SMERU Technical Report

Classroom Observation Tool for Assessing the Dimensions of Teaching Practices (CERDAS)

Asri Yusrina

Luhur Bima

smeru RESEARCH INSTITUTE

SMERU TECHNICAL REPORT

Classroom Observation Tool for Assessing the Dimensions of Teaching Practices (CERDAS)

Asri Yusrina

Luhur Bima

Reviewed by

Daniel Suryadarma, Menno Pradhan, and Thomas Coen

Editor

Alia An Nadhiva

The SMERU Research Institute

September 2020

Classroom Observation Tool for Assessing the Dimensions of Teaching Practices (CERDAS)

Authors: Asri Yusrina, Luhur Bima

Editor: Alia An Nadhiva
Cover photo: Novita Eka Syaputri

The SMERU Research Institute Cataloging-in-Publication Data

Asri Yusrina
Classroom Observation Tool for Assessing the Dimensions of Teaching Practices (CERDAS)/ Asri Yusrina, Luhur Bima; Editor, Alia An Nadhiva.
--Jakarta: SMERU Research Institute, 2020
--ix; 44 p; 29 cm.
ISBN 978-623-7492-37-5
ISBN 978-623-7492-38-2 [PDF]

1. Classroom Observation 2. Teaching Practices
I. Title

371.102 –ddc 23

Published by:
The SMERU Research Institute
Jl. Cikini Raya No. 10A
Jakarta 10330
Indonesia

First published in September 2020

A significant part of the research in this publication uses interviews and focus group discussions. All relevant information is recorded and stored at the SMERU office.

For further information on SMERU's publications, please contact us on 62-21-31936336 (phone), 62-21-31930850 (fax), or smeru@smeru.or.id (e-mail); or visit www.smeru.or.id.

RESEARCH TEAM

SMERU Researchers

Asri Yusrina

Luhur Bima

ACKNOWLEDGEMENTS

This report is produced by the RISE Programme in Indonesia to provide information on teaching processes that may contribute to student learning outcomes. Asri Yusrina and Luhur Bima authored the report under the operational leadership of Daniel Suryadarma and overall guidance of Sudarno Sumarto. We extend our gratitude as well to Menno Pradhan and Thomas Coen for reviewing the report.

We also acknowledge the efforts made by those involved in the development of this report. We thank Niken Rarasati and Goldy Dharmawan for their contribution in the pilot of the instrument and input on the data analysis. We are also grateful for the input provided by Shintia Revina in the scoring section. Without their support, this report would not have been possible.

We would also like to thank researchers involved in piloting the classroom observation instrument: Ulfah Alifia, Anisah Hafiszha Zulfa, I Wayan Arya Swarnata, Mirza Annisa Izati, Valentina Utari, and Awliya Rahimah. The pilot of the instrument was accomplished with the support of Totok Suprayitno, head of the Research and Development Agency, Ministry of Education and Culture, who facilitated access to schools. We also extend our gratitude to teachers in Kabupaten Bogor and Kota Bogor who participated in the pilot.

ABSTRACT

Classroom Observation Tool for Assessing the Dimensions of Teaching Practices (CERDAS)

Asri Yusrina and Luhur Bima

Observation of teacher practice may provide information on whether learning takes place in the classroom. We develop this observation instrument to describe and investigate teaching practices of primary and secondary school teachers. The design of our instrument is drawn upon aspects of teaching from two national teacher evaluations and selected international observation instruments. Our instrument consists of 8 key activities and 26 indicators. The eight key activities are (i) introducing a lesson, (ii) core activity in teaching, (iii) making connections in teaching, (iv) the use of learning materials by a teacher, (v) students' participation in learning, (vi) assessment of learning process and outcomes, (vii) control of the classroom environment by a teacher, and (viii) concluding a lesson. We present the most prevalent teaching practices in four themes: lesson structure, the structure and types of teacher-student interaction, questioning, and supportive classroom environment. These themes serve as observable proxies for learning. When the themes are used in conjunction with other data, such as student learning gain, student characteristics, and school level data, one can assess whether the absence of the themes would indicate poor student learning. We acknowledge limitation in our instrument, such that the instrument may not capture how a teacher provides practice in the classroom, and that teaching themes in the scoring section may not completely satisfy individual interest in specific teaching practices. When using this instrument in data collection, we suggest to have an observer in the classroom and videotape the teacher while teaching.

Keywords: classroom observation, teaching practice

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i

ABSTRACT ii

TABLE OF CONTENTS iii

LIST OF TABLES iv

LIST OF FIGURES iv

LIST OF APPENDICES iv

LIST OF ABBREVIATIONS v

EXECUTIVE SUMMARY vi

I. INTRODUCTION 1
- 1.1 Objective 1
- 1.2 Review of Teacher Evaluations at the National Level 2
- 1.3 Review of Observation Instruments Developed by NGOs/Development Partners 3

II. CLASSROOM OBSERVATION TOOL FOR ASSESSING THE DIMENSIONS OF TEACHING PRACTICES (CERDAS) 7
- 2.1 Design of the Instrument 7
- 2.2 Aspects of the Instrument 9

III. DEVELOPMENT OF THE INSTRUMENT 15
- 3.1 First Pilot 15
- 3.2 Second pilot 19

IV. OBSERVATION MANUAL 23
- 4.1 Observation Procedure 23

LIST OF REFERENCES 26

APPENDICES 28

LIST OF TABLES

Table 1. Preexisting or Established Observation Instruments 4

Table 2. A Comparison of CERDAS with the Two National-Level Teacher Formative Evaluations (PKG and the Instrument to evaluate PPL) 8

Table 3. Matching Agreement of Observers 22

LIST OF FIGURES

Figure 1. Stages of teaching in PKG and the instrument to evaluate PPL 3

Figure 2. The camera was set up in front of the class to record students' activities 17

Figure 3. The camera was set up at the center back of the classroom to record the teacher's activities 17

LIST OF APPENDICES

Appendix 1. Description of the Observation Sheet 29

Appendix 2. Scoring Manual 36

LIST OF ABBREVIATIONS

CAPI		Computer-Assisted Personal Interviewing
CLASS		Classroom Assessment Scoring System
EQUIP-Tanzania		Education Quality Improvement Programme-Tanzania
FFT		Framework for Teaching
ICALT		International Comparative Analysis of Learning and Teaching
LKS	*Lembar Kerja Siswa*	student worksheet/books
NGO		nongovernmental organization
PCA		Principal Component Analysis
Permendiknas	*Peraturan Menteri Pendidikan Nasional (sekarang Peraturan Menteri Pendidikan dan Kebudayaan)*	Minister of National Education Regulation (now Minister of Education and Culture Regulation)
PKG	*Penilaian Kinerja Guru*	Teacher Performance Evaluation
PPG	*Pendidikan Profesi Guru*	Preservice Teacher Professional Education
PPL	*Praktik Pengalaman Lapangan*	field experience practice
STIR		Schools and Teachers Innovating for Results
TIMSS		Trends in International Mathematics and Science Study

EXECUTIVE SUMMARY

Background

The RISE Programme in Indonesia seeks to produce knowledge of how education policies in Indonesia take effect at the national and *kota/kabupaten* levels. At the national level, we examine the impacts of education policies relating to teacher recruitment, teacher professional development, and preservice teacher training. We develop a classroom observation instrument, Classroom Observation Tool for Assessing the Dimensions of Teaching Practices (CERDAS), in an attempt to evaluate teacher practice by focusing on the teaching process. We believe observation of teacher practice may provide information on whether learning takes place in the classroom.

The design of CERDAS is drawn upon aspects of teaching from two national teacher evaluations, the Teacher Performance Evaluation (PKG) and the instrument to evaluate PPL, a field practice program performed by teacher candidates, and selected international observation instruments. We are not confident to solely refer to the two national evaluations as their instruments are not provided with supplementary descriptive information on its indicators, which will be difficult to use by observers who have no teaching background. Furthermore, after reviewing several international observation instruments, we found that some of them may not work exactly as intended for the Indonesian context. Ultimately we decided to develop an observation instrument that measures the frequencies of teaching practices.

The main objective of our observation instrument is to describe and investigate teaching practices of primary and secondary school teachers. Our target observation includes teachers and teacher candidates who teach Indonesian and mathematics. We also note that the effective teaching practice should be regarded in relation to the frequency of specific practices carried out, and how a teacher provides the practice in the classroom. To analyze the observation data using a mixed methods approach, we suggest to have an observer in the classroom and videotape the teacher while teaching.

Design and Aspects of the Instrument

Based on our reviews of the international observation instruments and the two national teacher evaluations, there are three points that we take into consideration when we develop CERDAS.

a) We incorporate three stages of activities into teaching i.e., introducing a lesson, core lesson, and concluding a lesson. The observation instrument that aligns with our interest is the one developed by the Education Quality Improvement Programme-Tanzania (EQUIP-Tanzania), which is similar to the two of Indonesia's teacher evaluations.
b) We use the two national teacher evaluations as our benchmark in developing our instrument's indicators, and incorporate "yes-no" response indicators from observation instruments developed by nongovernmental organizations (NGOs) or development partners that capture the national framework.
c) We measure frequencies of specific teaching practices and are likely to use observers without teaching background, thus, developing a low inference instrument is more applicable. Most of the low inference observation instruments developed by NGOs or development partners are somewhat a modification of the Stallings system. After analyzing the pros and cons of using the Stallings system, we decided to develop our own instrument, which is more appropriate to the Indonesian context.

In developing the indicators, we consider (i) indicators that can be incorporated by providing more contextual information; (ii) indicators that need to be derived into a simple rating or "yes-no" type; and (iii) indicators that we thought are better left out due to unclear wording or lack of feasibility to assess using the "yes-no" type.

We exclude the category of teacher's use of language in the national formative evaluations when determining aspects of teaching practices in the instrument. The reason is the category has unclear wording and lack of indicators definition. For example, it will be difficult to have a "yes-no" response for the "teacher delivers messages with a proper style" indicator. We also add another aspect in our instrument, core activity in teaching, to somewhat incorporate the teaching activities that we found in the Trends in International Mathematics and Science Study (TIMSS) Video Study.

The teaching practices observed in our instrument are categorized into eight key activities that reflect the teaching aspects assessed in the national evaluation. The eight key activities are (i) introducing a lesson, (ii) core activity in teaching, (iii) making connections in teaching, (iv) the use of learning materials by a teacher, (v) students' participation in learning, (vi) assessment of learning process and outcomes, (vii) control of the classroom environment by a teacher, and (viii) concluding a lesson. At the end of the observation, the observers fill in the "Overall observation of classroom and teaching practice" section.

More on the aspects of the classroom observation instrument, the observation sheet, and the scoring manual
• Introducing a lesson: Section 2.2.1
• Core activity in teaching: Section 2.2.2
• Making connections in teaching: Section 2.2.3
• The use of learning materials by a teacher: Section 2.2.4
• Students' participation in learning: Section 2.2.5
• Assessment of learning process and outcomes: Section 2.2.6
• Control of the classroom environment by a teacher: Section 2.2.7
• Concluding a lesson: Section 2.2.8
• Description of the observation sheet: Appendix 1
• Scoring manual—the most prevalent teaching themes: Appendix 2

Development of the Instrument

The stages of developing the instrument are as follows:

a) A series of internal discussion and consultation with an expert on teaching who was involved in using an observation instrument to evaluate teaching practices.

b) First pilot, which aimed at improving the instrument. The piloting is to compile possible examples of teaching and to check whether the observers have different interpretations of the indicator during live observation.

c) A two-hour internal try out of Computer-Assisted Personal Interviewing (CAPI) technique to resolve coding bugs using several videos from the first pilot.

d) Second pilot, which aimed at evaluating the structure of the instrument that had been modified into eight aspects of teaching practices. We wanted to see whether the instrument

is easy to follow during live observation. The second pilot was also meant to check the agreement among observers in one team in relation to the indicators.

