Vol. 14, No. 03, pp. 1443-1474, 2024 DOI: 10.23960/jpp.v14.i3.2024101

Jurnal Pendidikan Progresif

e-ISSN: 2550-1313 | p-ISSN: 2087-9849

http://jurnal.fkip.unila.ac.id/index.php/jpp/

Pedagogy Evaluations by Indonesian Teachers Implementing Quality Teaching Rounds: An Initial Exploratory Qualitative Study

Hannah-Charis O. O. Walker[134*], Niswanto[1], Nasir Usman[1], & Fahlida Harnita[2]

[1]Department of Education Administration, Universitas Syiah Kuala, Indonesia

[2]Department of Mathematics Education, Universitas Syiah Kuala, Indonesia

[3]Language Teachers Association, Tasmania, Australia member.

[4]British Educational Leadership, Management and Administration Society member

*Corresponding email: hannah.ach@mhs.usk.ac.id

Received: 27 September 2024 *Accepted: 12 November 2024* *Published: 25 November 2024*

Abstract: Pedagogy Evaluations by Indonesian Teachers Implementing Quality Teaching Rounds: An Initial Exploratory Qualitative Study. Objectives: QTR research in Indonesia has been limited to experimental research on selected aspects of the program. This preliminary study in a qualitative explanatory sequence aimed to describe Indonesian teachers' individual pedagogy evaluations during a high-fidelity application of the Quality Teaching Rounds (QTR) program, using comparisons with expert evaluations. A deductive, "top-down" approach to theme-finding provided hard data in this under-researched area, to be further investigated in later studies. **Methods:** Eight teachers from two schools were recruited to implement QTR. In school-based professional learning communities, teachers observed lessons, used the Quality Teaching Model to assign *kodes* (evaluations) to pedagogy using '*Koding* Sheets', then discussed their *kodes*. *Koding* Sheets were collected as primary data, and teacher initial *kodes* verified using discussion transcripts. Expert evaluations were created by an Australian internal expert using Quality Teaching Model rubrics, observation notes and consultations with Indonesian educators and an external expert. *Koding* Sheets were collated and colour-coded as 'the same', 'higher', or 'lower' than expert *kodes*. Frequencies were tabulated using Excel to find overarching themes or patterns addressing the research question. **Findings:** Themes included: (1) Dissimilarity of teacher *kodes* to expert *kodes*; (2) Prevalence of overestimating in *kodes*; (3) Trends of 'same' *kodes* increased for many and decreased for some; and (4) professional learning community-based differences in trends and counts. Results suggest social constructionism explains teacher learning leading to evaluation patterns, as well as the presence and influence of language, cultural, school contextual, relational factors on teacher pedagogy evaluations. **Conclusion:** Recommended research is outlined to confirm, identify, measure and determine how to mitigate these factors, to enable recommendations as to the suitability of QTR with the Indonesian context.

Keywords: evaluation, pedagogy evaluation, pedagogical knowledge, Quality Teaching Rounds, QTR.

To cite this article:

Walker, H. C. O. O. Niswanto., Usman, N., & Harnita, F. (2024). Pedagogy Evaluations by Indonesian Teachers Implementing Quality Teaching Rounds: An Initial Exploratory Qualitative Study. *Jurnal Pendidikan Progresif*, *14*(2), 1443-1474. doi: 10.23960/jpp.v14.i3.2024101.

INTRODUCTION

Problem

The goal of the Republic of Indonesia's education system (Constitutional Law: 1945/4) is "to make the nation intelligent", and in 79 years, the nation has "made significant achievements in education" including "steady improvements in enrolment rates and years of schooling" (Indrawati & Kuncoro, 2021: 31). However, the Organisation for Economic Co-operation and Development's (OECD, 2023) PISA testing of 80% of Indonesia's 15-year-old school students

showed they there were not yet on par with the rest of the world, ranking only 63 for science, 65 for mathematics and 66 for reading skills compared to 73 other countries. Although driven by a neoliberalism philosophy, treating education as tool to increase GDP (Silova, Rappleye & Auld, 2020), PISA results still have use as a baseline for Indonesia's educational output quality.

Numerous researchers and meta-analysts agree teachers are key factors in the quality of education (Darling-Hammond, 2013; Hattie, 2009; Hattie, 2023). Therefore, teachers are often prioritized in development efforts aimed to improve education quality (Gore, 2021; Gore & Rickards, 2021; Kennedy, 2016; Indrawati & Kuncoro, 2021; Musfah, 2015, Nazari, Nafissi, Estaji & Marandi, 2019). Successful development of teachers involves careful selection of professional development (PD) experiences for them. However, despite the current government and private sector driven PD regime in Indonesia (Rahman 2016), the nation's teachers are still lacking in subject knowledge and pedagogical skills (World Bank, 2015). Pedagogy is defined as "child education" (Wahyudi, Fakhrudiin & Ikmal, 2024) or the study of how to teach children and adolescents. One reason for this lack of gain in knowledge and skills from the current PD regime is lack of continuity: in 2008, Firman and Tola (80) noted that teachers were not implementing knowledge or skills gained through in-service PD, instead returning to "teaching in their conventional ways."

Proposed Solution

Quality Teaching Rounds (QTR) is one approach to teacher PD that has successfully improved teacher structures for thinking about pedagogy (Gore *et al*., 2016) as well as teachers' actual practice, thereby improving education quality in Australia. QTR is a PD package developed by Gore, Ladwig and the University of Newcastle for the Department of Education, NSW, Australia, and aims to improve teach*ing* rather than teach*ers* by avoiding top-down supervision approaches that reduce public and leadership trust in teachers as professionals (Gore, 2021; Gore & Rickards, 2021). By using the specific and objective language found in the Quality Teaching (QT) Model, Australian teachers were empowered as professionals to engage in rejuvenating discussions as they analysed and evaluated teaching and learning (Gore & Rickards, 2021; Gore, Rickards & Fray, 2023). The result was improved teacher practice and student outcomes, as shown by three longitudinal, quantitative studies (Amosa, Ladwig, Griffiths & Gore, 2007; Gore, Lloyd, Smith, Bowe, Ellis & Lubans, 2017; Miller, Gore, Wallington, Harris, Prieto-Rodriguez & Smith, 2019).

QTR deliberately situates PD amongst the teachers themselves, with formation of Professional Learning Communities (PLCs) as safe spaces for 3-5 teachers to learn and build knowledge together. One or more members of the PLC attend a two-day workshop to learn about the QT Model, then return to school to implement observational rounds. School administrative support is necessary to free teachers up for one day per PLC member, over 1-3 months. A "Rounds Day" is described (Gore, Miller, Fray, Harris, & Prieto, 2021) as including:

1. Professional Reading Discussion (30-60 minutes)
2. Lesson Observation (of one teacher's entire lesson)
3. Individual Coding (*Koding*) by all teachers, including the one who taught the lesson (30 minutes)
4. Professional Learning Community (PLC) Discussion (120 minutes)

As seen in point 2, QTR involves classroom visits by PLC member teachers to learn to pay attention to and evaluate pedagogy quality. They then use rubrics and detailed explanations from the Classroom Practice Guide (CPG) (NSW DET, 2006) to objectively give a score or code

(labelled *kode* for this paper) for each of the 18 Elements of the QT Model. The ensuing discussion involves teachers sharing their individual *kode*, then engaging in discussion as they come to consensus on the best *kode* for each Element, based on the example of pedagogy they have just witnessed. The aim of this exercise is not criticism or even supervision, which is described by Usman and Murniati (2019: 158), as aiming to "improve quality and performance" but in Indonesia is often (Anwar, Harun & Niswanto, 2022: 4043) "inspectional" in nature, where teachers' "mistakes" are noted by the supervisor "without coaching, justification" or the giving of "professional assistance to teachers through… fast and objective feedback". Instead, the aim of the coding process is the development of each teacher's ability to notice different aspects of pedagogy and recognise various levels of quality regarding those aspects.

A key feature of QTR is the Quality Teaching (QT) Model, developed from research in previous models of Authentic Pedagogy in USA and Productive Pedagogy in Australia. Newmann and Wehlage's (1995) Authentic Pedagogy framework improved student results, regardless of social background, in a study of over 1500 American schools. Newmann and Wehlage (1995:10) stressed: "a vision for high quality student learning is a necessary guide, but *not sufficient*. Teachers must *teach* according to the vision" (italics by the authors). To achieve this, teachers need positive support at school administrational, agency and policy levels. Lingard, Hayes and Mills (2003) built upon the Authentic Pedagogy framework to develop the Productive Pedagogies model. Lingard *et al.'s* large study (2001) in Australia from 1998-2000 identified twenty teacher practices that stimulated increased student results, both academically and socially, and re-affirmed that this teacher PD had positive effects on the quality of pedagogy. In 1993, Gore and Ladwig further refined and developed Newmann's and Lingard's models into the QT Model containing 18 Elements of classroom practice, grouped into three Dimensions, together forming a framework or lens regarding best practice in pedagogy.

Prior Study

Can the use of QTR improve education in Indonesia, too? Gore (2021) says further studies outside Australia are still needed to examine the level of suitability of QTR for improving quality of teacher pedagogy other nations. Rahman (2016:2) also warns that "it cannot be assumed" that recommendations for teacher PD "will work successfully in every context". The only relevant study in Indonesia is a doctoral study whose findings have not yet been published (Djulete, 2021). Djulete (2021) investigated in detail the impact of only two aspects of QTR, namely PLCs and the QT Model, on pedagogical content knowledge (PCK), as a cause of improving higher order thinking (HOT) skills. Her study focused on a rural location in the Eastern side of Indonesia, and aimed to both identify English teachers' PCK and develop the quality of their teaching through improving their PCK regarding HOT teaching in EFL classes.

Research Gap

It remains unknown how suitable QTR as a whole package is for application in an Indonesian context, especially for mixed groups of teachers from various disciplines, as per the original design. High fidelity to the original design would include the formation of inter-disciplinary PLCs, since QTR's original aim focuses on pedagogical knowledge (PK), not on pedagogical content knowledge (PCK).

Pedagogy, Knowledge and Evaluation

Pedagogical knowledge (PK) is part of the TPACK framework (Krisnaresanti, Ahman, Istikomah, Suwatno & Budiman, 2024) built from Shulman's (1987) seminal work on various kinds of teacher knowledge. The QT Model builds on

Lingard, Hayes and Mills' (2003: 404) concept of putting PK at the forefront of teacher PD in an effort to mitigate negative effects of top-down surveillance and accountability, highlighting instead "the role of teachers in the production of pedagogical knowledge." The QT Model (Lingard, Hayes & Mills, 2003: 405) enhances "teachers' professional language", highlighting "the role of teachers in the production of pedagogical knowledge" through the workshop, detailed Classroom Practice Guide and the iterative rounds of PLC members observing, evaluating, discussing and deciding on the quality of teaching observed. A connection is evident here between value judgements and the building of knowledge. During the Rounds Days (renamed PLC Day for this study), individual and group assigning of scores can lead to collective building of PK, and deeper PK leads to more accurate PEs.

Figure 1. PK and PE relationship proposition

The above diagram is proposed as the relationship between pedagogical knowledge (PK) and pedagogical evaluations (PE) during the QTR program. For example, lived experience in Australia, the QTR Foundational Workshop helps build teacher PK to enable teacher PE, and in the Quality Teaching Rounds, PLC discussions of teacher PEs tend to build teacher PK further. Whether or not PK and PE improvement are always directly proportional, they are intrinsically linked in the QTR program. Since it is harder to measure teacher pedagogical knowledge comprehensively before and after a QTR program, teacher pedagogy evaluations (PEs) are clear and easy to measure over time. Therefore, PK will not be the focus of this paper, although knowledge and understanding will be commented on at various points. For example, a comparison of Indonesian teacher evaluations with expert evaluations by an Australian-born researcher during the same program could reveal potential PK gain the more similar these evaluations were; and could reveal contextual influences the more dissimilar they were.

Evaluation has been defined as a HOT skill (Bloom et al., 1956; Krathwohl, 2002). In this paper, teacher pedagogy evaluation (PE) is defined narrowly as the act of giving a *kode* (score) to an example of teaching (pedagogy), based on the QT Model *koding* rubrics. Evaluation implies a value judgement, but the standard of valuing is not necessarily prescribed. However, since teacher pedagogy evaluations (PEs) are so inextricably linked to the QT Model and the PK associated with it, teacher PE in QTR is highly prescribed, and therefore open to evaluation in turn, or at the least, open to comparison with the standards, language and knowledge of the QT Model as understood by experts. Therefore, teachers' individual PEs can be further measured as being the 'same', 'lower', or 'higher' than an expert standard PE.

