



## OPEN ACCESS

## EDITED BY

Carola Manolino,  
Università della Valle d'Aosta, Italy

## REVIEWED BY

Viviane Hummes,  
University of Barcelona, Spain  
Micaela Martins,  
Universidade de Lisboa, Portugal

## \*CORRESPONDENCE

Daniela Caballero  
✉ caballed@mcmaster.ca

RECEIVED 31 October 2023

ACCEPTED 21 March 2024

PUBLISHED 10 April 2024

## CITATION

Caballero D, Mella-Defranchi E and  
Araya R (2024) The impact of observers'  
beliefs on the perceived contribution of a  
Research Lesson.  
*Front. Educ.* 9:1331293.  
doi: 10.3389/educ.2024.1331293

## COPYRIGHT

© 2024 Caballero, Mella-Defranchi and  
Araya. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# The impact of observers' beliefs on the perceived contribution of a Research Lesson

Daniela Caballero<sup>1,2\*</sup>, Eliacim Mella-Defranchi<sup>2</sup> and  
Roberto Araya<sup>2</sup>

<sup>1</sup>Experiential Learning Office—Engineering Faculty, McMaster University, Hamilton, ON, Canada, <sup>2</sup>CIAE, Institute of Education, University of Chile, Santiago, Chile

Recently, Lesson Study (LS) has gained popularity in countries worldwide because of its potentially positive effects on teachers' practices (e.g., reflection, cooperation, and pre-service development) and students' learning. However, despite global interest in LS's implementation, an important gap exists between Japanese LS and its implementation in other countries, which may be due to several reasons, such as differences in culture and educational systems or teachers' beliefs. In this study, we examined the effect of teachers' beliefs on their evaluation of LS in Chile. We administered a questionnaire to 94 teachers who participated in the Research Lesson (RL) as observers. The questionnaire assessed teachers' beliefs and RL's contributions to knowledge of the subject matter, instructional strategies, monitoring skills, lesson planning, and student understanding. Using a stepwise logistic regression, after controlling for sex and occupation, we found that the observers' beliefs influenced their perceptions of RL's contributions to monitoring, assessment, and instruction. Participants with student-focused beliefs were more likely to find that RL contributed to their monitoring skills and ability to assess students' understanding of content. The results regarding instructional strategies were mixed. Our findings can help devise strategies to increase the effectiveness of LS implementation. For example, by designing two types of LS, one adapted to teachers with student-focused teaching beliefs and the other to teacher-focused teaching beliefs. To the best of our knowledge, this dual strategy is not part of LS implementation, at least in Chile. LS teams could easily explore this dual strategy, which could improve teachers' professional development.

## KEYWORDS

lesson study, research lesson, peer observation, teacher professional development, teachers' beliefs

## 1 Introduction

Lesson Study (LS) began in 1880 in Japan with the goal of reproducing the best practices in teaching (Isoda, 2015) and is a collaborative teaching improvement process designed to build strong and productive communities of teachers who share and learn from each other. Many countries have introduced this methodology (Fujii, 2014). Evidence indicates that LS accelerates the production of effective lessons and, in particular, helps develop open-ended approaches with lessons that aim to develop higher-order skills (Inprasitha, 2015). Additionally, Vermunt et al. (2019) concluded that the LS approach to pedagogical development fosters meaning-oriented teacher learning, which can be explained by LS's strong focus on analyzing

and understanding pupils' learning. In particular, teachers in schools implementing LS have reported more meaning- and application-oriented learning than teachers from schools without LS experience (Vermunt et al., 2019).

Enhancing the dissemination and awareness of LS's benefits is important, as Chokshi and Fernandez (2004) found that schools allocate time for LS initiatives once teachers recognize the potential of LS. The potential benefits of LS include various pedagogical aspects and students' learning outcomes. Researchers have studied numerous different effects of LS, including teachers' increased self-efficacy (Mintzes et al., 2013; Schipper et al., 2018; Vermunt et al., 2019), teachers' content knowledge (Fernandez and Robinson, 2006; Lewis and Perry, 2014; Juhler, 2016), teaching practices and teachers' skills (Fernandez and Robinson, 2006; Fernandez, 2010; Vermunt et al., 2019), teachers' beliefs (Fernandez and Robinson, 2006; Lewis and Perry, 2015; Yakar and Turgut, 2017) and pre-service teachers' preparation (Marble, 2006, 2007; Suh and Fulginiti, 2012). In Chile, there used to be funds to implement this methodology, but due to a significant funding reduction, only a few institutions continue to work with the methodology (Estrella et al., 2018).

However, a notable gap in the implementation of LS exists between Japan and other countries owing to divergent cultures and educational systems. Countries implementing LS outside of Japan have been found to not fully capture its key components (Fujii, 2014). For instance, in the US, most LS teams do not share their learning with their peers, losing the opportunity to share and discuss their experiences with their educational community (Whitney, 2020). Similarly, in England, Seleznyov (2020) reported that LS practices were diluted over time and identified various impediments, such as a lack of time and excessive teacher workload, a focus on demonstrating short-term impact, and a lack of teacher research skills. Additionally, Godfrey et al. (2019) found that engagement and reported learning decreased when teachers in England had to commit their own time to the LS process.

## 1.1 Research lessons

Research lessons (RLs) are critical to LS, which involve teachers coming together to observe and discuss lessons. Some benefits of RLs include opportunities for teachers to observe experienced teachers as a form of modeling their own learning experiences. RLs provide valuable examples of meaningful curricula and standard discussions, encouraging teachers to align their teaching with policy- and research-based recommendations (Lumpe et al., 2014). Furthermore, Dudley et al. (2019) studied the implementation of RLs in mathematics within a school system in England and concluded that by focusing on student learning and recursive cycles of LS, teachers could develop the curriculum and raise standards while supporting the creation of the necessary conditions for learning. This transition to a student-centered education was also part of a two-year project in which teachers focused on students' mathematical reasoning and sense-making rather than on other teacher actions or teaching (Wessels, 2018).

Another potential benefit of RLs is teacher dialogue, consisting of more meaningful reflections. During the implementation of RLs in Ireland, teachers engaged in meaningful dialogues about pedagogy and student learning, fostering deeper levels of reflection on their understanding and practice knowledge (McSweeney and Gardner,

2018). Similarly, a study in Singapore concluded that after four RLs and discourse analyses, a balance was achieved in mathematical representation while escalating the construction of meaning-making (Fang et al., 2019).

