



OPEN ACCESS

EDITED BY

Brahm Norwich,
University of Exeter, United Kingdom

REVIEWED BY

Anne Mette Karlsen,
Western Norway University of Applied
Sciences, Norway
Adriana Richit,
Federal University of the Southern Frontier,
Brazil

*CORRESPONDENCE

Klara Kager
✉ klara.kager@uni-potsdam.de

RECEIVED 16 October 2023

ACCEPTED 15 March 2024

PUBLISHED 05 April 2024

CITATION

Kager K, Kalinowski E, Jurczok A and
Vock M (2024) A systematic review of
transparency in Lesson Study research: how
do we report on the observation and
reflection stages?
Front. Educ. 9:1322624.
doi: 10.3389/feduc.2024.1322624

COPYRIGHT

© 2024 Kager, Kalinowski, Jurczok and Vock.
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.

A systematic review of transparency in Lesson Study research: how do we report on the observation and reflection stages?

Klara Kager^{1*}, Eva Kalinowski¹, Anne Jurczok² and Miriam Vock¹

¹Department for Empirical Instruction and Intervention Research, University of Potsdam, Potsdam, Germany, ²Senatsverwaltung für Bildung, Jugend und Wissenschaft, Berlin, Germany

Lesson Study is a method of professional development for teachers that has gained traction in recent decades. However, publications routinely fail to describe crucial details of the implementation or to link the mechanisms that facilitate teachers learning in Lesson Study to theory. This makes it difficult to meaningfully synthesize and replicate research findings. Using a protocol based on three dimensions of transparency, this systematic review examines 129 articles on Lesson Study published between 2015 and 2020 to identify how transparent they were in their reporting of how teachers observed and reflected together. The findings indicate a lack of transparency across several dimensions of how the Lesson Study intervention is reported and highlight a current lack of theorization and coherence in the field. To address some of these issues, we propose a framing structure that empirical papers on Lesson Study should give critical attention to in order to ensure relevance and transferability.

KEYWORDS

professional development, Lesson Study, research transparency, observation, reflection, systematic review

1 Introduction

Lesson Study (LS) is a popular approach to the collaborative professional development (PD) of teachers that originated in Japan. Following its transfer to the US (Stigler and Hiebert, 1999), LS has experienced a surge of international interest over the past three decades and is currently practiced in some form in over 40 countries around the globe (Yoshida et al., 2021). Similarly, scientific publications on LS have increased rapidly during this time, particularly with regard to qualitative research (Xu and Pedder, 2014; Gülhan, 2021; Norwich et al., 2021). The PD approach consists of iterative cycles in which a group of teachers follows a series of core stages: study, plan, teach and observe, and reflect (Lewis, 2009). Specifically, teachers identify a problem or question relevant to their practice, they then study the curriculum and other materials about that issue and subsequently plan a lesson or series of lessons that address it. In a next step, one teacher teaches the lesson, while the remaining group members observe the lesson with a focus on student learning behavior. These observations are then analyzed, with teachers collaboratively reflecting on their lesson and negotiating alternative approaches for future teaching (Lewis, 2009).

LS is considered an effective and systematic approach to improving teaching, but the LS implementation itself does not follow standardized methods or unified protocols (Cerbin and Kopp, 2006; Norwich, 2018). Although it is generally acknowledged that LS includes the core stages of studying, planning, teaching and observing, and reflecting, these stages are flexible and can be tailored to various cultural contexts, professional climates, and the specific needs of individual schools or systems (Goei et al., 2021). This has led to various conceptualizations of LS that follow specific procedures, such as Collaborative Lesson Research in the United States (Takahashi and McDougal, 2016) or Lesson Study for a Learning Community in Japan (Saito et al., 2015). Other examples of adaptations relate to the focus of classroom observations (Dudley, 2013), the refinement and reteaching of a lesson (Quaresma and Da Ponte, 2021), or the use of specific protocols to structure teachers' observations and their reflection process (e.g., Færøyvik Karlsen, 2019; Kager et al., 2022). This adaptability of LS is definitely a strength: it allows educators and researchers from all over the world to explore questions in the classroom in a manner that is meaningful to them. On the other hand, however, this complexity also makes it challenging to investigate LS and its impact on teacher learning in a reliable way.

This double-sidedness is reflected in the type of research the field has produced. Randomized controlled trials would arguably provide the most reliable evidence on the effectiveness of LS (Lewis and Perry, 2017). While at least three randomized controlled trials of LS have been conducted (Lewis and Perry, 2015; Murphy et al., 2017; Wake et al., 2023), such experimental designs are difficult to conduct and tend to struggle with the issue of implementation fidelity (Seleznov, 2019). Furthermore, given the complexity and adaptability of LS, randomized controlled trials may not generate the kind of knowledge that would be crucial for others to replicate LS successfully in their own setting (Bryk, 2015). For these reasons, the majority of LS research is small-scale, qualitative, and highly contextualized (Hadfield and Jopling, 2016; Lewis and Perry, 2017), leading several systematic reviews to conclude that, to date, we still lack robust scientific evidence on the effectiveness of LS (Cheung and Wong, 2014; Seleznov, 2019; Willems and Van den Bossche, 2019).

Nevertheless, qualitative and small-scale research also holds significant value as it can contribute to a deeper understanding of the factors influencing the success of LS in specific contexts. Due to the flexible character of LS, it is reasonable to assume that researchers and practitioners have identified various useful procedures and features to enhance the effectiveness of LS adaptations within their unique settings. A systematic review by Larssen et al. (2018) demonstrates, however, that this presumed knowledge often remains tacit and underreported in publications. Examining 34 articles on LS in initial teacher education, Larssen et al. (2018) found that the majority of studies treated details about the LS implementations with a "taken-for-granted understanding" (17) that required readers to infer how a specific step, such as teachers' observations in the classroom, was conducted. These findings suggest a lost opportunity for knowledge generation within the qualitative research in the LS field.

Another avenue to better understand the link between LS features and their impact on outcomes in qualitative research is the theorization of the processes that facilitate teacher learning. LS includes several features that have been recognized as key in PD, such as the opportunities for teachers to collaborate, observe student learning, and engage in critical reflection (Darling-Hammond et al.,

2017). Færøyvik Karlsen (2019) showed that the observation of student learning in LS and their interpretation in the post-lesson discussion act as pivotal processes that can generate learning moments. Similarly, studies by Warwick et al. (2016) and Bae et al. (2016) indicate that the collaborative talk about student observations in the post-lesson reflection can promote teachers to reevaluate their beliefs and attitudes about teaching and learning. This aligns with the broader discourse on teacher learning, which has determined observation and reflection as key processes that facilitate the growth of professional competences (e.g., Schön, 1995; Van Es and Sherin, 2002; Korthagen, 2016). If teachers lack the knowledge of what and how to observe, or the ability to reflect on their observations critically, they will likely engage in superficial conversations with little opportunity for learning (Mynott, 2019; Kager et al., 2022).

Research by Vrikki et al. (2017) and Calleja and Formosa (2020) indicates that the theoretical understanding of the concepts of observation and reflection can influence their practical implementation by researchers and educators. For example, in the study by Vrikki et al. (2017), the theoretical background not only informed the LS procedures, such as the focus of teachers' observations on specific case students, but also the researchers' analytical approach to the data. Grounding observation and reflection processes in a strong theoretical framework can therefore inform both the design of LS protocols and the interpretation of research findings.

Stigler and Hiebert (2016), while describing LS as undertheorized, note that theoretical approaches to explaining teacher learning through LS do exist, but "much of the theory behind lesson study is implicit" (583). This is illustrated by the findings of the systematic reviews by Larssen et al. (2018) and Xu and Pedder (2014), which show that the majority of the publications reviewed made little or no use of theoretical frameworks. A more consistent and comprehensive theorization of LS and its mechanism is, however, of great importance to the field, as it could facilitate the successful adaptations of LS to new contexts and help interpret qualitative research findings against a common backdrop.

The challenges outlined above emphasize the need for a shared language as well as high transparency in scientific articles concerning the ways in which LS is implemented in practice and also conceptualized by researchers and educators. Only by acknowledging the sources of complexity of LS in scientific papers, such as different theoretical approaches and local adaptations, can the field converse critically about those aspects of LS that demonstrably enable and stimulate professional learning. The present review therefore sets out to investigate how the field currently theorizes and communicates teachers' observations and reflection processes in empirical LS research. We use Moravcsik's (2020) three dimensions of research transparency to first conceptualize how decisions taken by researchers and teachers can be communicated transparently. By means of a systematic review, we then examine how the observation and reflection stages of LS are reported in peer-reviewed in-service teacher LS literature published between 2015 to 2020. Specifically, we aim to answer the following questions: (1) How transparent are LS articles in reporting their observation and reflection stages? And (2) which theoretical frameworks are used in these studies to conceptualize the observation and reflection stages in LS? Based on our findings, we then propose a framework for how prospective empirical articles can best report on the observation and reflection stage in LS.

2 Theoretical background

The issue of transparent descriptions of the LS process in publications has received little attention to date within the field. Lewis and colleagues, however, argued already in 2006 that it is especially critical to build a rich descriptive knowledge base on the LS approach and to theorize its innovation mechanisms. Concerning the former, several studies echo Larssen et al.'s (2018) findings of incomplete LS descriptions and suggest that also research methodologies frequently remain underreported in articles (e.g., Cheung and Wong, 2014; Baumfield et al., 2022). As a result, crucial information is often missing in published research on LS. Concerning the latter, Xu and Pedder (2014) and Stigler and Hiebert (2016) argue that, while attempts have been made to theorize and explain the processes that features of LS effective, a coherent theory how teachers learn through LS is still missing or remains implicit in publications. As Hervas and Medina (2020) note, anchoring LS in a more solid theoretical framework would not only help the field to better understand LS and anticipate its impact, but it would enable facilitators and teachers to avoid misconceptions about LS and get the most out of their learning when putting LS into practice. Overall, it seems that Lewis's (2009) call for description and theorization of LS have not yet been sufficiently addressed by the field.

This hints at a bigger issue that goes beyond the field of LS: the importance of clearly communicating steps and decisions taken during an educational intervention, and more broadly during research. As several review articles have shown, the research literature in the social sciences, including the field of education, frequently falls short of this cornerstone of transparent communication (Mann and Walsh, 2013; DeLuca et al., 2015; Aguinis et al., 2018; Hardwicke et al., 2020; Brown et al., 2021).

Transparency in research refers to “the degree of detail and disclosure about specific steps, decisions, and judgment calls made during a scientific study” (Aguinis et al., 2018, 84). In other words, studies display high transparency if they explicitly communicate choices made by the researchers about design, data collection, and analysis, and if they make resources, such as protocols and materials, available. Especially in qualitative research, there seems to exist some confusion over how research transparency can be best achieved for different types of research (Moravcsik, 2020), which has resulted in “a serious neglect of transparency and reproducibility” in some parts of social sciences, including education (Hardwicke et al., 2020, p. 7). Wiggins and Christopherson (2019), approaching transparency from the angle of the replication crisis that has hit psychology in the last decade, note that the way in which data is collected and analyzed cannot be treated as a “secret recipe” (209), but has to be replicable to others. This need for transparency is not limited to a study's scientific research methods (i.e., case study, randomized controlled trial), but also includes any intervention that is pertinent to the study design (i.e., LS). Without clear descriptions of the specific intervention and how it was implemented, readers can neither compare results to those of other studies using an adaption of the same intervention, nor replicate the intervention in their own context (Rosenshine, 1994). This means that studies on an educational intervention such as LS need to transparently describe both their research methods and their LS intervention.

There are several reasons for why critical information concerning a study's methodology or intervention might get lost on the journey

to publication. First, the omission of information might be due to external circumstances, that is, some information may be subject to ethical or legal barriers, or has to be omitted due to the strict word limits that some journals have (Moravcsik, 2020). Second, researchers might expect their readers to understand certain terms or processes without further explanation. Frequently used constructs are often presumed to be understood universally, at least among researchers in a specific discipline (Eisenhart and DeHaan, 2005; Wolgemuth et al., 2017). As a result, these constructs tend to be underdefined in the literature, often lacking a theoretical underpinning. One such construct “riddled with inconsistencies” (Mann and Walsh, 2013, p. 292) in the field of education is “reflection” (e.g., Mann and Walsh, 2013; DeLuca et al., 2015; Brown et al., 2021). While articles generally identify reflection as being a vital part of teachers' inquiry processes, the reflection process itself remains largely undefined in publications (DeLuca et al., 2015) and descriptions of *how* reflective practice can be operationalized are routinely omitted (Mann and Walsh, 2013). These findings are reiterated in the recent meta-narrative literature review by Brown et al. (2021) on reflective professional inquiry, which shows that the undertheorization of reflection, while increasingly criticized, still very much exists in the field of education. The use of the term ‘reflection’ in a research article without a definition or theoretical grounding is therefore not particularly helpful to readers and challenges the works' transparency and replicability.

Guidelines by journals or, for instance, the *Guide to APA Style* (American Psychological Association, 2022), provide clear recommendations on how to report method sections in scientific papers. In addition, research has produced lists and recommendations for how transparency can be improved in different fields and in specific types of research papers (e.g., Meyrick, 2006; Hardwicke et al., 2020). Moravcsik (2020), focusing on social sciences, delineates three dimensions of research transparency that can help to better conceptualize which aspects contribute to a clear description of research.

The first dimension, *data transparency*, concerns access to data and evidence that researchers base their findings on. Access to data enables other researchers to fully understand the analysis at hand and to judge its validity, as well as to improve or extend that analysis (Moravcsik, 2020). The second dimension, *analytic transparency*, concerns the way in which data has been collected and analyzed. This dimension is especially critical in qualitative research, as “social scientific evidence does not speak for itself” (Moravcsik, 2020, 3), but has to be inferred. The third dimension, *production transparency*, concerns the wider contextual conditions that impacted the collection and analysis of data—in other words, the methodological choices and processes that led to these choices.

These three dimensions have been formulated for the assessment of the transparency of research methods. They also provide, however, a useful framework to determine which aspects of teachers' observation and reflection processes in LS might influence the interventions' outcome and thus need to be transparently recorded in publications.

Translated to the LS process, the dimension of *data transparency* concerns the observation stage, in which teachers observe and record student learning. Both Brosnan (2014) and Bjuland and Mosvold (2015) describe cases in which the overall quality of the LS cycle suffered in part from teachers' unstructured note-taking. Færøyvik Karlsen (2019b) and Callahan (2019), on the other hand, describe that

the use of specific observation protocols enhanced teachers' observations. In order for other researchers to reconstruct teachers' observation process, articles therefore need to be clear on *how* (e.g., unstructured notes, specific template) and *whom* (e.g., whole class, case pupils) teachers observed, as well as about the materials that scaffolded this process (e.g., lesson plan, video recordings, phones).

The second dimension, *analytic transparency*, concerns the reflection stage in LS, in which teachers aim to derive new approaches for their future practice based on their observations (Lewis et al., 2019). Several studies describe that LS groups found it challenging to reflect critically together (e.g., Myers, 2012; Bae et al., 2016), or create and develop potential moments of learning in their discussions (Mynott, 2019). LS literature should therefore explicitly report back on how the teachers approached the reflection stage and whether their collaborative reflection followed a specific structure or protocol.