Hence, as a measure of teacher growth and improvement, their evaluations of pedagogy can be measured and described within a qualitative framework, as the proposed solution to the problem of low teacher pedagogical knowledge and practice. This is in line with the five steps of exploratory research (George, 20 November, 2023). As exploratory research, the original design developed into three studies as per Mansourian's (2008) description of research for the first time in a new area. In this case, the research's focus remained on individual and group PEs, in order to produce 'quick learning' (Gore, 2014) regarding the suitability of QTR as a PD approach in Indonesia. This paper reports on the first study, which examines the individual,

numerical teacher outputs as a preliminary indicator of understanding or misunderstanding of this introduced, foreign-designed PD program. The purpose of study 1, reported on in this article, is to provide hard data as a diving block for deeper exploratory qualitative inquiry. The research question is: *"In the PLC Discussions as part of the first Quality Teaching Rounds at two urban State Junior High Schools in Sumatra, Indonesia, how do individual teacher evaluations of pedagogy compare with expert evaluations?"*

METHOD

Participants

Indonesian schools can be grouped into government and private schools. The government schools in a particular city were selected as the population in order to control the variables between Indonesian private schools regarding curriculum, leadership structure and teacher selection and standards. Government school permanent teachers are part of the Indonesian civil service, described by Aruan (2015: 50) as having "a traditional model of personnel administration; a highly-centralised HRM system, 'career service' and security of tenure, lifelong employment, lower-ranks-focused recruitment and an internal based promotion system." The authors were thus able to control many variables by selecting two urban state schools. The Junior High level was selected over Primary level because of the study's aim to implement QTR with cross-disciplinary PLCs. Cross-disciplinary PLCs are hard to form in Primary schools since most Primary level teachers are not specialists in just one discipline but tend to teach 'Thematically'. Senior High Schools were eliminated from the study because teachers there tend to be busier, with no time for participation in research, and because Senior High Schools are governed by a Department of Education jurisdiction situated at a provincial, not city level, leading to less attention on local city-wide schools, which can affect researcher access to urban schools. Therefore, authors chose to conduct a explanatory study focused on government schools at the Junior High level in a regional urban centre in Sumatra, situated in the western part of Indonesia. In a regional urban centre in Sumatra, a more prosperous island overall compared with ENT province, and situated in the opposite side of Indonesia, west of the central island of Java and the nation's capital, Jakarta.

The subjects of this research were eight teachers from two schools that were selected not as a sample representative of all urban junior high schools in this city or in the nation of Indonesia, rather (Bryk, 2014: 473) as the basis of a short study with a "learning-by-doing orientation". In order to 'learn more,' two contrasting schools were selected, one ranking in the city's top five schools and the other in the bottom five, according to student end-of-year exam data from the city Department of Education.

The following variation of snowballing strategy after Palinkas *et al*. (2015) was applied to identify four teachers from each of two state Junior High Schools. Firstly, the principal of School Z was recruited from the authors' contacts. This is in line with Green and Aarons' example (2011, cited in Palinkas *et al*., 2015) of first recruiting program managers who could then be asked to point out potential staff for recruitment for the research project. Patton (2021, cited in Palinkas *et al*., 2015: 17) asserts that this strategy "begins by asking … well-situated people" a question such as "Who knows a lot about... ?" The strategy of starting with a key leader in the school is even more important when studying such a time-intensive program as QTR, since principal cooperation is necessary for participants to be freed from duties for six total days. In fact, Djulete (2021: xxii) found school leadership support was an important factor influencing the "gains" from PD programs

using PLCs. School Z was in the top five of the city's Junior High Schools.

Secondly, in order to obtain access to a school that was statistically low-achieving, the lowest five Junior High Schools in the city based on student results were listed, and School Z principal was asked, "Who is another principal who might be willing to run QTR in their school, from this list of five schools?" They selected a principal with whom they had the most rapport, and acted as a liaison to explain the opportunities of the QTR program with School Y's principal. School Y's principal met with the first author and expressed their willingness to open their school to the program. Thirdly, both principals suggested candidate participant teachers to answer the matter of "Who might be interested in, willing and available to participate in the QTR program?" School Z's principal suggested two teachers, one of whom suggested two further teachers. School Y's principal suggested 4 teachers, but 3 indicated they weren't willing or available to join, then three more were suggested by the leadership team. Fourthly, the research team met with each school's suggested teachers, explaining in detail QTR's goals and methods, the time commitment and benefits, and the data collection process, checking for interest, willingness and personal availability by giving ample opportunity for teachers to opt in or out of the program. Each of the eight teachers described in the table below expressed willingness to join and read, understood and signed data collection permission forms, the schools expressed a willingness to make the teachers available for the QTR program provided that official permissions were granted from the city Department of Education, which was applied for, and received. The schools also requested letters of invite to the two-day off-site workshop, and legitimised certificates with number of PD hours provided to the teachers to help them administrative professional development goals.

Table 1. Participating teachers, by subject taught

Subject taught	Gender	Age Range	School Role
English (EFL)	Female	40-43	Teacher, Healthy Lifestyle mentor
English (EFL)	Female	44-47	Teacher
Guidance and Counselling	Female	28-31	Teacher
Indonesian	Female	28-31	Teacher, School treasurer
Indonesian	Female	32-35	Teacher, Pedagogy Team member
Indonesian	Female	36-39	Teacher, Head of Pedagogy
Mathematics	Female	48-51	Teacher
Science	Female	44-47	Teacher, Vice-Principal of Curriculum

Research Design and Procedures

Bryk (2014: 473) points out that before getting too excited about a new program, researchers should be "starting small" and "learning quickly", especially since schools are complex systems steeped in "ambiguity". Research in a new area lends itself to the qualitative exploratory method (Trenholm-Jenson, Burns, Trenholm & Hand, 2022). A first-time implementation of QTR as a package though on a small scale was chosen to be researched from a "learning-by-doing orientation", intentionally maintaining high fidelity to the processes used in Australia in order to provide a complete picture of 'what does not work'. This is in accord with George's (20 November, 2023) five steps of

exploratory research. Step 1, Identify your problem and step 2, Hypothesize a solution, have been addressed in the Introduction. Step 3, Design your methodology, is addressed in the Method section. In short, study 1, reported on in this paper, utilises *a priori,* deductive thematic coding of teacher evaluations in a "top-down" manner (Xu & Zammit, 2020), to produce numeric mapping of individual teacher pedagogical evaluations as jumping-off point for study 2. Study 2 employs inductive coding to produce themes and subthemes from thematic analysis of more complex, complementary data. Study 3 arose mid-way in response to questions emerging from the process of inductively analysing the data, a phenomenon described by Mansourian (2008), and comprising robust research because still in line with the overall purpose of the inductive thematic analysis and the original research design. George's (20 November, 2023) step 4, Collect and analyse data, will be reported on in the Results and Discussion, while step 5, Avenues for future research will be addressed in the Conclusion.

This research proceeded with obtaining Faculty approval of the research proposal, after which, approval from the city's Department of Education (Primary and Junior High School level) to conduct research in the two schools, was also requested and granted. After participant selection and agreement, the two-day workshop was prepared and supporting documents translated. The CPG was already translated by Djulete and some edits were made with permission: front and back matter was added; the '*koding* scale' terms improved and 3 Element names adjusted. The authors collaborated with their university department to provide a venue, letters of invitation to the workshops and official numbers for certificates of completion of the entire program of 2 workshop days and 4 PLC days. Besides field notes made by the head researcher and assistant researchers each PLC Day for subsequent studies, the following data were collected each week, during the four 'PLC Days, per school's PLC:

1. Quality photographs of teachers' *Koding* Sheets showing kodes and evidence given by teachers individually, as well as final kodes agreed to by the PLC collectively. Teachers were asked to use two different coloured pens, one for the Individual Coding sessions, and the other for use during the PLC Discussion sessions, to make their initial individual *kodes* clear. The photographs were then typed using two colours, producing 32 typed Coding Sheet documents and including any crossed-out numbers in the Individual Kode section as sometimes the teachers forgot to switch their pen colours and would change their individual opinion over time as they discussed with colleagues. Teacher initial numbers were obtained by verifying the data with the audio recordings in point 2, below. Data from these documents were collated into spreadsheets for ease of comparing teacher Coding Sheets per PLC Day.
2. Audio recordings with supporting Video were taken of each PLC Day's session 4 (PLC Discussion), then transcribed over several weeks by seven transcribers from Sumatra, Indonesia, to produce eight Intelligent Verbatim transcripts.
3. During each PLC Day's 3 (Individual Coding), the researcher and assistants made completed separate *Koding* Sheets using the CPG and their detailed Lesson Observations.
4. At the end of the day and for several subsequent meetings, the research team would meet to discuss and agree on the 'expert' *kodes*. This process took some time as while watching back videos sometimes new evidence was found, any questions taken to

Quality Teaching Academy expert and expert kodes adjusted if necessary, based on new information. The evaluation tool used was the *Koding* Sheet (NSW DET, 2020:39) alongside the *koding* scales found in the Classroom Practice Guide (NSW DET, 2006:12-51) The first author, having joined the Quality Teaching Foundational workshop two times, delivered it once, consulted several times with workshop trainers from the Quality Teaching Academy, University of Newcastle, and undertaken postgraduate teacher education from the same worldview (La Trobe University, 2012), was considered to have built the pedagogical knowledge and skills to be able to read, understand and apply the coding scales from the quality teaching model as an internal expert. During the lesson observations, the first author took detailed notes, then engaged in further meetings and discussion with the Indonesian teachers, to mitigate language barriers and cross-check 'what happened' from their standing point. Where there was any confusion about the interpretation of the coding scales, the first author sought clarification a Quality Teaching Academy trainer as an external expert.

The schools were recruited in January and the two-day workshop delivered in Indonesian to all eight teachers in late January 2024. Each schools formed a PLC of four teachers each, and each teacher took turns hosting a 'PLC Day' for four weeks over the month of February.

Instruments

The Classroom Practice Guide (CPG) contains a double-page spread for each of the 18 QT Model Elements, with examples of high and low occurrences in a lesson, a koding scale, rubric, suggestions and further notes. The QT Model "is derived from a comprehensive review of empirical studies… and was refined through hours of classroom observation and sophisticated statistical analysis involving multilevel modelling and factor analysis" (Gore, 2020; Ladwig, 2007). The *koding* scale and rubric describing a sample Element from the QT Model as shown in the CPG is given in Table 2, with a translation into Indonesian in Table 3. Overviews of Rubric in Indonesian are also attached in Appendix 1.

The Lesson Observation Sheet used by the researchees and researchers can be found in Appendix 2 and was created by the Quality Teaching Academy and translated into Indonesian by the first author. The researchers tended to use both sides of the page to record many fine details of what was observed.

The *Koding* Sheet can be found in Appendix 3, translated from that found in the Classroom Practice Guide, with clear columns for the Individual Code (*Kode Individu*) and the PLC Code (*Kode PLC*).

Table 2. Coding scale & rubric descriptors for element 2.6 student direction

2.6 Student Direction	**Coding Scale:** To what extent do students exercise some direction over the selection of activities related to their learning and the means and manner by which these activities will be done?
	Descriptors
5	High student direction. Students determine many significant aspects of the lesson either independent of, or dependent on, teacher approval.
4	Substantial student direction. Some deliberation or negotiation occurs between teacher and students over at least some significant aspects of the lesson.
3	Some student direction. Students exercise some control in relation to some significant aspects of the lesson.

2	Low student direction. Although students exercise some control over some aspect of the lesson (choice, time, pace, assessment), their control is minimal or trivial.
1	No evidence of student direction. All aspects of the lesson are explicitly designated by the teacher for students.

Source: Quality Teaching Classroom Practice Guide, Third Edition. © State of NSW, Department of Education, 2020, Sydney, Australia, p. 32

Table 3. *Skala pengkodean dan indikator rubrik untuk aspek 2.6 pengarahan oleh peserta didik (bahasa indonesia)*

2.6 *Pengarahan oleh Peserta Didik*	**Skala Pengkodean:** *Sejauh mana anak didik mendapat kesempatan untuk memilih aktifitas yang berkaitan dengan pembelajaran mereka serta memilih cara menyelesaikan aktifitas tersebut?**
	Indikator
1	*Tidak terlihat bukti kemandirian peserta didik. Segala aspek dari pelajaran secara eksplisit di rancang oleh guru untuk siswa.*
2	*Rendahnya kemandirian siswa. Walaupun siswa diberikan beberapa control dalam beberap (sic) aspek pelajaran (piliha, waktu, pase, penilaian), control yang mereka punya minimal atau sedikit.*
3	*Terdapat sedikit pengarkemandirian (sic) siswa. Siswa mendapat sedikit control yangberhubungan degang (sic) beberapa aspek penting dalam pelajaran.*
4	*Kemandirian siswa yang cukup banyak. Negosiasi terjadi diantara guru dan siswa mengenai beberapa aspek penting dalam pelajaran.*
5	*Tingginya kemandirian siswa. Siswa menentukan banyak aspek penting dalam pelajaran baik secara mandiri atau bergantung pada persetujuan guru.*

Source: Quality Teaching Classroom Practice Guide, Third Edition. © State of NSW, Department of Education, 2020, Sydney, Australia, p. 32

Data Analysis

A short codebook (Table 4) was created, *a priori,* with "descriptive" colour-codes (Miles & Huberman, 1994: 58) not based on a particular theory but showing a "measurement" (Bernard, 1994:193) of degrees of comparison to expert *kodes*.