Nonetheless, the field's understanding of RLs remains limited, especially regarding methods of sharing work with others. Whitney (2020) examined cases of LS teams in the US and found that most LS teams did not distribute their learning to the field. Individuals external to the team observed the RL and engaged in subsequent post-lesson discussions.

This study has two objectives. First, we gathered the evaluations of observers outside the LS team when they attended an RL in Chile. Second, we examined how observers' beliefs affected their evaluation of RL. Thus, our research question (RQ) was, "To what extent do teachers' beliefs affect the perceived contributions of an RL in the context of an LS process?"

## 1.2 The Chilean context and teachers' predisposition to observation

Our first study aim was to gather the evaluations of observers attending an RL in Chile. Participation as an observer in an RL may share similarities with teachers' own evaluation processes, particularly in the context of observing pedagogical practices and sharing best practices with peers. Potential apprehension could exist among educators because of the history of teacher performance evaluations in Chile. This subsection briefly describes the historical perspective of the Chilean teacher evaluation system.

Chile's experience with teacher performance evaluation is distinctive, particularly considering its past political developments. During Chile's dictatorship (1973–1989), the teaching profession experienced a significant weakening, resulting in a substantial unemployment rate among teachers. Following the return to democracy in the 1990s, concerted efforts were made by the national government, in collaboration with teachers' unions, to rectify the adverse consequences of past politics (Avalos and Assael, 2006; Assael and Cornejo, 2018).

One contentious aspect of these remedial efforts was the introduction of an annual grading system exclusively within public schools (municipality-dependent), which has frequently been perceived as punitive and arbitrary by educators (Avalos and Assael, 2006; Taut et al., 2011; Assael and Cornejo, 2018). Despite some modifications over the years, the prevailing sentiment among teachers was that these evaluations remained unjust and were penalizing. Moreover, these evaluations impose additional work burdens on teachers, exacerbated by the prevailing practice of unpaid overtime among educators in Chile. This evaluation is referred to as the "Teaching Evaluation" and is mandated by law.

In 2016, a significant change occurred when a second method of evaluating teachers was introduced, referred to as the "Teaching Career Evaluation." This evaluation encompassed all teachers associated with schools that received state funding. Consequently, educators from state-funded educational institutions were included in the evaluation process (Martínez, 2022; Colegio de Profesores, 2023). Importantly, the evaluation was not designed to supplement the existing evaluation system but introduced an additional evaluation more closely associated with the years of a teacher's career.

Consequently, teachers within the public system were subject to evaluations from two distinct systems, increasing their workload and the demands placed on them.

Both evaluations were designed to reinforce the teaching profession and contribute to improving the quality of education. The “Teaching Evaluation” comprises four key components: (1) a self-assessment rubric; (2) an interview conducted by a peer evaluator; (3) a third-party reference report; and (4) a portfolio showcasing an in-video pedagogical performance. Each component includes a rating assigned to the teacher, classifying teachers into different performance levels: outstanding, competent, basic, or unsatisfactory. The “Teaching Career Evaluation” includes (1) an evaluation of specific and pedagogical knowledge and (2) the same portfolio used in the “Teaching Evaluation.” According to [Alvarado \(2012\)](#), the teachers’ performance and their students’ academic performance on math- and language-standardized national tests have been positively correlated.

Teachers who received favorable evaluations progress to higher tiers within the Teaching Career framework. The upper tiers often entail augmented salary packages and improved working conditions ([Manzi et al., 2011](#)). Engagement in these evaluative assessments may be a financial necessity for some teachers due to the comparatively modest salaries that educators in Chile receive ([Assaél and Cornejo, 2018](#); [Avalos-Bevan, 2018](#)), rather than an opportunity to foster reflection between peers regarding their pedagogical practice and work ([Assaél and Cornejo, 2018](#); [Avalos-Bevan, 2018](#)).

Consequently, public school teachers in Chile have experience with being observed and engaging in discussions with others regarding their teaching practices. However, these experiences may be biased toward an emphasis on assessment and evaluation, which, in turn, translates into an economic impact.

## 2 Methods

### 2.1 Participants

The participants in this study were observers of at least one RL hosted by different LS teams. The survey and methodology were approved by our ethics committee. An open invitation to complete the survey was sent to the participants of three different RLs sponsored by the Center for Advanced Research in Education (CIAE). Participants answered the survey between March and May 2020. All RLs were Science, Technology, Engineering and Math (STEM)-related and planned for primary education, with an open invitation to elementary and secondary school teachers, academics, researchers, educators, future teachers, school administrators, and educational authorities. The teachers who taught the lesson were experienced teachers. Two were held in Santiago, Chile, in 2017 and 2018, and the third in Valparaíso, Chile, in 2019. The last RL was conducted within the context of the XI Regional Conference on Lesson Studies.

A total of 97 participants answered the survey, and 92 agreed to participate, resulting in a response rate of 94.84%. Of the 92 participants, 62 were women (67.39%), and 30 were men (32.61%). Regarding the participants’ occupations, 70 were in-service teachers (76.08%), five were student teachers (5.43%), and 17 (18.47%) had other roles related to education, such as academics, members of a school board, and principals.

Regarding the participants’ observation experience, the information available is limited. As explained in the previous section, teachers in Chile have experience being observed; however, we do not have evidence of the level of training that the participants received as observers before attending the RL.

### 2.2 Instruments

A questionnaire was developed to assess participants’ perceptions of the RLs contributions to their skills and beliefs.

#### 2.2.1 RL contribution to pedagogical practices, overall LS evaluation, and future participation

Participants answered five Likert-scale questions regarding the contribution level of RL to their (1) knowledge of the subject matter, (2) instructional strategies, (3) skills to monitor the level of understanding of the students, (4) lesson planning skills, and (5) abilities to assess the students’ understanding of the content. Participants evaluated the contribution of RL by choosing one of five options (from 1 = *not at all* to 5 = *extremely relevant*). This type of question, which includes scale with neutral option, has been used in previous studies to evaluate the impact of LS on teachers’ professional development ([Alamri, 2020](#)), participants’ perception of changes in intercultural competence for lesson study ([Sakai et al., 2022](#)).

Participants also answered a Likert-scale question about their overall evaluation of RL that was rated from *very poor* (1) to *very good* (5), as well as a question regarding whether they would participate again in an RL session, which was rated as *yes*, *maybe*, or *no*.