The third dimension, *production transparency*, includes a broader set of aspects that may influence the observation and reflection stage: outside expertise, the way in which LS groups document their LS process, and the duration and setting of the reflection stage. The involvement of external expertise, in the form of, say, knowledgeable others and external facilitators, is an integral part of LS and has often been shown to play a crucial role in how impactful the measure is on teachers (e.g., Bae et al., 2016; Amador and Carter, 2018). The extent of their involvement in the LS process is therefore an important factor that needs to be described in research studies. Furthermore, LS is not a one-time event, but relies on iteration (Stigler and Hiebert, 2016). In order for LS groups to be able to consolidate their learning, transfer it to their next LS cycle, or be able to communicate their findings to their school and wider community, it is vital that they keep some kind of a record of their learning (Lewis et al., 2019; Seleznyov et al., 2021). In more recent studies, time and space for teachers' reflection have been highlighted as being important preconditions for successful LS (Seleznyov et al., 2021).

To sum up, the idea of LS as an approach to teachers' professional development has traveled around the world and taken on many forms and shapes. Given that controlled experimental research is often difficult to conduct in complex environments such as schools (Bryk, 2015), the bulk of current LS research is small-scale, qualitative, and contextualized (Hadfield and Jopling, 2016; Seleznyov, 2019). Researchers and educators can still learn a lot from qualitative research findings despite the variability in LS implementations. This presupposes, however, that the field describes the intervention, its use, and its outcomes in a mutually intelligible and transparent manner. By framing the LS stages of observation and reflection in terms of Moravcsik's (2020) dimensions of transparency, we have identified several steps and measures that require such explicit communication.

3 The present study

The present study expands on earlier research that claims that LS procedures as well as theoretical frameworks that could explain how teachers learn through LS frequently remain implicit in research publications (e.g., Stigler and Hiebert, 2016; Larssen et al., 2018). We have two objectives. Firstly, we examine how the implementation of teachers' observations and reflections are reported. While Larssen et al.'s (2018) study concentrated on the initial teacher education literature, we investigate the in-service teacher literature, which represents the bulk of LS research (Xu and Pedder, 2014). We also

extend our research focus to include both the observation and reflection stage, given that reflection has been shown to be frequently undertheorized and underdescribed in research studies (DeLuca et al., 2015; Brown et al., 2021) and the LS literature (Stigler and Hiebert, 2016). Secondly, we aim to synthesize whether the stages of observation and reflection are connected to, or defined in relation to, a theoretical framework. We ask the following questions:

- (1) How transparent are LS articles in reporting their observation and reflection stages of LS?
- (2) Which theoretical frameworks are used in these studies to conceptualize the observation and reflection stages in LS?

The issue of transparency and theorization in LS articles has so far mainly been discussed on the sidelines. This article aims to change this by offering a systematization of those aspects that remain particular ambiguous in publications, by discussing possible reasons behind uninformative descriptions, and by highlighting consequences and ways forward. It is therefore a third aim of this article to propose a checklist based on our findings that delineates which information relating to the observation and reflection stages needs to be made explicit in publications.

4 Methods

We followed the stages of a systematic review as set out by Gough (2007, p. 218–219) and have structured this section accordingly. We first define the inclusion criteria, then delineate the search strategy, and finally describe the coding process and data analysis. Prior to data analysis, we developed a systematic review protocol based on the PRISMA checklist proposed by Moher et al. (2009). The review protocol, along with a version of the coding tool, were pre-registered on Open Science Framework (OSF) on November 22, 2021, and both are available at doi:10.17605/OSF.IO/5NXXGY (Kager et al., 2021).

4.1 Inclusion criteria

The review included an article if: (a) it reported on LS with in-service teachers in a general educational school or preschool (kindergarten to secondary school); (b) it was published in a peer-reviewed journal; (c) it was published between January 2015 and December 2020; (d) it was available in English; (e) it was an original and predominantly qualitative study; (f) it focused on LS (rather than on a PD approach that only includes elements of LS); and (g) it focused on either the whole LS process or specifically on the observation and reflection stages.

The review protocol published on OSF provides a detailed account of our rationale behind each of these inclusion criteria. Nevertheless, we want to highlight and explain some decisions we made during the culling process. To begin with, we initially focused on studies published within the last decade (2010 to 2020) and in doing so identified an overwhelming number of eligible studies (see Figure 1). To keep the body of studies to a manageable size, and given that the majority of identified studies had been published between 2015 and 2020, we shortened the time frame to this period.

Secondly, we focused on qualitative research as it represents the bulk of LS research (Xu and Pedder, 2014) and one would expect this

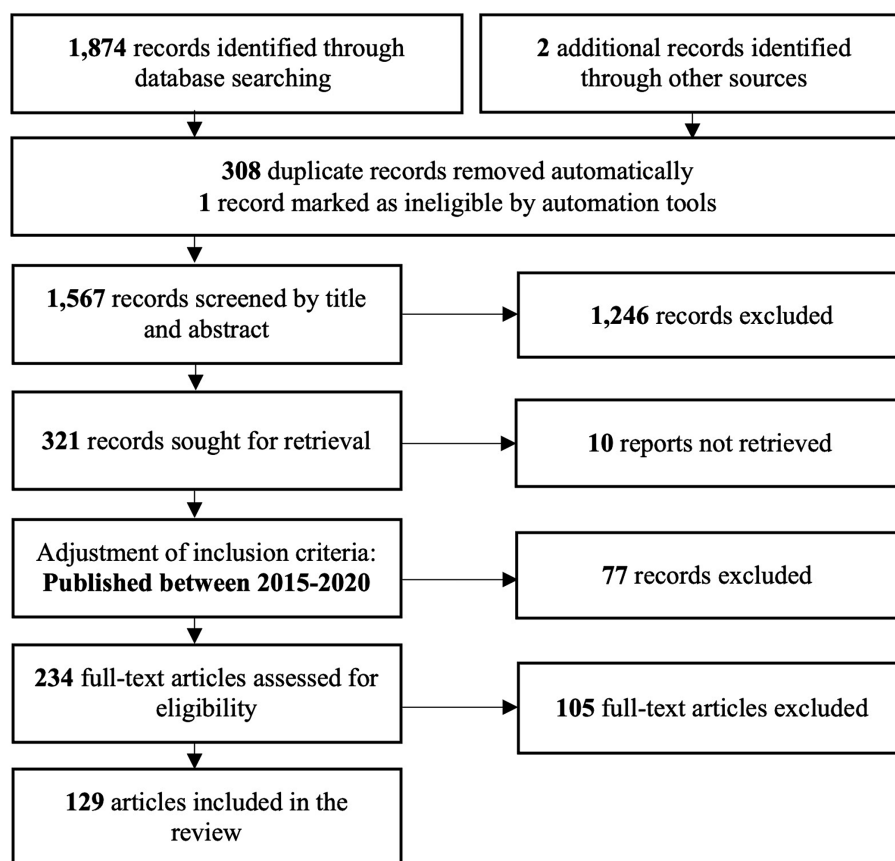


FIGURE 1

Flowchart of the literature search process adapted from Moher et al. (2009). Records identified per data base: ERIC $N = 913$; Scopus $N = 803$; APA PsychInfo $N = 100$; Academic Search Premier $N = 5$; Bibliography of Asian Studies $N = 5$; JTSOR $N = 7$; ProQuest $N = 41$.

kind of research to be most likely to include detailed descriptions of the LS interventions in question. We therefore included several qualitative designs, such as narrative research, case study, grounded theory, phenomenology, participatory action research, design-based research, and action research. Large-scale implementations of LS were therefore excluded, but have previously been reviewed in studies with a similar focus (Seleznov, 2019; Willems and Van den Bossche, 2019). We also excluded conceptual and theoretical articles, and end-of-project reports. We found that some cross-cultural articles, that compare LS processes from different countries, did not describe each LS implementation with the same thoroughness. Since the analytic rubric that we designed for our assessment cannot account for this, we also excluded this type of study.

Thirdly, we concentrated our analysis on LS with in-service teachers. This criterion was at times difficult to assess, as some articles report on in-service teachers that are enrolled in graduate courses (e.g., Pang, 2016), or on graduate students conducting LS with a group of in-service teachers (e.g., Csida and Mewald, 2016). In order to systematize our decisions, we included articles that self-identify their teachers as in-service teachers, as well as articles that report on a LS group that consisted predominantly of in-service teachers.

Finally, we had intended not to exclude any studies based on quality criteria as long as they were peer-reviewed. During full-text screening, however, we struggled to fully understand five eligible articles. While these studies provided key words that seemed relevant

to our analysis, they did so in inconsistent ways that challenged the coders' reliable and fair assessment. We therefore excluded these five studies on basis of their intelligibility. As specified in the pre-registered review protocol (Kager et al., 2021), we also excluded books and gray literature, as well as articles written in a language other than English, due to the authors' own language capabilities.

4.2 Search procedure and identification of studies

The literature search comprised several stages. Firstly, we searched the databases SCOPUS, ERIC, PsychInfo, Academic Search Premier, Bibliography of Asian Studies, JTSOR, and ProQuest for articles published between 2010 and 2020 that included the term "lesson study" in their title, abstract, or key words. In later stages, we identified two additional records through referential backtracking. Altogether, the search yielded 1,876 records, which were imported to the reference management software Zotero. After the automatic and manual removal of duplicates ($N = 308$) and records that had been retracted after publication ($N = 1$), we imported the remaining 1,567 records to Rayyan, a web-tool that supports the screening of literature (Ouzzani et al., 2016). The first author and a trained research assistant independently rated each abstract according to the set inclusion criteria. Disagreements on potential relevance of studies were

TABLE 1 The eight categories derived for the assessment of transparency.

Level	Category	Description
Data transparency	Means of data collection	How are teachers collecting data, and what type of data?
	Focus of observation	What are teachers observing?
	Scope of observation	Whom are teachers observing?
Analytic transparency	Interpretive process	How are teachers reflecting on the data?
	Procedure/structure of the post-lesson discussion	Does the reflection stage follow a specific procedure or structure?
Production transparency	Role of outside expertise	Are outside experts present and what role are they playing?
	Record-keeping	Is someone taking notes on the LS process?
	Setting of the post-lesson discussion	Where and for how long are teachers reflecting together?

TABLE 2 Overview over intercoder reliability between the three coders.

	Studies coded (N = 129)	Studies double coded by Coder 1 (overall 20%)	Cohen's κ
Coder 1	78	–	–
Coder 2	25	12	0.80 (strong agreement)
Coder 3	26	13	0.81 (strong agreement)

discussed and solved collaboratively. The Rayyan' app gave the raters a near perfect intercoder reliability of above 95%. This step reduced the set to 321 articles.

The full papers of these studies were imported to Zotero, with 10 potentially relevant studies excluded on the grounds that the texts were not publicly available. During the full-text screening, we made two changes to the inclusion criteria. Firstly, for the reasons outlined above, we adjusted the date range to only included articles published between 2015 and 2020. This led to the exclusion of 77 full texts. Secondly, we added inclusion criterion g, which specifies that the article had to focus on either the whole LS process or specifically on the observation and reflection stages. This criterion was added in order to ensure that all included studies could be expected to include relevant information about the observation and reflection stages. Overall, we excluded 105 studies in this phase. This left 129 studies in the review, which were subsequently coded in Excel. Figure 1 illustrates the stages of this culling process.

4.3 Data coding and analysis

The data coding and analysis followed five stages: 1) identifying categories; 2) developing the coding tool; 3) coding and assessing intercoder reliability; 4) extracting data of theoretical frameworks; and 5) data analysis.

4.3.1 Identifying categories

We began by reviewing the LS literature to identify a list of decisions taken by researchers and/or teachers that relate to the observation and reflection stages in LS. We piloted this list of categories by coding 25 randomly selected articles on LS. This took place before the systematic literature search and the piloted studies were not subject to our inclusion criteria. Based on our findings from the pilot coding, we refined the list and settled on eight categories for the assessment of transparency (Table 1). Each of these categories were assigned to one of Moravcsik's (2020) three dimensions of transparency.

4.3.2 Developing the coding tool

The design of the coding protocol is based on Hallinger's (2014) analytic rubric, which uses three levels of distinction (i.e., an article does not include information, includes partial information, includes detailed information). We developed definitions and anchor examples for each category and level. The final coding protocol (Supplementary Table S1) consisted of three parts: a *Quick Critical Appraisal Checklist* that reiterated the inclusion criteria; *Additional Information*, in which coders recorded general characteristics of the article, such as its research design and the label used to refer to the reflection stage (e.g., "post-lesson discussion"); and the *Analytic Rubric*, which included the eight categories outlined above for assessing transparency (see Table 1).

4.3.3 Coding and assessing interrater reliability

We coded the articles from the final set of studies according to a procedure adapted from O'Connor and Joffe (2020). The first author, who acted as the primary coder, coded a small amount of data during the development stage of the coding protocol to ensure its suitability. The first author then trained two research assistants by triple-coding studies, comparing results, discussing ambiguous examples and refining definitions in the coding tool. Satisfactory reliability was achieved after four rounds of coding and an updated version of the coding protocol was established. Subsequently, the three coders worked independently but met regularly to discuss problematic cases. During this stage, the coders collaboratively assembled a list of keywords for each category (i.e., words associated with the reporting of a certain category, see Supplementary Table S2). After the completion of the coding, we used the search function and the list of keywords to double check categories which we had rated with 0 (i.e., no information provided) to ensure that we had not missed any information. The first author double coded 20% of all studies, which has been suggested as an appropriate proportion for large data sets (O'Connor and Joffe, 2020). In order to account for the multiple coders, the first coder randomly selected and coded studies from each additional coder. Intercoder reliability (Table 2), calculated in R (R Core Team, 2013), was strong (McHugh, 2012).

4.3.4 Extracting data of theoretical frameworks

After coding was completed, each coder searched their allocated articles for any theoretical frameworks on observation and reflection. This process was also supported by the list of keywords. Findings were recorded in the form of notes in Excel.

4.3.5 Data analysis

Finally, we recorded our findings in an overview Excel sheet to organize the information and calculate frequencies. The terms used in articles to refer to the reflection stage had to be organized in thematic groups in order to be quantified. We imported the list of all labels extracted from the studies to MaxQda (VERBI Software, 2019) and created a Code Co-occurrence Modell with MaxMaps. We first grouped the labels according to themes and developed codes, such as “discussion” and “conversation.” To represent variations of the same concept (e.g., “reflection,” “reflecting,” and “reflective”), we grouped some words under a joint label (“reflect*”). We double-coded labels that included several themes. For example, the label “post-lesson reflection” was double-coded as “post-lesson” and “reflect*,” and the label “reflective debrief” as “reflect* and “debrief*.”

For the analytic rubric, we calculated raw frequencies and percentages for each category as well as the total score for each article in Excel. These frequencies were imported to R (R Core Team, 2013) to create graphs. We then selected several qualitative examples and quotations from the reviewed studies to illustrate our findings and complement the analysis.

5 Result

The results are organized into three main sections. We first describe the general characteristics of the studies included in this review and then report the findings on the transparency of the LS intervention. We then present the findings on the theorization of the observation and reflection stages. All findings are illustrated by examples; additional examples are provided in the [Supplementary materials](#).

5.1 General characteristics

The 129 studies included in this review took place in 33 countries. The countries represented with the highest numbers of studies are the US ($N=24$), Turkey ($N=13$), and the United Kingdom ($N=12$). As [Table 3](#) indicates, the number of publications being published on LS

has increased relatively consistently over the five years from 2015 to 2020.