Before commencing data analysis, the research team created 'expert evaluations' of all eight lessons observed. The *Koding* Sheet data were verified against transcripts where gaps existed. Whenever teachers had written two numbers with their first pen or scratched out a first guess using the first pen, the researcher triangulated with the transcripts, taking the first number the teacher spoke aloud as their original individual *kode*, unless they implied that they had adapted their kode as they were speaking. In cases when a teacher didn't explicitly state their *kode*, the non-crossed out *kode* was taken. In rare cases where a teacher wrote two *kodes* and read neither aloud, the one closer to the expert *kode* was taken. Then experts' and verified teachers' *Koding* Sheets were collated into spreadsheets in Excel for Mac 16.80 (Table 5) for ease of comparison.

Table 4. Codebook of colour-codes for comparing to experts

Category	Colour-code	Description
'Lower'	Lower-4	Give when first *Kode Individu* is lower by 4 than expert *kode*
	Lower-3	Give when first *Kode Individu* is lower by 3 than expert *kode*
	Lower-2	Give when first *Kode Individu* is lower by 2 than expert *kode*
	Lower-1	Give when first *Kode Individu* is lower by 1 than expert *kode*
'Same'		Give when first K*ode Individu* is the same as expert *kode*
'Higher'	Higher-1	Give when first *Kode Individu* is higher by 1 than expert *kode*
	Higher-2	Give when first *Kode Individu* is higher by 2 than expert *kode*
	Higher-3	Give when first *Kode Individu* is higher by 3 than expert *kode*
	Higher-4	Give when first *Kode Individu* is higher by 4 than expert *kode*

Then the first author was able to apply the above nine colour-codes deductively (Fereday & Muir-Cochrane, 2006) to the '*Kode Individu"*, or initial individual *kodes* in a simple thematic analysis process as per Miles and Huberman (1994) to answer the research question from this paper. Teacher *kodes* were systema tically colour-coded as 'same' (orange), 'lower' than (blue), or 'higher' than the expert *kode* (pink) Blue and pink colour-codes were further differen tiated by shades relative to the numerical differen from the expert *kode*. Spreadsheets were triple-checked to eliminate human error. Below is an extract from one PLC Day's spreadsheet showing colour-coding of individual as well as PLC *kodes* (not discussed in this article). Due to the sensitive nature of this information, dates and teacher and school identities have been removed:

Table 5. *Aspek pedagogi: 2.6 pengarahan oleh peserta didik* [element of pedagogy: 2.6 student direction]

Teacher A			Teacher B			Teacher C			Teacher D			Experts	
Information	Individual Kode	PLC Kode	Information	Individual Kode	PLC Kode	Information	Individual Kode	PLC Kode	Information	Individual Kode	PLC Kode	Information	Expert Kode
Siswa dapat sedikit kontrol dalam beberapa aspek pelajaran [Students had some control in several aspects of the lesson]	3	2	*Pemilihan bacaan , waktu tidak ada* [Choice of readings, time/pace was absent]	2	2	*Kemandiri-an siswa terbatas/ rendah* [Student independ-ence was limited/ low]	3	2	*Negoisasi terjadi antar guru dan siswa mengenai beberapa aspek* [Negotiations occurred between teacher and students regarding some aspects]	4	2	Student direction is low. Students had some choice of roles within the group and pace of work as individuals/ groups, but the impact of these choices was minimal. Groups, teaching materials, task, criteria and order of activities were determined by the teacher: students didn't negotiate.	2

Numerical data deriving from a qualitative study can be displayed in graphs and tables, however because qualitative data represents "individual (mutually exclusive) categories, the descriptive statistics that can be calculated are limited" (Australian Bureau of Statistics, n.d.). This numeric data therefore was aggregated in tables to show frequency of occurrences in order

to more easily identify overall themes in the data. Teacher kodes were grouped and displayed per week, per teacher, and per QT Model Element. Because a difference of 1 had been considered 'on the right track' during the Quality Teaching Academy workshop, the number of times teachers gave kodes that were same OR only 1 kode above or below was also summed. Running tallies built into Excel in order to reveal any human errors, which were corrected immediately.

■ RESULT AND DISCUSSION

This study surmised that if indeed involvement in QTR could build teachers' pedagogical knowledge (PK), this growth in PK could be reflected in the teachers' pedagogy evaluations (PEs), here interpreted as the numeric *kodes*. Hence, the researchers evaluated QTR participant PEs against the expert PEs, using colour-coding, to answer the research question: *"In the PLC Discussions as part of the first Quality Teaching Rounds at two urban State Junior High Schools in Sumatra, Indonesia, how do individual teacher evaluations of pedagogy compare with expert evaluations?"*

Theme 1: Teacher PEs were more often dissimilar than 'same'

Surprisingly, teacher PEs according to lesson observations and using the QT Model rubrics were not as similar to expert PEs as may have been expected, either per PCL Day, per PLC or overall (for all eight teachers). The authors will refer to PLC Z and PLC Y hereon in. Figure 3 below provides an overview of all teachers' evaluations.

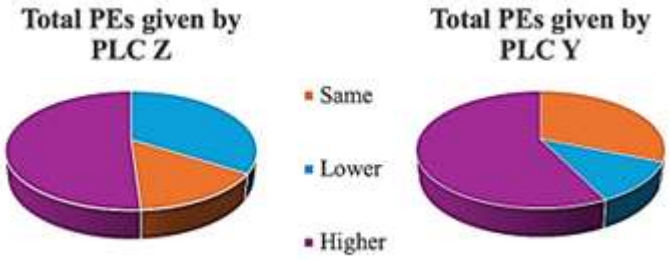


Figure 2. Total PEs given by teachers in PLC Z and PLC Y

More than half of the time, teacher PEs were not the 'same' as expert PEs. It is possible that similarity with expert PEs is normally low during first-time implementing of QTR if averaged across all teachers in a first time PLC, though such a low degree of similarity suggests that teacher understanding of the rubrics or of the QT Model was not optimal, and further suggests the presence of outside factors, skewing the results of evaluations, which will be discussed under the subsequent themes. As can be seen in Table 3, there were found some small errors in the rubric descriptors in Indonesian. These translation errors could have had some impact on teacher under standing or even give rise to misunderstandings. Further analysis of the discussion transcripts is necessary in order to ascertain how much impact language errors had on teacher PEs.

Theme 2: Teacher PEs more often 'higher' than 'lower'

During the QTR program, teachers were asked to individually evaluate 18 Elements of the QT Model, for four lessons, resulting in a total of 72 evaluations per teacher over the four weeks. Teacher evaluation *kodes* were expected to be at times higher and at times lower than those of the experts, reflecting an organic process of constructing understanding of the QT Model and

how to apply it, however, teacher PE was consistently higher than that of the research team. While the number of 'same' evaluations made over the four weeks ranged from 18 times (by Teacher Z2) to 31 times (by Teacher Z4), all except for Teacher Z4 made more PEs that were 'higher' than 'same'. In addition, all teachers without exception made more 'higher' than 'lower' PEs over the whole program, up to seven times more, as shown in Table 6:

Table 6. Summary: 'lower' versus 'higher' PEs, per PLC

		Total 'lower' PEs	Total 'higher' PEs	Ratio lower: higher
PLC Z	**Teacher Z1**	9	37	1:4
	Teacher Z2	13	40	1:3
	Teacher Z3	9	41	1:5
	Teacher Z4	12	29	1:2
PLC Y	**Teacher Y1**	10	41	1:4
	Teacher Y2	8	40	1:5
	Teacher Y3	6	43	1:7
	Teacher Y4	9	41	1:5

It had been imagined before delivering the QTR program that the distribution of teacher PE similarity would resemble a normal distribution, perhaps slightly skewed to the right or left. However, comparing teacher PEs to expert PEs revealed a strong skew to the right, as shown in Figure 4, which displays Teacher Y2's total PEs over four weeks:

Even viewed per week, a tendency to give 'higher' over 'lower' PE than experts is evident, as in this extract showing Teacher Z2 and Teacher Y2's PEs:

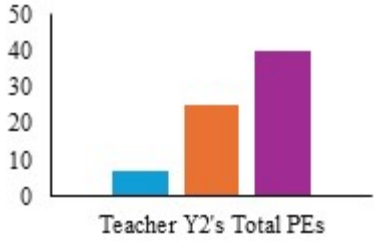


Figure 3. Sample teacher's total overall PEs

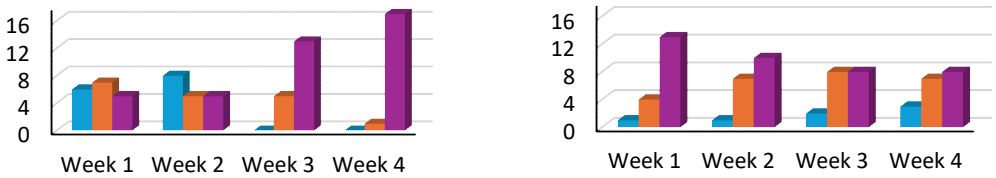


Figure 4. Sample of teacher PEs per week

Although in some weeks, a teacher might have given more 'same' or 'lower' PEs than the expert PEs, more often than not (Figure 5), the majority of *kodes* assigned by teachers for the lesson they were observing were colour-coded as 'higher' than expert *kodes*. In other words, teachers often overestimated rather than underestimated in their pedagogy evaluations. This theme might reflect a similar cultural tendency as the Japanese as described by Benedict (2005: 223), who feel shame at "criticism" and seek to guard (147) the "honour… to one's name". If so, this theme could indicate that participants had negative perceptions of low scores or positive perceptions of high scores that consciously or subconsciously affected their assigning of *kodes* for the instances of pedagogy that were witnessed. Further studies into the prevalence of this cultural factor in Indonesian school contexts are required, including perceptions of shame or prestige, and perceptions of high and low scores.

Theme 3: Increase in number of 'same' PEs over time.

As shown in Figure 3, Teacher Z4, Y2, Y3 and Y4's PEs displayed a clear incline, and Teacher Y1's PEs were slightly inclining in number of 'same' PEs over time:

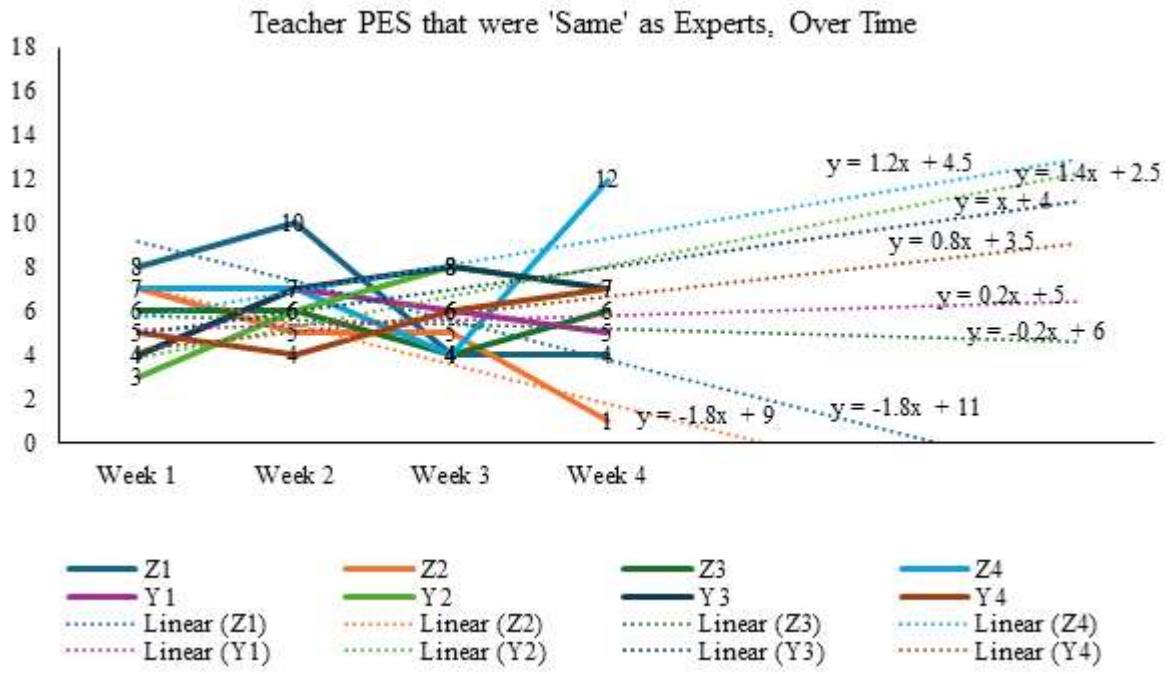


Figure 5. All teachers' 'same' PEs, over time

Conversely, three teachers' linear trendlines according the limited 4-point data seemed to be declining over time, as shown by the negative sign in the linear equations. Teacher Z2 and Y3's numbers of 'same' PEs were clearly declining over time, and Z3's 'same' PE totals were slightly declining over time. This finding is both encouraging and surprising. It is clear that teachers were able to make 'accurate' evaluations, and able to learn and improve their evaluation 'accuracy' over time. This was encouraging since an incline indirectly reflects an increasing teacher ability to pay attention to pedagogy quality in an observed lesson, understand the 18 Elements of the QT Model, and insightfully draw parallels between the lesson and rubrics in order to

'accurately' gauge the quality of the pedagogy observed (see Jacobs, Lamb & Philipp, 2010 for a theory of attending to what happens in classrooms). Factors leading to teacher increase in 'same' PEs could include personal aptitude. It is noted that all teachers suggested for participation from this school were considered by their leaders as 'good teachers' who were either 'movers and shakers' or highly motivated to learn. In regards to the two surprising findings, it is also clear that teachers who had one week evaluated pedagogy such that they came to the same conclusion as the research team, at times did not do so at a later date. Some reasons for this latter occurrence are discussed in Pattern 4.