#### 2.2.2 Teachers’ beliefs about learning

The Teacher Beliefs Interview (TBI), a semi-structured interview protocol, was developed to describe and map various epistemological beliefs held by science teachers ([Luft and Roehrig, 2007](#)). Based on the TBI, teachers’ beliefs range from teacher-focused to student-focused. The TBI includes seven items assessing their knowledge and learning beliefs. Depending on the teachers’ responses, their beliefs were coded into five categories: (1) *traditional*, with a focus on information, transmission, structure, or sources; (2) *instructive*, with a focus on providing experiences, teacher focus, or teacher decisions; (3) *transitional*, with a focus on teacher/student relationships, subjective decisions, or affective responses; (4) *responsive*, with a focus on collaboration, feedback, or knowledge development; and (5) *reform-based*, focusing on mediating student knowledge or interactions. The traditional and instructive categories are teacher-focused, whereas the responsive and reform-based categories are student-focused. This instrument has been employed in prior studies about LS and teachers’ beliefs. In the study by [Tupper \(2022\)](#), the TBI was utilized alongside other instruments to assess the relationship between teachers’ beliefs, self-efficacy, and professional noticing with multiple lesson study experiences. Similarly, [Fortney \(2009\)](#) incorporated the TBI, among other instruments, to investigate how teachers reconcile student-centered instruction with their preexisting traditional beliefs about teaching. Finally, [Yakar and Turgut \(2017\)](#) also used the TBI and found that through lesson study preservice teachers changed their beliefs toward more student-centered.

Of the seven questions on TBI, four were related to learning beliefs. The other three items were related to knowledge beliefs. Three questions regarding learning were included in the questionnaire (the

fourth question was excluded because it asked about science learning). Rather than adhering to the open-ended format, we modified the approach by incorporating the three questions that asked the participants to choose between two real-life strategies using concrete examples. These three questions were related to the participants' beliefs about learning. This adaptation led to a structured multiple-choice format, providing participants with clear examples while maintaining the integrity of the TBI's underlying principles. Table 1 shows the questions, examples, and mapping suggested by Luft and Roehrig (2007) for assessing the participants' beliefs.

The Cronbach's alpha was calculated to measure the reliability of the instrument. The results showed that the instrument had an estimated  $\alpha$  equal to 0.68.

## 2.3 Analysis

To answer the research question, a descriptive analysis and stepwise logistic regression model were performed for each dependent variable. Five independent models were tested, one for each dependent variable.

### 2.3.1 Dependent variable—RL contribution

The five dependent variables were the RL contributions to knowledge, instruction, monitoring, planning, and assessment. Each variable was binary, with value 1 being assigned when the participant evaluated the RL's level of contribution as either *very* or *extremely* and a value of 0 being assigned for all other ratings.

TABLE 1 Questions in the survey and the corresponding response mapping based on TBI.

Question—Option	Examples	TBI's belief mapping
<i>Q: Which of the following options represents the strategy you most frequently use to maximize student learning in your classroom? (Choose 2 options)</i>		
By providing information in a structured environment	I carefully plan my lessons, using PowerPoint presentations, making students to face toward me, or following the textbook and study guides	Teacher-Focused, Traditional
By monitoring student actions or behavior during instruction	I look at the students' responses, and I watch the students closely as they complete their tasks.	Teacher-Focused, Instructive
By creating a classroom environment that involves the students	I use different types of activities; I build a good relationship with students outside of class; I encourage them to do their own thinking.	Transitional
By designing a classroom environment to enable students to interact with each other	I use small group activities where students hypothesize, question and share; I give students the opportunities to defend their ideas.	Student-Focused, Responsive
By considering student responses in order to design the classroom environment	I know that not all students learn the same; I organize different scenarios taking into account the students; I allow students to choose their own ways to learn.	Student-Focused, Reform-based
<i>Q: Which of the following options represents the strategy you most frequently use to know when your students understand? (Choose 2 options)</i>		
When they receive the information	When they have heard several times when we have covered in class, or when I covered the content in different ways.	Teacher-Focused, Traditional
When they can reiterate or demonstrate what has been presented	When they can use their own words to explain a concept or when they can repeat the answer on a written test, and the answer is correct.	Teacher-Focused, Instructive
When they give an explanation or response that is related to the presented information	When they talk about the content in different ways; when they can ask a basic question; when they are excited or when they talk about the lesson outside of class.	Transitional
When they can utilize the presented knowledge	When they clearly defend their ideas using evidence and examples they have experienced.	Student-Focused, Responsive
When they can apply knowledge in a novel setting, or construct something novel that is related to the knowledge	When they can come up with questions or comments about the topic in situations that have not been experienced in class or when they use other subjects to solve my subject problems.	Student-Focused, Reform-based
<i>Q: Which of the following options represents the strategy you most frequently use to know when learning is occurring in your classroom? (Choose 2 options)</i>		
By the students' actions during instruction. Emphasizing students' order and attention	Students are paying attention or they are quiet.	Teacher-Focused, Traditional
By the answers students give to my questions, especially when the answers are correct	I ask questions from time to time to see if they are understanding, they follow instructions to exercises or I look at their notebooks.	Teacher-Focused, Instructive
By my own conclusions; based on my experience	The classroom gets noisy; I see the students' faces; students are engaged and not passive, or through students' reflections.	Transitional
By the way students interact with their peers or with me	Students interact to solve problems; when the students are helping each other or when they defend their ideas through evidence and examples	Student-Focused, Responsive
By the significant interactions with one another and/or with me	Students can formulate thoughtful questions about the content; students seek other student's opinions or when students challenge one another.	Student-Focused, Reform-based

### 2.3.2 Independent variables

Participants' beliefs were considered independent variables. Thus, there were three independent variables, each corresponding to one of the three questions regarding how participants (1) maximized learning in their classrooms, coded as *max\_learning*; (2) knew when their students understood, coded as *understanding*; and (3) knew when learning occurred in their classrooms, coded as *learning*.

As with the dependent variables, each independent variable was binary, with the variable being scored as Student Focused (SF) when participants chose strategies that were mapped as student-focused and/or transitional on the TBI (Luft and Roehrig, 2007) and Not Student Focused (NSF) in any other case. In other words, if a participant chose a responsive (R) and/or reform-based (RB) strategy alongside a transitional strategy, the participant was coded as SF. In all other cases, the participant was coded as NSF. Table 2 shows the scenarios in which participants were classified as having SF. The participants were classified as having NSF for all other combinations of selected strategies.