The majority of studies described their PD approach as LS ($N=79$), Japanese LS ($N=12$), or Chinese LS ($N=6$). Some studies used modifying words (i.e., participatory LS, blended LS), and three studies used an established acronym to refer to their LS adaptation, such as CLR (i.e., Collaborative Lesson Research). Most studies were conducted either in secondary school ($N=61$) or primary school ($N=44$). Almost half of all studies ($N=60$) reported using some sort of case study design as their research methodology. Forty-nine studies reported that they employed a type of qualitative research design without further specifying their approach. Detailed tables for these general characteristics are included in the [Supplementary Tables S3–S6](#) for this article.

We documented a wide array of labels used to refer to the reflection stage. We also found variation within articles, with 25 studies using at least two different labels to refer to the reflection stage within the text. However, 15 studies did not make use of any specific label at all. The map in [Figure 2](#) illustrates how often terms occurred by themselves or were used in combination with one another. The largest group consists of the phrase “post-lesson” ($N=47$), followed by “discuss*” ($N=42$) and “reflect*” ($N=42$). The map also demonstrates that the by far most common combination was “post-lesson discussion” ($N=28$), followed by “post-lesson reflection” ($N=9$).

5.2 Assessment of the transparency in the observation and reflection stages

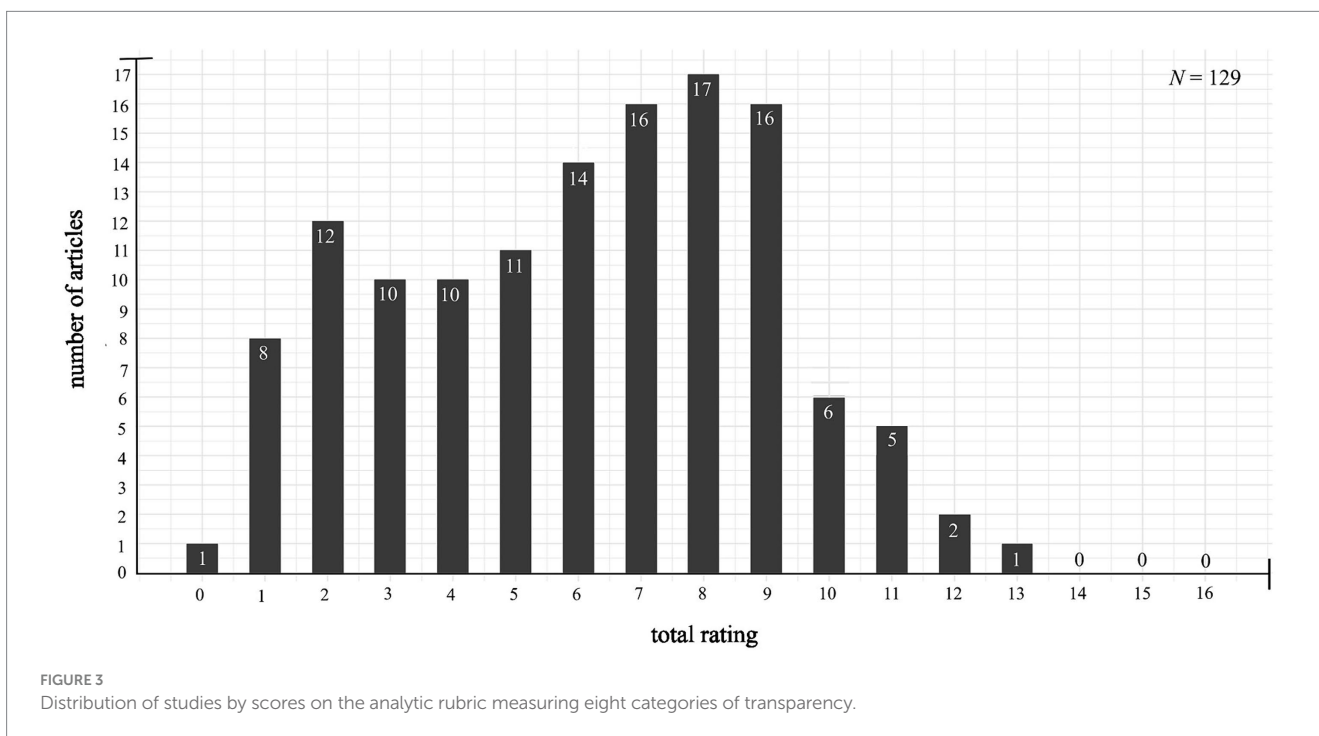
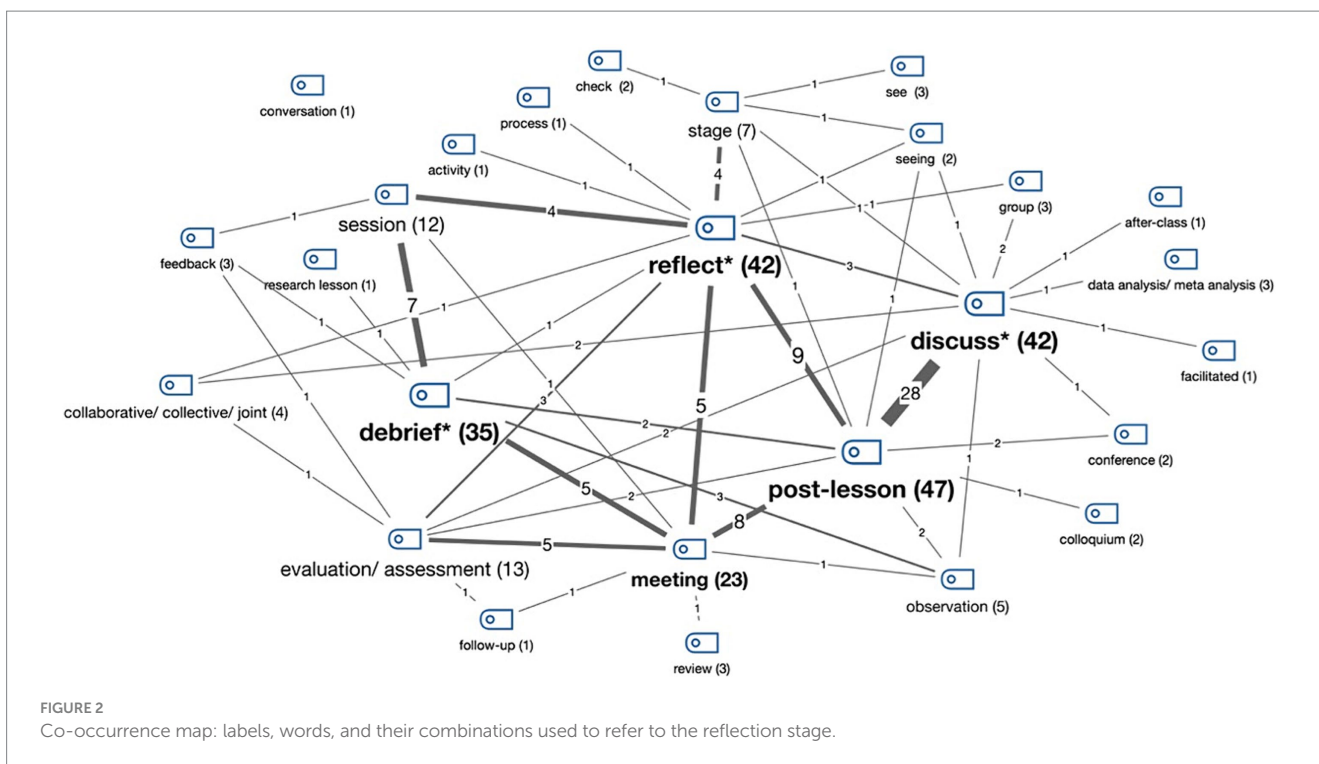
In this section we report the results of the analytic rubric, which was used to assess the transparency of articles when reporting the observation and reflection stages of LS. We will first present an overview of the total scores and then address each category individually.

5.2.1 Overall rating

[Figure 3](#) presents the distribution of the 129 studies included in this review by scores on the analytic rubric measuring eight categories for transparency. The categories were assessed with scores of 0 (does not include information), 1 (includes partial information), and 2 (includes detailed information). The maximum score would yield a rating of 16. The highest rated article scored 13 points ([Aydogan Yenmez et al., 2017a](#)), followed by two articles that scored 12 points ([Warwick et al., 2016](#); [Færøyvik Karlsen, 2019](#)). Almost 50% of articles were scored between 6 and 9 points, the most frequently scored rating

TABLE 3 Number of publications per year and according to region.

	2015	2016	2017	2018	2019	2020	Total
Asia	7	10	6	9	14	16	62
Europe	2	4	5	5	5	11	32
North America	5	5	4	5	4	5	28
Australia		1	1		1		3
Africa				1	1	1	3
South America						1	1
Total	14	20	16	20	25	34	129



being 8 ($N=17$). On the lower end, several articles met almost none of the eight criteria, with 21 articles scoring 2 or lower.

Figure 4 displays the assessment of transparency according to each category and indicates stark differences between the categories. The category *Role of Outside Expertise* was the most transparently communicated category by a large margin. Some categories, such as *Scope of Observation*, *Interpretative Process*, *Documentation of Reflection Stage*, and *Setting of Reflection Stage*, were rated with 0

across the majority of the articles. In the following, we will discuss each category separately.

5.2.2 Means of data collection

The majority of studies (61%) included some information on the type of data collected by teachers. The most common type of data was notes ($N=40$), followed by videos or audio-visual recordings ($N=19$), and student work ($N=17$). A complete list of data types is presented

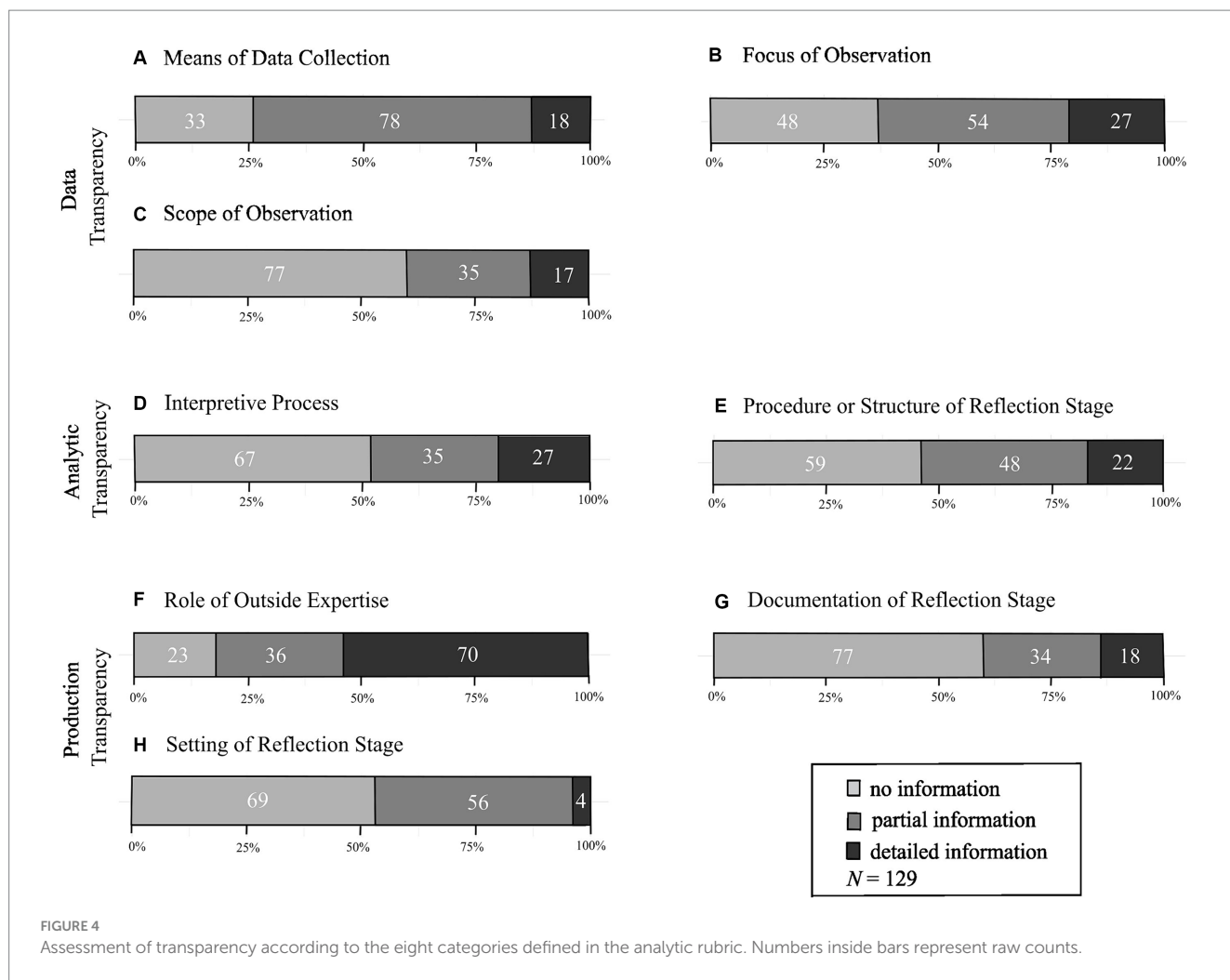


TABLE 4 Type of data collected by teachers during the research lesson.

Type of data	N	Type of data	N
notes	40	Private memos/ reports	2
Videos, audio-visual recordings, photos	20	Assessment instrument	1
student work /artifacts	17	Rubric	2
student interview	13	Mobile phones	1
observation form/ sheet/tool/template	8	Written feedback from students	1
pre-post test	8	Blackboard writing	1
Observation protocol/ log/ notebook	8	Articles stating that observation followed no protocol	1
lesson plan	4	No information given in article	33

A total of 38 articles described collecting several types of data.

in Table 4. Only a few articles (13%) explained the rationale behind the means of data collection or provided additional information about the process. Articles that did include this information described, for

instance, that LS groups developed their own observation forms or rubrics (e.g., Bruce et al., 2016; Craney et al., 2020), or referenced existing templates or material from a specific LS handbook (e.g., Lucenario et al., 2016; Khokhotva and Elexpuru Albizuri, 2019). A list of all articles rated with 2 for this category and their approaches to data collection can be found in the Supplementary Table S7.

5.2.2.1 Challenges encountered with this category

Several articles briefly mentioned “notes,” “systematic observations,” or “field notes from lesson observation,” but failed to unambiguously state whether these notes had been taken by teachers, facilitators, or researchers, and whether these notes were analyzed by teachers during the reflection stage, rather than by researchers as part of their research study. Other articles mentioned LS handbooks or work-books, but did not specify which they were or provide any references.

5.2.3 Focus of observation

The majority of articles provided some information (43%) or even detailed information (21%) on the focus of teachers’ data collection during the research lesson—that is to say, *what* teachers observed. For example, Won (2017) explains how teachers discussed the focus of their observations in the planning process and noted down expected or desired student responses in the lesson plan to guide their

observations. Gilissen et al. (2020, p. 1261) describe teachers focusing their observations on students' ability for systems thinking in biology education. During the research lesson, teachers observed how the case students behaved, communicated, and performed during certain key activities in the lesson, which are also detailed in the article.