Theme 4: Differences between PLCs in Findings per Week and per Element

Surprising and counter-intuitive differences between teacher trends and results per Element of the QT Model were found between the two PLCs.

4.1 PLC-based Trends in PE Similarity and Dissimilarity per Week

From experience attending the QTR Foundational Workshop, the first author surmised that during the one-month QTR program, teacher pedagogy evaluation (PE) sameness or 'accuracy' might rise, due to the deep collegial dialogue in the PLC Discussions intended to aid teacher gain of a more rigorous understanding of pedagogy, leading to more and more accurate PEs. It was expected that in the first few weeks, teachers' individual PEs might reflect a lower level of understanding, by being more dissimilar to the expert PEs. Correspondingly, it was expected that in the last few weeks, if any change, an increase in similarity with the expert PEs might occur as a result of the two-hour PLC discussions each week. However, as shown in Table 7, less than half of the teachers' PEs clearly improved in similarity to the expert PEs over time (three out of eight). Two other teachers' PE similarity with expert PEs increased slightly over time and one teacher's 'same' PEs decreased in number over time. Surprisingly, two teachers' 'same' PEs decreased in number over time as compared to the expert PEs.

It is clear from Table 8 that more of the teachers with the 'increasing' trend were in PLC Y, and conversely, PLC Z contained both teachers whose number of 'same' PEs were declining over time. The graphs below are given to increase transparency, showing that teacher PE similarity was not straightforward. While the dotted trendlines help the observation of trends and themes, they do not show the whole picture (Figures 6 and 7).

Table 7. Trends in teacher PE similarity to expert PE over time, per school

Colour-code	Trend	No. of PLC Z teachers	No. of PLC Y teachers
'same'	decreases	2	
	slightly decreases	1	
	slightly increases		2
	increases	1	2

When the instances of teachers giving 'same' *kodes* are added with the times they gave higher and lower evaluations with a difference of just 1, the contrast between PLC Z and PLC is greater. Figure 8 shows that PLC Z teachers' evaluations which were the 'same' or close to being the same are still not seen to improve over four weeks, while PLC Y's teacher evaluations in the same category experience a definite incline:

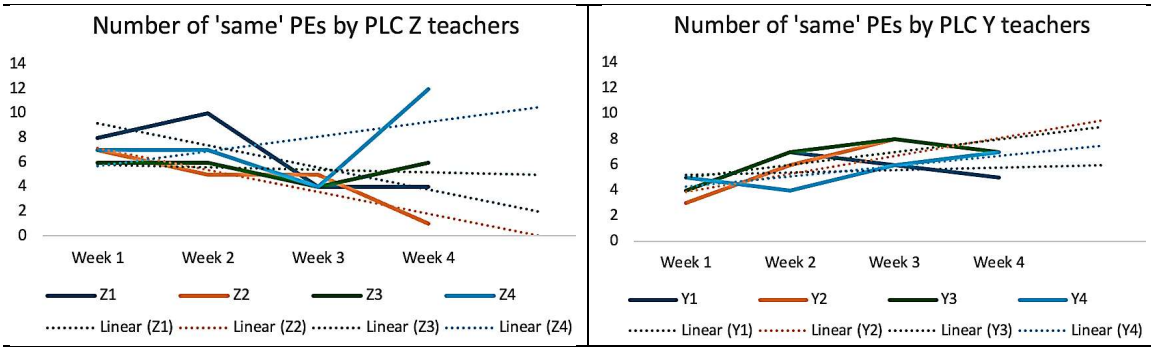


Figure 7. Number of 'same' PEs by PLC Z and Y teachers, per week

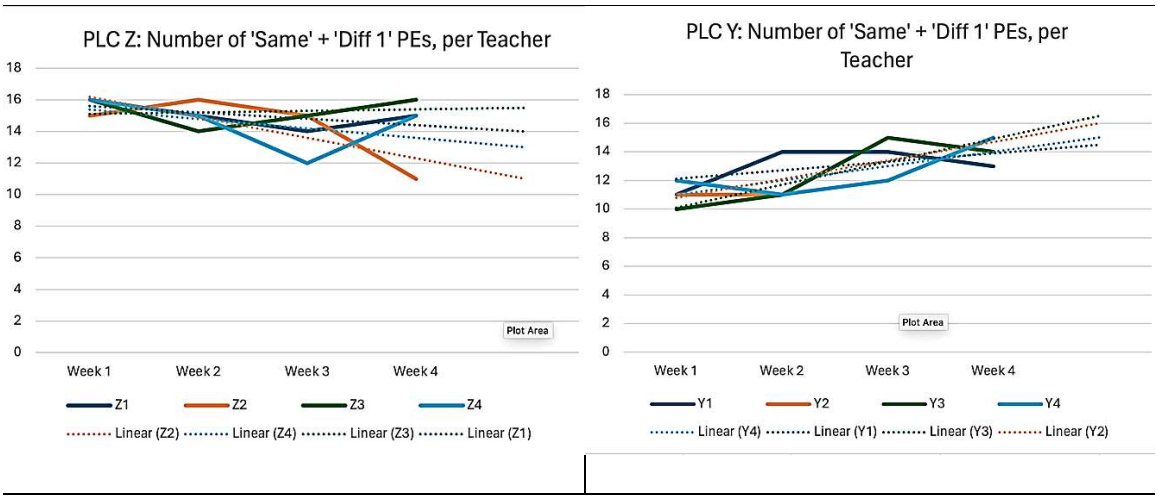


Figure 8. PLC Z and PLC Y Total 'same' and 'Difference of 1' PEs per week

As seen in the two bar charts in Figure 5 for Pattern 2, Teacher Z2's 'lower' PE totals were decreasing each week, while Teacher Y2's 'lower' PEs were slightly increasing each week. The overall story of these bar charts is that Y2's dissimilar PEs per week were trending towards 'same', while the Z2's story is that her dissimilar PEs were trending towards 'higher', in this short implementation of QTR. When tabulated, it is clear that more teachers in PLC Y had trends encroaching on 'same' over time, suggesting dis playing a clear trend towards accuracy in under stand the QT Model as a basis for evaluating pedagogy. Meanwhile, in PLC Z, more teachers' numbers of lower & igher individualPEs increased in tandem, showing trend towards overestimating.

Table 8. Trends in teacher PE dissimilarity to expert PE over time, per school

Colour-Code and Trend	No. of PLC Z teachers	No. of PLC Y teachers
'lower' decreases & 'higher' increases.	3	
'lower' decreases & 'higher' is flat	1	
'higher' decreases & 'lower' is flat		1
'lower' increases & 'higher' decreases.		3

Perhaps these differences in trends between PLCs are simply due to the short-term nature of this study, since significant teacher and student improvement are normally noted after 2 years of implementation of QTR (Burke *et al*., 2019, in Ryan, Prieto-Rodriguez, Miller & Gore, 2024). Djulete's (2021) dissertation abstract corroborates this, stating it takes quite a long time to influence or change the beliefs, attitudes and practices of a teacher. Besides this, three categories of explanations exist for PLC Z teachers' declines in PE similarity and tendencies towards overestimating. Firstly, internally, the teachers might be genuinely becoming confused for some reason. Secondly, they might be making the right evaluations yet externally, writing down higher ones due to intrinsic or extrinsic factors. Thirdly, the research team might be influenced by outside factors in those instances. In the first instance, assuming teacher evaluation-giving were not random, as based on clear trends obvious in Figure 5, confusion is a slight possibility and will be discussed in Pattern 4.2, below. To address the third option, the research team met several times to argue and counter-argue each 'expert' PE, and even adjusting towards the PE of teachers when the teachers were producing valid justifications for their evaluations, thus reducing expert PE subjectivity.

The second point seems more likely in this case, *viz* that subjectivity, whether conscious or subconscious, affected the teachers in PLC Z. It is interesting to note that three of the four teachers from PLC Z all gave abnormally low PEs on week 3, as shown in Figure 8, further corroborating the influence of certain factors on their PEs. Possible factors influencing teacher objectivity in pedagogy evaluations, based on initial perusals of field notes, are discussed below.

When details of each PLC environments are considered, inconclusive but intriguing patterns start to emerge. School Z, with a history of high student results in end-of-year state exams, is ranked in the top 5 state schools of the city and colloquially known as a 'favourite school'. All teachers suggested for participation from this school were considered excellent teachers with responsibility to influence other teachers. A school culture of prestige was observed, with teachers expected to be involved in many school-based and even government-based programs. Field notes recorded that three of the four PLC members delivered a 'special' lesson in a non-normal room, using some different techniques than usual. Two of these, including one whose lesson was observed in week 3, admitted they were killing two birds with one stone by having their lesson video-recorded at the same time, to submit as an exemplary lesson for one of the multiple national teacher training programs they were also enrolled in concurrently. This may have resulted in a desire to showcase teacher abilities and could be connected to the school culture of prestige. Connected to this was what seemed to be a residual framework of top-down supervision derived from the national education system's performance appraisal regime, since the teacher in week 3 was very nervous, and one of the PLC members who had a higher position than the others consistently referred to QTR as 'supervision'. The school prestige culture and the supervision culture could be related to internal values coming from the broader Eastern cultures of saving face or avoiding shame, but this connection will need to be corroborated or disproved through the more in-depth inductive study 2.

Another external factor could be relationships between the members of this PLC, since there was sometimes resistance observed between the two non-leaders and the two school leaders. The leaders' practice was evaluated more critically, while the non-leaders were close friends and seemed to be evaluating each other 'higher' than anyone else. Finally, PLC Z's members were decidedly more busy and tended to rush more,

sometimes in relation to school expectations of attendance at other concurrent teacher training programs. This seemed to affect the group's building of PK since they didn't always spend the time to listen deeply to each other's reflections.

Meanwhile, School Y's history of low student results ranks it in the bottom five of the city's public schools, and it is considered by many as a 'non-favourite school'. While the participant teachers maintained good rapport with students, a more relaxed teacher culture was noted. Leaders were not present in the PLC, but outside the PLC Days, leaders sometimes referred to a gap between the 'older generation' of teachers who might be less focused on quality work, and the 'younger generation', who were considered more motivated to develop in their practice and apply new technologies and ideas to their teaching. These leaders expressed a desire to prioritise the second group of teachers for participation in QTR as they might take the program more seriously and take initiative to share their learning with others in the school, making the QTR program more sustained. Hence, most of the participants from this school were in this under-40-year-old category.

While several teachers mentioned nervousness about their practice being evaluated and a dislike of receiving *kode* 1 or 2 out of 5, as a group, they seemed to become more comfortable with giving 1s and 2s over time, suggesting a level of comfort and safety (Gore & Bowe, 2015) had been achieved in this group. Few teachers were observed to mention supervision, and the hierarchy in PLC Y was flatter, with no-one having a rank above another. No 'side-taking' behaviour was noted; although two of the teachers were slightly closer with each other than the others, these tended to be more critical of each other and spur each other on towards objectivity, rather than subjectivity. All teachers remarked on their feeling of close friendship with the others in the group.