### 2.3.3 Other explanatory variables

Two other independent variables are also considered. The first was sex (woman or man), while the second was the participant's occupation (*in-service teacher, student teacher, or other*).

## 2.4 Models

The stepwise logistic regression selection considered a full model with all the variables involved (*max\_learning*, *understanding*, *learning*, *sex*, and *occupation*) and a base model with only sex as an independent variable (Figure 1).

The notation for the full model is:

$$\text{contribution} \leftarrow \text{max\_learning} + \text{understanding} + \text{learning} + \text{sex} + \text{occupation}$$

For the base model, it is:

$$\text{contribution} \leftarrow \text{sex}$$

Model selection used the R function *stepAIC* (Zhang, 2016), which iteratively adds and removes predictors to the base model. Thus, the best model was selected based on its lower AIC value (Akaike, 1974). This process was performed independently for each dependent variable: the level of contribution of RL to (1) knowledge, (2) instruction, (3) monitoring, (4) planning, and (5) assessment. Finally, when a variable was significant, the odds ratio (OR) was reported as a measure of the association between exposure and outcome (Szumilas, 2010).

## 3 Results

### 3.1 LS evaluation and future participation

Overall, the LS received a positive evaluation: 76.09% of teachers evaluated RL as either *good* or *very good* ( $N=70$ ). When asked if they

would participate again in RL, 86.95% reported that they would. The distribution of the answers is shown in Figure 2.

### 3.2 Participants' beliefs

Table 3 summarizes the distribution of teachers' beliefs after categorizing them as either SF or NSF.

### 3.3 Reported RL impact on pedagogical aspects

Table 4 shows the level of contribution of RL to pedagogical practices as reported by the participants. The aspects with the highest levels of impact were planning and assessment, while the aspects with the lowest impact on teachers' pedagogical practices were monitoring and instruction.

### 3.4 Impact of participants' beliefs on RL contribution

Table 5 summarizes the distribution of the level of contribution by belief for the three independent variables. Results in Table 5 suggest that participants' beliefs might influence the perception of the RL contribution. The data imply a consistent trend where observers with SF beliefs tend to value the RL contribution more than those with NSF beliefs. Moreover, the differences are more or less pronounced depending on how those underlying beliefs are measured (i.e., max learning, learning, or understanding).

Differences in the RL's contribution based on the observers' beliefs were observed. Regarding monitoring skills, 72.34% of those with SF beliefs (measured with the variable *max\_learning*) valued the contribution of RL as *very* or *extremely*, while less than half of observers with NSF beliefs shared the same valuation.

In terms of assessing skills, 76.60% of participants with SF beliefs (measured with the variable *max\_learning*) rated the contribution as *very* or *extremely*, compared to 55.56% of those who were NSF.

The results for the instructional skills varied. When considering the observers' beliefs measured by the variable *max\_learning*, 74.47% of the observers with SF beliefs valued the contribution of RL positively, in contrast to 48.89% of those with NSF. In contrast, when measuring beliefs by the variable *learning*, 71.05% of observers with NSF beliefs valued the contribution as positive, whereas 55.56% of observers with student-focused SF beliefs shared the same view.

#### 3.4.1 Knowledge

The summary of the logistic regression model is shown in Table 6.

The selected model was the base model ( $AIC=121.87$ ), with sex as the only independent variable. However, sex was not a significant predictor of the probability of evaluating contribution to RL ( $p=0.14$ ; Table 6). Thus, the participants' beliefs did not add information when explaining the contribution of RL to their knowledge skills.

#### 3.4.2 Monitoring

The summary of the logistic regression model is shown in Table 7.

TABLE 2 Categorization of participants' answers based on number of strategies selected.

Number of strategies selected	Traditional (teacher-focused)	Instructive (teacher-focused)	Transitional	Responsive (student-focused)	Reform-based (student-focused)
2				X	X
2			X		X
2			X	X	
1				X	
1					X

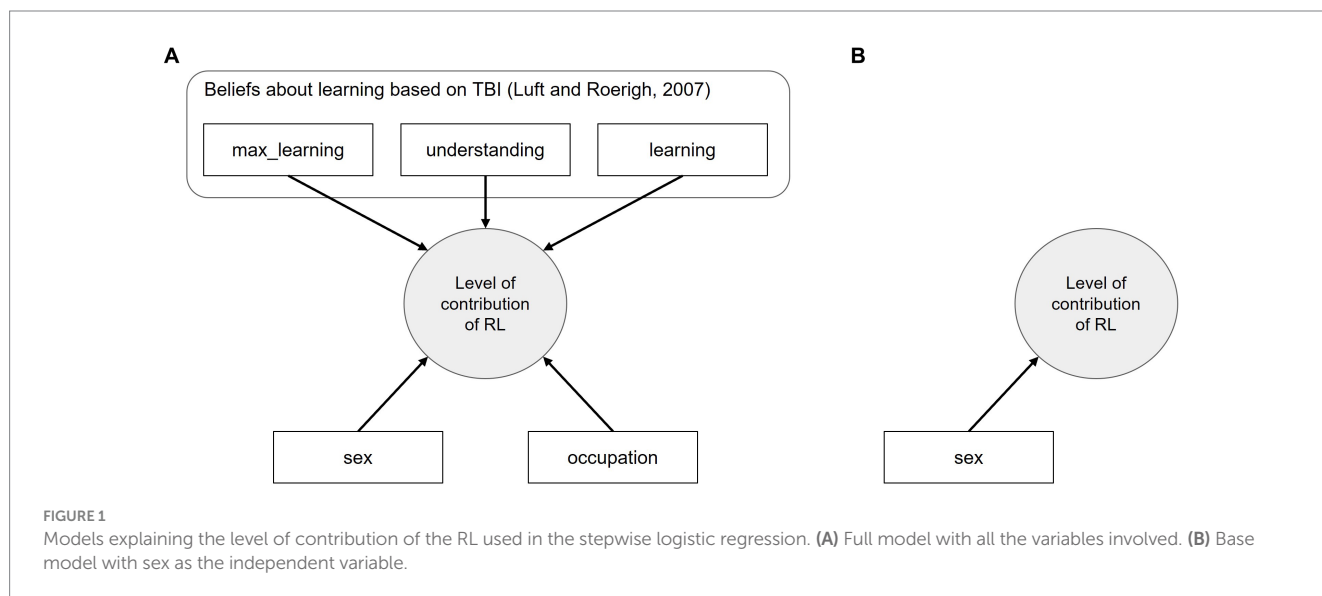


FIGURE 1 Models explaining the level of contribution of the RL used in the stepwise logistic regression. (A) Full model with all the variables involved. (B) Base model with sex as the independent variable.