Piloting the instrument resulted in changes of indicators; revision on the procedural of the observation and the videotaping; and changes in the section of overall observation filled in once the lesson ended. The results of the pilots are used as a basis for developing the observation manual.

Observation Manual

Our observation instrument is segmented into three stages of lesson: introducing a lesson, core lesson, and concluding a lesson. Each of both introducing-a-lesson and concluding-a-lesson stages has two indicators that use simple ratings, in which definitions and possible examples are provided. The core lesson includes 22 "yes-no" indicators that record the indicators' presence or absence based on the observation. At the end of observation and videotaping, the observers fill in the overall observation of classroom and teaching practice.

a) Observations should be conducted for at least two sessions. In primary schools, the first session equals to 35 minutes of teaching, while in secondary schools it is 45 minutes. To know which teachers are teaching the mathematics and Indonesian, we suggest to have an advance visit so that the observer can take notes on the school timetable.
b) The observation is carried out by one observer in one classroom. The role of the observer is to watch and listen for signs of observable teacher practices listed and to record it without making judgments as to relative importance or relevance of those practices.
c) The observer, who carries out live observation, may include people without teaching background trained specifically for the project.
d) The observation data is collected using SurveyCTO software in an Android tablet.
e) The live observation is also videotaped for a viewing session held separately, either for another checking or for implementing other observation instruments when needed.
f) The location of the observer and the position of the tablet should be at the back corner (left or right) of the class. Placing the tablet at the back corner allows for a wider range of the camera to follow teacher activities in the classroom.

On the day of the observation and prior to the observed class, we suggest to hold a preobservation meeting between the observer and the teacher. This would allow the observer to explain to the teacher the observation protocols, clarify issues, and have the teacher fill in the consent form. List of areas to cover during this meeting are as follows:

a) Protocols regarding confidentiality as explained in the consent form
b) Where the observer will sit
c) Whether the observer's presence needs to be explained to students
d) How the observation will be recorded during the session
e) The completion of non-observable items in the teaching observation instrument, such as school location, teacher's ID, teacher's gender, information about the lesson's subject, and the number of hours of the lesson

The observer should enter the class before the teacher begins the lesson in order to:

a) Explain the purpose of the observer's presence to students (as agreed by the teacher)
b) Place themselves in the agreed position

c) Be as discrete and unobtrusive as possible (minimize contact/communication with students)
d) Record the time they enter the class
e) Collect information on the teacher's activity before the teacher starts the lesson
f) Record the time the observation begins

At the beginning of the lesson, the observer spends a maximum of 10 minutes to mark the practice of the teacher introducing the lesson. During the lesson, the observer spends each 5-minute period to observe the teacher's behavior and mark if any observable practices occur. Leave blank the space of indicators/practices which did not occur. The observer should consider 22 indicators/practices listed in the core lesson, mark them, or leave them blank. One indicator can only be marked once in a provided checklist box, regardless of how many times that practice occurs within the 5-minute observation period. The observer should record each of the 5-minute observations. A marking period is indicated by a question of whether the lesson continues or has ended.

The observer should also take some notes on teaching situation during the lesson, such as how long the teacher leaves the classroom, and how many times guests visit or talk to the teacher inside or in front of the classroom.

The observer should make additional notes when an unexpected situation occurs. For instance, the observed teacher suddenly must leave the classroom for an urgent reason and then replaced by another teacher. Furthermore, the observer can also make a note when there is a condition which can affect normal classroom activities, such as the school principal observing the classroom.

I. INTRODUCTION

Shifting the focus of an education system from access to learning requires evidence of the extent to which the current education policies shape learning outcomes. The RISE Programme in Indonesia aims to produce knowledge of how education policies in Indonesia take effect at the national and *kota/kabupaten* levels. At the national level, we examine the impacts of education policies relating to teacher recruitment, teacher professional development, and preservice teacher education.

Two approaches are widely used to understand how education policies affect teacher quality. The first approach attempts to evaluate teachers by focusing on student learning outcomes, which is measured by the value-added of student achievement—generally known as Value Added Model. This approach may reflect a teacher's contribution to the changes in the test scores of the same student in two different times (Hanushek and Rivkin, 2010). We carried out this approach using the Comprehensive Reading and Mathematics Assessment Tool (CERMAT), our self-designed student learning assessment tool (see Rarasati et al., 2019).

The second approach also attempts to evaluate teachers by focusing on the teaching process. The evaluation is measured using a classroom observation instrument.[1] Observation of teacher practice may provide information on whether learning takes place in the classroom. This approach receives equal attention as a robust method as the Value Added Models (Martinez, Taut, and Schaaf, 2016). It may also serve as a good proxy of teacher quality, which contributes to student learning achievement (Bill and Melinda Gates Foundation, 2012). Moreover, the evaluation of teacher practice may contribute to the improvement of teacher training methods as well as provide feedback to the individual teacher (Pianta and Hamre, 2009; Bill and Melinda Gates Foundation, 2012).

1.1 Objective

In our opinion, the education system behind each instrument is as important as the instrument. We initially looked at two of Indonesia's national teacher evaluations, the Teacher Performance Evaluation (PKG) and the instrument to evaluate PPL, a field practice program performed by teacher candidates, as well as several international observation instruments. We then decided to develop our own classroom observation instrument because there are some aspects of teaching in the two national teacher evaluations that are not assessed or included in the international observation instruments. This suggests that some of the international observation instruments may not work exactly as intended for the Indonesian context. Meanwhile, to use the two national teacher evaluations alone also limits us since the instruments require observers with teaching experiences. We review the two national teacher evaluations and sets of teacher observation instruments developed by nongovernmental organizations (NGOs) as well as development partners in the next two sections before describing the classroom observation instrument that we develop.

We use our observation instrument to describe and investigate teaching practices of primary and secondary school teachers. Our target observation includes teachers and teacher candidates who teach Indonesian and mathematics. Through observation of teaching practices, we also intend to

[1]Another suggested instrument for the second approach is a survey of student's perception regarding their classroom learning experience (Waxman, 1991). However, the survey is argued to be used only for students in 9th grade and above. Since the RISE Programme in Indonesia focuses on basic education, the instrument relevant to our study is the classroom observation.

explore which practices are used and how they relate to student learning outcomes. We expect that the relationship between teaching practices and student learning outcomes, holding other things constant, provide insights of what may be relatively more effective teaching practices in Indonesia. We also note that the effective teaching practice should be regarded in relation to the frequency of specific practices carried out, and how a teacher provides the practice in the classroom. However, our observation instrument is limited as it only provides the frequencies of teaching practices.

In preparing our classroom observation instrument, we consider findings from the Trends in International Mathematics and Science Study (TIMSS) Video Study on teaching practices in Indonesia, in addition to drawing upon aspects of teaching from the national teacher evaluations and the international observation instruments. The longitudinal study from 2007 to 2011 revealed that in mathematics subject, lecture continues to be the dominant approach in teaching practice, while group work decreases (Ragatz et al., 2015). We are interested to see whether teacher's approach to teaching has changed since the implementation of the 2013 Curriculum, Indonesia's latest curriculum, which focuses on the aspect of student-centered teaching.

1.2 Review of Teacher Evaluations at the National Level

Evaluation of teacher quality heavily depends on local context and culture; however, some aspects of quality teaching might appear on various observation instruments with different local contexts (Martinez, Taut, and Schaaf, 2016). Similarly, we intend to design our instrument based on the formative evaluation of teachers of the national framework. We chose PKG and the instrument to evaluate PPL as our references in designing our observation instrument because both are used nationally. The two evaluations also meet the expected teaching standard of teachers in Indonesia and align with the standard of teacher competence set by the government.[2]

PKG is carried out regularly by school supervisors to assess in-service teachers across primary and secondary schools in Indonesia. The result of this assessment is usually used to pave the way of a teacher's professional development and promotion. PPL, on the other hand, is only available at universities with teacher preparation programs.[3] PPL is mandatory only for preservice teachers undergoing Preservice Teacher Professional Education (PPG). Lecturers of the PPG program observe and assess teacher candidates when they practice teaching in partner schools of the respective teacher training institute. The PPL aims to assess the knowledge teacher candidates gained from the workshops of the PPG program.

In general, the two formative evaluations share some similarities. First, both assessments are carried out by an observer with a teaching background (lecturers and school supervisors), and the results are expected to provide feedback to teachers or teacher candidates. Second, the structure of the assessments is arranged into three stages of activities: introducing a lesson, core lesson, and concluding a lesson (Figure 1). Unlike the first and third stages, where each only has one aspect, the second stage consists of six aspects, each of which has indicators. Third, the indicators are assessed using a rating score. However, guidelines for both assessments do not provide rubric or examples for rating each indicator.

[2]To improve the quality of teachers, the government issued the Minister of National Education Regulation (*Permendiknas*) Number 16/2007 on the Standards of Teacher's Academic Qualification and Competence.

[3]Universities that have teacher education programs also use other forms of evaluation for their teacher candidates. Typically, that evaluation is tailored to each university's values.

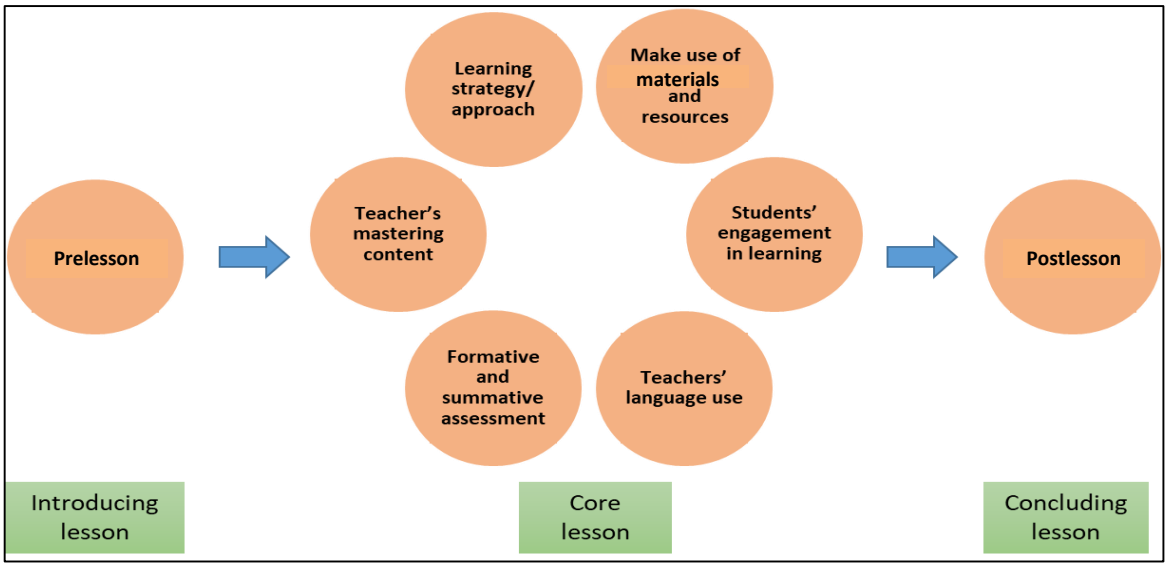

Figure 1. Stages of teaching in PKG and the instrument to evaluate PPL

Note: In PKG, the aspect of learning strategy/approach is differentiated into scientific approach and instructional approach.