5.2.3.1 Challenges encountered with this category

Studies that did not include any explicit information on the focus of teachers' observations usually mentioned student learning as the general focus of LS at some point in the article. We consider the notion of student learning, in this context, as vague and nondescriptive, as the term could potentially refer to almost any pedagogical activity that occurs in the classroom. The variety of observation foci that we found among the reviewed studies demonstrates the fact that only being given the information that teachers focused their observations on student learning is not sufficient to understand how this part of LS was executed and, as such, is certainly not replicable.

5.2.4 Scope of observation

The majority of articles (60%) did not include information on whom teachers observed during the research lesson. Only 13% of articles provided explicit and detailed information on this topic. One example comes from Norwich et al. (2016, p. 183), who specify that each LS group in their study chose two students for observation. Teachers in this study based their selection on learning performance, observing both a student who usually struggled with the lesson's content and a student who represented a level that teachers felt was "typical" for this class. Liu (2016, p. 106), on the other hand, tells us that the teacher who implemented the lesson asked the other team members to form groups and each observe a subgroup of students. Their goal was to learn something about each student.

5.2.4.1 Challenges encountered with this category

The information about whom teachers observed was sometimes disclosed between the lines. For example, some articles mentioned at some point the number of students in the class, inviting the conclusion that teachers observed all students. The majority of articles did not, however, communicate this in an unambiguous way that did not require the reader to make any inferences. Furthermore, most articles that focused their analysis on student work rather than observational notes did not indicate whether or not the work of all students was considered in the reflection stage, or rather just the work of specific students.

5.2.5 Interpretative process

About half of the articles (52%) did not clearly explain how teachers analyzed and reflected on the collected data. Twenty-seven percent of articles included partial information, that is, they provided examples of, or original quotations from, the interpretative process. And 21% of the articles included a definition or conceptualization of the reflection stage or of teachers' interpretative processes. Some of these articles did so in passing, while others dedicated more time to the issue.

5.2.5.1 Challenges encountered with this category

A variety of studies briefly referred to concepts or terminology in connection to the reflection stage of LS. These articles did not, however, provide a definition or explanation for the relevant terms.

Similarly, some articles mentioned reflection in connection to concepts such as the community of inquiry or professional learning communities, but did not explicitly conceptualize or define reflection itself.

5.2.6 Procedure/structure of the reflection stage

Almost half of all studies (46%) did not specify how the reflection stage of LS was structured, specifically whether or not teachers followed a specific procedure. Only 17% of articles provided detailed descriptions of these processes. These usually included a chronological component. For example, Huang et al. (2017) relate how teachers first shared their reactions to the lesson, discussed the learning outcomes, and then talked about their concerns. Kanellopoulou and Darra (2018a, p. 71), on the other hand, describe teachers following a research lesson review protocol adopted from Stepanek et al. (2007), and list several chronological steps followed by these teachers.

5.2.6.1 Challenges encountered with this category

Several studies that were coded as providing no information on this step did still include some indication of what a typical post-lesson discussion might include in a general sense. We usually found this information in the studies' literature review when the specifics of LS were introduced. These studies did not, however, define what their own implementation of LS looked like and they failed to clearly state whether or not their adaptation included any or all of these typical steps, and in what order those steps were taken.

5.2.7 Role of outside expertise

With only 18% of articles not including this information, this category was communicated in a largely transparent manner in most studies and the majority of articles included detailed information on the roles outside experts took. Pang (2016), for instance, reports on LS based on the collaboration between a university professor and in-service teachers at a Korean primary school who were enrolled in a graduate course. Pang informs the readers that she took on the role of the "knowledgeable other" and shared her expertise with the LS group, for example by commenting on the lesson plans and providing feedback during the reflection stage.

5.2.7.1 Challenges encountered with this category

The analysis indicated that researchers take on a variety of roles in the LS process. Pang (2016), for example, explicitly states that researchers acted as external facilitators, providing expertise and guidance to the LS group. The studies by Norwich et al. (2016, 2018) report that the LS group was joined by both the researchers and additional experts. Some researchers accompanied the process as active participants in the LS process and simultaneously acted as authors of the research paper (e.g., Leong et al., 2016; Ni Shuilleabhain and Seery, 2018), while others described their role as researchers being that of invisible observers (Moghaddam et al., 2015). This diversity made it difficult to clearly understand the role of researchers and external experts in articles that mentioned external instructors or experts, but neither identified them nor explained their role in the LS process.

5.2.8 Documenting the reflection stage

The majority of articles (60%) did not include any explicit information about whether someone documented the group's

reflection process and their take-aways in any way. Only 14% of articles provided detailed information about this. [Watanabe et al. \(2019\)](#), for example, include an appendix with documentation from the LS process that could serve as templates for others. [Özaltun Çelik and Bukova Güzel \(2018\)](#), p. 182 describe teachers keeping individual reflective diaries after each LS cycle to record their experiences and thoughts in regard to specific questions they faced. Another example comes from [Moss et al. \(2015\)](#), who report teachers documenting the LS process in a so-called “LS package and iBook,” which can be accessed online by anyone interested in learning more about their study.

5.2.8.1 Challenges encountered with this category

Several articles referred to notes or records but failed to clearly describe who took those notes and at what point in time, nor even whether the purpose of the notes was to document the LS process.

5.2.9 Setting of the reflection stage

This category examined whether articles included information on the duration of and/or setting for the reflection stage. The majority of articles (54%) did not include any explicit information on this. Across the remaining articles, 28 included details on the length of the reflection stages, the most common duration being 1 hour ($N=8$), followed by up to 1 hour ($N=7$), up to 2 hours ($N=7$), and longer than 2 hours ($N=6$). Concerning timings, articles usually specified whether the reflection stage had taken place immediately after the research lesson, or some time later. [Bradshaw and Hazell \(2017\)](#), for example, report that the teachers’ reflection stage followed soon after the teaching session so that “ideas and observations from the lessons were strong in the minds of the observers” (34). Whereas [Aydogan Yenmez et al. \(2017a\)](#), p. 321 tell us that the students’ reports—which were the basis for data analysis—were copied after the research lesson so that each teacher would have their own copy available to them during the reflection stage.

5.2.9.1 Challenges encountered with this category

This category was easy to code, as the vast majority of articles did not provide any information on this issue.

5.3 Theoretical frameworks for the observation stage and reflection stage

We found 10 studies (8%) that explicitly connected the observation stage to a theoretical framework or to concepts of observation that already exist in the literature. Five of these articles referred to the notions of “(professional) noticing” and “professional vision” (based on, e.g., [Sherin and Han, 2004](#); [Jacobs et al., 2010](#); [Van Es, 2011](#)). In [Karlsen and Helgevold \(2019\)](#), professional noticing was in fact the focus of their research objectives, exploring the depth of teachers’ observations and their analytic stance in the post-lesson reflection. They conclude that teachers’ professional noticing in LS should be supported by observation forms designed explicitly to capture student learning. Other articles referenced more general frameworks, such as active learning ([Garet et al., 2001](#)) or theories of teacher learning ([Penuel et al., 2007](#); [Marton, 2015](#)), while explicitly highlighting observation and its role within these frameworks. [Koutsouris et al. \(2017\)](#) used [Dyke et al.’s \(2006\)](#) notion of “tunnel

vision” to elaborate on difficulties with videotaping the research lesson and to describe the effect classroom videos might have on its observers and those being observed. A list of these studies and their approaches can be found in the [Supplementary Table S8](#).

We found 20 studies (16%) that explicitly theorized teachers’ reflection processes. In general, reflection was identified as an important aspect in teacher learning and several articles ground their understanding of reflection in the works of [Dewey \(1933\)](#) and [Zimmerman \(2000\)](#). The most frequently cited scholar was [Schön \(1983, 1995\)](#), with six articles referring to his notion of the reflective practitioner, as well as reflection-in-action and reflection-on-action. Another reoccurring framework was rooted in the theory of cognitive conflict (e.g., [Posner et al., 1982](#); [Piaget, 1985](#); [Limon, 2001](#)). A list of the studies that theorized reflection can be found in the [Supplementary Table S9](#).

In addition to the frameworks discussed above, we found that seven studies (5%) grounded their understanding of the reflection stage in alternative theoretical perspectives. [Brown et al. \(2016\)](#), for instance, referred to theoretical perspectives on “learning conversations,” and [Lee and Tan \(2020\)](#) on “professional conversations.” [Warwick et al. \(2016\)](#) and [Bae et al. \(2016\)](#) both connect the reflection stage to the notions of dialog, interthinking, and modes of talk ([Mercer, 2000](#); [Littleton and Mercer, 2013](#)). A list of these studies and their approaches can be found in the [Supplementary Table S10](#).

6 Discussion

This systematic review set out to examine two research questions. Firstly, we asked how transparent in-service LS articles are in reporting on their observation and reflection stages of LS. Secondly, we asked which theoretical frameworks are currently being used to conceptualize these two stages. In regard to the first question, our analysis of 129 articles indicates that several categories across all three dimensions of transparency ([Moravcsik, 2020](#)) were either omitted completely or described only partially in the majority of studies. In line with [Cheung and Wong \(2014\)](#) and [Larsen et al.’s \(2018\)](#) previous assessments, these findings provide broad evidence of a lack of transparency on two crucial stages of the LS interventions in the current literature. In regard to the second question, we discovered that only a small minority of studies theorized the observation and reflection stages of LS. These findings are also consistent with previous assessments from the field of education and social sciences ([Mann and Walsh, 2013](#); [DeLuca et al., 2015](#)), reporting that frequently used concepts, such as reflection, often remain undertheorized in publications.

The major consequence of insufficient descriptions in research articles is that these articles may not be particularly helpful or useful for researchers and practitioners. When studies omit procedural details about the LS intervention, then researchers and educators are not able to fully comprehend which aspects might have contributed to teachers’ learning, or how positive learning outcomes might be replicated and emulated in other settings. We now want to look at three aspects of this issue that emerged from our analysis: the omission of information; the complexity of LS; and the lack of a shared theoretical framework for LS.

6.1 Omission of information

The analysis found that, in the studies we looked at, some categories were communicated more clearly than others due to information being completely omitted. In general, the reasons for this omission appeared to be the presumption of a shared understanding of LS, a lack of awareness that certain kinds of information might be important to understanding findings on LS, and an unbalanced focus on LS outcomes over LS processes.

Concerning a shared understanding of LS, it makes sense to first look at the category *Role of Outside Expertise*. This was the most transparently communicated category and addresses a topic that is also increasingly the primary subject of research (Takahashi, 2014; Lewis, 2016; Amador and Carter, 2018; Hauge, 2021; Mynott and Michel, 2022). The scientific discourse on outside expertise in LS seems to be driven by researchers' own interest in how to best initiate, lead, and sustain LS. As our review showed, researchers are frequently also personally involved in LS and take on various roles, such as project leaders, coordinators, and educators. We assume that the researchers' active roles in LS have translated to the high transparency in the communication of this versatile role. Articles that provide this kind of information therefore contribute significantly to the building of a knowledgebase on the manifold ways in which outside experts can shape teachers' learning in LS.

This shared understanding concerning a certain part of the LS intervention—and its importance to LS—was largely lacking for the remaining categories assessed in this review. Four categories (*Interpretative Process*, *Scope of Observation*, *Documentation of Reflection Stage*, and *Setting of Reflection Stage*) were scored as “not included” in the vast majority of articles. If mentioned at all, these categories were frequently described in vague or general terms that left out crucial details, such as how teachers carried out the reflection process.

The low transparency of these categories might stem from researchers' belief that it is enough to indicate how LS is “commonly” or “usually” conducted. This belief, however, is refuted by the abundance of distinct approaches to LS that were described in a number of the articles assessed in this review. For example, the documentation of a LS process and the sharing of findings is a crucial part of Japanese LS and has been argued to be important in order to sustain and grow LS in schools outside of Japan (Seleznyov, 2018). Our analysis found, however, that only 18 out of 129 articles clearly communicated whether LS groups documented their learning in any way—and if, then in what way. In addition, all 18 articles approached the step of documentation differently. Approaches included, among others, reflective diaries or journals kept by teachers (Özaltun Çelik and Bukova Güzel, 2018; Calleja and Formosa, 2020), a specific template for note taking (Lee and Tan, 2020), meeting calendars (Kanellopoulou and Darra, 2018a), a report prepared by either the whole LS group (Özdemir, 2019) or a designated group member (Chua, 2019), as well as teachers' individual documentation of the process in an online space (Joubert et al., 2020). This variety of ways in which LS groups document and mobilize their learning demonstrates that the research community cannot and should not presume that there is a standard process of documenting teachers' learnings in LS that requires no further communication in research articles. Only by explicitly reporting details about these steps can others

learn from successful methods or avoid avenues that have been tried and abandoned.

Another source of low transparency was the predominance of articles that focused on LS outcomes over LS processes. This underreporting of information concerning the production of research, or in this case an educational intervention, can stem, for instance, from researchers preferring a clear “storyline” over descriptions of trial and error (Aguinis et al., 2018), or—especially in qualitative research—from trying to keep to strict word limits imposed by journals (Moravcsik, 2020). Abridged descriptions, however, sideline valuable information about judgment calls and choices crucial if others are going to be able to replicate the LS adaptation and emulate its outcomes (Aguinis et al., 2018). In addition, steps that appear trivial to the authors may be necessary for readers to understand LS procedures from a different context. Our analysis suggests that even just a short statement or description of *how* a certain step was conducted or conceptualized can greatly enhance the transparency of publications in this regard.

Several of the reviewed articles provided innovative solutions for the problem of strict word limits and restrictive formats. By including links or references to [Supplementary material](#) stored on journal websites, online repositories, or school- or project-specific websites, these articles found an effective way of making their materials and their approach to LS widely accessible to others (e.g., Moss et al., 2015; Watanabe et al., 2019). Sharing data and materials openly in order to enhance transparency is central to the Open Science movement (Nosek et al., 2012), which is becoming increasingly important in educational science (Van Dijk et al., 2021). The examples of open sharing of LS materials we found in our sample suggests that Open Science practices can also advance and deepen discourses in the field of LS.

6.2 The complexity of Lesson Study

Our findings show that the transparency of the articles we reviewed was further complicated by the complexity of LS and its various conceptualizations as a research method, as teacher-led research, or as a research object. Some articles stated that LS itself was used as a research method by researchers (akin to action-based research) to explore, for example, how to best teach fractions in math. Researchers therefore conducted research *through* LS, rather than *on* LS. Other articles conceptualized LS as teacher-led research, with the researcher(s) taking on an active part in the LS group and frequently focusing their articles on relating their experiences. The vast majority of articles, however, viewed LS as a teacher-led PD approach and research object (i.e., an intervention) that was investigated through the use of a separate methodology, such as a case study approach or design-based research.

This versatility of LS makes it an appealing approach to classroom research and teachers' PD, and also provides a framework for a rich exchange between research and practice as well as the co-construction of knowledge. Given that LS is embedded in teachers' practice, a LS cycle will always be influenced by a multitude of local factors and subsequently differ from conventional educational interventions that can be tightly controlled by researchers. Even if controlled experimental research designs are challenging to conduct in the field of LS, scientific articles of any type should nevertheless adhere to the

principle of research transparency, that is, to clearly report their evidence, analysis, and overall research design (Moravcsik, 2020). Part of this is to describe LS in enough detail so that others can understand or replicate it.