The model that best explained the contribution of RL to monitoring skills was the independent variable `max_learning` after controlling for sex and occupation (AIC = 121.44). The variable `max_learning` had a significant positive estimated value ( $p = 0.035$ ), whereas the other variables (i.e., sex and occupation) were not significant. In other words, when holding the variables of sex and occupation constant, participants with SF beliefs (measured through the variable `max_learning`) were more likely to rate the contribution of RL on their monitoring skills as *very* or *extremely* (OR = 1.99, 95%CI [1.06, 3.85]).

### 3.4.3 Planning

The selected model considered `max_learning` and `understanding` as independent variables, controlling for sex (AIC = 108.29). Table 8 summarizes the corresponding logistic regression results.

In the selected model, the three independent variables had nonsignificant estimated values. The variables `max_learning` and `sex` had positive values, while `understanding` had an estimated negative value.

### 3.4.4 Assessing

The model that best explains the contribution to assessment is the independent variable `max_learning`, after controlling for sex and occupation (AIC = 117.66). Table 9 summarizes the corresponding logistic regression results.

In the selected model, `max_learning` had a positive and significant estimated value ( $p = 0.039$ ). Thus, when holding the variables of sex and occupation constant, observers with SF beliefs, measured through

the variable `max_learning`, were more likely to value the contribution on assessing as *very* or *extremely*, in contrast to the observers with NSF beliefs (OR = 2.01, 95%CI [1.05, 3.98]).

### 3.4.5 Instruction

The model that best explains the contribution to instruction had the independent variables `max_learning` and `learning` controlled for by sex (AIC = 114.68). Table 10 summarizes the corresponding logistic regression results.

In this model, the variable `max_learning` had a significant, positive estimated value ( $p = 0.002$ ), whereas the variable `learning` had a significant, negative estimated value ( $p = 0.007$ ). Thus, when holding the variable of sex constant, observers with beliefs which were SF, measured through the variable `max_learning`, were more likely to value the contribution on assessing as *very* or *extremely*, in contrast to the observers with NSF beliefs. Further when controlling for the variable of sex, observers with beliefs that were NSF, measured through the variable `learning`, were more likely to value the contribution on instruction as *very* or *extremely*.

## 4 Discussion and conclusion

This study has two aims. First, we aimed to collect evaluations of observers outside the LS team when attending the LS Research Lesson stage. Second, we examined how the observers' beliefs affected their evaluations of RL. Thus, the research question (RQ) was, "To what

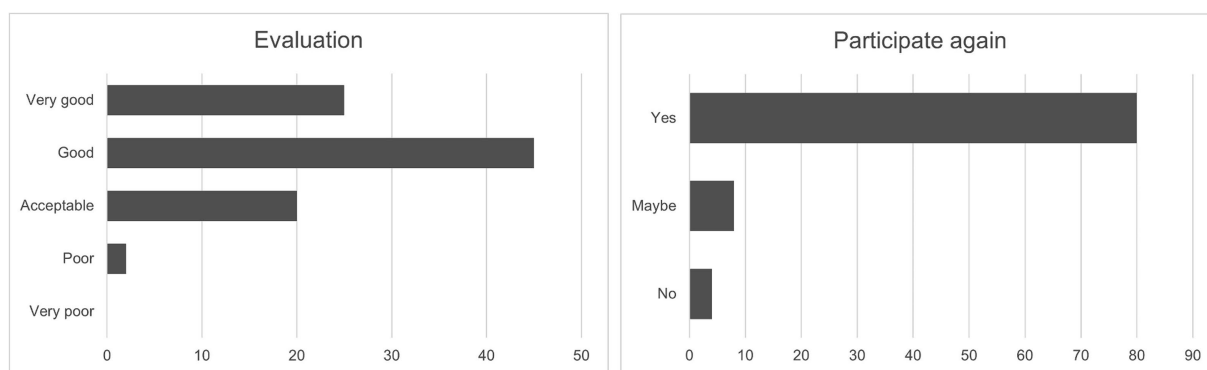


FIGURE 2 Number of answers of the participants' overall evaluation of RL and future participation.

TABLE 3 Distribution of the observers' beliefs for each independent variable, based on the TBI categories.

Variable label	Student-focused (SF)		Not only student-focused (NSF)	
	<i>n</i>	%	<i>n</i>	%
max_learning	45	48.91	47	51.09
learning	38	41.30	54	58.70
understanding	29	31.52	63	68.48

extent do teachers' beliefs affect the perceived contribution of a Research Lesson in the context of a lesson-study process?"

Evidence indicates that LS teams do not share their learning with their peers, thereby losing opportunities to share and discuss their experiences with the educational community. An RL provides teachers new opportunities to observe experienced teachers to model their own learning experiences and discuss curricula and standards, helping persuade teachers to align their teaching with policy and evidence-based recommendations. However, little is known about the RS stage, particularly regarding how to share work with others. Whitney (2020) studied cases of LS study teams in the US and found that most LS teams did not distribute their learning to the field, and instances of people outside the team watching the research lesson and participating in post-lesson discussions existed.

Nonetheless, a relationship has been observed between beliefs and classroom practices (Skott, 2001; Stipek et al., 2001; Schoenfeld, 2011). Moreover, these beliefs could affect teachers' changes in the context of professional development programs (Maass, 2011). de Vries et al. (2013) found a relationship between teachers' participation in Continuous Professional Development programs and their beliefs.

This paper aims to reduce the existing gap between the implementation of LS in Japan and in other countries. When considering the implementation of LS in Chile, it is critical to look at teachers' predisposition to observation. Historically, teachers of public schools have experience in being observed; however, class observation has been associated with assessment and evaluation. Although this research does not specifically inquire about teachers' beliefs regarding observation, it is an important element to consider in the teaching culture of Chile. For example, by conducting RL observation, Chilean teachers could focus more on evaluating aspects that are important in

TABLE 4 Evaluation of the contribution of RL to different aspects.