Two teachers seemed to be introducing 'non-normal' elements such as the use of deep questioning, story-telling and extra printed resources, but other than this, the lessons were held in normal classrooms and seemed to portray usual teaching and learning in those classrooms. Finally, PLC Y seemed more relaxed about time; only once did they cut the time shorter than recommended. Indeed, on one day instead of going home at the end, the teachers stayed behind to continue sharing, even after the researcher left.

4.2 PLC-based 'accuracy' and 'inaccuracy' of evaluating some QT Model Elements

The picture painted above is not intended to suggest that PLC Z's teachers were less adept at evaluating than PLC Y's teachers. In contrast, PLC Y teacher PEs strayed further from expert PEs, more often, suggesting PLC Z teachers also had the required skills and knowledge to evaluate pedagogy critically, shown in Table 9:

Table 9. Total PEs given by teachers with a difference of 3 or 4

	School Z	**School Y**
Teacher 'difference of 3' + 'difference of 4' *kodes* compared to expert *kodes*	8	22

The final pattern noted from the tabulation of deductive colour-codes is more directly related to pedagogical knowledge (PK) because it is linked to the QT Model, detailed in Gore (2007) and having 18 Elements in three Dimensions, as outlined below:

Table 10. Elements of the quality teaching model

Intellectual Quality	Quality Learning Environment	Significance
Deep knowledge	Explicit quality criteria	Background knowledge
Deep understanding	Engagement	Cultural knowledge
Problematic knowledge	High expectations	Knowledge integration
Higher order thinking	Social support	Inclusivity
Metalanguage	Students' self-regulation	Connectedness
Substantive communication	Student direction	Narrative

Source: NSW Department of Education and Training (2003), cited in Gore, Smith, Bowe, Ellis, Lloyd and Lubans (2015:4)

While no trends per Element of the QT Model were evident over time, when the number of 'same' PE were simply tallied, it was found that some Elements were often evaluated 'same' or not the same by the majority of a PLC's members, even though they were working individually. In other words, teacher PE tended to be more similar between fellow PLC members, per element of pedagogy. These counts reveal potential strengths and gaps in participant teacher understanding of these Elements. This supports the constructivist theory, that teachers were constructing knowledge about pedagogy together in the learning communities, week after week.

Teachers in a school's PLC sometimes consistently evaluated certain elements of pedagogy with a strong degree of similarity to the experts overall, though this was a rarity and limited to *Problematic Knowledge* (Table 11) and *Narrative* by PLC Y, as well as *Engagement* as evaluated by PLC Z's members. For all three of these Elements, the other PLC's members conversely displayed a strong degree of dissimilarity with expert PEs. These 'negative' finding also support the concept of social constructionism in that confusions or misunderstandings can also be constructed into a faulty knowledge base and communally strengthened, week after week.) Houston (2021: 846) summarises Berger and Luckman's social constructionism argument that understanding and even civilization itself is the "product of people engaging with one another, where such interactions become externalized, objectified and then internalized". This approach highlights "the importance of human subjectivity… 'as a sense making activity'" (White, 1997, cited in Houston, 2021: 848). In the light of this theory, the differing 'senses' constructed in the two discrete engagement sites: PLCs Z and Y, are explained as contingent on the human subjects within the communities themselves, engaging with one another and building understanding together. This underscores the power of the PLC as a venue for teacher learning and growth, but suggests that to maximise understanding of 'standard' model such as the QT Model and to maximise evaluation objectivity, teachers should be re-oriented back to the CPG regularly.

To further illustrate this point, the below table shows an example of PLC Y teacher PEs being the 'same' as expert PEs, more often than not, suggesting their understanding of this element of pedagogy was strong. In contrast, PLC Z's teachers rarely evaluated *Problematic Knowledge* the 'same' as the expert PE, suggesting low understanding of this element throughout the QTR program.

Table 11. PLC Y and Z teacher colour-code totals for problematic *knowledge*

1.3 Problematic Knowledge	Teacher Y1	Y2	Y3	Y4	PLC Y Count	*1.3 Problematic Knowledge*	Teacher Z1	Z2	Z3	Z4	PLC Z Count
'lower'				1	1	**'lower'**					0
'same'	2	2	2	3	**9**	**'same'**	1	1			**2**
'higher'	2	2	2		6	**'higher'**	3	3	4	4	**14**
'diff 1'	1			1	2	**'diff 1'**	1	1	1	2	5
'diff 2'		2	2		4	**'diff 2'**	2		3	1	**6**
'diff 3'	1				1	**'diff 3'**		1		1	2
'diff 4'					0	**'diff 4'**		1			1

Problematic Knowledge involves teaching in a way that helps students understand the social construction of knowledge. This had proved a difficult Element for the first author to translate and explain for the workshop in a way that made sense to the eight participants who were all raised in a context where questions of 'how do you know?' are not always encouraged. Terminology and cultural considerations led to this Element being re-named as *Asal-Usul Pengetahuan* (*Origins of Knowledge)*, while the *koding* rubrics, translated with no changes, still refer to students asking questions about the knowledge they are being taught. Indeed, the findings in Table 11 are encouraging, since even though half of the teachers seemed to not understand *Problematic Knowledge,* the other half had the highest similarity to experts for this Element, suggesting those teachers understood it well enough to evaluate it correctly. Another prominent difference between PLCs along this vein was with *High Expectations*, where conversely, PLC Y members consistently mis-evaluated this Element, compared with the PLC Z members who evaluated *High Expectations* at a moderate rate of sameness, giving the same *kode* the experts had given, nearly half of the time.

A variance in a different vein consisted of the evaluations for the Element *Metalanguage* (Table 12). This Element was regularly poorly evaluated by members of both PLCs, resulting in being evaluated the 'same' as experts only six times. However, PLC Y consistently gave this Element 'lower' PEs than the expert PEs, while interestingly, PLC Z consistently evaluated *Metalanguage* 'higher' than the experts.

Table 12. PLC Y and Z teacher code comparison totals for *metalanguage*

1.5 Meta-language	Teacher Y1	Y2	Y3	Y4	PLC Y Count	*1.5 Meta-language*	Teacher Z1	Z2	Z3	Z4	PLC Z Count
'lower'	3	2	2	3	**10**	**'lower'**					0
'same'		1	1		2	**'same'**	1	1	1	1	**4**
'higher'	1	1	1	1	4	**'higher'**	3	3	3	3	**12**
'diff 1'	2	1	1	2	6	**'diff 1'**	1	3	1	1	6
'diff 2'	2	2	2	2	8	**'diff 2'**	2		2	2	6

The salient difference between PLC Y and Z in Table 12 above reveals a strong likelihood of low understanding of the Element at play. Since the teachers were all given the same two-day training at the start of the program, and the same access to the CPG in Indonesian, the social constructionism theory again can attempt to explain this result, in that the teachers'

collaborative discussions from the week before were informing their current week's individual evaluations.

Besides the fact that this was a small, short study unable to prove definite impacts of the QTR program, the results discussed so far in this section highlight the probable influence of various factors on teacher PEs during QTR. These factors could include *language factors* such as translation errors; Asian *cultural factors* such as an orientation to avoid shame or guard honour, resulting in negative perceptions of low scores and positive perceptions of high scores; teacher *personal factors* such as aptitude and motivation to learn; *school contextual factors* such as prestige, busyness and attitudes to supervision; *relationship factors* such as power play, closeness and trust; and the *communal factor* of knowledge construction in community which aligns with the social constructionism view.

Cultural and relational factors seemed to stymy the design of QTR (Gore, Rickards & Fray, 2023: 465) to move "the locus of control from externally imposed surveillance to teacher judgment and collaborative engagement with relevant criteria" in School Z's PLC especially. However, the success of School Y's PLC to create (Gore, Rickards & Fray, 2023: 465) "spaces of freedom – for teacher agency to flourish" seems to be linked to a greater group similarity of PEs and a greater number of individuals displaying PE growth. A deeper and broader study to ascertain the degree to which all factors above influenced teacher individual coding is in order. Perhaps the second study of this project will shed some light on these factors as it will focus on teacher communication patterns within the PLC Discussions.

Limitations of this study are delineated as follows. (1) Despite efforts, this was not a perfectly 'pure' application of QTR. For example, in Australia, emphasis is put on teacher desire to improve practice and QTR is ideally driven by a 'coalition of the willing'. In this study, the impetus to implement QTR came from the researchers, instead of from the didn't come from the schools or teachers. Also, the first author sat in on PLC Discussions and sometimes explained Elements or corrected teachers, which was observed to sometimes help and sometimes hinder (due to language barriers) teacher understanding which may have impacted PK growth. In addition, some Indonesian-language translation mistakes probably led to a few teacher misunderstandings of the QT model, while the researchers themselves were seen to have mis-conceptualised Element *1.1 Deep Understanding* and to some extent, *1.2 Deep Knowledge.* While the 'expert' PEs were later edited to reflect the proper concepts as re-explained by a Quality Teaching Academy trainer, the teacher PEs for *Deep Understanding* and *Deep Knowledge* were therefore skewed and hence this data was excluded from the results per Element. Although these results could be 'realistic' in that Australian school teachers can also struggle with evaluating these Elements, these results were considered invalid and not reported on in the 'per element' section.

(2) Assumptions had been made by the workshop facilitator regarding participant prior knowledge, which was in fact was lower in Indonesia than in Australia regarding the use of rubrics. In the last week it was discovered that teachers in School Z had had little prior training in the act of grading using rubrics. Teachers could use rubric descriptors well enough, but when arguing about one or two lines of text in a descriptor, a shared understanding of 'what to do' had not yet been built in their prior experience. (3) This study is of such a short-term nature that teacher development might not be measurable by any method, and may not even be a valid expectation. (4) Many qualitative studies, even of multiple cases, are "not immediately generalisable" (Miles & Huberman, 1994: 29).

Therefore, the authors do not mean to imply in reporting results of this study that all teachers from State Junior High Schools in Indonesian cities would produce results like these (after Bogdan & Biklen, 1998: 32). (5) In addition, permission and access limitations meant that other methods of sampling than what were used were very difficult.

Delimitations of this study include the fact that during the PLC Discussions, the teachers often felt free to disregard advice given to them by the first author. Also, she was careful not to suggest any *kode* to them, at any point, and mostly asked open questions referring teachers back to the lesson or the CPG. In addition, the confusion regarding Elements 1.1 and 1.2 is fairly normal in Australia; teacher selection does not always follow the concept of the 'coalition of the willing'; and in first-time applications of QTR PLC members often start with low trust between each other, meaning that this application of the whole QTR program was reflective of 'real' or 'normal' implementation circumstances in the original context.

■ CONCLUSION

Eight teachers from two State Junior High Schools in an urban area of Sumatra, west Indonesia, engaged in a two-month Quality Teaching Rounds (QTR) program designed to increase teacher pedagogical knowledge and practice through lesson observations, pedagogical evaluations (PEs) using the Quality Teaching (QT) Model and followed by rigorous professional learning community (PLC) discussions in each school. The Indonesian teachers' 72 PEs were compared to external expert PEs by the research team led by an Australian educator in a simple qualitative thematic analysis.

Comparative, *a priori* coding of teacher data produced four salient themes to answer the research question. Teacher pedagogy evaluations followed these patterns when compared to experts in the field: (1) teacher PEs were more often dissimilar than similar to the expert PE; (2) teachers tended to overestimating in their PE giving; (3) about half of the teachers an increase in teacher 'same' PEs over time for and a decrease in the same for some teachers; and (4) differences between PLCs in teacher trends and Element tallies.

While the results of this study are clear, there are few implications for immediate action in the education world, since exploratory qualitative research tends to produce an output of recommendations for further research (George, 20 November, 2023; Trenholm-Jenson, Burns, Trenholm & Hand, 2022). The themes or patterns of teacher pedagogy evaluation assigning described in the previous section imply the presence and influence of various language, cultural, school context, relational and communal factors on teacher PEs during the implementation of high-fidelity QTR in Indonesia. However, further study is required to inquire: (a) whether or not these factors are the causes of the patterns observed; (b) whether or not there are other causes or explanations for these patterns; (c) to what extent these factors influenced teachers in this instance of QTR; (d) whether QTR employed in other Indonesian contexts would be also influenced by the same factors; (e) whether these factors negatively or positivity impact on teacher practice and thereby on student learning; and (f) whether these factors can be mitigated against without touching the integrity of the QTR design. The answers to these questions will enable recommendations as to the suitability of QTR with the Indonesian context.