Our findings indicate that the majority of articles did not provide such descriptions regarding the observation and reflections stages of LS and did not adhere to any discernable reporting standard. In fact, information related to the LS intervention was sometimes reported in unexpected places, such as the theory section or discussion. Other articles scattered the information across multiple sections, with relevant information sometimes appearing only late in the text. Another problem was that articles that conducted research on LS frequently failed to separate the descriptions of their research method from those of the LS intervention. Some articles, for example, reported the data collected by both researchers and teachers in the same chapter, sentence, or even bullet list, making it unclear who had collected which data for what purpose.

In order to avoid confusion, we recommend that articles clearly position themselves as either research *through* or *on* LS. Research *on* LS needs to clearly separate descriptions of their research method from descriptions of the LS intervention. We suggest to report the LS intervention in a separate subchapter within the method section. It is further important to use unambiguous terminology. For example, if both researchers and teachers collected observational notes during and of the LS process, these different types of notes need to be clearly identifiable through the use of consistent language.

6.3 Lack of a shared theoretical framework

Another source of low transparency in our sample was the frequent use of the terms ‘observation’ and ‘reflection’ without providing clear definitions or situating these constructs within a theoretical framework. In fact, only a small minority of articles clearly defined these terms and 92% and 79% of articles did not connect the processes of observation and reflection to any theoretical framework, respectively. These findings provide new and concrete insights into previous assessments of the level of undertheorization in LS research (Elliott, 2012; Stigler and Hiebert, 2016) and demonstrate that the LS community uses a diverse set of terminology and labels while assuming that there is a shared understanding of these concepts. Our findings demonstrate that this shared understanding cannot be guaranteed and that the lack of definitions and theorization renders terms such as “observation” and “reflection” untransparent in LS research.

The problem with missing theorization is that it raises concerns around whether or not results can be sufficiently accounted for and whether practices can be recommended based on confirmed relationships (Fleetwood and Hesketh, 2006). In other words, the high complexity of innovation mechanisms (such as an educational intervention) makes it difficult for researchers to reliably identify which practices cause certain outcomes and may predict outcomes in the future. The coherent use of theories of teacher learning in the field of LS could therefore not only lend greater explanatory power, but it could provide guidance as to how learning mechanisms (e.g., classroom observations, critical reflection) can be effectively structured in practice.

Figure 2 illustrates the wide variety of different labels that are used in the reviewed articles to refer to the reflection stage. While some articles, for instance, speak of the reflection stage in terms of a “post-lesson reflection,” others use the labels “debrief,” “evaluation,” or “data analysis.” This variety highlights the heterogeneity of LS, but it can also be viewed as a symptom of a missing theory or model of LS that could ensure intelligibility in spite of LS’s complexity. Given that the majority of articles neither defined their terminology nor used these terms with reference to a specific theoretical framework, it remains ambiguous to the readers whether teachers engage in similar processes in a ‘debrief’, a ‘post-lesson reflection’, or a ‘data analysis’. While stronger theorization cannot substitute transparent descriptions of the reflection stage in publications, it could greatly support the field to derive coherent terminology, establish a common point of departure, and thereby ensure that everyone is in fact investigating and talking about similar things.

It has to be noted at this point, that theorization is by no means absent from the field of LS. Various theoretical perspective have been used to inform research on LS, such as self-determination theory, self-efficacy theory, and knowledge integration theory (Lewis et al., 2019). Empirical studies frequently underpin their LS research with models of PD and teacher growth, such as those by Clarke and Hollingsworth (2002) and Guskey (2002), or, in the case of Huang et al. (2016), they develop their own theory-based LS model. We have also seen the generation of new theories from empirical research on LS, such as Mynott’s (2019) theoretical outcome model of LS. In addition, this review identified a number of useful theories for the conceptualization of specific learning processes in LS, such as cognitive conflict, modes of teacher talk, and professional noticing. Some of these approaches have been picked up and further investigated in recent studies, such as Dick et al. (2022), Hrastinski (2021), and Karlsen and Ohna (2021) for the professional noticing of teachers, and Ustuk and De Costa (2021) and Kager et al. (2022) for critical and collaborative reflection.

This development indicates that the theorization of learning mechanisms in LS is actively being pursued by researchers in the field. Our findings nevertheless suggest that, to date, these theoretical perspectives are still being negotiated and advanced within the field and have not yet been adopted by the broader research community in their empirical research. In light of this finding, we argue that the field of LS has reached a point at which it would benefit from some standardization in order to negotiate what Kim (2021) refers to as a “conceptual grid” for LS outside of Japan. Importantly, we are not suggesting to standardize LS itself, but rather to standardize the way we talk about it.

6.4 A framework for reporting the observation and reflection stages

We want to conclude our review by making the following recommendations concerning the reporting of the observation and reflection stages research publications (Table 5). Firstly, researchers should aim to communicate their specific LS intervention in a concise way within the article, such as a subchapter as part of the method section. Secondly, researchers should strive to employ clearer terminology. This means that the specific use of terms such

TABLE 5 Checklist of items recommended to include when reporting a LS intervention.

	Checklist item
Observation stage	
Theoretical framework	How did researchers (and the LS group) understand the observation process from a theoretical perspective?
Type of data	What kind of data did teachers collect (e.g., structured notes, videos, student work, ...)?
Process of data collection	How did teachers collect this data? Was data collection guided by a specific protocol?
Focus of observation	What did teachers focus on in their observations (e.g., what aspect of student learning)?
Scope of observation	Did teachers observe the entire class, a subset of students, or case students? What guided this decision?
Outside expertise	How were outside experts involved in the observation stage?
Materials	Can materials used in the observation stage be accessed somewhere?
Reflection stage	
Theoretical framework	How did researchers (and the LS group) understand the reflection process from a theoretical perspective?
Process of reflection	How did teachers enact the collaborative reflection?
Structure of reflection	How was the reflection stage structured chronologically and what activities were involved?
Length of reflection stage	How long did teachers reflect together?
Setting of reflection stage	How was the reflection stage influenced by other contextual factors or decisions?
Outside expertise	How were outside experts involved in the reflection stage?
Record keeping	How was the reflection stage (or LS process) documented?
Materials	Can materials, such as reflection protocols, be accessed somewhere?

as “observation” or “reflection” need to be explained and ideally derived from or embedded in a theoretical framework. It also means that researchers should be aware that, without sufficient explanation, readers are likely to draw their own conclusions concerning terminology or labels used in the text. Thirdly, we recommend the use of the following checklist based on the findings of this review. The checklist can be used by researchers to evaluate the transparency of their manuscript and decide which aspects of their LS intervention need to be communicated to guarantee the usability of their research.

6.5 Limitations

Our methodology is subject to a number of limitations. Firstly, it is important to note that the study focused solely on how transparent descriptions of LS’s observation and reflection stages were. We recognize that a multitude of additional factors contribute to an LS outcome, including social and cultural contexts, hierarchical structures within the LS groups, the groups’ motivation, and teachers’ experience (Bocala, 2015; Hadfield and Jopling, 2016; Seleznyov et al., 2021). Secondly, we did not assess an article’s quality or overall research transparency, but specifically the degree of transparency with which an article communicated the observation and reflection stages of LS. The total rating given to an article does therefore not provide any assessment about the overall quality or scientific value of the article. Such evaluations have, however, been reported elsewhere (e.g., Rzejak, 2019; Seleznyov, 2019). Thirdly, we treated all categories assessed in the analytic rubric equally in our analysis and did not assign any weight to them. This choice might skew the results in so far as not all categories are likely to have the same impact on the outcome of an LS cycle. In order to assign weight to the categories, however, we would need further research that can provide a justification for this weighting. We would like to propose this as an avenue for future research.

There might be additional eligible articles that were not included in this review, as no database has complete coverage. Likewise, the list of categories assessed in this review were derived from the research literature, yet there might be additional categories of interest that we did not cover. The assessment of transparency, while guided by an analytic rubric, demanded definitive choices by the coders. These choices were not always easy, as they required coders not to try and read between the lines or make inferences. Nevertheless, we achieved high intercoder reliability and our findings are consistent with previous evidence. We have detailed further challenges that we faced in the assessment process in the findings section to enhance the transparency of this analysis process.

Lastly, we recognize Ishii’s (2017) concern that research in LS frequently focuses on reflection at the expense of LS’s first two phases—identifying a research question and planning instruction. Our analysis adds to this bias in so far as we only examine the observation and reflection stages of LS. We hope, however, that the present review can act as a springboard for future research into the transparency of each of the core stages of LS.

7 Conclusion

The present review demonstrates that the field of LS is currently marred by low transparency in how the observation and reflection stages are communicated in research articles. These findings build on similar observations about the underdescription and undertheorization of LS (e.g., Cheung and Wong, 2014; Stigler and Hiebert, 2016; Larssen et al., 2018). We broaden their analyses by outlining reasons for these issues and subsequently recommending specific communication practices for empirical research on LS. The proposed checklist can, in the first instance, support practitioners in their implementation of LS and, in the second, motivate researchers to rigorously and comprehensively question and document their decisions on the implementation of LS, even when it appears trivial.

We draw a range of practical and theoretical implications from these findings. Our review underlines Lewis et al.'s (2006) argument that in order to make LS effective we need to identify its crucial underlying processes and implementation steps. We saw that explicit descriptions of the intervention can greatly contribute to the building of just such a knowledge base. In this sense, we hope that the lists and examples provided in the current review and its [Supplementary material](#) of articles that explicitly communicated their interventions can act as a resource on how to conduct and establish standards and on how to report the observation and reflection stages in LS. Our research further implies that Open Science practices, such as providing open access to resources and making data publicly available, can positively impact knowledge generation in the field of LS and ensure the usefulness and replicability of research.

Turning to theoretical implications, our review highlights the need for further theoretical development for LS in general, and the observation and reflection stages in particular. The theorization of these stages was consistently absent in the articles reviewed, though some articles presented promising avenues to stronger theorization. While a complete theory of LS might be too ambitious due to its variable and complex character, it does seem possible to advance these existing theories in the field of LS and to increasingly integrate them into empirical research in a more comprehensive, extensive, and thus potentially valuable way.

Author contributions

KK: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. EK: Writing – review & editing. AJ: Conceptualization, Writing – review & editing. MV:

References

- References marked with an asterisk indicate studies included in the systematic review.
- *Abdella, A. S., Reddy, C., and Carl, A. (2018). Lesson study in Eritrea: its impact on middle school science teachers' learning and classroom practice. *J. Turk. Sci. Educ.*, 15, 1–26. doi: 10.12973/tused.10234a
- *Abu-Alghayth, K., Jones, P., Pace-Phillips, D., and Meyers, R. (2020). Through the looking glass: lesson study in a center school. *Int. J. Educ. Methodol.*, 6, 423–433. doi: 10.12973/ijem.6.2.423
- Aguinis, H., Ramani, R. S., and Alabduljader, N. (2018). What you see is what you get? Enhancing methodological transparency in management research. *Acad. Manag. Ann.* 12, 83–110. doi: 10.5465/annals.2016.0011
- *Allan, D., O'Doherty, E., Boorman, D., and Smalley, P. (2020). Lesson study and the construction of capital: empowering children through dialogic engagement. *Education* 49, 730–744. doi: 10.1080/03004279.2020.1781224
- *Alwadi, H. M., Mohamed, N., and Wilson, A. (2020). From experienced to professional practitioners: a participatory lesson study approach to strengthen and sustain English language teaching and leadership. *Int. J. Lesson Learn. Stud.*, 9, 333–349. doi: 10.1108/IJLLS-10-2019-0072
- Amador, J. M., and Carter, I. S. (2018). Audible conversational affordances and constraints of verbalizing professional noticing during prospective teacher lesson study. *J. Math. Teach. Educ.* 21, 5–34. doi: 10.1007/s10857-016-9347-x
- American Psychological Association. (2022). Publication manual of the American psychological association. Available at: <https://apastyle.apa.org/products/publication-manual-7th-edition>
- *Anif, S., Sutopo, A., and Prayitno, H. J. (2020). Lesson study validation: model for social and natural sciences teacher development in the implementation of national curriculum in Muhammadiyah Schools Indonesia. *Univ. J. Educ. Res.*, 8, 253–259. doi: 10.13189/ujer.2020.080132
- Funding acquisition, Project administration, Supervision, Writing – review & editing.
- ## Funding
- The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Our research was supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF).
- ## 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.
- ## Supplementary material
- The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2024.1322624/full#supplementary-material>
- *Asari, S., Fauziyah, N., and Uchtiawati, S. (2018). Improving teacher pedagogic competences in remote areas through lesson study activity. *Int. J. Edu. Lit. Stud.*, 6, 53–62. doi: 10.7575/aiac.ijels.v6n.2p.53
- *Austin, L. (2017). Leading the introduction and development of lesson and learning study in an English secondary academy. *Int. J. Lesson Learn. Stud.*, 6, 80–96. doi: 10.1108/IJLLS-10-2016-0036
- *Aydogan Yenmez, A., Erbas, A. K., Alacaci, C., Cakiroglu, E., and Cetinkaya, B. (2017a). Evolution of mathematics teachers' pedagogical knowledge when they are teaching through modeling. *Int. J. Educ. Math. Sci. Technol.*, 5, 317–332. doi: 10.18404/ijemst.296552
- *Aydogan Yenmez, A., Erbas, A. K., Cakiroglu, E., Alacaci, C., and Cetinkaya, B. (2017b). Developing teachers' models for assessing students' competence in mathematical modelling through lesson study. *Int. J. Math. Educ. Sci. Technol.*, 48, 895–912. doi: 10.1080/0020739X.2017.1298854
- *Bae, L. C., Hayes, K. N., Seitz, J., O'Connor, D., and DiStefano, R. (2016). A coding tool for examining the substance of teacher professional learning and change with example cases from middle school science lesson study. *Teach. Teach. Educ.*, 60, 164–178. doi: 10.1016/j.tate.2016.08.016
- *Barber, K. (2018). Developing teachers' mathematical-task knowledge and practice through lesson study. *Int. J. Lesson Learn. Stud.*, 7, 136–149. doi: 10.1108/IJLLS-09-2017-0042
- Baumfield, V., Bethel, A., Boyle, C., Katene, W., Knowler, H., Koutsouris, G., et al. (2022). How lesson study is used in initial teacher education: an international review of literature. *Teach. Dev.* 26, 356–372. doi: 10.1080/13664530.2022.2063937
- *Baz, M. A. (2020). Improving visual literacy skills through news photos. *J. Inquiry Based Activ.*, 10, 61–82. Available at: <https://ated.info.tr/ojs-3.2.1-3/index.php/ated/article/view/12>