Aspect	Not at all—slightly—somewhat		Very—extremely	
	<i>n</i>	Percentage	<i>n</i>	Percentage
Knowledge	33	35.87	59	64.13
Monitoring	36	39.13	56	60.87
Planning	24	26.09	68	73.91
Assessing	31	33.70	61	66.30
Instruction	35	38.04	57	61.96

their own teaching evaluation, such as class planning and student assessment.

The results show that, in the Chilean context, when explaining the perceived contribution of the RL, observers' beliefs improved the models' performance, contributing to instruction, monitoring, planning, and evaluation. The current findings are beneficial, as they provide insights into how different participants value their contributions to RL. Furthermore, LS teams should consider this when issuing open invitations to external individuals to observe RL. Understanding the influence of participants' beliefs on their evaluation of a situation is valuable. In addition, these findings can help develop strategies for more effective LS implementation, such as designing two types of LS, one adapted to teachers with beliefs about teaching focused on students and the other adapted to teaching beliefs focused on teachers. To the best of our knowledge, this dual strategy is currently not a part of LS implementation, at least in Chile. LS teams could easily explore this dual strategy, which could improve teachers' professional development. In terms of practical implications, we believe that by adapting different types of LS, the whole group collaboration within each RL could lead to a richer discussion on aspects of the lesson. It would bring to the forefront those aspects that might not initially appear to contribute significantly. As a result, it would promote reflection among observers based on their own beliefs. Moreover, this dual strategy could help emphasize the benefits of the methodology and strengthen the overarching goals of the LS by highlighting those aspects that are less appreciated within each group.

LS encompasses multiple stages and requires dedication from the LS team. This paper provides initial insights into a single stage of LS based on a limited sample. Nonetheless, future research should focus on the

TABLE 5 Distribution of evaluation of the contribution of the RL based on the category of observers' beliefs.

Contribution		max_learning		learning		understanding	
		NSF N = 45	SF N = 47	NSF N = 38	SF N = 54	NSF N = 29	SF N = 63
Knowledge	Not at all/slightly/ somewhat	16 (35.56%)	17 (36.17%)	16 (42.11%)	17 (31.48%)	9 (31.03%)	24 (38.10%)
	Very/extremely	29 (64.44%)	30 (63.83%)	22 (57.89%)	37 (68.52%)	20 (68.97%)	39 (61.90%)
Monitoring	Not at all/slightly/ somewhat	23 (51.11%)	13 (25.66%)	15 (39.47%)	21 (38.89%)	12 (41.38%)	24 (38.10%)
	Very/extremely	22 (48.89%)	34 (72.34%)	23 (60.52%)	33 (61.11%)	17 (58.62%)	39 (61.90%)
Planning	Not at all/slightly/ somewhat	14 (31.11%)	10 (21.28%)	9 (23.68%)	15 (27.78%)	5 (17.24%)	19 (30.16%)
	Very/extremely	31 (68.89%)	37 (78.72%)	29 (76.32%)	39 (72.22%)	24 (82.76%)	44 (69.84%)
Assessing	Not at all/slightly/ somewhat	20 (44.44%)	11 (23.40%)	13 (34.21%)	18 (33.33%)	10 (34.48%)	21 (33.33%)
	Very/extremely	25 (55.56%)	36 (76.60%)	25 (65.79%)	36 (66.67%)	19 (65.52%)	42 (66.67%)
Instruction	Not at all/slightly/ somewhat	23 (51.11%)	12 (25.53%)	11 (28.95%)	24 (44.44%)	11 (37.93%)	24 (38.10%)
	Very/extremely	22 (48.89%)	35 (74.47%)	27 (71.05%)	30 (55.56%)	18 (62.07%)	39 (61.90%)

TABLE 6 Logistic regression summary for the selected model illustrating the contribution of RL to knowledge.

Coefficients	Estimate	Std. Error	z-value	p-value
Intercept	0.8168	0.2755	2.965	0.00303**
Sex: M	-0.6832	0.4581	-1.491	0.13581

Significance codes at 95% level.  
\*\*\*p < 0.0001; \*\*p < 0.001; \*p < 0.05.

TABLE 7 Logistic regression summary for the selected model showing the contribution of RL to monitoring.

Coefficients	Estimate	Std. error	z-value	p-value
Intercept	17.4122	1689.5725	0.010	0.9918
max_learning:	0.6886	0.3270	2.106	0.0352*
occupation: other	-17.5124	1689.5726	-0.010	0.9917
occupation: in-service teacher	-16.7802	1689.5725	-0.010	0.9921
Sex: M	-0.3208	0.4758	-0.674	0.5002

Significance codes at 95% level.  
\*\*\*p < 0.0001; \*\*p < 0.001; \*p < 0.05.

whole process, considering a more holistic perspective of the LS implementation. Similarly, future research could explore the beliefs of the LS team, rather than solely focusing on the observers of the RL. Additionally, considering other contextual factors such as school culture, previous observation experiences, and administrative resources to participate in Professional Development activities. For instance, it

TABLE 8 Logistic regression summary for the selected model showing RL's contribution to planning.

Coefficients	Estimate	Std. error	z-value	p-value
Intercept	1.0868	0.3270	3.324	0.000887 ***
max_learning	0.5868	0.3640	1.612	0.106943
understanding	-0.6806	0.4187	-1.625	0.104078
Sex: M	0.6051	0.5527	1.095	0.273567

Significance codes at 95% level.  
\*\*\*p < 0.0001; \*\*p < 0.001; \*p < 0.05.

TABLE 9 Logistic regression summary for the selected model showing RL's contribution to assessing.

Coefficients	Estimate	Std. error	z-value	p-value
Intercept	17.2536	1741.5101	0.010	0.992
max_learning	0.6971	0.3377	2.064	0.039*
type: Other	-17.2914	1741.5101	-0.010	0.992
type: in-service Teacher	-16.5373	1741.5101	-0.009	0.992
Sex: M	0.2061	0.4947	0.417	0.677

Significance codes at 95% level.  
\*\*\*p < 0.0001; \*\*p < 0.001; \*p < 0.05.

would greatly enrich the analysis of future research to know the level of expertise participants have in conducting observations. This would allow us to understand more deeply which aspects teachers highlight depending on their experience in observation. Finally, the students'



TABLE 10 Logistic regression summary for the selected model showing RL's contribution to instruction.