Of both evaluations, we see one characteristic as detrimental. Each indicator in the two teacher evaluations is not categorized into or associated to a certain domain or element of teaching as in the international observation instruments. For example, in the assessment, there is a category of teacher makes use of instructional or learning materials, which could fall in the element of instruction, but is not informed in the guidelines. Both evaluations also use ratings without rubric or examples (the indicators only measure elements of teaching globally). The tool uses three rating levels to evaluate teacher's performance: "0" (does not meet the standard), "1" (partly meets the standard), and "2" (meets the standard). There is no information on the length of observation, but the structure of both observations seem to require observers (school supervisors or lecturers) to evaluate an entire lesson.

The expected teaching practices in the formative evaluations also reflect the 2013 Curriculum, which emphasizes a student-centered approach. This requires a key focus area on the teacher-student interaction. However, both assessments tend to measure behaviors that focus merely on teachers and not necessarily on teacher-student interactions. For example, there is an indicator of whether the teacher encourages students to be actively involved, but there is not any that capture what they do.

1.3 Review of Observation Instruments Developed by NGOs/Development Partners

We acknowledge several observation instruments of which reliability and validity have been tested in various disciplines, grades, or countries (see Table 1). We categorized the instruments into two groups. The first group includes Classroom Assessment Scoring System (CLASS), the 2013 edition of Framework for Teaching (FFT), and the International Comparative Analysis of Learning and Teaching (ICALT). We view these instruments as high inference as they require the observers to assess the quality or the coherence of teaching practices in the classroom. To be able to rate teaching quality with the instruments, observers must undergo a high standard training beforehand. Each instrument is provided with rubric and possible examples to limit the subjectivity of observers when scoring its indicators. Duration of observation is typically 20 minutes to 40 minutes (an event-

sampling observation instrument) or one period of lesson. Since both the design and observers are of high quality, we consider the instruments costly to use in a developing country setting such as Indonesia.

The second group consists of instruments that we view as low inference: Stallings, Schools and Teachers Innovating for Results (STIR), and Education Quality Improvement Programme (EQUIP)-Tanzania. All three instruments are used to identify activities in the classroom and do not assess the suitability or coherence of the teaching practices. The instruments describe the type or elements of teaching and the frequency of teaching practices conducted by a teacher, and less likely assess how a teacher delivers a lesson. The instruments' method of observation are in a specified time interval. For example, the observer records every five minutes or at a certain interval, depending on the length of the lesson. These types of instrument require shorter training for the observers to comprehend and do the coding. The most widely used observation instrument in developing countries is Stallings, while the others are somewhat modification of Stallings.

Overall, we note overlaps regarding domains or aspects observed. Instructional approach, classroom management, and student behavior occur in both high and low inference observation instruments. We could not conclude the ideal length of an observation as it depends on the context and the objective of the observation.

Table 1. Preexisting or Established Observation Instruments

No.	Instrument Name	Dimension/ Aspects Observed	Coding/Scoring Method	Observation Method	Observation Target
			High Inference		
1.	**Classroom Assessment Scoring System (CLASS)[a]**	There are three domains: 1. Emotional Support, includes three dimensions 2. Organizational Support, includes three dimensions 3. Instructional Support, includes three dimensions	Coders learned to rate each of the CLASS dimensions along a 1–7 scale, with • "1" or "2" indicating low quality; • "3", "4", or "5" indicating midrange quality; and • "6" or "7" indicating high quality.	Observers typically select two 20–25 minute segments of the class to apply the instrument. A 40-minute class equals to a coding of two 20-minute CLASS segment. Requires training and certification of observers for the purpose of rating the quality of teaching in the lessons.	Dimensions may vary with grade level but within the same hypothesized three-domain structure for preK to Grade 5. Somewhat different set of dimensions have been specified for Grades 6 to 12.
2.	**Framework for Teaching (FFT), the 2013 edition[b]**	There are four domains: 1. Planning and Preparation, includes six components 2. Classroom Environment, includes five components 3. Instruction, includes five components 4. Professional Responsibilities,	Teachers' teaching practices are evaluated based on four levels of performances: • Level 1 (Unsatisfactory) • Level 2 (Basic) • Level 3 (Proficient) • Level 4 (Distinguished)	The length of observation is typically 30 minutes or require that an observation occur over the course of an entire lesson. Domain 1 and 4 are off stage while Domain 2 and 3 are on stage. Requires training and certification of observers for the purpose of rating the quality of	The FFT could be applied to all disciplines from kindergarten to Grade 12. It is considered as grounded framework since teaching, in whatever context, requires the same basic tasks.

No.	Instrument Name	Dimension/ Aspects Observed	Coding/Scoring Method	Observation Method	Observation Target
			High Inference		
		includes six components In the 2013 edition, the four domains includes several components and each component consists of several smaller elements. The FFT protocols provides possible examples of teaching practices for each level of performance for each elements.		teaching in the lessons.	
3.	**International Comparative Analysis of Learning and Teaching (ICALT)[c]**	There are six domains: 1. Safe Learning Climate 2. Classroom Management 3. Clear Instruction 4. Activating Teaching Methods 5. Learning Strategies 6. Differentiation Observation instrument has six quality characteristics, which together comprises 24 indicators. Within each indicators, there are possible examples of teaching as guidance for scoring.	Inspectors/ observers can score these indicators: (1) predominantly weak; (2) more weaknesses than strengths; (3) more strengths than weaknesses; and (4) predominantly strong.	ICALT can be described as an event-sampling observation instrument. This instrument is brief and straightforward and can be completed easily by an inspector/observers during a lesson period of approximately 40 minutes. Requires training of observers for the purpose of rating the quality of teaching in the lessons despite there are relatively low-inference indicators.	ICALT has been applied in primary schools and secondary schools, particularly in European countries.
	Instrument Name	**Dimension/ Aspects Observed**	**Coding/Scoring Method**	**Observation Method**	**Observation Target**
			Low Inference		
4.	**Stallings[d]**	The instrument aims to estimate teachers' use of instructional time, teachers' use of materials, core pedagogical practices, and teachers' ability to keep students engaged. The activities are grouped into four categories: 1. Academic Activities/ Instruction	Checklist	Observations are coded at ten different moments in every class, at exact intervals which spacing depends on the length of the class; every 4 minutes in a 40-minute class, every 6 minutes in a 60-minute class, etc. Each observation consists of a 15-second scan of the classroom, starting	The instrument is language and curriculum-neutral. It could be applied to all disciplines and grade level.

	Instrument Name	Dimension/ Aspects Observed	Coding/Scoring Method Low Inference	Observation Method	Observation Target
		2. Classroom Management 3. Teacher Off-Task 4. Student Off-Task		with the teacher and proceeding clockwise around the room. Each 15-second observation is coded on a single sheet, thus each class observed generates ten coded sheets.	
5.	**Schools and Teachers Innovating for Results (STIR) - Education**[e]	What teachers do in the classroom were categorized into three activities: 1. Teaching 2. Classroom Management 3. Teacher Off-Task	Checklist, a modification of Stallings instrument.	Specified time of observation: observers to record every 5 minutes. The classroom snapshot section: questions are answered based on the first minute of observation. The classroom overview section: questions are answered based on the remaining four minutes.	The instrument has been applied in Grades 1–8.
6.	**Quality Improvement Programme (EQUIP) - Tanzania**[f]	Fifteen pedagogical strategies contribute to effective classroom practices, which include analysis on introductory stage and end-of-lesson stage. Key skills observed include: (i) introducing a lesson; (ii) questioning students; (iii) giving feedback to students; (iv) using instructional materials; (v) building rapport with students; and (vi) concluding a lesson.	Introductory and concluding-a-lesson stage: uses simple ratings of "0" (No), "1" (Partly), and "2" (Yes). Middle stage: uses checklist	Introductory stage is filled during the first 5 minutes of lesson. Middle stage: observer to record the main teaching activities from a list of prompts at every 5-minute interval in the lesson. Concluding-a-lesson stage is filled during 5–10 minutes of lesson.	The instrument has been applied in Grades 1–3 for teachers of Tanzania's national language and mathematics.

Source:

[a]CLASS: (Bruns, De Gregorio, and Taut, 2016)

[b]FFT: (Danielson, 2013)

[c]ICALT: (van de Grift, 2007)

[d]Stallings: (Bruns, De Gregorio, and Taut, 2016)

[e]STIR: (IDinsight, 2015)

[f]EQUIP-Tanzania: (Pettersson et al., 2016)

The high inference instruments commonly use video recording rather than live observation. Video recordings can be watched repeatedly; therefore, it may improve accuracy. However, it was argued that video recording may also lose valuable information on the classroom, mainly when recorded from a single or limited perspective.

II. CLASSROOM OBSERVATION TOOL FOR ASSESSING THE DIMENSIONS OF TEACHING PRACTICES (CERDAS)

2.1 Design of the Instrument

In this section, we describe the final observation instrument that went through several stages of development. We call our instrument the Classroom Observation Tool for Assessing the Dimensions of Teaching Practices (CERDAS). The process of developing CERDAS is elaborated in Section 3. Our observation instrument aims to describe and investigate teaching practices of primary and secondary school teachers in Indonesia. Based on the reviews in the previous section, there are some points that we consider when developing our observation instrument.

First, investigating whether teachers carried out three stages of activities during their teaching. The observation instrument that aligns with our interest is the one developed by EQUIP-Tanzania. It divides the pattern of class into three stages, which is similar to PKG. We adopted some indicators in EQUIP that are relevant to the teaching aspects assessed in both the PKG and the instrument to evaluate PPL, and modified the indicators to fit the Indonesian context.

Second, despite their relevancy to the Indonesian context, PKG indicators lack rubric or examples on how to rate teaching performance. We did not study how PKG's instruments were used, so we do not have the information on how the indicators are interpreted. The wording of each indicator does not allow observers without teaching background to easily identify the teaching practices that take place in the classroom. We used PKG and the instrument to evaluate PPL as our benchmark in developing our instrument, and incorporate "yes-no" response indicators from observation instruments developed by NGOs/development partners that capture the national framework.

Third, as we measure frequencies of specific teaching practices and are likely to use observers from nonteaching background, developing a low inference instrument is more applicable. Since most of the low inference observation instruments developed by NGOs/development partners are somewhat a modification of the Stallings system, we analyzed the pros and cons of using the system. The result of the discussion led us to develop our own instrument, which we regard to be more appropriate to Indonesian context. The followings are the conclusion of the discussion:

Pros:

a) The data collection process can make use of a group of observers with little or from nonteaching background; however, they would also have to take notes on the teaching activities in addition to the coding. Although the use of a paper-based Stallings instrument is considered to be cumbersome, the issue is addressed by the Stallings iPad-based application (MacKinnon et al., 2016) or Android-based application (World Bank, 2017).
b) Snapshot indicators generate aggregate feedback, enable benchmarking, and can be traced over time (Bruns, De Gregorio, and Taut, 2016).

Cons:

a) The observation takes one session of teaching throughout one lesson. It does not cover the lesson stages from the beginning to the end.

b) The observation focuses on the number of students engaged and the activities of students who are not engaged in the lesson rather than their involvement or active participation.

c) There are no indicators in Stallings that reflect teacher mastering the content, such as incorporating students' daily life experience or common knowledge when teaching. Whereas, this is one aspect of teaching practices that we found in the two national teacher evaluations that we reviewed.

The teaching practices observed in our instrument are categorized into eight key activities that reflect the teaching aspects assessed in the two teacher formative evaluations that we reviewed, as presented in Table 2.