- Bjuland, R., and Mosvold, R. (2015). Lesson study in teacher education: learning from a challenging case. *Teach. Teach. Educ.* 52, 83–90. doi: 10.1016/j.tate.2015.09.005
- *Bocala, C. (2015). From experience to expertise: the development of teachers' learning in lesson study. *J. Teach. Educ.*, 66, 349–362. doi: 10.1177/0022487115592032
- *Boran, E., Tarım, K., and Özsezer, M. Sencer Bulut. (2018). The effect of lesson study on the perceptions of mathematics teachers about subject area competencies. *Eur. J. Educ. Stud.*, 5, 371–394. doi: 10.5281/ZENODO.2529510
- *Bozkurt, E., and Yetkin-Özdemir, İ. E. (2018). Middle school mathematics teachers' reflection activities in the context of lesson study. *Int. J. Instr.*, 11, 379–394. doi: 10.12973/iji.2018.11126a
- *Bradshaw, Z., and Hazell, A. (2017). Developing problem-solving skills in mathematics: a lesson study. *Int. J. Lesson Learn. Stud.*, 6, 32–44. doi: 10.1108/IJLLS-09-2016-0032
- Brosnan, A. (2014). Introducing lesson study in promoting a new mathematics curriculum in Irish post-primary schools. *Int. J. Lesson Learn. Stud.* 3, 236–251. doi: 10.1108/IJLLS-09-2013-0050
- Brown, C., Poortman, C., Gray, H., Groß Ophoff, J., and Warf, M. (2021). Facilitating collaborative reflective inquiry amongst teachers: what do we currently know? *Int. J. Educ. Res.* 105, 1–11. doi: 10.1016/j.jer.2020.101695
- *Brown, C., Taylor, C., and Ponambalum, L. (2016). Using design-based research to improve the lesson study approach to professional development in Camden (London). *London Rev. Educ.*, 14, 4–24. doi: 10.18546/LRE.14.2.02
- *Bruce, C. D., Flynn, T. C., and Bennett, S. (2016). A focus on exploratory tasks in lesson study: the Canadian “math for young children” project. ZDM: the international journal on *Math. Educ.*, 48, 541–554. doi: 10.1007/s11858-015-0747-7
- *Bruce, C. D., and Hawes, Z. (2015). The role of 2D and 3D mental rotation in mathematics for young children: what is it? Why does it matter? And what can we do about it? *ZDM*, 47, 331–343. doi: 10.1007/s11858-014-0637-4
- Bryk, A. S. (2015). 2014 AERA distinguished lecture: accelerating how we learn to improve. *Educ. Res.* 44, 467–477. doi: 10.3102/0013189X15621543
- *Bussi, M. G. B., Bertolini, C., Ramploud, A., and Sun, X. (2017). Cultural transposition of Chinese lesson study to Italy: an exploratory study on fractions in a fourth-grade classroom. *Int. J. Lesson Learn. Stud.*, 6, 380–395. doi: 10.1108/IJLLS-12-2016-0057
- *Bütün, M. (2019). Mathematics teachers' early lesson study experiences in Turkey: challenges and advantages. *World J. Educ.*, 9, 51–62. doi: 10.5430/wje.v9n5p51
- *Cajkler, W., Wood, P., Norton, J., Pedder, D., and Xu, H. (2015). Teacher perspectives about lesson study in secondary school departments: a collaborative vehicle for professional learning and practice development. *Res. Pap. Educ.*, 30, 192–213. doi: 10.1080/02671522.2014.887139
- *Callahan, C. (2018). Promoting second-order historical domain knowledge through recursive and collaborative professional development. *J. Soc. Sci. Stud. Res.*, 42, 287–297. doi: 10.1016/j.jssr.2017.06.001
- *Callahan, C. (2019). Middle school geography teachers' professional development centered around historical photographs. *J. Soc. Stud. Res.*, 43, 375–388. doi: 10.1016/j.jssr.2018.11.003
- Calleja, J., and Formosa, L. (2020). Teacher change through cognitive conflicts: the case of an art lesson study. *Int. J. Lesson Learn. Stud.* 9, 383–395. doi: 10.1108/IJLLS-05-2020-0028
- *Cammarata, L., and Haley, C. (2018). Integrated content, language, and literacy instruction in a Canadian French immersion context: a professional development journey. *Int. J. Biling. Educ. Biling.*, 21, 332–348. doi: 10.1080/13670050.2017.1386617
- *Canonigo, A. M. (2016). Using a non-coercive process to engage mathematics teachers in lesson study. *Int. J. Lesson Learn. Stud.*, 5, 329–347. doi: 10.1108/IJLLS-02-2016-0004
- Carbin, W., and Kopp, B. (2006). Lesson study as a model for building pedagogical knowledge and improving teaching. *Int. J. Teach. Learn. High. Educ.* 18, 250–257. Available at: <https://www.isetl.org/ijtlhe/pdf/ijtlhe110.pdf>
- *Chang, C.-H., Pascua, L., and Ess, F. (2018). Closing the “hole in the sky”: the use of refutation-oriented instruction to correct students' climate change misconceptions. *J. Geogr.*, 117, 3–16. doi: 10.1080/00221341.2017.1287768
- Cheung, W. M., and Wong, W. (2014). Does lesson study work?: a systematic review on the effects of lesson study and learning study on teachers and students. *Int. J. Lesson Learn. Stud.* 3, 137–149. doi: 10.1108/IJLLS-05-2013-0024
- *Chin Mon, C., Dali, M. H., and Chap Sam, L. (2016). Implementation of lesson study as an innovative professional development model among Malaysian school teachers. *Malays. J. Learn. Instr.*, 13, 83–111. doi: 10.32890/mjli2016.13.1.5
- *Chua, V. C. G. (2019). Going the distance: a lesson study on deriving the distance formula. *Int. J. Lesson Learn. Stud.*, 8, 149–159. doi: 10.1108/IJLLS-08-2018-0052
- Clarke, D., and Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18, 947–967. doi: 10.1016/S0742-051X(02)00053-7
- *Coenders, F., and Verhoef, N. (2019). Lesson study: professional development (PD) for beginning and experienced teachers. *Prof. Dev. Educ.*, 45, 217–230. doi: 10.1080/19415257.2018.1430050
- *Collet, V. S. (2017). Lesson study in a turnaround school: local knowledge as a pressure-balanced valve for improved instruction. *Teach. Coll. Rec.*, 119, 1–58. doi: 10.1177/016146811711900605
- *Collet, V. S., and Greiner, A. C. (2020). Revisioning grammar instruction through collaborative lesson study: a new apprenticeship of observation. *Lit. Res. Instr.*, 59, 95–120. doi: 10.1080/19388071.2019.1709927
- *Craney, C. L., Lau, T., Nelson, W. T., Ghomeshi, A., Rust, J. M., de Groot, R. M., et al. (2020). Collaborative middle school science outreach project using the Japanese lesson study model. *J. Chem. Educ.*, 97, 1256–1265. doi: 10.1021/acs.jchemed.9b00780
- *Csida, S., and Mewald, C. (2016). PrimarWebQuest in foreign language education. *Int. J. Lesson Learn. Stud.*, 5, 45–59. doi: 10.1108/IJLLS-09-2015-0029
- Darling-Hammond, L., Hyler, M. E., and Gardner, M. (2017). Effective teacher professional development. Learning Policy Institute. Available at: <https://learningpolicyinstitute.org/product/teacher-prof-dev>
- DeLuca, C., Shulha, J., Luhanga, U., Shulha, L. M., Christou, T. M., and Klinger, D. A. (2015). Collaborative inquiry as a professional learning structure for educators: a scoping review. *Prof. Dev. Educ.* 41, 640–670. doi: 10.1080/19415257.2014.933120
- Dewey, J. (1933). *How we think (original work published in 1910)*. Buffalo, NY: Prometheus Books.
- Dick, L. K., Appelgate, M. H., Gupta, D., and Soto, M. M. (2022). Continuous improvement lesson study: a model of MTE professional development. *Math. Teach. Educ.* 10:19. doi: 10.5951/MTE.2020.0077
- Dudley, P. (2013). Teacher learning in lesson study: what interaction-level discourse analysis revealed about how teachers utilised imagination, tacit knowledge of teaching and fresh evidence of pupils learning, to develop practice knowledge and so enhance their pupils' learning. *Teach. Teach. Educ.* 34, 107–121. doi: 10.1016/j.tate.2013.04.006
- Dyke, M., Harding, A., and Lajeunesse, S. (2006). Digital observation of teaching practice. Annual conference of the American Educational Research Association (AERA), San Francisco, CA. Available at: <https://eprints.soton.ac.uk/24124/>
- Eisenhart, M., and DeHaan, R. L. (2005). Doctoral preparation of scientifically based education researchers. *Educ. Res.* 34, 3–13. doi: 10.3102/0013189X034004003
- Elliott, J. (2012). Developing a science of teaching through lesson study. *Int. J. Lesson Learn. Stud.* 1, 108–125. doi: 10.1108/20468251211224163
- *Estrella, S., Zakaryan, D., Olfos, R., and Espinoza, G. (2020). How teachers learn to maintain the cognitive demand of tasks through lesson study. *J. Math. Teach. Educ.*, 23, 293–310. doi: 10.1007/s10857-018-09423-y
- *Færøyvik Karlsen, A. M. (2019). Investigating teacher learning in lesson study: the important link between reported observations and change of plans. *Prof. Dev. Educ.*, 48, 53–69. doi: 10.1080/19415257.2019.1685564
- Fleetwood, S., and Hesketh, A. (2006). HRM-performance research: under-theorized and lacking explanatory power. *Int. J. Hum. Resour. Manag.* 17, 1977–1993. doi: 10.1080/09585190601041026
- *Fox, A., and Poultney, V. (2020). Teacher professional learning through lesson study: teachers' reflections. *Int. J. Lesson Learn. Stud.*, 9, 397–412. doi: 10.1108/IJLLS-03-2020-0011
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., and Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* 38, 915–945. doi: 10.3102/00028312038004915
- *Gee, D., and Whaley, J. (2016). Learning together: practice-centred professional development to enhance mathematics instruction. *Math. Teach. Educ. Dev.*, 18, 87–99. doi: 10.3316/aeipt.214245
- *Gilissen, M. G. R., Knippels, M.-C. P. J., and van Joolingen, W. R. (2020). Bringing systems thinking into the classroom. *Int. J. Sci. Educ.*, 42, 1253–1280. doi: 10.1080/09500693.2020.1755741
- Goei, S. L., van Joolingen, W. R., Goetsch, F., Khaled, A., Coenen, T., in 't Veld, S. G. J. G., et al. (2021). Online lesson study: virtual teaming in a new normal. *Int. J. Lesson Learn. Stud.* 10, 217–229. doi: 10.1108/IJLLS-09-2020-0078
- *Goh, R., and Fang, Y. (2017). Improving English language teaching through lesson study: case study of teacher learning in a Singapore primary school grade level team. *Int. J. Lesson Learn. Stud.*, 6, 135–150. doi: 10.1108/IJLLS-11-2015-0037
- Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Res. Pap. Educ.* 22, 213–228. doi: 10.1080/02671520701296189
- *Groves, S., Doig, B., Vale, C., and Widjaja, W. (2016). Critical factors in the adaptation and implementation of Japanese lesson study in the Australian context. *ZDM*, 48, 501–512. doi: 10.1007/s11858-016-0786-8
- Gülhan, F. (2021). How is the lesson study model implemented in Turkey, what does it mean?: a systematic review. *Karaelmas J. Educ. Sci.* 9, 230–242. Available at: <https://dergipark.org.tr/tr/download/article-file/2031701>
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers Teach* 8, 381–391. doi: 10.1080/135406002100000512
- *Gutiérrez, S. B. (2015a). Collaborative professional learning through lesson study: identifying the challenges of inquiry-based teaching. *Issues Educ. Res.*, 25, 118–134. doi: 10.3316/informit.376688277371502