Coefficients	Estimate	Std. error	z-value	Pr (> z )
Intercept	0.8704	0.3086	2.820	0.00480 **
max_learning	1.2312	0.3956	3.112	0.00186 **
learning	-1.0998	0.4091	-2.689	0.00717 **
Sex: M	-0.5478	0.4990	-1.098	0.27224

Significance codes at 95% level.  
 \*\*\* $p < 0.0001$ ; \*\* $p < 0.001$ ; \* $p < 0.05$ .

behavior, which was outside of the scope of this work, may be affected by changes in their routine, such as the physical space allowing for a large number of observers or different teachers than those they are used to. On the other hand, we used the TBI (Luft and Roehrig, 2007) in a novel manner, which was developed to be used as an interview. We used the open-ended TBI questions and provided the participants with concrete examples based on the framework developed by Luft and Roehrig (2007), thereby framing the questions in a multiple-choice format instead of an open-ended approach. While this method proved helpful in explaining the differences in the perceived contribution, future research should study alternative instruments for assessing participants' beliefs. In particular, the questionnaire instrument has a Cronbach's alpha value of 0.68, which is close to the threshold of reliability; the reliability of this instrument might impact the measurement properties for this study, and the results should be taken with caution. Furthermore, a qualitative approach could provide insights into how and why those beliefs influence the participants' perception of RL's contributions.

Finally, in the Chilean context, opportunities for collaboration and peer observation that extend beyond evaluation purposes are important. The current study is an example of how people outside the LS team can benefit from different pedagogical aspects of this methodology. Future research should assess changes in teachers' beliefs to determine whether the reported value of RL affects their daily practices.

## 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 Comité de Ética de la Investigación (CEI) de la Facultad de Ciencias Sociales de la

## References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19, 716–723. doi: 10.1109/TAC.1974.1100705
- Alamri, N. M. (2020). The implementation of the lesson study strategy in teaching mathematics: teachers' perspectives. *Educ. Res. Int.* 2020, 1–8. doi: 10.1155/2020/1683758
- Alvarado, M. (2012). "La Evaluación Docente y Sus Instrumentos: Discriminación Del Desempeño Docente y Asociación Con Los Resultados de Los Estudiantes."
- Assaél, J., and Cornejo, R. (2018). "Work regulations and teacher subjectivity in a context of standardization and accountability policies in Chile." In R. Normand, M. Liu, L. Carvalho, D. Oliveira, and L. LeVasseur (eds.), *Education policies and the*

Universidad de Chile <https://facso.uchile.cl/facultad/comites/comite-de-etica-de-la-investigacion>. 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

DC: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing. EM-D: Formal analysis, Writing – original draft, Writing – review & editing, Investigation. RA: Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by ANID Postdoctoral Fellowship Nb. 3180590 and ANID/PIA/Basal Funds for Centers of Excellence FB0003.

## Acknowledgments

Authors gratefully acknowledge the support from ANID Postdoctoral Fellowship Nb. 3180590 and ANID/PIA/Basal Funds for Centers of Excellence FB0003.

## 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.

## 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.

*restructuring of the educational profession: global and comparative perspectives*, Singapore: Springer, 245–257.

Avalos, B., and Assael, J. (2006). Moving from resistance to agreement: the case of the Chilean teacher performance evaluation. *Int. J. Educ. Res.* 45, 254–266. doi: 10.1016/j.ijer.2007.02.004

Avalos-Bevan, B. (2018). Teacher evaluation in Chile: highlights and complexities in 13 years of experience. *Teach. Teach.* 24, 297–311. doi: 10.1080/13540602.2017.1388228

Chokshi, S., and Fernandez, C. (2004). Challenges to importing Japanese lesson study: concerns, misconceptions, and nuances. *Phi Delta Kappan* 85, 520–525. doi: 10.1177/003172170408500710