The authors strongly recommend revising the Indonesian version of the Classroom Practice Guide and making it available to Indonesian initial teacher education providers (ITEs) such as universities so that the QT Model can be socialised with the Indonesian education community and school- or teacher-initiated trials

of QTR be more likely to occur naturally. To gain more conclusive results, the authors specifically suggest, subsequent studies could more use quantitative measures to closely monitor teacher knowledge growth as well as teaching practice, over a longer period of time, per Element of the QT Model, with more schools. Deeper, inductive thematic analysis or grounded theory could using transcripts, observations, interviews, or focus groups could be applied to aspects of Indonesian school workplace culture that remain under-studied, Supplementary studies could include qualitative measures of opinions and attitudes towards QTR and the QT Model, perceptions of high-scale (*kodes* 4 and 5) and low-scale (kodes 1 and 2) evaluations, and levels of trust between and within leader and non-leader PLC members. Interview methods could help draw further inferences on inter-PLC relationship impact on objectivity in *koding*. Other variables of interest may include contrasts between state civil servant participants and contract teachers; between or across public and state schools; and implementations of QTR driven by school leaders; in order to more fully grasp and explore the suitability of QTR with the Indonesian context.

■ REFERENCES

Amadhannn, A. (24 March 2023). Map of western indonesia. wikimedia commons. [Image]. Retrieved from: https://commons.wikimedia.org/w/index.php?curid=129942483 Accessed 25 September 2024.

Amosa, W., Ladwig, J., Griffiths, T., & Gore, J. (2007). Equity effects of quality teaching: closing the gap. Paper presented at the australian association for research in education conference, fremantle. Retrieved from: https://www.aare.edu.au/data/publications/2007/amo07284.pdf

Anwar, A., Harun, C. Z, dan Niswanto, N. (2022). Academic supervision of the school principal in improving the performance of elementary teachers in clusters i-ii in bandar dua, pidie jaya district, indonesia. Traektoria Nauki = Path of Science, 8(12), 4042-4050. Retrieved from: https://doi.org/10.22178/pos.88-13

Aruan, C. D. (2015). HRM reform in the indonesian civil service: towards strategic fit? 2nd global conference on business and social science-2015, GCBSS-2015, 17-18 September 2015. Bali: Elsevier. *Procedia - Social and Behavioral Sciences, 211*(2015): 49–54. Retrieved from: https://doi.org/10.1016/j.sbspro.2015.11.008

Australian Bureau of Statistics. (no date). Quantitative and qualitative *data.* ABS website. Retrieved from: https://www.abs.gov.au/statistics/understanding-statistics/statistical-terms-and-concepts/quantitative-and-qualitative-data

Benedict, R. (2005). The chrysanthemum and the sword: patterns of japanese culture. 1st Mariner Books ed. Boston, Houghton Mifflin.

Bernard, H. R. (1994). Research methods in anthropology: qualitative and quantitative approaches, second edition. Walnut Creek: AltaMira Press.

Bloom, B., Englehart, M. Furst, E., Hill, W. & Krathwohl, D. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York: Longmans, Green. Retrieved from: https://web.archive.org/web/20201212072520id_/https://www.uky.edu/~rsand1/china2018/texts/Bloom%20et%20al%20-Taxonomy%20of%20Educational%20Objectives.pdf

Bogdan, R. C. & Biklen, S. K. (1997). Qualitative research for education: an introduction to theory and methods. Needham Heights: Allyn & Bacon

Bryk, A. S. (2014). Improving: Joining improvement science to networked communities. *Distinguished Lecture presented at the American Educational Research Association, Philadelphia, Pennsylvania.* Retrieved from: http://convention2.allacademic.com/one/aera/aera14/index.php?cmd=Online+Program+View+Event&selected_box_id=165082&PHPSESSID=ehvpf2svmltu7u3dhvee8c9ul7

Darling-Hammond, L. (2013). Getting teacher evaluation right: What really matters for effectiveness and improvement. New York: Teachers College Press.

Djulete, W. (2021). The influence of professional learning communities employing quality teaching framework on English as a foreign language teachers' pedagogical content knowledge for promotion of higher order thinking in rural Indonesia. [Dissertation Abstract]. Flinders University, College of Education, Psychology and Social Work. Retrieved from: https://theses.flinders.edu.au/view/f077bdf5-95eb-41ca-a805-a2473c33ed8f/1

Fereday, J. & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods, 5*(1), 80-92. Retrieved from: https://doi.org/10.1177/160940690600500107

Harry Firman, H. & Tola, B. (2008). The future of schooling in indonesia. cice hiroshima university, *Journal of International Cooperation in Education.* 11(1): 71-84. Retrieved from: https://doi.org/10.15027/34287

George, T. (20 November, 2023). Exploratory research: definition, guide, & examples. *Scribbr.* Retrieved from: https://www.scribbr.com/methodology/exploratory-research/ (Accessed October 28, 2024).

Gore, J. M. (2007). Improving pedagogy: Challenges of moving teachers toward higher *levels of Quality Teaching.* In J. Butcher & J. Asha (Eds). Making a difference: Challenges for teachers, teaching, and teacher education (pp. 15–33). New York: Sense. Retrieved from: https://doi.org/10.1163/9789087901332_003

Gore, J. (2020). Why isn't this empowering? The discursive positioning of teachers in efforts to improve teaching. In A. Brown & E. Wisby (eds.) *Knowledge, policy and practice: The struggle for social justice in education* (199-216). London, UK: UCL IOE Press.

Gore, J. M. & Bowe, J. M. (2015). Interrupting attrition? Re-shaping the transition from preservice to inservice teaching through Quality Teaching Rounds. *International Journal of Educational Research, 73*(2015), 77-88. Retrieved from: https://doi.org/10.1016/j.ijer.2015.05.006

Gore, J., Smith, M., Bowe, J., Ellis, H., Lloyd, A., & Lubans, D. (2015). Quality Teaching Rounds as a professional development intervention for enhancing the quality of teaching: Rationale and study protocol for a cluster randomised controlled trial. *International Journal of Educational Research, 74*(2015): 82–95. Retrieved from: https://doi.org/10.1016/j.ijer.2015.08.002

Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H., & Lubans, D. (2017). Effects of professional development on the quality of teaching: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education*, *68*(November 2017): 99–113. Retrieved from: https://doi.org/10.1016/j.tate.2017

Gore, J. (2021). The quest for better teaching. *Oxford Review of Education*, *47*(1): 45-60. Retrieved from: https://doi.org/10.1080/03054985.2020.1842182.

Gore, J., Miller, A., Fray, L., Harris, J., & Prieto, E. (2021). Improving student achievement through professional development: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education,* 101(103297): 1-12. Retrieved from: https://doi.org/10.1016/j.tate.2021.103297

Gore, J. & Rickards, B. (2021). Rejuvenating experienced teachers through Quality Teaching Rounds professional development, *Journal of Educational Change,* 22(2021): 335-354. Retrieved from: https://doi.org/10.1007/s10833-020-09386-z.

Gore, J., Rickards, B. and Fray, L. (2023). From performative to professional accountability: re-imagining 'the field of judgment' through teacher professional development, *Journal of Education Policy, 38*(3): 452-473. Retrieved from: https://doi.org/10.1080/02680939.2022.2080274

Hattie, J. A. C. (2009). Visible Learning: A synthesis of over 800 meta-analyses relating to achievement. New York: Routledge.

Hattie, J. A. C. (2023). Visible learning: the sequel: a synthesis of over 2,100 meta-analyses relating to achievement. London: Routledge.

Houston, S. (2001). Beyond social constructionism: critical realism and social work. the british journal of social work. 31(6, December 2001), 845-861. Retrieved from: https://www.jstor.org/stable/23716466

Indrawati, S. M. & Kuncoro, A. (2021). Improving competitiveness through vocational and higher education: indonesia's vision for human capital development in 2019–2024. *Bulletin of Indonesian Economic Studies,* 57(1, 2021), 29–59. ANU Indonesia Project. Routledge. Retrieved from: http://dx.doi.org/10.1080/00074918.2021.1909692

Jacobs, V. R., Lamb, L. L. C. & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*(2, March 2010): 169-202. National Council of Teachers of Mathematics. Retrieved from: https://www.jstor.org/stable/20720130

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research,* 86(4): 945-980. American Educational Research Association. Retrieved from: https://www.jstor.org/stable/44668240

Krathwohl, D. R. (2002) A revision of bloom's taxonomy. *Theory into Practice. 41*(4, Autumn, 2002): 212-218. Ohio State University. Retrieved from: https://doi.org/10.1207/s15430421tip4104_2

Krisnaresanti, A., Ahman, E., Istikomah, N., Suwatno & Budiman, A. (2024). Research trends in technological pedagogical content knowledge (TPACK) Research (2014-2024): A scientometric analysis. *Jurnal Pendidikan Progresif, 14(*1): 352-369. Retrieved from: https://doi.org/10.23960/jpp.v14.i1.202426

Ladwig, James G. (2007). 'Modelling pedagogy in australian school reform', pedagogies: an international Journal 2 (2): 57–76. Retrieved from: https://doi.org/10.1080/15544800701343919

Lingard, B., Ladwig, J., Mills, M., Bahr, M., Chant, D., Warry, M., Ailwood, J., Capeness, R., Christie, P., Gore, J., Hayes, D. & Luke, A. (2001). The queensland

school reform longitudinal study. Brisbane: education queensland.

Lingard, B., Hayes, D. and Mills, M. (2003). Teachers and productive pedagogies: contextualising, conceptualising, utilising. *Pedagogy, Culture and Society, 11*(3): 399-424. Retrieved from: https://doi.org/10.1080/14681360300200181

Mansourian, Y. (2008). Exploratory nature of, and uncertainty tolerance in, qualitative research. *New Library World. 109*(5/6): 273–286. Retrieved from: https://doi.org/10.1108/03074800810873614

Miles, M. B. and Huberman, A. M. (1994) Qualitative data analysis: an expanded sourcebook. (Second Edition). Thousand Oaks: Sage Publications.

Miller, A., Gore, J., Wallington, C., Harris, J., Prieto-Rodriguez, E., & Smith, M. (2019). Improving student outcomes through professional development: Protocol for a cluster randomised controlled trial of Quality Teaching Rounds. *International Journal of Educational Research,* 98, 146–158. Retrieved from: https://doi.org/10.1016/j.ijer.2019.09.002

Musfah, J. (2015). Manajemen pendidikan: teori, kebijakan, dan praktik. Jakarta: Kencana.

Nazari N., Nafissi Z., Estaji, M. & Marandi, S. S. (2019). Evaluating novice and experienced EFL teachers' perceived TPACK for their professional development. *Cogent Education*. *6*(1): Article 1632010, 1-26. Retrieved from: https://doi.org/10.1080/2331186X.2019.1632010

Newmann, F. M. & Wehlage, G. G. (1995). Successful school restructuring: a report to the public and educators by the center on organization and restructuring of schools. madison: board of regents of the university of wisconsin system. Retrieved from: http://files.eric.ed.gov/fulltext/ED387925.pdf

New South Wales Department of Education and Training (NSW DET). (2006). Quality teaching in NSW public schools: A classroom practice guide (Second Edition). Retrieved from: https://maryland-p.schools.nsw.gov.au/content/dam/doe/sws/schools/m/maryland-p/localcontent/qt_classroom_practiceguide.pdf

New South Wales Department of Education and Training (NSW DET). (2020). Quality teaching classroom practice guide. (Third Edition). [Digital version].

OECD. (2023). PISA 2022 Results Volume I: The State of Learning and Equity in Education. Paris: OECD Publishing. Retrieved from: https://doi.org/10.1787/53f23881-en

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and policy in mental health, 42(5), 533–544. Retrieved from: https://doi.org/10.1007/s10488-013-0528-y

Rahman, A. (2016). Teacher professional development in Indonesia: The influences of learning activities, teacher characteristics and school conditions. [Doctoral Thesis, University of Wollongong]. Retrieved from: https://ro.uow.edu.au/theses/4721

Republic of Indonesia Constitutional Law. 1945/4.

Ryan, A., Prieto-Rodriguez, E., Miller, A. and Gore, J. (2024). What can Implementation Science tell us about scaling interventions in school settings? A scoping review. *Educational Research Review.* 44(2024): Article 100620. Retrieved from: https://doi.org/10.1016/j.edurev.2024.100620

Shulman, L. S. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review, 57*(1): 1-22.

Retrieved from: http://dx.doi.org/10.17763/haer.57.1.j463w79r56455411

Silova, I., Rappleye, J. & Auld, E. (2020). Beyond the western horizon: rethinking education, values, and policy transfer. In Handbook of Education Policy Studies: Values, Governance, Globalization, and Methodology (Volume 1, Chapter 1). Guorui Fan and Thomas S. Popkewitz (Eds). Singapore: SpringerOpen. [eBook]. Retrieved from: https://rdcu.be/dVhWM

Trenholm-Jenson, E. A., Burns, L., Trenholm, J. L. & Hand, C. J. (2022) Beyond tingles: An exploratory qualitative study of the Autonomous Sensory Meridian Response (ASMR). *PLoS ONE, 17*(12): e0277962. Retrieved from: https://doi.org/10.1371/journal.pone.0277962

Usman, Nasir & A.R, Murniati. (2019). *Pengantar manajemen pendidikan*. Online: An1mage.