Table 2. A Comparison of CERDAS with the Two National-Level Teacher Formative Evaluations (PKG and the Instrument to evaluate PPL)

No.	CERDAS	PKG and the Instrument to evaluate PPL
1.	Introducing a lesson	Introducing a lesson
	Core lesson	Core lesson
2.	Core activity in teaching	
3.	Making connections in teaching	Mastery of learning content by a teacher
4.	The use of learning materials by a teacher	The use of learning materials by a teacher
5.	Students' participation in learning	Teacher encouraging and engaging students in learning
6.	Assessment of learning process and outcomes	Assessment of learning process and outcomes
7.	Control of the classroom environment by a teacher	Learning strategy/approach
	Concluding a lesson	
8.	Concluding a lesson	

Prior to developing the eight key activities, we sorted out the indicators in each teaching aspect of the national teacher assessments. During the process, we categorized the indicators into (i) indicators that can be incorporated by providing more contextual information; (ii) indicators that need to be derived into a simple rating or "yes-no" type; (iii) indicators that we thought are better left out due to unclear wording or lack of feasibility to assess using the "yes-no" type.

We excluded the category of teacher's use of language in the national formative evaluations because most of the indicators are relatively challenging to incorporate due to unclear wording and lack of definition. Examples of the said indicators are "Teacher's verbal language is clear, good, and right" and "Teacher delivers messages with a proper style". We also added one aspect of observation, the core activity in teaching, to somewhat incorporate the teaching activities that we found in the TIMSS Video Study.

The stages of introducing and concluding a lesson use a simple rating to assess teacher's practices. The indicators required observers to give a rating of "0" when there is no evidence or practice is not delivered, "1" when more or less is proven, and "2" when the practice is clearly delivered. We provided explanation and examples of practices for each rating and indicator. Teaching practices during the middle of a lesson (the core lesson) is assessed using a "yes-no" answer in a specified time observation.

2.2 Aspects of the Instrument

2.2.1 Introducing a Lesson

In general, the aim of observing the introducing-a-lesson stage or prelesson is to see whether the teacher generates interest or attention from the students. In PKG, this first stage includes the indicators of teacher prepares student to learn and teacher does apperception setting. As mentioned earlier, the formative evaluation guideline lacks explanation on what practices a teacher does to provide the basis for students to learn new content. Observers with limited teaching as well as nonteaching backgrounds will find difficulties in differentiating what practices meet the standard of teaching and which do not. Thus, we introduce two indicators that reflect what a teacher does to engage students at the beginning of a lesson:

a) The teacher clearly outlines learning objectives of the new content to students.
b) The teacher checks students' knowledge.

When a teacher introduces new content, he or she should also communicate its learning objectives. When the teacher clearly outlines the learning objectives, students become aware of the type of knowledge they should be advancing in (Siraj and Taggart, 2014). This is also to avoid students making different interpretations of what the teacher intended (Park, Song, and Abrahams, 2015). For example, students may be able to solve a mathematics problem using the formula given, but are less likely to understand the principles of that solution. Preparing and introducing learning objectives also allows a teacher to make some benchmark of which learning strategy will take place (Klauer, 1988). Thus, teacher conveying the learning objectives of a new content is evidence of practices that aim to engage students.

Teacher can check students' knowledge by quoting previous lesson materials or daily activities relevant to today's learning. A teacher who prepares the new content of today's lesson well will make an effort to attract students' attention to the topic. Students can get conceptual knowledge of the new content when they can connect it to their experience or prior knowledge (Park, Song, and Abrahams, 2015).

We are aware that these two indicators only reflect parts of what a teacher does to provide the basis for students to learn the topic introduced. There are criticism of using the "teacher outlining learning objectives" indicator. Park, Song, and Abrahams (2015) argue that students might not achieve the teacher's intended learning objectives when it is not at their interest. Overall, we assume that at the very least those practices provide some convenience for students to learn. Furthermore, to investigate whether the teacher had any intention of articulating their learning objectives prior to teaching a lesson, we could incorporate the indicator into the teacher questionnaire. Thus, we could confirm their statement and the actual teaching practices in the observation instrument.

2.2.2 Core Activity in Teaching

Core activity in teaching is typically what a teacher does to make students understand or learn the new content mentioned in the learning objectives. We developed three indicators to capture this practice:

a) Teacher explains learning materials.
b) Teacher explains how to solve problems and/or gives examples of how to solve problems.
c) Teacher gives assignments on the new content to students in the classroom. The assignment is given in group work setting or individual seatwork.

The three indicators show the typical teacher-student interaction that occurs in the classroom. The TIMSS Video Study found that activities in the classroom which encourage teacher-student interaction leads to improvement in student learning (Ragatz et al., 2015). However, the three indicators are not self-explanatory in showing teacher-student interaction when the instrument is low inference. This suggests for other indicators, such as “student participation”, “teacher asking questions”, and “teacher circulating the classroom”.

We consider a teacher is lecturing if, during the observation, they focus on explaining new content and do not involve student participation. Similarly, when a teacher explains the correct answer to a problem and does not engage the students, that means the teacher-centered approach is still dominant. Typically, a teacher gives assignments to students once they finish explaining a new content. Teacher circulating the classroom and giving feedback to students when students are working on the assignments is evidence of teacher-student interaction. A teacher being active and involved throughout a lesson has a positive relationship with student learning outcomes, all else being equal (Ragatz et al., 2015). We include information about whether a teacher assigns students in group work or individual seatwork, but not the purpose of the setting.

2.2.3 Making Connections in Teaching

The third key activity that we observed comes from the aspect of mastery of learning content by a teacher in PKG. We adopted three indicators from that aspect and put them under the aspect of making connections in teaching in our instrument. We excluded two other indicators from the aspect of mastery of learning content by a teacher because their descriptions of the assessed teaching practices are unclear. It will be difficult to derive “yes-no” indicators that can capture whether a teacher carries out different learning strategies for weaker students. Moreover, observers with little teaching background will find difficulties in identifying the extent to which a teacher is knowledgeable about a new content. Thus, in our instrument, we regard that a teacher’s effort of making connections with students to gain a deeper understanding includes these indicators:

a) Teacher connects today’s content with other contents that have been taught before.
b) Teacher connects today’s content with situations or daily activities that are commonly experienced or carried out by students.
c) Teacher reiterates learning objectives in relation to today’s material.

According to Danielson’s framework (2013), incorporating students’ interests and daily experiences into a lesson is evidence of a teacher’s skill in managing the lesson’s course. Teachers who are connecting their lesson with other subjects or daily-life information can strengthen students’ understanding of the content; however, this practice is less likely to be performed by teachers in poor schools (Siraj and Taggart, 2014). Since we incorporate those indicators as a “yes-no” response, and also considering our observers’ background, there is no need to justify the relevancy of information mentioned by a teacher with the content of a lesson.

Restating the learning objectives is somewhat associated to the introducing-a-lesson stage, i.e., when a teacher is outlining learning objectives. However, the goal of restating the objectives in the middle of a lesson is to remind students about what they have learned as well as remind the teacher about the lesson’s course.

2.2.4 The Use of Learning Materials by a Teacher

In both the PKG and the instrument to evaluate PPL, the indicators require observers to rate the efficacy and the efficiency of learning materials used by a teacher and the appeal of teaching messages. As it would be difficult to modify those indicators into "yes-no" indicators, we explore the extent to which a teacher uses a variety of learning materials. It is also important to find whether a teacher actually uses learning materials at all. A study in Latin America and Caribbean found that teachers continue to use only blackboard/whiteboard and no other learning materials (Bruns and Luque, 2014).

To explain the use of learning materials by a teacher, we developed indicators as follows:

a) Teacher explains today's lesson through writing and/or diagrams that can be clearly viewed by the students.
b) Teacher uses electronic equipment, such as laptop and projector.
c) Teacher uses instructional materials that include, but not limited to, written or printed learning media (not the compulsory textbook for the lesson) prepared by the teacher; maps, charts, and cards.

In terms of writing and/or diagrams that can be clearly viewed by the students, the observers are to identify whether the students have difficulties in seeing or reading the teacher's handwriting on the board, or words displayed on the projector. In Danielson's 2013 framework, if students cannot see the writings, words, or pictures presented on the board, then learning cannot take place.

Effective use of electronics can also promote student learning (Danielson, 2013). Our indicator on the use of electronic equipment captures whether the words or pictures displayed on the projector can be seen clearly by the students. It merely shows that the electronics can be used by students and teachers, but there is no assessment on whether the projector displays suitable or proper content necessary for the lesson. Thus, the indicator tends to indicate a low level of effective use of electronic equipment. The use of a projector and a laptop may also depict a teaching practice that does not rely on a single learning material, such as a black/whiteboard.

Teacher's use of learning materials, whether it is technology-based or others, is important to keep students of lower grade engaged in learning and for teachers to deliver a well-organized lesson to higher grade students (Bruns and Luque, 2014). The richer and the more varied learning materials used in the classroom means the more likely a teacher is able to meet the individual needs of their students (Siraj and Taggart, 2014). Therefore, identifying the use of learning materials by a teacher is important because it may indicate that learning takes place. Our indicator does not measure the variety of the learning materials; it is, however, able to capture if a teacher uses learning materials other than the compulsory text book.

2.2.5 Students' Participation in Learning

We modified an instrument from both the PKG and the instrument to evaluate PPL, what a teacher is doing, into what the students are doing to ensure that our instrument is not focusing too much on the teacher so that we could better identify a teacher-student interaction. By doing so, we try to address the critical aspects of the observation instrument used by EQUIP-Tanzania, which, unlike Stallings, was not designed to capture students' participation. We are also measuring the level of student participation, whether it involves small or large group. Our approach in student participation has a more general meaning of taking part in a classroom activity. We believe that active student participation in class is expected. Participating in a classroom activity helps students

to form ideas and gain understanding, which are useful to their ongoing learning. These are the developed "yes-no" indicators:

a) Students' involvement in using learning materials provided by the teacher
b) Active participation of the students
c) Students copy the material provided into their notebooks

The first indicator attempts to capture students' involvement in using learning materials employed by a teacher; whether it is teacher's request or students' self-initiative. For the second indicator, we capture any form of students' participation that is recognizable by nonteaching background observers. The form of active participation is ranging from a low participation, such as students responding to teacher's questions with a single word, to a high level participation, such as students being involved in a student-teacher discussion about the lesson. We also capture whether students self-initiatively or willing respond to a teacher's request of working on a problem in front of the class, and whether students are asking questions in reference to the lesson.

In Stallings, copying materials into notebooks is part of a teacher's passive instruction and is used frequently in teaching activities (Bruns and Luque, 2014). We reframed it to address such case, where students are copying teacher's writing even when the teacher does not instruct them to. Typically, copying has become habitual for students. Students also consider copying when they have to write in their book the materials read out loud by the teacher.

Our three indicators become more and more meaningful in depicting trends in teaching when we combine the analysis with what teacher does within the specified time of observation. For example, when a teacher explains new content or asks close-ended questions, and students respond to the questions, that means teacher-student interaction occurs. A higher teacher-student involvement has a positive relationship with students test scores, all else being equal (Ragatz et al., 2015).

2.2.6 Assessment of Learning Process and Outcomes

This section focuses on practices carried out by a teacher to monitor students' understanding. In Danielson (2013), the monitoring of students is part of using assessment in instruction. The teacher can measure students' comprehension through a single method or more. Here, we only capture teachers' assessment on student learning of today's lesson through the checking of class assignments and giving questions. The indicators are as follows:

a) Teacher receives and checks students' assignments.
b) Teacher checks students' understanding regarding the assignment.
c) Teacher asks closed-ended questions to students.
d) Teacher asks open-ended questions to students.