- *Gutierrez, S. B. (2015b). Teachers' reflective practice in lesson study: a tool for improving instructional practice. *Alberta J. Educ. Res.*, 61, 314–328. doi: 10.11575/ajer.v61i3.56087
- *Gutierrez, S. B. (2016). Building a classroom-based professional learning community through lesson study: insights from elementary school science teachers. *Prof. Dev. Educ.*, 42, 801–817. doi: 10.1080/19415257.2015.1119709
- Hadfield, M., and Jopling, M. (2016). Problematizing lesson study and its impacts: studying a highly contextualised approach to professional learning. *Teach. Teach. Educ.* 60, 203–214. doi: 10.1016/j.tate.2016.08.001
- Hallinger, P. (2014). Reviewing reviews of research in educational leadership: an empirical assessment. *Educ. Adm. Q.* 50, 539–576. doi: 10.1177/0013161X13506594
- *Handayani, R. D., Wilujeng, I., Prasetyo, Z. K., and Triyanto, . (2019). Building an indigenous learning community through lesson study: challenges of secondary school science teachers. *Int. J. Sci. Educ.*, 41, 281–296. doi: 10.1080/09500693.2018.1548789
- *Hapidin, Y., Hartati, S., and Nurani, Y. (2020). The continuous professional development for early childhood teachers through lesson study in implementing play based curriculum (case study in Jakarta, Indonesia). *Int. J. Innov.*, 12, 17–25.
- Hardwicke, T. E., Wallach, J. D., Kidwell, M. C., Bendixen, T., Crüwell, S., and Ioannidis, J. P. A. (2020). An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014–2017). *R. Soc. Open Sci.* 7:190806. doi: 10.1098/rsos.190806
- Hauge, K. (2021). "The researcher's role: an intervention study using lesson study in Norway" in *Pedagogy—Challenges, recent advances, new perspectives, and applications*. ed. H. Senol (London: IntechOpen)
- Hervas, G., and Medina, J. L. (2020). Key components of lesson study from the perspective of complexity: a theoretical analysis. *Teach. Teach. Theory Pract.* 26, 118–128. doi: 10.1080/13540602.2020.1745174
- *Howell, J. B., and Saye, J. W. (2016). Using lesson study to develop a shared professional teaching knowledge culture among 4th grade social studies teachers. *J. Soc. Stud. Res.*, 40, 25–37. doi: 10.1016/j.jssr.2015.03.001
- *Howell, J. B., and Saye, J. W. (2018). Integrating theory and practice: factors shaping elementary teachers' interpretation of an inquiry model for teaching social studies. *J. Soc. Stud. Res.*, 42, 201–214. doi: 10.1016/j.jssr.2017.04.003
- Hrastinski, S. (2021). Digital tools to support teacher professional development in lesson studies: a systematic literature review. *Int. J. Lesson Learn. Stud.* 10, 138–149. doi: 10.1108/IJLLS-09-2020-0062
- *Huang, R., Gong, Z., and Han, X. (2016). Implementing mathematics teaching that promotes students' understanding through theory-driven lesson study. *ZDM: the international journal on Math. Educ.*, 48, 425–439. doi: 10.1007/s11858-015-0743-y
- *Huang, R., Kimmins, D., Winters, J., and Rushton, G. (2020). Does a technology-assisted lesson study approach enhance teacher learning while eliminating obstacles of traditional lesson study? *Contemp. Issues Technol. Teach. Edu.*, 20, 618–659.
- *Huang, R., Prince, K. M., and Barlow, A. T. (2017). Improving mathematics teaching as deliberate practice through Chinese lesson study. *Math. Educ.*, 26, 32–55.
- *Huang, R., Zhang, Q., Chang, Y., and Kimmins, D. (2019). Developing students' ability to solve word problems through learning trajectory-based and variation task-informed instruction. *ZDM*, 51, 169–181. doi: 10.1007/s11858-018-0983-8
- Ishii, T. (2017). "Historical overview of lesson study" in *Curriculum, instruction and assessment in Japan: Beyond lesson study*. eds. K. Tanaka, K. Nishioka and T. Ishii. 1st ed (London: Routledge)
- Jacobs, V. R., Lamb, L. L., and Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *J. Res. Math. Educ.* 41, 169–202. doi: 10.5951/jresmetheduc.41.2.0169
- *Jiang, H., Choy, B. H., and Lee, C. K.-E. (2020). Refining teaching expertise through analysing students' work: a case of elementary mathematics teacher professional learning during lesson study in Singapore. *Prof. Dev. Educ.*, 46, 731–750. doi: 10.1080/19415257.2019.1634624
- *Joubert, J., Callaghan, R., and Engelbrecht, J. (2020). Lesson study in a blended approach to support isolated teachers in teaching with technology. *ZDM: the international journal on Math. Educ.*, 52, 907–925. doi: 10.1007/s11858-020-01161-x
- Kager, K., Jurczok, A., Bolli, S., and Vock, M. (2022). "We were thinking too much like adults": examining the development of teachers' critical and collaborative reflection in lesson study discussions. *Teach. Teach. Educ.* 113, 1–13. doi: 10.1016/j.tate.2022.103683
- Kager, K., Kalinowski, E., Jurczok, A., and Vock, M. (2021). *A systematic review of methodological transparency in lesson study: how do we report on the observation and reflection stages?* Open Science Framework. Charlottesville, VA
- *Kanellopoulou, E.-M., and Darra, M. (2018a). The planning of teaching in the context of lesson study: research findings. *Int. Educ. Stud.*, 11, 67–82. doi: 10.5539/ies.v11n2p67
- *Kanellopoulou, E.-M., and Darra, M. (2018b). The implementation of the lesson study approach to secondary education in Greece: the case of the literature lesson. *Int. J. Learn. Teach. Educ. Res.*, 17, 94–105. doi: 10.26803/ijlter.17.7.6
- *Kanellopoulou, E.-M., and Darra, M. (2018c). Making use of the lesson study approach to secondary education in Greece—the contribution to the professional development of teachers. *Int. Educ. Stud.*, 11, 78–85. doi: 10.5539/ies.v11n12p78
- *Karabağ, F., and Ilin, G. (2019). Practicing lesson study in a Turkish education context: considering the challenges, suggestions and benefits from EFL teachers' perspectives. *Int. J. Lesson Learn. Stud.*, 8, 60–78. doi: 10.1108/IJLLS-05-2018-0036
- *Karlsen, A. M. F., and Helgevd, N. (2019). Lesson study: analytic stance and depth of noticing in post-lesson discussions. *Int. J. Lesson Learn. Stud.*, 8, 290–304. doi: 10.1108/IJLLS-04-2019-0034
- Karlsen, A. M. F., and Ohna, S. E. (2021). Pupils' voices in teachers' collaborative professional learning in lesson study. *Int. J. Educ. Res.* 110:101877. doi: 10.1016/j.ijer.2021.101877
- *Khokhotva, O. (2018). Lesson study in Kazakhstan: case study of benefits and barriers for teachers. *Int. J. Lesson Learn. Stud.*, 7, 250–262. doi: 10.1108/IJLLS-04-2018-0021
- *Khokhotva, O., and Elexpuru Albizuri, I. (2019). Student voice in lesson study as a space for EFL teachers' learning: a case study in Kazakhstan. *Int. J. Lesson Learn. Stud.*, 9, 153–166. doi: 10.1108/IJLLS-06-2019-0054
- Kim, J. (2021). "Through foreign eyes: a critical understanding of lesson study-based teacher education in Japan" in *Lesson study-based teacher education*. eds. J. Kim, N. Yoshida, S. Iwata and H. Kawaguchi (London and New York: Routledge), 9–28.
- *Kincal, R. Y., Ozan, C., and İleritürk, D. (2019). Increasing students' English language learning levels via lesson study. *Engl. Lang. Teach.*, 12, 88–94. doi: 10.5539/elt.v12n9p88
- *Klammer, S., and Hanfstringl, B. (2019). A first time lesson study that turned into a learning study. *Int. J. Lesson Learn. Stud.*, 8, 305–319. doi: 10.1108/IJLLS-04-2019-0029
- *Klefbeck, K. (2020). Lesson study for students with intellectual disability. *Int. J. Lesson Learn. Stud.*, 9, 245–259. doi: 10.1108/IJLLS-12-2019-0082
- *Kohlmeier, J., Howell, J., Saye, J., McCormick, T., Shannon, D., Jones, C., et al. (2020). Investigating teacher adoption of authentic pedagogy through lesson study. *Theory Res. Soc. Educ.*, 48, 492–528. doi: 10.1080/00933104.2020.1751761
- Korthagen, F. (2016). Inconvenient truths about teacher learning: towards professional development 3.0. *Teachers Teach.* 23, 387–405. doi: 10.1080/13540602.2016.1211523
- *Koutsouris, G., Norwich, B., Fujita, T., Ralph, T., Adlam, A., and Milton, F. (2017). Piloting a dispersed and inter-professional lesson study using technology to link team members at a distance. *Technol. Pedagog. Educ.*, 26, 587–599. doi: 10.1080/1475939X.2017.1364290
- Kunseeda, P., Inprasitha, M., and Changsri, N. (2018). Developing students' writing skills in mathematical problem-solving classroom using lesson study and open approach. *Asia Pac. J. Sci. Technol.* 23, 1–8. doi: 10.14456/kkurj.2018.3
- Larssen, D. L. S., Cajkler, W., Mosvold, R., Bjuland, R., Helgevd, N., Fauskanger, J., et al. (2018). A literature review of lesson study in initial teacher education: perspectives about learning and observation. *Int. J. Lesson Learn. Stud.* 7, 8–22. doi: 10.1108/IJLLS-06-2017-0030
- *Lawrence, M., Yang, L.-L., Briggs, M., Hession, A., Koussa, A., and Wagoner, L. (2016). Breathing life into engineering: a lesson study life science lesson. *Sci. Act. Classr. Prof. Curric. Ideas*, 53, 137–146. doi: 10.1080/00368121.2016.1211079
- *Lee, V., and Madden, M. (2019). "We're in this together": principals and teachers as partners and learners in lesson study. *NASSP Bull.*, 103, 51–64. doi: 10.1177/0192636519826717
- *Lee, L. H. J., and Tan, S. C. (2020). Teacher learning in lesson study: affordances, disturbances, contradictions, and implications. *Teach. Teach. Educ.*, 89, 102986. doi: 10.1016/j.tate.2019.102986
- *Leifler, E. (2020). Teachers' capacity to create inclusive learning environments. *Int. J. Lesson Learn. Stud.*, 9, 221–244. doi: 10.1108/IJLLS-01-2020-0003
- *Lelinge, B., and Svensson, C. (2020). Teachers' awareness and understanding of students' content knowledge of geometric shapes. *Probl. Edu. 21st Cent.*, 78, 777–798. doi: 10.33225/pec/20.78.777
- *Leong, S. S. M., Said, H. M., Shahrill, M., and Perera, J. S. H. Q. (2016). Using lesson study to enhance meaningful understanding on the topic of pressure. *nt. J. Environ. Sci. Educ.*, 11, 8425–8437.
- Lewis, C. (2009). What is the nature of knowledge development in lesson study? *Educ. Action Res.* 17, 95–110. doi: 10.1080/09650790802667477
- Lewis, C. (2015). What is improvement science? Do we need it in education? *Educ. Res.* 44, 54–61. doi: 10.3102/0013189X15570388
- Lewis, J. (2016). Learning to lead, leading to learn: how facilitators learn to lead lesson study. *ZDM: the international journal on. Math. Educ.* 48, 527–540. doi: 10.1007/s11858-015-0753-9
- Lewis, C., Friedkin, S., Emerson, K., Henn, L., and Goldsmith, L. (2019). How does lesson study work? Toward a theory of lesson study process and impact. In R. Huang, A. Takahashi and Ponte, P. da (Eds.), *Theory and practice of lesson study in mathematics* (pp. 13–37). Springer International Publishing, London
- 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. Research in mathematics education*. eds. J. Middleton, J. Cai and S. Hwang (Cham: Springer), 133–158.
- Lewis, C., and Perry, R. (2017). Lesson study to scale up research-based knowledge: a randomized, controlled trial of fractions learning. *J. Res. Math. Educ.* 48, 261–299. doi: 10.5951/jresmetheduc.48.3.0261

- Lewis, C., Perry, R., and Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educ. Res.* 35, 3–14. doi: 10.3102/0013189X035003003
- *Lewis, J., and Potts, P. (2019). Collaborative practitioner research: a case of lesson study in a small Jewish school. *J. Jew. Educ.*, 85, 209–220. doi: 10.1080/15244113.2019.1599609
- *Lim, C. S., Kor, L. K., and Chia, H. M. (2016). Revitalising mathematics classroom teaching through lesson study (LS): a Malaysian case study. *ZDM*, 48, 485–499. doi: 10.1007/s11858-016-0779-7
- Limon, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: a critical appraisal. *Learn. Instr.* 11, 357–380. doi: 10.1016/S0959-4752(00)00037-2
- Littleton, K., and Mercer, N. (2013). *Interthinking: Putting talk to work*. London: Routledge.
- *Liu, S.-H. (2016). Development of lesson study approach in three rural elementary schools of Taiwan. *Asia Pac. Soc. Sci. Rev.*, 16, 101–111.
- *Loc, N. P., Uyen, B. P., Tong, D. H., and Ngoi, H. T. (2020). A teaching process of fostering students' problem-solving skills: a case study of teaching the equation of a line. *Univ. J. Educ. Res.*, 8, 1741–1751. doi: 10.13189/ujer.2020.080510
- *Lomibao, L. S. (2016). Enhancing mathematics teachers' quality through lesson study. *Springerplus*, 2016, 1590. doi: 10.1186/s40064-016-3215-0
- *Lucenari, J. L. S., Yangco, R. T., Punzalan, A. E., and Espinosa, A. A. (2016). Pedagogical content knowledge-guided lesson study: effects on teacher competence and students' achievement in chemistry. *Educ. Res. Int.*, 2016, 1–9. doi: 10.1155/2016/6068930
- *Lundbäck, B., and Egerhag, H. (2020). Lesson study as a bridge between two learning contexts. *Int. J. Lesson Learn. Stud.*, 9, 289–299. doi: 10.1108/IJLLS-02-2020-0006
- *Ma'rufi, M. R., and Muhammad, I. (2019). Pedagogical content knowledge (PCK) among mathematics teachers on function materials through lesson study in junior high school. *Int. J. Innov. Creat. Change*, 9, 135–156.
- Makel, M. C., and Plucker, J. A. (2014). Facts are more important than novelty: replication in the education sciences. *Educ. Res.* 43, 304–316. doi: 10.3102/0013189X14545513
- Mann, S., and Walsh, S. (2013). RP or 'RIP': a critical perspective on reflective practice. *Appl. Linguist. Rev.* 4, 291–315. doi: 10.1515/applrev-2013-0013
- Marton, F. (2015). *Necessary conditions of learning*. London: Routledge.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem. Med.* 22, 276–282.
- Mercer, N. (2000). *Words and minds: How we use language to think together*. London: Routledge.
- Meyrick, J. (2006). What is good qualitative research?: a first step towards a comprehensive approach to judging rigor/quality. *J. Health Psychol.* 11, 799–808. doi: 10.1177/13591053060666643
- *Midgett, A. J., Ilten-Gee, R., Powers, D. W., Murata, A., and Nucci, L. (2018). Using lesson study in teacher professional development for domain-based moral education. *J. Moral Educ.*, 47, 1–21. doi: 10.1080/03057240.2018.1445982
- *Moghaddam, A., Sarkar Arani, M. R., and Kuno, H. (2015). A collaborative inquiry to promote pedagogical knowledge of mathematics in practice. *Issues Educ. Res.*, 25, 170–186.
- Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. PRISMA Group (2009). Reprint—preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Phys. Ther.* 89, 873–880. doi: 10.1093/ptj/89.9.873
- Moravcsik, A. (2020). *Transparency in qualitative research*. London: SAGE Publications.
- *Moss, J., Hawes, Z., Naqvi, S., and Caswell, B. (2015). Adapting Japanese lesson study to enhance the teaching and learning of geometry and spatial reasoning in early years classrooms: a case study. *ZDM: the international journal on Math. Educ.*, 47, 377–390. doi: 10.1007/s11858-015-0679-2
- Murphy, R., Weinhardt, F., Wyness, G., and Rolfe, H. (2017). Lesson study: evaluation report and executive summary. *Edu. Endowment Found.*:60. doi: 10.1016/0273-2297(83)90007-2
- Myers, J. (2012). The effects of lesson study on classroom observations and perceptions of lesson effectiveness. *J. Effect. Teach.* 12, 94–104.
- Mynott, J. P. (2019). Lesson study outcomes: a theoretical model. *Int. J. Lesson Learn. Stud.* 8, 117–134. doi: 10.1108/IJLLS-08-2018-0057
- Mynott, J. P., and Michel, D. (2022). The invisible leader: facilitation in lesson study. *Educ. Process Int. J.* 11, 48–61. doi: 10.22521/edupij.2022.113.3
- *Ni Shuilleabhain, A., and Seery, A. (2018). Enacting curriculum reform through lesson study: a case study of mathematics teacher learning. *Prof. Dev. Educ.*, 44, 222–236. doi: 10.1080/19415257.2017.1280521
- Norwich, B. (2018). Making sense of international variations in lesson study and lesson study-like practices: an exploratory and conceptual perspective. *Int. J. Lesson Learn. Stud.* 7, 201–216. doi: 10.1108/IJLLS-02-2018-0007
- Norwich, B., Benham-Clarke, S., and Goei, S. L. (2021). Review of research literature about the use of lesson study and lesson study-related practices relevant to the field of special needs and inclusive education. *Eur. J. Spec. Needs Educ.* 36, 309–328. doi: 10.1080/08856257.2020.1755929
- *Norwich, B., Fujita, T., Adlam, A., Milton, F., and Edwards-Jones, A. (2018). Lesson study: an inter-professional collaboration approach for educational psychologists to improve teaching and learning. *Educ. Psychol. Pract.*, 34, 370–385. doi: 10.1080/02667363.2018.1468733
- *Norwich, B., Koutsouris, G., Fujita, T., Ralph, T., Adlam, A., and Milton, F. (2016). Exploring knowledge bridging and translation in lesson study using an inter-professional team. *Int. J. Lesson Learn. Stud.*, 5, 180–195. doi: 10.1108/IJLLS-02-2016-0006
- Nosek, B. A., Spies, J. R., and Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspect. Psychol. Sci.* 7, 615–631. doi: 10.1177/1745691612459058
- O'Connor, C., and Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *Int J Qual Methods* 19:160940691989922. doi: 10.1177/1609406919899220
- *Obara, S., and Bikai, N. (2019). Promoting math teacher active learning with the lesson study approach: a case study of in-service teachers' perspectives. *Int. J. Lesson Learn. Stud.*, 8, 135–148. doi: 10.1108/IJLLS-11-2018-0088
- *Ogegbo, A. A., Gaigher, E., and Salagaram, T. (2019). Benefits and challenges of lesson study: a case of teaching physical sciences in South Africa. *S. Afr. J. Educ.*, 39, 1–9. doi: 10.15700/saje.v39n1a1680
- Ouzzani, M., Hammady, H., Fedorowicz, Z., and Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Syst. Rev.* 5:210. doi: 10.1186/s13643-016-0384-4
- *Özaltun Çelik, A., and Bukova Güzel, E. (2018). Describing lesson study designed for improvement of mathematics teachers' knowledge of student thinking. *Int. J. Math. Teach. Learn.*, 19, 176–204. doi: 10.4256/ijmt.v19i2.53
- Özaltun Çelik, A., and Guzel, E. B. (2020). How to improve a mathematics teacher's ways of triggering and considering divergent thoughts through lesson study. *Int. Electron. J. Math. Educ.* 15, 1–14. doi: 10.29333/iejme/8461
- *Özdemir, S. M. (2019). Implementation of the lesson study as a tool to improve students' learning and professional development of teachers. *Particip. Educ. Res.*, 6, 36–53. doi: 10.17275/per.19.4.6.1
- *Pang, J. (2016). Improving mathematics instruction and supporting teacher learning in Korea through lesson study using five practices. *ZDM: the international journal on Math. Educ.*, 48, 471–483. doi: 10.1007/s11858-016-0768-x
- Parks, A. N. (2008). Messy learning: preservice teachers' lesson-study conversations about mathematics and students. *Teach. Teach. Educ.* 24, 1200–1216. doi: 10.1016/j.tate.2007.04.003
- *Pehlivan, F. C., and Guzel, E. B. (2020). Development of mathematics teachers' moves that support students' reasoning through lesson study. *Turk. J. Comput. Math. Educ.* doi: 10.16949/turkbilmat.683535
- *Pella, S. (2015). Pedagogical reasoning and action: affordances of practice-based teacher professional development. *Teach. Educ. Q.*, 42, 81–111.
- Penuel, W. R., Fishman, B. J., Yamaguchi, R., and Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *Am. Educ. Res. J.* 44, 921–958. doi: 10.3102/0002831207308221
- Piaget, J. (1985). *The equilibrium of cognitive structures*. University of Chicago Press. Chicago
- Posner, G. J., Strike, K. A., Hewson, P. W., and Gertzog, W. A. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Sci. Educ.* 66, 211–227. doi: 10.1002/see.3730660207
- Quaresma, M., and Da Ponte, J. P. (2021). Developing collaborative relationships in lesson study. *PNA. Revista de Investigación En Didáctica de La Matemática* 15, 93–107. doi: 10.30827/pna.v15i2.16487
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>
- Radovic, D., Black, L., Williams, J., and Salas, C. E. (2018). Towards conceptual coherence in the research on mathematics learner identity: a systematic review of the literature. *Educ. Stud. Math.* 99, 21–42. doi: 10.1007/s10649-018-9819-2
- *Rahim, S. S. A., Sulaiman, S., and Sulaiman, T. (2015). Teacher professional development through lesson study in secondary schools. *Adv. Sci. Lett.*, 21, 2360–2364. doi: 10.1166/asl.2015.6276
- *Rochintaniawati, D., Riandi, R., Kestiniaty, J., Kindy, N., and Rukayadi, Y. (2019). The analysis of biology teachers' technological pedagogical content knowledge development in lesson study in West Java Indonesia. *J. Pendidikan IPA, Indonesia*, 8, 201–210. doi: 10.15294/jpii.v8i2.19303
- Rosenshine, B. (1994). The conduct and reporting of intervention studies. *Educ. Psychol. Rev.* 6, 245–254. doi: 10.1007/BF02213186
- *Rozimela, Y. (2020). Developing teachers' professionalism through school initiative-based lesson study. *Eur. J. Educ. Res.*, 9, 1513–1526. doi: 10.12973/eu-jer.9.4.1513
- Rzejak, D. (2019). "Zur Wirksamkeit von lesson study [on the effectiveness of lesson study]" in *Lesson study. Das Handbuch für kollaborative Unterrichtsentwicklung und Lernforschung [the manual on collaborative development of instructions and research into learning]*. eds. C. Mewald and E. Rauscher (Innsbruck: Studien Verlag), 97–111.