Wahyudi, W. E., Fakhrudiin, A. H. & Ikmal, H. (2024). Educational verses in the qur'anic text: pedagogical and andragogy perspectives. *Jurnal Pendidikan Progresif, 14*(1): 650-663. Retrieved from: https://doi.org/10.23960/jpp.v14.i1.202447

World Bank (2015) Indonesia: Teacher certification and beyond, Report No. 94019-ID, Washington DC: World Bank.

Xu, W. and Zammit, K. (2020). Applying thematic analysis to education: a hybrid approach to interpreting data in practitioner research. *International Journal of Qualitative Methods. 19*(2020). 1-9. Retrieved from: https://doi.org/10.1177/1609406920918810

Appendix 1: Coding Scale Overviews

From Quality Teaching in NSW Public Schools: A Classroom Practice Guide, Second Edition. Copyright © 2006 by the State of NSW, Department of Education and Training, Professional Support and Curriculum Directorate, Sydney, Australia, (pp. 57-59). Translated by W. Djulete, July 2017, formatted by F. Harnita, January 2024.

Garis Besar Skala Pengkodean Kualitas Intelektual

Aspek	1	2	3	4	5
1.1 Pengetahuan mendalam	*Hampir semua pengetahuan konten dari pelajaran dangkal karena tidak membahas konsep dan ide penting.*	*Beberapa konsep utama dan ide utama disebutkan atau dibahas oleh guru atau siswa, tetapi tidak secara mendalam.*	*Pengetahuan diperlakukan secara tidak merata dalam proses pengajaran. Pokok bahasan yang penting mungkin dibahas sebagai bagian dari pelajaran, tetapi secara umum pelajaran tidak berfokus secara terus menerus pada konsep utama atau ide utama sepanjang pelajaran.*	*Hampir seluruh pengetahuan konten dari pelajaran dibahas secara mendalam. Konsep utama atau ide utama jarang diinterupsi oleh ide dan konsep yang tidak mendalam atau tidak berhubungan dengan pokok bahasan.*	*Pengetahuan diajarkan secara mendalam karena fokus di pertahankan pada pokok bahasan utama atau konsep utama sepanjang pelajaran.*
1.2 Pemahaman secara mendalam	*Siswa mendemonstrasikan pemahaman yang dangkal*	*Sebagian siswa mempunyai pemahaman yang dangkal selama hampir seluruh pelajaran, dengan hanya satu atau dua siswa yang memahami pelajaran.*	*Pemahaman mendalam tidak merata. Siswa menunjukkan pemahaman yang dangkal dan pemahaman yang lebih dalam pada sesi yang berbeda selama pelajaran. Satu dari konsep utama dipahami oleh beberapa orang siswa, namun siswa yang lain mungkin tidak memahaminya.*	*Sebagian besar siswa menyampaikan informasi, argumen, atau alasan yang menunjukkan pemahaman mendalam selama sebagian besar pelajaran.*	*Hampir semua siswa menunjukkan pemahaman yang mendalam sepanjang pelajaran.*

1.3 Asal-usul pengetahuan	*Seluruh ilmu pengetahuan dipresentasikan hanya sebagai fakta dan tidak terbuka untuk dipertanyakan.*	*Beberapa ilmu pengetahuan diperlakukan terbuka bagi pandangan yang berbeda.*	*Pengetahuan diperlakukan sebagai sesuatu yang terbuka bagi pandangan yang berbeda, dilihat sebagai sesuatu yang terbentuk secara sosial dan karena itu terbuka untuk dipertanyakan.*	*Pengetahuan dilihat sebagai sesuatu yang terbentuk secara sosial dan terdiri dari berbagai pandangan yang berbeda bukan hanya dipresentasikan, tetapi dieksplorasi lewat mempertanyakan asumsi dasar pengetahuan tersebut.*	*Pengetahuan dilihat sebagai sesuatu yang terbentuk secara sosial, dengan berbagai pandangan yang berbeda, dan/atau interpretasi yang bertolak belakang ditunjukkan dan dieksplorasi sampai pada tahap di mana sebuah penilaian mengenai ketepatan pengetahuan itu berdasarkan konteks pada situasi tersebut.*
1.4 Pemikiran tingkat tinggi	*Siswa hanya menunjuk-kan (mendemonstrasikan) kemampuan berpikir tingkat rendah. Mereka dapat menerima atau menjejali pengetahuan yang sudah disiapkan atau hanya berpartisipasi dalam praktek rutin, dan dalam aktifitas belajar siswa tidak belajar untuk lebih dari sekedar menghasilkan ulang pengetahuan yang mereka pelajari.*	*Siswa pada umumnya menunjukkan kemampuan berpikir tingkat rendah, tapi pada beberapa kesempatan, setidaknya beberapa siswa menunjukkan kemampuan berpikir tingkat tinggi sebagai pengalihan kecil dalam pelajaran.*	*Siswa pada umumnya menunjukkan kemampuan berpikir tingkat rendah secara rutin dalam kegiatan belajar di mana pelajaran di bagikan secara baik. Setidaknya terdapat satu pertanyaan yang penting atau aktifitas di mana kebanyakan siswa menampilkan beberapa kemampuan berpikir tingkat tinggi.*	*Hampir semua siswa menunjukkan kemampuan berpikir tingkat tinggi dalam setidaknya satu aktifitas utama yang merupakan bagian utama dalam pelajaran.*	*Seluruh siswa, hampir sepanjang pelajaran, menunjukkan kemampuan berpikir tingkat tinggi.*
1.5 Metabahasa	*Tidak ada metabahasa. Pembelajaran berlangsung tanpa guru atau siswa berhenti dan berkomentar pada bahasa yang sedang digunakan.*	*Rendah dalam metabahasa. Ketika istilah dalam pel-ajaran dijelaskan baik guru maupun siswa memberikan penilaian atau kom-entar atas bahasa. Akan tetapi tidak ada klarifikasi atau bantuan yang disedia-kan menyangkut bahasa.*	*Beberapa kali metabahasa digunakan. Pada awal pelajaran, atau pada beberapa saat penting dalam pelajaran, guru atau siswa berhenti dan menjelaskan atau membuat 'pelajaran kecil' menyangkut beberapa aspek Bahasa, seperti genre, vocabulary (kosakata), tanda atau simbol.*	*Metabahasa digunakan secara berkala. Guru maupun siswa memberikan komentar pada aspek Bahasa di beberapa kesempatan selama pelajaran berlangsung.*	*Penggunaan metabahasa yang tinggi. Pelajaran berlangsung dengan komentar yang sering dalam penggunaan Bahasa.*
1.6 Komunikasi yang substantif	*Hampir tidak ada komunikasi substantif terjadi selama pembelajaran.*	*Komunikasi substantif di antara para siswa dan/atau di antara guru dan siswa terjadi secara singkat.*	*Komunikasi substantif di antara para siswa dan/atau di antara guru dan siswa terjadi kadang-kadang dan melibatkan setidaknya dua interaksi yang bertahan.*	*Komunikasi substantif, dengan interaksi yang lama, terjadi hampir setengah dari pelajaran dengan guru dan/atau siswa dalam mem-bangun pembicaraan.*	*Komunikasi substantif, dengan interaksi yang lama, terjadi sepanjang pelajaran, dengan guru dan/atau siswa membangun komunikasi.*

Garis Besar Skala Pengkodean Lingkungan Belajar yang Berkualitas

Aspek	1	2	3	4	5
2.1 Kriteria kualitas eksplisit	*Tidak ada pernyataan eksplisit mengenai kualitas pekerjaan yang dibuat. Hanya kriteria prosedur dan teknis tugas yang dibuat eksplisit.*	*Hanya pernyataan umum yang dibuat mengenai kualitas pekerjaan yang diinginkan.*	*Kriteria yang detail mengenai kualitas pekerjaan yang dibuat secara eksplisit selama pelajaran, tetapi tidak ada bukti bahwa siswa menggunakan kriteria untuk menilai kualitas pekerjaan siswa.*	*Kriteria detail mengenai kualitas pekerjaan dibuat eksplisit atau di tekankan selama pelajaran dan ada bukti dari beberapa siswa, beberapa kali, menilai kualitas pekerjaan mereka yang berkaitan dengan kriteria ini.*	*Kriteria detail mengenai kualitas pekerjaan dibuat eksplisit atau ditekankan selama pelajaran dan ada bukti yang konsisten dari siswa yang menilai kualitas pekerjaan mereka menggunakan kriteria ini.*
2.2 Keterlibatan	*Keterlibatan yang rendah atau tidak-terlibatan. Siswa secara rutin tidak mengerjakan tugas, kemungkinan berperilaku yang mengganggu, sebagai bukti banyak siswa yang tidak menaruh perhatian pada kelas atau merupakan gangguan yang serius.*	*Keterlibatan yang jarang. Kebanyakan siswa, hampir sepanjang waktu, kelihatan apatis dan tidak tertarik atau siswa jarang aktif dalam mengerjakan tugas yang diberikan. Beberapa siswa terlihat tidak melakukan tugas.*	*Keterlibatan yang tidak konsisten. Hampir semua siswa terlibat secara serius di sebagian pelajaran, tetapi beberapa terlihat tidak tertarik selama beberapa bagian dalam pelajaran dan beberapa terlihat tidak mengerjakan tugas.*	*Keterlibatan yang lebih luas. Kebanyakan siswa, hampir sepanjang waktu, mengerjakan tugas yang menyangkut pokok penting dalam pelajaran. Hampir semua siswa mengerjakan tugas dengan serius dan mencoba dengan keras.*	*Keterlibatan serius. Semua siswa terlibat secara dalam, hampir di sepanjang waktu pelajaran, dalam mempelajari pokok penting dalam pelajaran.*
2.3 Harapan tinggi	*Tidak ada siswa, atau hanya beberapa saja, yang berpartisipasi dalam mengerjakan tugas yang menantang.*	*Beberapa siswa berpartisipasi dalam tugas yang menantang selama beberapa waktu dalam pelajaran. Mereka didorong (secara eksplisit atau lewat proses pelajaran) untuk berusaha keras dan mengambil resiko dan dihargai karena melakukannya.*	*Banyak siswa yang berpartisipasi dalam tugas yang menantang selama setidaknya setengah dari pelajaran. Mereka didorong (secara eksplisit atau lewat proses pelajaran) untuk berusaha keras dan mengambil resiko dan dihargai karena melakukannya.*	*Hampir semua siswa berpartisipasi dalam pekerjaan yang menantang selama hampir sepanjang pelajaran. Mereka didorong (secara eksplisit atau lewat proses pelajaran) untuk berusaha keras dan mengambil resiko dan dihargai karena melakukannya.*	*Seluruh siswa berpartisipasi dalam pekerjaan yang menantang sepanjang pelajaran. Mereka didorong (secara eksplisit atau lewat proses pelajaran) untuk berusaha keras dan mengambil resiko dan dihargai karena melakukannya.*
2.4 Dukungan sosial	*Dukungan sosial rendah. Tindakan atau komentar dari guru atau siswa menyebabkan rasa malu atau merendahkan dan suasana kelas negatif.*	*Dukungan sosial yang bercampur. Terlihat adanya kedua perilaku atau komentar baik yang merendahkan dan memberi dukungan dalam observasi.*	*Dukungan sosial netral atau agak positif. Walaupun tidak terlihat ada perilaku yang merendahkan, perilaku atau komen yang positif diarahkan kepada siswa yang paling aktif dalam pelajaran, dibandingkan dengan siswa yang kurang aktif.*	*Dukungan sosial secara jelas positif. Perilaku dan komen yang penuh dukungan ditujukan kepada kebanyakan siswa, termasuk usaha yang jelas dalam mendukung siswa yang tidak ragu-ragu atau kurang aktif.*	*Dukungan sosial sangat kuat. Perilaku atau komentar yang penuh dukungan dari siswa dan guru diarahkan bagi seluruh siswa, termasuk mengajak dan menghargai masukan atau kontribusi dari semua.*
diri peserta didik	*Sedikit siswa menunjukkan otonomi dan inisiatif dalam mengatur perilaku mereka sendiri. Guru menghabiskan waktu lebih banyak untuk mendisiplinkan dan mengatur perilaku siswa daripada belajar*	*Beberapa siswa menunjukkan otonomi dan inisiatif dalam mengatur perilaku mereka sendiri., tetapi tetap ada gangguan yang cukup banyak terhadap proses pembelajaran dalam hal ini mendisiplin atau hal*	*Banyak siswa menunjukkan otonomi dan inisiatif dalam mengatur tingkah laku mereka sendiri dan pelajaran berlangsung dengan terkordinir. Akan tetapi, guru mengatur sikap siswa beberapa kali, membuat*	*Hampir seluruh siswa, sepanjang pelajaran, menunjukkan otonomi dan inisiatif dalam mengatur tingkah laku mereka dan terdapat hanya sedikit gangguan dalam pembelajaran. Sekali atau dua kali*	*Semua siswa, hampir sepanjang jam pelajaran, menunjukkan otonomi dan inisiatif dalam mengatur perilaku diri sendiri dan pelajaran berlangsung tanpa interupsi.*