Checking student assignments in class can be considered a low-level type of assessment as it only gives students information about what is true or false. The indicator may also incorporate criteria such as attendance or participation that are not direct measures of learning. Students cannot automatically reflect their understanding of the class assignment result (Baas et al., 2014). After marking students' assignments, the common practice is that the teacher continues without knowing whether students understand the concept.

Students need more support from their teacher; this requires a teacher to give a quality feedback. Once the students receive feedback, they know what they should prepare or develop in the next learning (Baas et al., 2014). Our approach in identifying this practice is by observing whether a teacher checks students' comprehension through giving corrective comments or provide further

explanation about the content of the class assignments to students, individually or collectively. If a teacher does not, at the very least, check students' comprehension, students spend more time in working on assignment and less in learning.

In TIMSS Video Study (Ragatz et al., 2015), questioning is an attempt made by a teacher to check students' understanding. The study found that most of the questions given in mathematics class are close-ended that require a single word response. Teachers rarely ask open-ended questions. We tried to incorporate indicators on whether a teacher gives open-ended questions. The pilot result shows that it was difficult for observers with nonteaching background to acknowledge whether the teacher is asking an open-ended question. Furthermore, it will require follow-up questions using a teacher questionnaire to check whether a teacher specifically formulated questions to seek evidence of student understanding.

2.2.7 Control of the Classroom Environment by a Teacher

Managing classroom environment is part of teaching practices that aims to mitigate the loss of instructional time. In both the PKG and the instrument to evaluate PPL, teacher classroom management is regarded as one of the indicators in the aspect of learning strategy. Other indicators in that aspect are teachers who are fostering positive behavior and teaching within the allocated time. We derived indicators relevant to classroom management into several indicators in the form of "yes-no" response. The indicators are as follows:

a) Teacher circulates the class to monitor students.
b) Teacher gives praise or positive responses to students for their answers.
c) Teacher uses cheers, handclaps, or songs to control the class.
d) Teacher conveys the expected behavior of students during class.
e) Teacher overcomes students' misbehavior in the class.
f) Students' misbehavior (filled when the teacher corrects student behavior)

On the "teacher circulating the classroom to monitor students" indicator, we do not identify whether a teacher communicates with them. At the very least, when a teacher circulates the class, they could keep the classroom environment in control. It will be difficult to differentiate whether a teacher gives feedback or unrelated comment to teaching when they are circulating the class because, on account of our observation design, the observers are holding the video camera while doing the live observation.

The "teacher gives praise or positive responses to students for their answers" indicator aims to capture practices that create or maintain classroom climate. Teacher's positive responses and/or praises to student's answer is a way to build a good rapport with the student; it can create an interactive discourse that helps extend student understanding (Siraj and Taggart, 2014).

The last four of our indicators focus on practices in managing student behavior. Typically, a teacher uses handclaps, songs, and cheers to gain students' attention in class. The "teacher conveys the expected behavior of students during class" indicator does not necessarily have to be followed by the teacher correcting students' behavior. Teacher articulates behavior as what students are allowed and not allowed to do to keep students engage with the lesson (Danielson, 2013).

When a teacher corrects student's misbehavior in the class, observers also take notes of the observable behavior that the teacher considers as disruptive. To reduce observers' subjectivity in capturing students' misbehavior, we only take note if the teacher is correcting it. There is a possibility that the observers would think that the class is in disorder, but the teacher does not

respond. On the other hand, it could also be that the teacher is unaware or does not consider the disorder to be disruptive. The more time a teacher spends to manage student behavior would suggest that they are unlikely able to control the class environment.

2.2.8 Concluding a Lesson—Closing Activity

PKG emphasizes an end of lesson on two indicators. First, a teacher makes a recap or reflection of today's lesson by involving students. Second, a teacher gives homework or directs students to the next stage of learning. Our instrument incorporates the first indicator as we believe that ending a lesson by presenting key takeaways is useful to student learning. A summary of today's lesson may make students aware of what learning they have gained, enable them to internalize it, and make them think of why it matters (Ganske, 2017). We maintain the coding of activities on a rating scale and provide explanation and possible examples for each rating. We also add other indicators using a "yes-no" response to see whether a teacher assesses students' understanding.

The indicators in the concluding-a-lesson stage are as follows:

a) Teacher checks/assesses students' knowledge gained from today's learning.
b) Teacher presents a summary of today's lesson by involving students.

III. DEVELOPMENT OF THE INSTRUMENT

In designing our classroom observation instrument, we adopted the models of other observation instruments, i.e., the national-level PKG and the instrument to evaluate PPL and Stallings' observation instrument.

The stages of developing the instrument were as follows:

a) A series of internal discussions and consultations with an expert on teaching who was involved in using an observation instrument to evaluate teaching practices.
b) First pilot, which aimed at improving the instrument. The pilot was conducted to compile possible examples of teaching and to check whether the observers have different interpretations of the indicators during live observation.
c) A two-hour internal try out of Computer-Assisted Personal Interviewing (CAPI) technique to resolve coding bugs using several videos from the first pilot.
d) Second pilot, which aimed at evaluating the structure of the instrument that had been modified into eight aspects of teaching practices. We wanted to see whether the instrument is easy to follow during live observation. The second pilot was also meant to check the agreement among observers in one team in relation to the indicators.

We conducted the observation live and with video recording, so that the video could be assessed using other observation instruments if needed. The observers collected the data using the CAPI technique.

Both pilots took place in Kota Bogor and Kabupaten Bogor. The two regions were selected for the pilot because their locations are adjacent to Jakarta. The selection of schools is based on a school quality index that we have previously constructed. The index reflects the quality of schools measured by some indicators, and it was used to select schools which represent different categories: (i) high quality, (ii) mid quality, and (iii) low quality.[4]

Before piloting the observation, we did a prepilot; another team had previously visited the schools to record schools' timetable or teachers' schedules. This allows the observers to identify which teacher gave mathematics or Indonesian lesson on the days of the visit. In a one-day school visit, a team of two observers were responsible for recording two teachers.

3.1 First Pilot

The first pilot was carried out in 23–26 July 2018, which was the second week of effective teaching in the 2018 school calendar. Nevertheless, some of the teachers had not prepared their subject teaching schedule because a week before they attended a teacher training that prevented them from preparing the lesson. Another reason was the meeting between teachers and the principal had yet to be arranged, so the schedule of the subject of teaching or school timetable had not been developed. This created difficulties for the prepilot team to arrange school visits and which classes were likely to be observed. Through the first pilot, we concluded that school visit in the first two

[4]Both pilots were also used to test CERMAT. The number of schools, as well as their quality, in the first and second pilot, was selected by the team who developed CERMAT. See Rarasati et al. (2019) for a more detailed explanation of the school selection index.

weeks of school's schedule is less visible because most of them are not yet prepared with their lesson schedule.

The school sample size for observation in this pilot study was five primary schools and three junior secondary schools. Two teams of observers were able to videotape ten primary school teachers and five junior secondary school teachers during four days of the pilot.

In the first pilot, the constructed observation instrument was tested for completeness and its usability. The piloting aimed to check whether each question in the instrument may be interpreted differently by the observers. To facilitate subsequent in-depth analysis of our observations, we added a note sheet to our observation instrument. Observers made some notes on the indicators in the instrument which needed to be clarified, modified, or removed; and identified activities in class which had not been captured by the instrument. The observers' agreement on the changes was reviewed once the pilot finished. Whenever necessary, the descriptions of the indicators were modified, for example, to better distinguish the categories of rating or the "yes-no" indicators.

3.1.1 Observation Method

There are six distinct methods of observation in the first pilot:

a) Classroom activities were recorded using tablets placed on two sides of the class. One tablet was set in front of the class to document students' activities; another tablet was set at the back of the class to record teaching practices, including teacher's interaction with students.

b) The tablet positioned at the back of the class was not set to follow the teacher's movement around the class. We intended to shoot in different angles to see the differences of what can be captured. Therefore, it was up to the observers where to place the tablet; in the middle of the class or at one corner of the class.

c) The observers were researchers of the Indonesia CRT involved in developing the instruments. There were two teams of observers, and each consisted of two researchers or observers. They were sitting at the back in each corner of the class while conducting the observation. Each observer is equipped with one tablet to fill the instrument and a sheet of paper to note activities in the classroom.

d) The core lesson stage was observed in separate 10-minute intervals. In every 10-minute observation, each indicator was marked once in a given checklist box, no matter how many times the practices occurred.

e) Twelve indicators in the core lesson stage were not arranged into certain aspects of teaching practices, as were those in both the PKG and the instrument to evaluate PPL.

f) In the first pilot, our class observation was considered to be unannounced. Although we informed the school of the days of the visit, we did not announce which teacher would be selected for the observation.

Figure 2. The camera was set up in front of the class to record students' activities

Figure 3. The camera was set up at the center back of the classroom to record the teacher's activities

3.1.2 Findings

Some notes were made during the pilot study, particularly in relation to the main teaching activities. These notes include the following issues.

a) Technical observation or documentation issues:

   (1) The instrument was set to observe either Indonesian or mathematics sessions. Nevertheless, Indonesia's latest curriculum for primary school adopts a thematic approach which combines all subjects in two to three sessions.

   (2) The duration of observation varied across schools. In one school, one learning session lasted for two sessions. Meanwhile, in another school, it took three sessions. Consequently, the number of observation periods in the main activity section also varied.

(3) Observers felt that the 10-minute interval in observing the main teaching activity was too long and could be shortened.

(4) At a low-quality primary school, setting a tablet in front of the class to record class activities greatly distracted the students because it made them pay more attention to the tablet and could not focus on the learning activities.

b) Issues relating to the indicators:

(1) In lower grade classes of primary school, teachers used cheers, handclaps, and songs very often to control the class.

(2) When teachers used projector or wrote on the whiteboard, there was no indicator in the instrument that documented the clarity of the display.

(3) The instrument did not capture students' participation in class.

(4) There was no indicator that captured a situation where students copy questions or lesson materials from the whiteboard or textbook. This type of activity is important to document since it commonly occurs, and it may take a significant amount of time, particularly in low-quality schools.

(5) Teachers were likely to carry out assessment on the assignment or practical work given in the class. However, the assessment was merely to check whether students participated in doing the work in class or whether their answer was correct. It was rare for the teacher to ask further questions on how students were able to come up with the answer. Teachers did not check the extent to which students understand the topic.

(6) Nonteaching work (classroom management) occurred frequently, but the instrument did not sufficiently capture that type of activity.

(7) When teachers gave students questions, observers faced difficulties in classifying the questions as an open- or close-ended question.

c) Issues relating to overall observation, filled in once the lesson ended:

(1) External disruption, such as visits from other teachers or students' parents, was not appropriately captured using the instrument.

(2) The condition of a classroom which can affect learning process, such as lighting and air circulation, was sufficiently captured by the existing instrument. However, the team considered that the explanation of the indicator need to be improved.

Based on the descriptive analysis of the observation results and evaluation among observers, the observation instrument's indicators had not captured the variety of activities that occurred in the classroom. Observers found types of teaching practices or activities repeatedly occurring in the classroom, but were not included in the observation list. The description for each indicator was not yet sufficient. Consequently, observers were likely to choose different indicators in the instrument to document similar activities based on their interpretation. For example, indicators on the type of question asked by a teacher to the students.

Some revisions made in the instrument are as follows.

a) Technical observation or documentation issues addressed:

(1) The template of SurveyCTO used to record the observation was modified to allow the thematic approach session in addition to Indonesian and mathematics sessions. The modification also allowed for more observation periods in the main activity section.

