinson, C., et al. (2022). Methods for capturing and analyzing adaptations: implications for implementation research. *Implement. Sci.* 17:51. doi: 10.1186/s13012-022-01218-3

Janssen, F., Westbroek, H., Doyle, W., and Van Driel, J. (2013). How to make innovations practical. *Teachers College Record* 115, 1–43. doi: 10.1177/016146811311500703

Kellermann, C., Nachbauer, M., Gaertner, H., and Thiel, F. (2022). The effect of school leaders' feedback on teaching quality perceived by students – results of an intervention study. *Sch. Eff. Sch. Improv.* 34, 24–42. doi: 10.1080/09243453.2022.2075018

Kennedy, M. M. (2016). How does professional development improve teaching? *Rev. Educ. Res.* 86, 945–980. doi: 10.3102/0034654315626800

Kraft, M. A., Blazar, D., and Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: a Meta-analysis of the causal evidence. *Rev. Educ. Res.* 88, 547–588. doi: 10.3102/0034654318759268

Lai, M. K., and McNaughton, S. (2016). The impact of data use professional development on student achievement. *Teach. Teach. Educ.* 60, 434–443. doi: 10.1016/j.tate.2016.07.005

Mandinach, E. B., and Gummer, E. S. (2016). What does it mean for teachers to be data literate: laying out the skills, knowledge, and dispositions. *Teach. Teach. Educ.* 60, 366–376. doi: 10.1016/j.tate.2016.07.011

McChesney, K., and Aldridge, J. M. (2021). What gets in the way? A new conceptual model for the trajectory from teacher professional development to impact. *Prof. Dev. Educ.* 47, 834–852. doi: 10.1080/19415257.2019.1667412

McKenney, S. E., and Reeves, T. C. (2019). Conducting educational design research. *Second* Edn. London, UK: Routledge/Taylor & Francis Group.

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., and Earl, L. (2014). State of the art – teacher effectiveness and professional learning. *Sch. Eff. Sch. Improv.* 25, 231–256. doi: 10.1080/09243453.2014.885451

Mundelsee, L., and Jurkowski, S. (2021). Think and pair before share: effects of collaboration on students' in-class participation. *Learn. Individ. Differ.* 88:102015. doi: 10.1016/j.lindif.2021.102015

O'Connor, C., and Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *Int J Qual Methods* 19:1609406919899220. doi: 10.1177/1609406919899220

OECD (2019). TALIS 2018 Results (Volume I). Paris, France: OECD Publishing.

OECD. (2020). *TALIS 2018 Results* (Volume II). doi: 10.1787/19cf08df-en

OECD. (2023). *Chile*. Education at a Glance 2023: OECD Indicators.

Poortman, C. L., and Schildkamp, K. (2016). Solving student achievement problems with a data use intervention for teachers. *Teach. Teach. Educ.* 60, 425–433. doi: 10.1016/j.tate.2016.06.010

Potvin, A. S. (2021). "Students speaking to you": teachers listen to student surveys to improve classroom environment. *Learn. Environ. Res.* 24, 239–252. doi: 10.1007/s10984-020-09330-1

R Core Team (2024). *R: A language and environment for statistical computing* [computer software]. Vienna, Austria: R Foundation for Statistical Computing.

Rogers, E. M. (2003). Diffusion of innovations. New York: Free Press.

Röhl, S. (2021). Effects of Student Feedback on Teaching and Classes: An Overview and Meta-Analysis of Intervention Studies. In *Student Feedback on Teaching in Schools. Using Student Perceptions for the Development of Teaching and Teachers.* eds. W. Rollett, H. Bijlsma and S. Röhl (Springer International Publishing), pp. 139–156.

Röhl, S., Bijlsma, H., and Schwichow, M. (2025). Can feedback from students to teachers improve different dimensions of teaching quality in primary and secondary education? A hierarchical meta-analysis: Educational Assessment, Evaluation and Accountability.

Roulston, K. (2010). Considering quality in qualitative interviewing. *Qual. Res.* 10, 199–228. doi: 10.1177/1468794109356739

Ryan, A., Prieto-Rodriguez, E., Miller, A., and Gore, J. (2024). What can implementation science tell us about scaling interventions in school settings? A scoping review. *Educ. Res. Rev.* 44:100620. doi: 10.1016/j.edurev.2024.100620

Schildkamp, K., Handelzalts, A., Poortman, C. L., Leusink, H., Meerdink, M., Smit, M., et al. (2018). The data team™ procedure: a systematic approach to school improvement. Cham, Switzerland: Springer International Publishing.

Schildkamp, K., Poortman, C. L., Ebbeler, J., and Pieters, J. M. (2019). How school leaders can build effective data teams: five building blocks for a new wave of data-informed decision making. *J. Educ. Chang.* 20, 283–325. doi: 10.1007/s10833-019-09345-3

Sun, J., Przybylski, R., and Johnson, B. J. (2016). A review of research on teachers' use of student data: from the perspective of school leadership. *Educ. Assess. Eval. Account.* 28, 5–33. doi: 10.1007/s11092-016-9238-9

Thompson, M., and Wiliam, D. (2007). Tight but loose: Scaling up teacher professional development in diverse contexts. Chicago, US: American Educational Research Association (AERA).

van den Boom-Muilenburg, S. N., Poortman, C. L., Schildkamp, K., de Vries, S., and van Veen, K. (2023). Sustaining data use professional learning communities in schools: the role of leadership practices. *Stud. Educ. Eval.* 78:101273. doi: 10.1016/j. stueduc.2023.101273

van der Lans, R. M., van de Grift, W. J. C. M., and van Veen, K. (2015). Developing a teacher evaluation instrument to provide formative feedback using student ratings of teaching acts. *Educ. Meas. Issues Pract.* 34, 18–27. doi: 10.1111/emip.12078

van Geel, M., Keuning, T., Visscher, A., and Fox, J.-P. (2017). Changes in educators' data literacy during a data-based decision making intervention. *Teach. Teach. Educ.* 64, 187–198. doi: 10.1016/j.tate.2017.02.015

van Merriënboer, J. J. G., and Kirschner, P. A. (2018). Ten steps to complex learning. *Third* Edn. London, UK: Routledge Taylor & Francis Group.

Visscher, A. J., Dmoshinskaia, N., and Pellegrini, M. (2023). Effects of teacher professional development on student academic achievement: A meta-analysis of 143 studies. Thessaloniki, Greece: EARLI.

Wiliam, D., and Leahy, S. (2014). Sustaining Formative Assessment with Teacher Learning Communities. *Learning Sciences Dylan Wiliam Center.* https://www. commoncorediva.com/wp-content/uploads/2015/08/sustaining-tlcs-20140829.pdf

Yoon, I., and Kim, M. (2023). Determinants of teachers' positive perception on their professional development experience: an application of LASSO-based machine learning approach. *Prof. Dev. Educ.*, 1–15. doi: 10.1080/19415257.2023.2264296

Zhao, Y., Pugh, K., Sheldon, S., and Byers, J. L. (2002). Conditions for classroom technology innovations. *Teach. Coll. Rec.* 104, 482–515. doi: 10.1111/1467-9620.00170