- Colegio de Profesores. (2023). "Evaluación Docente." Available at: <https://www.colegiodeprofesores.cl/wp-content/uploads/2023/05/EVALUACION-DOCENTE-001.pdf>.
- de Vries, S., Jansen, E. P. W. A., and van de Grift, W. J. C. M. (2013). Profiling teachers' continuing professional development and the relation with their beliefs about learning and teaching. *Teach. Teach. Educ.* 33, 78–89. doi: 10.1016/j.tate.2013.02.006
- Dudley, P., Haiyan, X., Vermunt, J. D., and Lang, J. (2019). Empirical evidence of the impact of lesson study on students' achievement, teachers' professional learning and on institutional and system evolution. *Eur. J. Educ.* 54, 202–217. doi: 10.1111/ejed.12337
- Estrella, S., Mena-Lorca, A., and Olfo, R. (2018). "Lesson study in Chile: a very promising but still uncertain path" in *Mathematics lesson study around the world: theoretical and methodological issues*. eds. M. Quaresma, C. Winslow, S. Clivaz, J. da Ponte, A. Ni Shúilleabháin and A. Takahashi (Cham: Springer), 105–122.
- Fang, Y., Wang, X., and Kim-Eng, C. L. (2019). "Representing instructional improvement in lesson study through principled analysis of research lessons in Singapore: a case of equivalent fractions" in *Theory and practice of lesson study in mathematics: an international perspective*. eds. R. Huang, A. Takahashi, and J. P. da Ponte (Switzerland: Springer), 393–418.
- Fernandez, M. L. (2010). Investigating how and what prospective teachers learn through microteaching lesson study. *Teach. Teach. Educ.* 26, 351–362. doi: 10.1016/j.tate.2009.09.012
- Fernandez, M. L., and Robinson, M. (2006). Prospective teachers' perspectives on microteaching lesson study. *Education* 127:203+.
- Fortney, B. S. (2009). "The impact of Japanese lesson study on preservice teacher belief structures about teaching and learning science" (Order No. 3360299). Available from Pro Quest Dissertations & Theses Global. (305005930).
- Fujii, T. (2014). Implementing Japanese lesson study in foreign countries: misconceptions revealed. *Math. Teach. Educ. Dev.* 16:n1
- Godfrey, D., Seleznyov, S., Anders, J., Wollaston, N., and Barrera-Pedemonte, F. (2019). A developmental evaluation approach to lesson study: exploring the impact of lesson study in London schools. *Prof. Dev. Educ.* 45, 325–340. doi: 10.1080/19415257.2018.1474488
- Inprasitha, M. (2015). "Prospective teacher education in mathematics through lesson study." In *Series on mathematics education, vol. 3*, by M. Inprasitha, M. Isoda, P. Wang-Iverson and B.-H. Yeap, eds., 185–196. Singapore: World Scientific.
- Isoda, M. (2015). "The science of lesson study in the problem solving approach." In *Series on mathematics education, vol. 3*, by M. Inprasitha, M. Isoda, P. Wang-Iverson and B.-H. Yeap, eds., 81–108. Singapore: World Scientific.
- Juhler, M. V. (2016). The use of lesson study combined with content representation in the planning of physics lessons during field practice to develop pedagogical content knowledge. *J. Sci. Teach. Educ.* 27, 533–553. doi: 10.1007/s10972-016-9473-4
- Lewis, C., and Perry, R. (2014). Lesson study with mathematical resources: a sustainable model for locally-led teacher professional learning. *Math. Teach. Educ. Dev.* 16, 22–42.
- Lewis, C. C., and Perry, R. R. (2015). "A randomized trial of lesson study with mathematical resource kits: analysis of impact on teachers' beliefs and learning community" in *Large-scale studies in mathematics education*. eds. J. Middleton, J. Cai and S. Hwang (Cham: Springer), 133–158.
- Luft, J. A., and Roehrig, G. H. (2007). Capturing science teachers' epistemological beliefs: the development of the teacher beliefs interview. *Electron. J. Res. Sci. Math. Educ.* 11, 38–63.
- Lumpe, A., Vaughn, A., Henrikson, R., and Bishop, D. (2014). "Teacher professional development and self-efficacy beliefs" in *The role of science teachers' beliefs in international classrooms* (Brill), 49–63. Rotterdam: Sense Publishers.
- Maass, K. (2011). How can teachers' beliefs affect their professional development? *ZDM* 43, 573–586. doi: 10.1007/s11858-011-0319-4
- Manzi, J., González, R., and Sun, Y. (2011). *La Evaluación Docente En Chile*.
- Marble, S. T. (2006). Learning to teach through lesson study. *Action Teach. Educ.* 28, 86–96. doi: 10.1080/01626620.2006.10463422
- Marble, S. (2007). Inquiring into teaching: lesson study in elementary science methods. *J. Sci. Teach. Educ.* 18, 935–953. doi: 10.1007/s10972-007-9071-6
- Martínez, D. (2022). "Fin a La Doble Evaluación a Docentes. [Master's Thesis, University of Chile]." Available at: <https://repositorio.uchile.cl/handle/2250/194670>.
- McSweeney, K., and Gardner, J. (2018). "Lesson study matters in Ireland." In *Rural Environment. Education. Personality (REEP 2018) 2018* (Vol. 11, pp. 304–313). Latvia University of Life Sciences and Technologies. doi: 10.22616/REEP.2018.037
- Mintzes, J. J., Marcum, B., Messerschmidt-Yates, C., and Mark, A. (2013). Enhancing self-efficacy in elementary science teaching with professional learning communities. *J. Sci. Teach. Educ.* 24, 1201–1218. doi: 10.1007/s10972-012-9320-1
- Sakai, T., Akai, H., Ishizaka, H., Tamura, K., Lee, Y.-J., Choy, B. H., et al. (2022). Changes in qualities and abilities of Japanese teachers through participation in global lesson study on mathematics. *Int. J. Lesson Learn. Stud.* 11, 290–304. doi: 10.1108/IJLLS-04-2022-0058
- Schipper, T., Goei, S. L., De Vries, S., and Van Veen, K. (2018). Developing teachers' self-efficacy and adaptive teaching behaviour through lesson study. *Int. J. Educ. Res.* 88, 109–120. doi: 10.1016/j.ijer.2018.01.011
- Schoenfeld, A. H. (2011). Toward professional development for teachers grounded in a theory of decision making. *ZDM* 43, 457–469. doi: 10.1007/s11858-011-0307-8
- Seleznyov, S. (2020). Lesson study: exploring implementation challenges in England. *Int. J. Lesson Learn. Stud.* 9, 179–192. doi: 10.1108/IJLLS-08-2019-0059
- Skott, J. (2001). The emerging practices of a novice teacher: the roles of his school mathematics images. *J. Math. Teach. Educ.* 4, 3–28. doi: 10.1023/A:1009978831627
- Stipek, D. J., Givvin, K. B., Salmon, J. M., and MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teach. Teach. Educ.* 17, 213–226. doi: 10.1016/S0742-051X(00)00052-4
- Suh, J. M., and Fulginiti, K. (2012). 'Situating the learning' of teaching: implementing lesson study at a professional development school. *School Univ. Partnerships* 5, 24–37.
- Szumilas, M. (2010). Explaining odds ratios. *J. Can. Acad. Child Adolesc. Psychiatry* 19, 227–229.
- Taut, S., Santelices, M. V., Araya, C., and Manzi, J. (2011). Perceived effects and uses of the national teacher evaluation system in Chilean elementary schools. *Stud. Educ. Eval.* 37, 218–229. doi: 10.1016/j.stueduc.2011.08.002
- Tupper, D. (2022). Elementary science lesson study: Relationships among efficacy, beliefs, and professional noticing (Order No. 29993352). Available from Pro Quest Dissertations & Theses Global. (2741309611).
- Vermunt, J. D., Vrikkki, M., Van Halem, N., Warwick, P., and Mercer, N. (2019). The impact of lesson study professional development on the quality of teacher learning. *Teach. Teach. Educ.* 81, 61–73. doi: 10.1016/j.tate.2019.02.009
- Wessels, H. (2018). *Noticing in pre-service teacher education: research lessons as a context for reflection on learners' mathematical reasoning and sense-making*. In: Kaiser, G., Forgasz, H., Graven, M., Kuzniak, A., Simmt, E., Xu, B. (eds) Invited Lectures from the 13th International Congress on Mathematical Education. ICME-13 Monographs. Cham: Springer. 731–748.
- Whitney, S. R. (2020). Are lesson study participants sharing their professional knowledge? *Int. J. Lesson Learn. Stud.* 9, 57–66. doi: 10.1108/IJLLS-11-2018-0090
- Yakar, Z., and Turgut, D. (2017). Effectiveness of lesson study approach on preservice science teachers' beliefs. *Int. Educ. Stud.* 10, 36–43. doi: 10.5539/ies.v10n6p36
- Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. *Ann. Transl. Med.* 4:136. doi: 10.21037/atm.2016.03.35