(2) The main teaching activity was each observed in a 5-minute period.

(3) Setting a tablet in front of the class was no longer used.

b) Revision on issues relating to the indicators:

(1) Indicators in the core lesson stage were rearranged and categorized into six aspects of teaching (see Table 2).

(2) Two indicators in relation to teacher probes or comments on student's answers (adapted from the EQUIP indicators) were dropped. Observers found difficulties in identifying practices when the teacher was in front of the class, let alone if the teacher were circulating the class and communicating with individual students distant from the observer's position.

(3) Students' copying activities, students' active participation, teacher assessment on class assignment, teacher incorporates daily life experiences or common knowledge in teaching, and additional classroom management (teacher uses of handclaps, songs, and cheers) were added.

(4) Additional explanation on the indicators and some minor changes in their wording.

c) Revision on issues relating to overall observation, filled in once the lesson ended:

Additional indicators, such as visits from other teachers, lighting condition and air circulation in the class, and type of sitting arrangement of students were included in the section of overall observation of class. The overall evaluation is carried out once the lesson ended.

3.2 Second pilot

The second pilot of classroom observation was conducted from 30 September to 4 October 2018 in Kota Bogor. Since we also conducted the first pilot there, permit from the local government was easier to process. The selection of schools for the second pilot was similar to the selection in the first pilot. We selected three primary schools and three junior secondary schools based on a school quality index that we had constructed before. In the second pilot, we observed nine teachers in six schools.

The purpose of the second pilot was to test whether indicators in the revised instrument can capture the classroom activities better. Observers were asked to test whether the instrument's indicators can be understood easily and the instrument's structure is easy to follow in conducting the classroom observation. The second pilot also aimed to review the agreement of observers in one team related to the indicators.

3.2.1 Observation Method

The observation method of the second pilot was arranged as follows.

a) As with the first pilot, each classroom observation in the second pilot was also carried out by a team of two observers. However, the composition of the team was different from the first pilot. In the second pilot, the team consisted of one SMERU researcher and one hired enumerator. We used this composition to test whether the instrument could be easy to understand by people who were not involved in the development of the instrument.

b) Classroom activities were recorded using one tablet camera positioned at the back corner of the class (left or right) to record teaching practices. Based on the first pilot, the tablet

positioned at the middle back of the class was relatively limited in capturing the teacher's movement.

c) One of the observers was responsible for holding the tablet while doing the live observation. The tablet should follow the teacher's movement around the class.

d) The core lesson stage was observed in separate 5-minute intervals. Similarly, each indicator was marked only once in a given checklist box, regardless of how many times the practices occurred.

e) There were 22 indicators in the core lesson stage and they were arranged into certain aspects of teaching practices following the national teacher formative evaluation.

f) The previous team who arranged the schedule of school visit had listed the name of teachers or classes that would be observed. We did acknowledge the consequence; some teachers might have made some extra preparation and might have acted unnaturally in the class.

Before starting the second pilot, the enumerators who would conduct the classroom observation had gone through training. We held a one-day training session to train the enumerators to use the classroom observation instrument in a tablet and set the tablet camera. Classroom activities recorded in the first pilot were used as examples in the training. Overall, the training had enabled enumerators to understand the instrument and to operate the application on the tablet.

3.2.2 Findings

The observation team, which consisted of RISE researchers and hired enumerators, had a follow-up meeting at the end of each observation day to communicate findings useful for improving the instrument. The notes of the meetings are as follows.

a) Technical observation or documentation issues:

The hired enumerators still found difficulties moving the tablet to follow the teacher's movement in the class and doing the live observation. They had to thoroughly memorize the instrument to handle two tasks. They felt that the next training of observers should allow them to try out the observation method in a real classroom, and not only in using the camera.

b) Issues relating to the indicators:

(1) The observers could not be sure whether or not teachers connected today's learning material with (the same learning material) previously taught.

(2) Students' participation in learning can be observed clearly. The variety could be in the form of answering the teacher's close-ended questions to individual initiatives to solve problems on the whiteboard.

(3) Observers found difficulties identifying whether the teacher's questions to students should be classified as either open- or close-ended questions. Close-ended questions were easier to identify compared to open-ended questions. The observers tended to classify mathematics questions as close-ended questions. Furthermore, teachers rarely explored students' answers, particularly in primary schools.

(4) Regarding the aspect of control of the classroom environment by a teacher, sometimes it was difficult to see whether or not teachers carried out activities which were not related to teaching, especially when they were sitting behind their desk. For example, when a teacher was working on her or his computer, the observers could not be sure whether or not that activity was related to teaching.

(5) The hired enumerators requested for more description on the indicator of teacher is giving assignment. When a teacher gave instructions in doing an assignment, it was unclear whether it was part of the "teacher is giving assignment" indicator.

(6) Observation of the concluding-a-lesson stage found that teachers rarely do it, both in primary and junior secondary schools. One of the reasons is that teachers were still busy with the main activities when the lesson session ended, while the class should had already continued with the next session. Another reason is that because teachers preferred to use the remaining time of his or her lesson for other activities; for example, to distribute the results of the previous daily examination to students. However, observers were still able to identify teaching practices requested in the indicators of concluding-a-lesson stage.

c) Issues relating to overall observation, filled in once the lesson ended:

The overall observation of the class section, filled by observers once the lesson ended, resulted in a relatively matching agreement on the indicators. Nevertheless, some indicators might be answered differently among observers since the indicators are quite subjective. Those indicators are the questions on the lighting and air circulation of the class. Furthermore, the observers could not be sure whether all students were able to see the writing on the board clearly from their position.

We have revised the instrument based on the above analysis and the team's evaluation on the pilot. Some of the revisions which were tailored to address issues in the second pilot are discussed below:

a) Technical observation or documentation issues addressed:

The team provided one to two days of trying out the observation method in a real classroom in addition to the training of observers in the classroom using video.

b) Revision on issues relating to indicators:

(1) An indicator asking about nonteaching activity carried out by the teacher in the classroom (included in the aspect of control of the classroom environment by a teacher) was dropped since the observers found difficulties in capturing what teachers do when they are at their desk.

(2) Although the description of the indicator had been improved after the first pilot, the indicator on the open-ended questions was still difficult to identify, particularly when observing the Indonesian language lesson. Similarly, the observers also found that teachers rarely gave open-ended questions to their students in mathematics lessons. Consequently, observers tended to define all questions asked in the classroom as close-ended questions. In addition to improving the indicator's explanation and more examples on the indicator, the team enhanced the observers' training by asking them to create close-ended questions and then change it to open-ended questions. By doing so, the observers have a better understanding of identifying the two types of questions.

(3) Additional explanation on indicators and more examples of teaching practices relating to indicators showed that issues were resolved.

c) Revision on issues relating to overall observation, filled in once the lesson ended:

Overall, the interpretation of the instrument's indicators among observers was relatively the same; there was less contradiction or difference in filling the instrument between observers in one team. In Table 3 below, we see that observers in each team produced similar observation results. Furthermore, the consistency of matching agreement between two observers in one team is relatively high. A matching agreement is the frequency of two observers in identifying

the same teaching practices in the same interval of observation. It means that in each 5-minute period, if both observers in one classroom repeatedly marked the same indicator, then it leads to a 100% agreement or matching rate. Based on the second pilot result, the observers matching rate of each classroom observation is between 80% and 95%.

Table 3. Matching Agreement of Observers

School level	Teacher	Average Matching Agreement Between Two Observers (%)
Primary	Teacher 1	88.07
	Teacher 2	92.45
	Teacher 3	94.48
	Teacher 4	86.89
Junior secondary	Teacher 1	91.36
	Teacher 2	81.40
	Teacher 3	90.93
	Teacher 4	91.55
	Teacher 5	94.15

We found that differences may occur across school qualities in the activities prior to preliminary activity.[5] In high-quality schools, teachers tended to act naturally. They started their teaching by explaining the topic to be discussed during the lesson session. Meanwhile, teachers in low-quality schools seemed to act unnaturally, realizing that they were being observed. For example, in one low-quality primary school, a teacher that we observed started her class by asking the children to sing the Indonesian national anthem and the Indonesian Red Cross song.

[5]The school index is generated using Principal Component Analysis (PCA) based on the school's characteristics that potentially have linkage to the learning outcomes. The high-quality schools are those with the 20% highest index. See Rarasati et al. (2019) for a more detailed explanation of the school selection index.

IV. OBSERVATION MANUAL

Our observation instrument provides a framework for observing and recording teacher practice in the classroom. The instrument contains 26 indicators or the observable teaching practices; some are positive and some are negatives. The instrument are segmented into three stages of lesson: introducing a lesson, core lesson, and concluding a lesson. Each of both introducing-a-lesson and concluding-a-lesson stages has two indicators that use simple ratings, in which definitions and possible examples are provided. The core lesson includes 22 indicators. For each indicator, check "Yes" or "No" to record the presence or absence of that indicator based on the observation.

Observations should be conducted for at least two sessions. In primary schools, the first session equals to 35 minutes of teaching, while in secondary schools it is 45 minutes. The observation is carried out by one observer in one classroom. The role of the observer is to watch and listen for signs of observable teacher practices listed and to record whether or not they were observed, without making judgments as to relative importance or relevance of those practices. The observer, who carries out a live observation, may include people with no teaching background trained specifically for the project. The observer will collect the observation data using SurveyCTO software in an Android tablet. We suggest the live observation is also videotaped for a viewing session held separately, either for another checking or for implementing other observation instruments when needed. The location of the observer and the position of the tablet should be at the back corner (left or right) of the class. Placing the tablet at the back corner allows for a wider range of the camera to follow teacher activities in the classroom.

To know which teachers teach mathematics and which teach Indonesian, it is necessary to carry out an advance visit one day before the observation session or prior to selecting the observed class. That would allow the observer to take notes on the school's timetable. Lesson observations are not randomly sampled but are instead chosen by convenience based on the school's timetable in effect on the day of the survey. This is to achieve variety in teaching style, grade level, and subject taught.

4.1 Observation Procedure

4.1.1 Planning the Observation

Before the observation takes place, the day before or right before it starts, it is necessary to have a preobservation meeting between the observer and the teacher to set the scene for the observation, clarify issues, and fill the consent form. Below is a list of areas to be covered at the meeting.

The observer and the teacher should agree or confirm on the following:

a) Protocols regarding confidentiality as explained in the consent form. The teacher needs to be assured that the observations are being conducted solely for the purpose of research, that the observers are not there to evaluate them, and that the observations would not be shared with the school or district personnel.
b) Where the observer is to sit
c) Whether or not the observer's presence is to be explained to students
d) How the observation is to be recorded during the session
e) The completion of non-observable items in the teaching observation instrument (relating to the teacher's information)

Before starting the observation, complete the cover sheets of the observation instrument indicating the following:

a) School location
b) Date of the observation
c) Observer's identification name
d) Teacher's identification name
e) Teacher's gender
f) Information about subject of the lesson and number of hours of the lesson

4.1.2 During the Lesson

The observers should enter the class before the teacher begins the lesson to

a) explain the purpose of the observers' presence to students (as agreed by teacher);
b) locate themselves in the agreed position;
c) be as discrete and unobtrusive as possible (minimize contact/communication with students);
d) record the time they enter the class;
e) collect information on teacher's activity before the teacher starts the lesson; and
f) record the time the observation starts.