2.5 Regulasi	*dan mengajar.*	*yang berkaitan dengan aturan kelas, sebagai usaha untuk mengurangi perilaku yang tidak baik, memperbaiki perilaku buruk dari waktu lalu atau sebagai reaksi langsung terhadap perilaku buruk siswa.*	*pernyataan mengenai tingkah laku bagi siswa di kelas, atau mungkin berfokus pada siswa yang bertingkah laku tidak pantas.*	*sepanjang pelajaran, guru berkomentar atau memperbaiki perilaku siswa atau pergerakan siswa.*	
2.6 Pengarahan oleh peserta didik	*Tidak terlihat bukti kemandirian peserta didik. Segala aspek dari pelajaran secara eksplisit di rancang oleh guru untuk siswa.*	*Rendahnya kemandirian siswa. Walaupun siswa diberikan beberapa kontrol dalam beberapa aspek pelajaran (pilihan, waktu, kecepatan, penilaian), kontrol yang mereka punya minimal atau sedikit.*	*Terdapat sedikit kemandirian siswa. Siswa mendapat sedikit kontrol yang berhubungan dengan beberapa aspek penting dalam pelajaran.*	*Kemandirian siswa yang cukup banyak. Negosiasi terjadi di antara guru dan siswa mengenai beberapa aspek penting dalam pelajaran.*	*Tingginya kemandirian siswa. Siswa menentukan banyak aspek penting dalam pelajaran baik secara mandiri atau bergantung pada persetujuan guru.*

Garis Besar Skala Pengkodean Signifikansi

Aspek	1	2	3	4	5
3.1 Latar belakang pengetahuan	*Pengetahuan dasar siswa tidak disebutkan atau disampaikan.*	*Pengetahuan dasar siswa disebutkan atau disampaikan, tetapi tidak secara mendalam dan tidak berhubungan dengan inti pelajaran.*	*Pengetahuan dasar siswa disebutkan atau disampaikan secara garis besar, berhubungan dengan inti pelajaran, dan terdapat beberapa hubungan dengan pengetahuan siswa di luar sekolah.*	*Pengetahuan dasar siswa disebutkan dan disampaikan beberapa kali, berhubungan dengan inti pelajaran, dan terdapat beberapa keterkaitan dengan pengetahuan luar sekolah siswa sebelumnya.*	*Pengetahuan dasar siswa secara konsisten tergabung dalam pelajaran, dan ada hubungan yang cukup banyak dengan pengetahuan dari luar sekolah siswa sebelumnya.*
3.2 Pengetahuan budaya	*Terbukti tidak ada pengakuan secara eksplisit atau penghargaan terhadap budaya selain budaya dari budaya yang dominan sepanjang pelajaran.*	*Beberapa pengetahuan budaya terbukti dalam pelajaran, tetapi di perlakukan dengan cara yang tidak mendalam.*	*Beberapa pengetahuan budaya diakui dan dihargai dalam pelajaran, tetapi dalam cara berpikir dari budaya yang dominan.*	*Pengetahuan budaya yang cukup banyak diakui/diketahui dan dihargai dalam pelajaran dengan beberapa tantangan terhadap cara berpikir dari budaya yang mendominasi.*	*Pengetahuan budaya secara substansi diakui dan dihargai sepanjang pelajaran dan pengetahuan ini diterima secara sama dengan budaya yang dominan.*
3.3 Integrasi pengetahuan	*Tidak ada koneksi yang bermakna. Seluruh pengetahuan yang diajarkan terbatas hanya mencakup satu topik atau pokok bahasan.*	*Terdapat beberapa koneksi yang tidak signifikan yang dibuat. Pengetahuan pada umumnya terbatas pada satu topik spesifik saja atau pokok bahasan.*	*Setidaknya ada satu koneksi yang berkaitan dan bermakna yang dibuat di antara topik dan pokok bahasan/pelajaran oleh guru dan/atau siswa selama pelajaran.*	*Beberapa koneksi yang bermakna dan berkaitan dibuat antara topik dan pokok bahasan/pelajaran oleh guru dan/ atau siswa selama pelajaran.*	*Koneksi yang bermakna dibuat secara tetap di antara topik atau area pelajaran oleh guru dan/ atau siswa selama pelajaran.*
3.4 Keterinklusifitas	*Beberapa siswa tidak dilibatkan, atau tidak melibatkan diri mereka, dalam kegiatan pembelajaran sepanjang pelajaran.*	*Beberapa siswa tidak dilibatkan, atau melibatkan diri mereka, selama hampir seluruh kegiatan belajar, namun terlibat dalam kegiatan yang kecil satu atau dua kali selama pelajaran.*	*Siswa dari seluruh kelompok terlibat dalam hampir semua aspek dalam pembelajaran, tetapi keterlibatan beberapa siswa dari kelompok tertentu terbatas atau tidak banyak dibandingkan dengan kelompok lainnya.*	*Siswa dari seluruh kelompok terlibat secara signifikan dalam hampir seluruh aspek pembelajaran, tetapi masih terdapat beberapa ketidaksamaan dalam keterlibatan oleh beberapa kelompok sosial yang berbeda.*	*Siswa dari seluruh kelompok terlibat dalam seluruh aspek dari pelajaran dan keterlibatan mereka signifikan dan sama rata dengan siswa yang berasal dari kelompok sosial yang berbeda.*

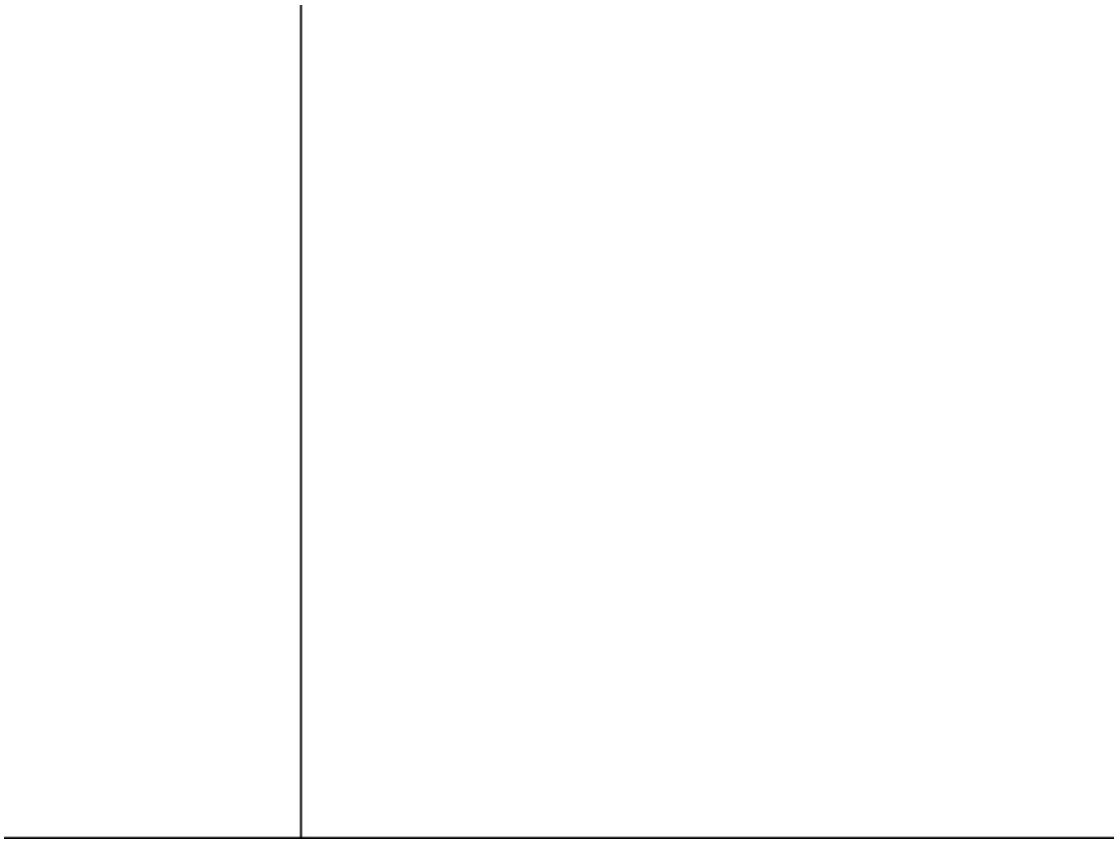
3.5 Keterhubungan	*Pelajaran tidak ada hubungan jelas dengan hal lain di luar pelajaran. Baik guru maupun siswa tidak memberikan alasan mengapa pelajaran tersebut mempunyai hubungan dengan lingkungan di luar sekolah.*	*Guru atau siswa mencoba menghubungkan apa yang dipelajari dengan lingkungan di luar kelas, tetapi hubungannya sangat lemah dan superfisial atau jarang.*	*Siswa mengenali beberapa hubungan dengan pengetahuan dan situasi di luar ruangan kelas, yang mungkin termasuk berbagi pekerjaan mereka dengan audiens di luar kelas, tetapi mereka tidak mengekspor dampak dari hubungan dan pengetahuan tetap berupa pengetahuan abstrak dan dugaan.*	*Siswa mengenal dan mengeksplorasi koneksi diantar pengetahuan yang dipelajari di kelas dan situasi di luar kelas dengan cara menciptakan makna personal dan menekankan pentingnya pengetahuan tersebut. Mungkin akan ada usaha untuk mempengaruhi audiens di luar kelas.*	*Siswa mengenal dan mengeksplorasi koneksi di antara pengetahuan dikelas dan situasi di luar kelas dengan cara yang menciptakan makna personal dan menekankan pentingnya pengetahuan tersebut Makna dan signifikansi dari pengetahuan tersebut cukup kuat agar siswa dibimbing sehingga terlibat dalam usaha mempengaruhi audiens yang di luar kelas.*
3.6 Narasi/cerita	*Narasi/cerita tidak digunakan dalam tahap apa pun di pelajaran, atau cerita yang digunakan tidak memiliki hubungan dengan pokok pelajaran.*	*Narasi/cerita digunakan namun merupakan bagian kecil dari pelajaran dan/atau tidak terlalu berhubungan dengan pokok pelajaran.*	*Narasi/cerita digunakan pada beberapa bagian pelajaran untuk meningkatkan pentingnya pokok pelajaran.*	*Narasi/cerita digunakan pada bagian besar pelajaran untuk meningkatkan pentingnya pokok pelajaran.*	*Narasi/cerita digunakan sepanjang pelajaran untuk meningkatkan pentingnya pokok pembelajaran.*

Appendix 2: Lesson Observation Sheet (Indonesian Version)

From Lesson Observation Sheet, © QT Academy, 2023. Translated by H.-C. Walker, January 2024.

Lembar Observasi Kelas

Tanggal:	Kelas:	MaPel:
Dimulai pada jam:	Selesai pada jam:	Jumlah Peserta Didik:
CP/Tujuan Pembelajaran:		Topik Pembelajaran:
Waktu	**Catatan Observasi**	
		Kotak catatan angka (Opsional)



Appendix 3: *Koding* Sheet (Indonesian Version)

From 'Quality Teaching in NSW Public Schools: A Classroom Practice Guide,' Third Edition. Copyright © The State of NSW, Department of Education and Training, Professional Support and Curriculum Directorate, 2020, p. 49: Coding Sheet. Translated by H.-C. Walker, January 2024.

Lembar Pengkodean

Nama Anda: _____ *Tanggal:* _____ *Kelas:* _____ *Mata Pelajaran:*_____

*Topik Pembelajaran:*_____

*Tujuan Pembelajaran:*_____

	Aspek	***Keterangan***	***Kode Individu***	***Kode PLC***
Kualitas Intelektual	*1.1 Pengetahuan mendalam*			
	1.2 Pemahaman secara			
	1.3 Asal-usul pengetahuan			
	1.4 Pemikiran tingkat tinggi			
	1.5 Metabahasa			

	1.6 Komunikasi yang substantif
Lingkungan Belajar yang Berkualitas	*2.1 Kriteria kualitas eksplisit*
	2.2 Keterlibatan
	2.3 Harapan tinggi
	2.4 Dukungan sosial
	2.5 Regulasi diri peserta didik
	2.6 Pengarahan oleh peserta
Signifikansi	*3.1 Latar belakang pengetahuan*
	3.2 Pengetahuan budaya
	3.3 Integrasi pengetahuan
	3.4 Keterinklusifitas
	3.5 Keterhubungan
	3.6 Narasi/cerita