- Saito, E., Watanabe, M., Gillies, R., Someya, I., Nagashima, T., Sato, M., et al. (2015). School reform for positive behaviour support through collaborative learning: utilising lesson study for a learning community. *Camb. J. Educ.* 45, 489–518. doi: 10.1080/0305764X.2014.988684
- *Sato, T. (2020). Japanese elementary teachers' professional development experiences in physical education lesson study. *Phys. Educ. Sport Pedagog.* 25, 137–153. doi: 10.1080/17408989.2019.1692808
- 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
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schön, D. A. (1995). *The reflective practitioner: How professionals think in action*. London: Ashgate.
- *Seino, T., and Foster, C. (2020). Analysis of the final comments provided by a knowledgeable other in lesson study. *J. Math. Teach. Educ.*, 24, 507–528. doi: 10.1007/s10857-020-09468-y
- Seleznov, S. (2018). Lesson study: an exploration of its translation beyond Japan. *Int. J. Lesson Learn. Stud.* 7, 217–229. doi: 10.1108/IJLLS-04-2018-0020
- Seleznov, S. (2019). Lesson study beyond Japan: evaluating impact. *Int. J. Lesson Learn. Stud.* 8, 2–18. doi: 10.1108/IJLLS-09-2018-0061
- Seleznov, S., Goei, S. L., and Ehren, M. (2021). International policy borrowing and the case of Japanese lesson study: culture and its impact on implementation and adaptation. *Prof. Dev. Educ.* 50, 1–15. doi: 10.1080/19415257.2021.1973069
- *Shanmugam, S. K. S., Chinnappan, M., and Leong, K. E. (2020). Action research in examining the enquiry approach of lesson study in mathematics. *Pertanika J. Soc. Sci. Hum.*, 28, 1675–1693.
- Sherin, M. G., and Han, S. Y. (2004). Teacher learning in the context of a video club. *Teach. Teach. Educ.* 20, 163–183. doi: 10.1016/j.tate.2003.08.001
- *Shi, L., and Cheng, E. C. K. (2020). Developing metacognitive teaching in Chinese language through conducting lesson study in Shanghai. *Int. J. Lesson Learn. Stud.*, 10, 75–88. doi: 10.1108/IJLLS-09-2020-0065
- *Sjunnesson, H. (2020). Initializing phase of lesson study: communication a special didactic tool in mathematics. *Int. J. Lesson Learn. Stud.*, 9, 261–275. doi: 10.1108/IJLLS-02-2020-0007
- *Skott, C. K., and Møller, H. (2017). The individual teacher in lesson study collaboration. *Int. J. Lesson Learn. Stud.*, 6, 216–232. doi: 10.1108/IJLLS-10-2016-0041
- *Skott, C. K., and Møller, H. (2020). Adaptation of lesson study in a Danish context: displacements of teachers' work and power relations. *Teach. Teach. Educ.*, 87, 102945. doi: 10.1016/j.tate.2019.102945
- Stepanek, J., Appel, G., Leong, M., Turner Mangan, M., and Mitchell, M. (2007). *Leading lesson study: A practical guide for teachers and facilitators*. Corwin Press. Thousand Oaks, CA
- Stigler, J. W., and Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Summit Books. Mandaluyong
- Stigler, J. W., and Hiebert, J. (2016). Lesson study, improvement, and the importing of cultural routines. *ZDM: the international journal on Math. Educ.* 48, 581–587. doi: 10.1007/s11858-016-0787-7
- *Suh, J., and Seshaiyer, P. (2015). Examining teachers' understanding of the mathematical learning progression through vertical articulation during lesson study. *J. Math. Teach. Educ.*, 18, 207–229. doi: 10.1007/s10857-014-9282-7
- *Suhirman, L. (2018). Lesson study-based instruction for enhancing EFL teacher's pedagogical competence. *Asian EFL J.* 20, 154–174. doi: 10.5430/wjel.v13n7p10
- *Suhirman, L. (2019). Building professional development for peer coaching and collaborating to plan research lesson: case of Japanese lesson study approach. *Asian EFL J.*, 21, 9–32.
- *Suhirman, L. (2020). Lesson study—a Japanese instructional approach: empowering student – centered learning. *Asian EFL J.*, 27, 209–229.
- Takahashi, A. (2014). The role of the knowledgeable other in lesson study: examining the final comments of experienced lesson study practitioners. *Math. Teach. Educ. Dev.* 16, 4–21. doi: 10.3316/aeipt.205651
- Takahashi, A., and McDougal, T. (2016). Collaborative lesson research: maximizing the impact of lesson study. *ZDM* 48, 513–526. doi: 10.1007/s11858-015-0752-x
- *Tan, R. M. (2015). Improving the use of physical manipulatives in teaching science concepts through lesson study. *Int. J. Lesson Learn. Stud.*, 4, 328–342. doi: 10.1108/IJLLS-01-2015-0001
- *Thinwianthong, S., Eddy, C. M., and Inprasitha, M. (2020). Mathematics teachers' abilities in developing formative assessment after the introduction of lesson study and open approach innovations. *Malays. J. Learn. Instr.*, 17, 101–132.
- *Tsukui, A., and Saito, E. (2018). Stroll into students' learning: acts to unload teachers' values through the practices of lesson study for learning community in Vietnam. *Improv. Sch.*, 21, 173–186. doi: 10.1177/1365480217717530
- Ustuk, Ö., and De Costa, P. I. (2021). Reflection as meta-action: lesson study and EFL teacher professional development. *TESOL J.* 12:e531. doi: 10.1002/tesj.531
- *Vale, C., Widjaja, W., Doig, B., and Groves, S. (2019). Anticipating students' reasoning and planning prompts in structured problem-solving lessons. *Math. Educ. Res. J.*, 31, 1–25. doi: 10.1007/s13394-018-0239-5
- Van Dijk, W., Schatschneider, C., and Hart, S. H. (2021). Open science in education sciences. *J. Learn. Disabil.* 54, 139–152. doi: 10.1177/0022219420945267
- Van Es, E. (2011). “A framework for learning to notice student thinking” in *Mathematics teacher noticing. Seeing through teachers' eyes*. eds. M. G. Sherin, V. R. Jacobs and R. A. Philipp (London: Routledge), 164–181.
- Van Es, E., and Sherin, M. (2002). Learning to notice: scaffolding new teachers' interpretations of classroom interactions. *J. Inf. Technol. Teach. Educ.* 10, 571–596.
- VERBI Software (2019). MaxQda 2020. VERBI Software. Available at: <https://www.maxqda.com>
- *Verhoef, N. C., Coenders, F., Pieters, J. M., van Smaalen, D., and Tall, D. O. (2015). Professional development through lesson study: teaching the derivative using GeoGebra. *Prof. Dev. Educ.*, 41, 109–126. doi: 10.1080/19415257.2014.886285
- Vrikki, M., Warwick, P., Vermunt, J. D., Mercer, N., and van Halem, N. (2017). Teacher learning in the context of Lesson Study: A video-based analysis of teacher discussions. *Teach. Teach. Educ.* 61, 211–224. doi: 10.1016/j.tate.2016.10.014
- Wake, G., Adkins, M., Dalby, D., Hall, J., Joubert, M., Lee, G., et al. (2023). Centres for excellence in maths teaching for mastery randomised controlled trial: Evaluation report. Center for excellence in maths; nationalarchives.gov.uk. Available at: <https://www.et-foundation.co.uk/wp-content/uploads/2023/04/cfem-tfm-report.pdf>
- *Wang, H., Gibson, C., and Zander, U. (2020). Editors' comments: is research on corporate social responsibility undertheorized? *Acad. Manag. Rev.*, 45, 1–6. doi: 10.5465/amr.2019.0450
- *Warwick, P., Vrikki, M., Færøyvik Karlsen, A. M., Dudley, P., and Vermunt, J. D. (2019). The role of pupil voice as a trigger for teacher learning in lesson study professional groups. *Camb. J. Educ.*, 49, 435–455. doi: 10.1080/0305764X.2018.1556606
- *Warwick, P., Vrikki, M., Vermunt, J. D., Mercer, N., and Van Halem, N. (2016). Connecting observations of student and teacher learning: an examination of dialogic processes in lesson study discussions in mathematics. *ZDM*, 48, 555–569. doi: 10.1007/s11858-015-0750-z
- *Watanabe, T., Takahashi, A., and Barham, A. I. (2019). Implementing school-wide collaborative lesson research in Qatar. *Edu. Sci.* 17, 47–70. Available at: <http://hdl.handle.net/10576/14352>
- *Widjaja, W., Vale, C., Groves, S., and Doig, B. (2017). Teachers' professional growth through engagement with lesson study. *J. Math. Teach. Educ.*, 20, 357–383. doi: 10.1007/s10857-015-9341-8
- Wiggins, B. J., and Christopherson, C. D. (2019). The replication crisis in psychology: an overview for theoretical and philosophical psychology. *J. Theor. Philos. Psychol.* 39, 202–217. doi: 10.1037/teo0000137
- Willems, I., and Van den Bossche, P. (2019). Lesson study effectiveness for teachers' professional learning: a best evidence synthesis. *Int. J. Lesson Learn. Stud.* 8, 257–271. doi: 10.1108/IJLLS-04-2019-0031
- *Williams, J., and Ryan, J. (2020). On the compatibility of dialogism and dialectics: the case of mathematics education and professional development. *Mind Cult. Act.*, 27, 70–85. doi: 10.1080/10749039.2019.1686026
- Wolgemuth, J. R., Hicks, T., and Agosto, V. (2017). Unpacking assumptions in research synthesis: a critical construct synthesis approach. *Educ. Res.* 46, 131–139. doi: 10.3102/0013189X17703946
- *Won, N. (2017). Inner-city teachers' perceptions in a lesson study for critiquing mathematical reasoning. *Prof. Educ.*, 42, 58–79.
- *Wood, K., Jaidin, H., Jawawi, R., Perera, J. S. H. Q., Salleh, S., Shahrill, M., et al. (2017). How and what teachers learn from collaborative professional development. *Int. J. Lesson Learn. Stud.*, 6, 151–168. doi: 10.1108/IJLLS-09-2016-0028
- *Woranetsudathip, N., and Yuenyong, C. (2015). Enhancing grade 1 Thai students' learning about mathematical ideas on addition through lesson study and open approach. *Mediterr. J. Soc. Sci.*, 6, 28–33. doi: 10.5901/mjss.2015.v6n2s1p28
- Xu, H., and Pedder, D. (2014). “Lesson study: an international review of the research” in *Lesson study: Professional learning for our time*. ed. P. Dudley (London: Routledge), 29–59.
- *Yildiz, A., and Baltaci, S. (2017). Reflections from the lesson study for the development of techno-pedagogical competencies in teaching fractal geometry. *Eur. J. Educ. Res.*, 6, 41–50. doi: 10.12973/eu-jer.6.1.41
- *Yoke, W. W., and Phang, F. A. (2017). Implementation of lesson study among physics teachers: a preliminary study. *Man India*, 97, 265–274.
- Yoshida, N., Matsuda, M., and Miyamoto, Y. (2021). “The landscape of lesson study: a methodology for teachers' professional development and educational research” in *Lesson study-based teacher education: The potential of the Japanese approach in global settings*. eds. J. Kim, N. Yoshida, S. Iwata and H. Kawaguchi (London: Routledge), 29–50.
- *Zhang, Y. (2015). Sustaining lesson study in schools with positive peer leadership: a case study in Hong Kong. *Int. J. Lesson Learn. Stud.*, 4, 140–154. doi: 10.1108/IJLLS-07-2014-0018
- *Zhang, H., Yuan, R., and Liao, W. (2019). EFL teacher development facilitated by lesson study: a Chinese perspective. *TESOL Q.* 53, 542–552. doi: 10.1002/tesq.480
- Zimmerman, B. J. (2000). “Attaining self-regulation: a social cognitive perspective” in *Handbook of self-regulation*. eds. M. Boekaerts, P. R. Pintrich and M. Zeidner (New York: Academic Press), 13–39.