The observation stages in our instrument are in line with the teaching stages in the two formative evaluations, PKG and the instrument to evaluate PPL. Typically, the transition from introducing a lesson to the second stage, core lesson, is clear as the teacher begins delivering the lesson's content. The core lesson contains several separate 5-minute observations, and marking period is indicated by a question of whether the lesson continues or ends. The third stage, concluding a lesson, commonly occurs 5–10 minutes before the lesson hour ends. The observer can mark the third stage when the teacher states that the lesson has ended, then gives a conclusion or an assignment. When the teacher assigns seatwork until the hour lesson ends and the teacher does not deliver a closing remark, it is marked that the concluding-a-lesson stage does not occur.

The observer must record each of the 5-minute observations. A timetable of the 5-minute observation period will appear at the bottom of the SurveyCTO application. The observer spends each 5-minute period to observe the teacher's behavior and mark the observable practices that occurred. The observer should leave blank the space beside indicators/practices which did not occur. The observer should take into consideration the 22 indicators/practices listed in the core lesson, mark them or leave them blank. A particular item is marked only once in a given checklist box, regardless how many times the practice occurs within the 5-minute observation period. A practice that occurs a dozen times gets one check mark, the same as an item that occurs only once. This process is repeated for the next 5-minute period. The transition from core lesson to concluding a lesson is indicated by the teacher conveying the summary of the lesson or giving homework to students or mentioning the content of the next lesson.

The observer should take notes on teaching situations related to how long the teacher leaves the class and the number of times guests visit or talk to the teacher inside or in front of the classroom.

4.1.3 After the Lesson

Once the teacher finishes her or his teaching within the observed hours of the lesson, the observer notes the time the observation ended, completes the classroom details, and responds to questions about the overall teaching situation.

4.1.4 Additional Notes

The observer should make additional notes when an unexpected situation occurs. For instance, the observed teacher suddenly must leave the classroom for an urgent reason and then is replaced by another teacher. Furthermore, the observer can also make a note when there is a condition which can affect normal classroom activities, such as the school principal observing the classroom.

LIST OF REFERENCES

Baas, Diana, Jos Castelijns, Marjan Vermeulen, Rob Martens, and Mien Segers (2014) 'The Relation between Assessment for Learning and Elementary Students' Cognitive and Metacognitive Strategy Use.' *British Journal of Educational Psychology*, 85 (1): 33–46.

Bill and Melinda Gates Foundation (2012) *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains*. Research Paper MET Project. Bill and Melinda Gates Foundation [online] <https://eric.ed.gov/? id=ED540960> [3 September 2018].

Bruns, Barbara and Javier Luque (2014) *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean (Advance Edition)* [ebook] Washington, DC: The World Bank [online] <https://www.worldbank.org/content/dam/Worldbank/document/LAC/Great_Teachers-How_to_Raise_Student_Learning-Barbara-Bruns-Advance%20Edition.pdf> [8 April 2018].

Bruns, Barbara, Soledad De Gregorio, and Sandy Taut (2016) *Measures of Effective Teaching in Developing Countries*. RISE-WP-16/009. RISE Programme [online] <https://www.riseprogramme.org/sites/www.riseprogramme.org/files/publications/RISE_WP-009_Bruns.pdf> [30 September 2018].

Danielson, Charlotte (2013) *The Framework for Teaching Evaluation Instrument, 2013 Edition*. [ebook] The Danielson Group [online] <http://www.loccsd.ca/~div15/wp-content/uploads/2015/09/2013-framework-for-teaching-evaluation-instrument.pdf> [23 August 2018].

Ganske, Kathy (2017) 'Lesson Closure: An Important Piece of the Student Learning Puzzle.' *The Reading Teacher* [online] 71 (1): 95–100 [online] <https://eric.ed.gov/?id=EJ1146102> [14 May 2018].

Hanushek, Eric A. and Steven G. Rivkin (2010) 'Generalizations about Using Value-Added Measures of Teacher Quality.' *American Economic Review*, 100 (2): 267–271.

IDinsight (2015) *Quantitative Assessment of Teacher Motivation, Classroom Practices, and Student Learning*. IDinsight.

Klauer, K. J. (1988) 'Teaching for Learning-To-Learn: A Critical Appraisal with Some Proposals.' *Instructional Science*, 17 (4): 351–367.

MacKinnon, Gregory, Lourens Schep, Lisa Lunney Borden, Anne Murray-Orr, Jeff Orr, and Paula MacKinnon (2016) 'Evaluating classroom interaction with the iPad®: An Updated Stalling's tool.' *International Journal of Education and Development using ICT*, 12 (3): 93–107.

Martinez, Felipe, Sandy Taut, and Kevin Schaaf (2016) 'Classroom Observation for Evaluating and Improving Teaching: An International Perspective.' *Studies in Educational Evaluation*, 49: 15–29.

Maulana, Ridwan, Marie-Christine Opdenakker, Kim Stroet, and Roel Bosker (2012) 'Observed Lesson Structure during the First Year of Secondary Education: Exploration of Change and Link with Academic Engagement.' *Teaching and Teacher Education*, 28 (6): 835–850.

Park, Jisun, Jinwoong Song, and Ian Abrahams (2015) 'Unintended Learning in Primary School Practical Science Lessons from Polanyi's Perspective of Intellectual Passion.' *Science & Education*, 25 (1–2): 3–20.

Pettersson, Gunilla, Georgina Rawle, Rachel Outhred, Stephanie Brockerhoff, Gabrielle. Wills, Dita Nugroho, Paul Jasper, Andrej Kveder, and Adrian Beavis (2016) *EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report, Volume I: Results and Discussion*. Oxford Policy Management [online] <https://www.researchgate.net/publication/289504936_EQUIP-Tanzania_Impact_Evaluation_Final_Baseline_Technical_Report_Volume_I_Results_and_Discussion> [7 April 2018].

Pianta, Robert C. and Bridget K. Hamre (2009) 'Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity.' *Educational Researcher*, 38 (2): 109–119.

Ragatz, Andy, Susiana Iskandar, Ratna Kesuma, Susie Sugiarti, Andrew B. Ragatz (2015). *Indonesia - A Video Study of Teaching Practices in TIMSS Eighth Grade Mathematics Classrooms: Understanding What Teaching Practices are Used, Why They are Used and How They Relate to Student Learning*. Main Report. Washington, D.C.: The World Bank [online] <http://documents.worldbank.org/curated/en/886911472471847117/Indonesia-A-video-study-of-teaching-practices-in-TIMSS-eighth-grade-mathematics-classrooms-understanding-what-teaching-practices-are-used-why-they-are-used-and-how-they-relate-to-student-learning> [4 April 2018].

Richards, Jack C. and Charles Lockhart (1994) *Reflective Teaching in Second Language Classrooms*. Mku.edu.tr. [online] <http://www.mku.edu.tr/files/132-93d901d4-2c17-49d2-83a9-45796c691e58.pdf> [7 April 2020].

Siraj, Iram and Brenda Taggart (2014) *Exploring Effective Pedagogy in Primary Schools: Evidence from Research*. Pearson [online] <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/ExploringEffectivePedagogy.pdf> [16 October 2018].

van de Grift, Wim. (2007) 'Quality of Teaching in Four European Countries: A Review of the Literature and Application of An Assessment Instrument.' *Educational Research*, 49 (2): 127–152.

Waxman, Hersholt C. (1991) 'Investigating Classroom and School Learning Environments: A Review of Recent Research and Developments in the Field.' *The Journal of Classroom Interaction*, 26 (2): 1–4.

World Bank (2017) *The Stallings Classroom Observation System* [online] <https://www.worldbank.org/en/programs/sief-trust-fund/brief/the-stallings-classroom-snapshot> [8 November 2018].

APPENDICES

APPENDIX 1

Description of the Observation Sheet

The observation sheet comprises three main sections (i.e., preliminary, main, and closing activity) and one form of overall class and teaching practices observation. Each item in every section has explanation and examples to provide a clear description so that the observers would have similar understanding and interpretation.

Table A1. Preliminary Activity

Beginning of Lesson – Preliminary Activity		There are two learning practices observed using 0, 1, and 2 rating scale		
Time of observation starts (e.g., 07:35 or 13:10)		___ : ___		
1.	Teacher clearly outlines learning objectives of the new content to students.	(0) Not delivered	(1) More or less delivered	(2) Clearly delivered
2.	Teacher checks students' knowledge (which can be obtained from past learning or related to everyday life) and connects them to today's learning material.	(0) Not proven	(1) More or less proven	(2) Clearly proven

1	**Scale**	Not delivered	More or less delivered	Clearly delivered
	Explanation	Teacher only introduces the topic of today's lesson without mentioning the purpose of the lesson.	Teacher introduces today's learning materials and conveys the learning objectives without connecting the materials to how or when students can apply them.	Teacher introduces today's learning materials and conveys the learning objectives, and connects them to the benefits/knowledge that students will gain.
	Example	Today, we will learn about simple narrative story texts. Open book … page …	Today, we will learn about simple narrative story texts. The purposes of learning about story texts are to understand the shape/structure of a text and practice your writing skills.	Today, we will learn about simple narrative story texts. The purposes of this learning are to understand the shape/structure of a text and practice your writing skills. Later, you can share your ideas/experiences through writing.
2	**Scale**	Not proven	More or less proven	Clearly proven
	Explanation	Teacher does not connect the topic of today's learning with past learning or daily lives.	Teacher connects current learning materials with past learning materials and/or daily lives, but does not ask students about their past learning knowledge.	Teacher mentions past learning materials (or learning materials from some time ago) along with today's learning. And then, the teacher asks about the students' understanding about the past learning materials or general knowledge.

Table A2. Main Activity

Main activity	There are several teaching practices observed
Mark [X] each time you find a teaching practice happens during a five-minute observation.	
Time of observation starts	

	Observed aspects	Explanation
	A. Main activity in teaching	
A1.	Teacher explains learning materials.	Teacher explains today's learning material or what the students will learn.
A2.	Teacher provides example problems or problems relating to the learning materials, gives students examples on how to solve the problems.	Observer does not assess the quality or whether the diagram or the writing is in accordance with the given materials.
A3.	Teacher gives assignments to students in the classroom relating to today's learning materials.	DOES NOT include copying the teacher's writing relating to the explanation of the learning materials on the board. • Teacher gives assignments that students have to do in the class (e.g., do the exercises written on the board or from the textbook and student worksheet or prepared by the teacher).
	a. In pairs or in groups	Teacher asks students to do their assignments in pairs or in groups.
	b. Individual	Teacher asks students to do their assignments given in the class individually based on the rules set by the teacher.
	B. Making connections in teaching	
B1.	Teacher connects today's content with other content that has been taught before.	Teacher mentions that the material taught today is still related to a different material that has been taught in the previous schedule.
B2.	Teacher connects today's content with situations or daily activities that are commonly experienced or done by students.	Teacher connects learning material into real life contexts experienced by students. These contexts include the use of objects, an event, a place/location or people who are commonly/familiarly known or listened to by students.
B3.	Teacher reiterates learning objectives relating to today's material.	During the observation period, teacher reminds students (verbally and by pointing to a list of goals written on the board/in the book) the learning objectives to be achieved today.
	C. The use of learning materials by a teacher	
C1.	Teacher uses electronic equipment.	Projector and laptop
C2.	Teacher explains the material to students through clear writing and/or diagrams.	• Seen from the teacher's writing on the board • Seen from the clarity of the projector's display *Observer does not assess the quality or whether the diagram or writing is in coherent with the material provided.*
C3.	Teacher uses instructional media that includes, but not limited to, written or printed learning media (excluding the compulsory textbook for	For example: • Learning media

	Observed aspects	Explanation
	the lesson) prepared by teacher; map, chart, and card.	• Learning media complementary to text books, such as novel and poetry. • Learning media created/printed by teacher, which is used to help explain the material or assignment. *Student worksheet/books (LKS) are not included as learning media.*
	D. Student participation in learning process	
D1.	Students' involvement in using the learning materials provided by the teacher.	Students are involved when • teacher demonstrates the use of learning media or props, which could be by engaging one or several students; and • teacher provides learning media or props to students for group or individual use.
D2.	Students' active participation	• Student participation occurs because of teacher appointment or student's initiative. • Student is willing to come to the front of the class or speak from his/her seat to work on the problem/question (relating to the material) presented by the teacher. • Student asks the teacher about the material/assignment; both when the teacher is teaching and when the teacher is checking students' assignment in the class.
D3.	Students copy the material provided into their notebooks.	• This happens when the teacher writes the material on the board and then students copy it into their notebooks. • Students copy the material in the textbook into their notebooks. DOES NOT include working on questions in the textbook because this would be part of an assignment.
	E. Assessment of learning process and outcomes	
E1.	Teacher receives and checks students' assignments. NOT to measure the extent to which students understand the material provided.	Directed to the assessment of student participation, such as • whether or not students work on the questions; and • whether students give the correct answer (without knowing/checking further if they did it on their own or if they looked at their friends' answer). Other examples: • Teacher checks the number of wrong and correct questions, but does not explain further to the students. • Teacher asks students to assess their seatmates' work, then the teacher checks the students' answers. Teacher simply wants to ensure that the students gave the correct answer.
E2.	Teacher checks students' understanding regarding assignment.	Teacher monitors students while giving questions or further explanation or feedback based on students' comments or questions. Examples:

	Observed aspects	Explanation
		• Students are asked to copy the reading text; whoever has finished takes it to the teacher's desk and then reads it. • Teacher checks the students' work and explains further if something is incorrect or if there are students who did not understand [the assignment/reading]. ➢ Teacher is circulating the classroom and checking how the students are doing with the assignment. ➢ Teacher is calling each student by name to submit their assignments or is asking each of them whether they have completed their assignments.
E3.	Teacher asks closed-ended questions to students.	• Teacher asks questions with only one correct answer. This question is useful for checking students' understanding when the teacher explains or provides repetition. • Teacher asks questions that require short one-word answers, such as "yes", "no", "correct", "understand". • Teacher asks multiple-answer questions relating to the material previously delivered.
	- Questions are intended for the whole class.	Teacher asks the questions in front of the class; students give their answers simultaneously or one by one if they answer voluntarily.
	- Questions are intended for groups.	Teacher asks questions to several groups of students. If one or two representatives of the group answer the question, it is still considered as a question intended for the group.
	- Questions are intended for individuals.	• Teacher asks questions to individual students by mentioning their name or asks the students to raise their hand if they want to answer the question. • Teacher asks questions in front of the class, but the teacher mentions that she/he only wants answers from students who raise their hand.
E4.	Teacher asks open-ended questions to students.	Examples: • Why does the mother in that story (from a textbook/paper) feel sad? Which part of the story shows 'the mother's feeling'? • How many ways (solutions/operations) can you think of that will help you solve the problem of calculating the sales profit in Store X? Which solution/operation is the easiest?
	- Questions are intended for the whole class.	Teacher asks questions in front of the class; the students give their answers simultaneously or one by one if they answer voluntarily.
	- Questions are intended for groups.	Teacher asks questions to several groups of students. If one or two representatives of the group answer the question, it is still considered as a question intended for the group.
	- Questions are intended for individuals.	• Teacher asks questions to individual students by mentioning their name or asks

Observed aspects		Explanation
		the students to raise their hand if they want to answer the question. • Teacher asks questions in front of the class, but the teacher mentions that she/he only wants answers from students who raise their hand.
F. Control of the classroom environment by a teacher		
F1.	Teacher gives praises or positive responses to students for their answers.	Teacher gives praises or positive responses to encourage other students to answer the questions.
F2.	Teacher circulates the classroom to monitor students.	During the observation, teacher circulates the classroom while interacting or without making any interaction with the students.
F3.	Teacher uses cheers, handclaps, or songs to control the class.	
F4.	Teacher conveys the expected behavior of students during class.	Teacher conveys the rules of behaving in the classroom (verbal or written); what students can and cannot do. This is done both when students misbehave in class and when they behave accordingly.
F5a.	Teacher disciplines student's unexpected behavior in class.	Teacher observes students' behavior during class and attempts to respond to students who behave mischievously by reprimanding them or giving punishment. F5 is filled if F4 occurs
F5b.	Students behave the way teacher did not expect in class. (open-ended question—filled if the teacher corrects student behavior)	Students violate or disobey class rules/standard behavior rules. Observer writes student behavior that is not in accordance with the teacher's expectations in the classroom.

Table A3. Closing Activity

End of Lesson—Closing Activity	There are two learning practices observed using 0, 1, and 2 rating scale		
The time observation starts	___ : ___		
1. Teacher checks/assesses student knowledge gained from today's learning material.	(0) Not proven	(1) More or less proven	(2) Clearly proven
2. Teacher presents a summary of today's learning material.	(0) Not delivered	(1) More or less delivered	(2) Clearly delivered
The time observation ends	___ : ___		

1	**Scale**	Not proven	Proven	
	Explanation	Teacher does not assess what students have learned from today's material. Teacher makes the students do the assignment until the end of the lesson.	Teacher checks/assesses the knowledge gained from today's learning material of **some of the students**. Teacher appoints some students to answer or waits for some students to answer voluntarily.	The assessment could be in writing or verbally. • Verbal assessment: teacher asks questions and asks **all students** to answer by raising their hands if they agree. Teacher then asks some students to explain their answers. Teacher gives feedback to the students. • Written assessment: teacher asks questions that **all students** should answer in a piece of paper (with their names) to be given to the teacher once they have finished. With written assessment, the teacher could not provide direct feedback and correct information at the same time. However, the teacher would be able to identify each student's understanding.
	Example		What have we learned today? Does everyone understand today's topic? Who could answer the following questions?	What have we learned today? Can anyone mention what is a set?
2	**Scale**	Not delivered	More or less delivered	Clearly delivered
	Explanation	Teacher closes the class without giving any conclusions. Teacher only mentions the material described today or asks students about what the teacher has taught today. Teacher can continue by giving assignments to students or not.	Teacher presents a summary of today's learning material without indicating that that was a conclusion of the learning material.	Teacher mentions that the conclusion would be presented as the closing of today's learning. Teacher delivers the conclusion of today's learning material by including the students' responses or summaries during the lesson (if any).
	Example		So, today we have learned XXX, which are AAA, BBB, CCC.	To close today's lesson, I will repeat some of the things we can conclude. The first point is ... As XX said, in this material .... The second point is …

Overall observation of classroom and teaching practice				
Fill this form after lessons are finished.				
1.	In general, did the teacher call each student by their name?	(0) No	(1) Yes	
2.	Did the teacher use a local language to improve the students' understanding of the learning materials?	(0) No	(1) Sometimes	(2) Often
3.	Did the teacher leave the classroom during the lesson?	(0) No	(1) Yes	
4.	In total, how many minutes did the teacher leave the classroom during the lesson?	____ minute(s)		
5.	During the lesson, did the teacher receive a visit from and speak to other parties (other teachers/student's parents/principal)?	(0) No	(1) Yes	
6.	How many times did the teacher receive a visitor?	____ time(s)		
7.	Was the lighting in the classroom sufficient? *E.g., sunlight brightened the classroom without missing a spot or the lamp lighted the whole classroom.*	(0) No	(1) Yes	
8.	Was the air circulation sufficient? *E.g., some students fanned themselves, or the observer felt the air was stuffy.*	(0) No	(1) Yes	
9.	How was the classroom seating arrangement?			
	a. Arranged in columns (common classroom seating)	(0) No	(1) Yes	
	b. Arranged in the letter 'U'	(0) No	(1) Yes	
	c. Arranged in groups	(0) No	(1) Yes	
10.	Were **all students** able to see the writing on the board clearly from their sitting position?	(0) No	(1) Yes	
11.	Was the number of tables and chairs sufficient for the number of students? *E.g., if some students sit in a group of three in one seat, then the number of tables and chairs was not sufficient.*	(0) No	(1) Yes	

APPENDIX 2

Scoring Manual

This section assesses the most prevalent teaching practice themes in the classroom observation data: lesson structure, the structure and types of teacher-student interaction, questioning, and supportive classroom environment. These themes serve as observable proxies for learning. The themes are consistent with the expected teacher practices assessed in both PKG and the instrument to assess PPL. We use the themes to analyze trends in several teaching practices mentioned in Indonesia TIMSS Video Study 2007 and 2011.

We use observation sample data of 163 secondary school teachers from 42 schools in Kota Yogyakarta to present the assessment of each theme. The sample teachers consists of 40 mathematics teachers of the 7th grade; 43 mathematics teachers of the 8th grade; 40 Indonesian teachers of the 7th grade; and 40 Indonesian teachers of the 8th grade.

(1) Lesson Structure

Every lesson should have a recognizable structure. A study comparing lesson structures of Indonesian and Dutch students categorizes the distribution of class time into lesson introduction, review, introducing new content, and student work time (Maulana et al., 2012). Our observation instrument does not capture the distribution of time of teacher introducing the lesson or doing the review. The instrument only looks at whether the teacher is introducing the lesson. However, our instrument can be used to analyze the proportion of instructional time spent on the introduction of new content and student work time. For the lesson structure, we use terminologies, teacher's presentation, and teacher's giving practice session, respectively.

a. Introducing and closing a lesson

In our observation instrument, we use a rating scale of "0" to "1" to score indicators of teacher's practices in introducing and closing a lesson. We snapshot these indicators one time at the beginning of lesson, and another time at the end of lesson, rather than at every five-minute observation interval. The description of the indicators and the scale is available in Section 1 of the Appendix (Description of the Observation Sheet).

In general, the lesson plan developed by the teacher should include opening and closing routines. Implementation of both routines indicates the lesson is carried out in accordance with the plan and is in a timely manner. When the teacher introduces the lesson, it helps to focus students' attention (Richards and Lockhart, 1994). When the teacher closes the lesson rather than simply stops, it helps to integrate what the students have learned from the lesson (Richards and Lockhart, 1994). Thus, both routines characterize good teaching practices. In our observation instrument, if the teacher practices both opening and closing routines, then the data is marked as "1".

Opening a lesson. The observation begins with identifying whether introducing the lesson includes practices of

- teacher delivers the learning objective relating to the specified material clearly (coded as AWP1) and
- teacher checks students' knowledge and connects them to today's learning material (coded as AWP2).

We use a binary variable in scoring. A score of "1" means the teacher begins the lesson with performing at least one indicator, while a score of "0" means that the teacher does not carry out any of the indicators when introducing the lesson.

Closing a lesson. The observation ends with marking whether the closing of the lesson shows practices of

- teacher checks/assesses students' knowledge gained from today's learning material (coded as AKP1) and
- teacher presents a summary of today's learning material (coded as AKP2).

The scoring for lesson closing is the same as the opening, where a teacher gets a score of "1" if she or he does either AKP1 or AKP2, and "0" if she or he does not do either.

To understand how the teacher opens or closes the lesson, one can analyze further using the Likert scale from the observation instrument data and/or through the videotapes.

The following graph can be produced once the raw observation data is processed based on the themes of lesson structure in introducing and/or closing the lesson. What we are able to show in the graph below is the proportion of teachers, from the Kota Yogyakarta sample, who indeed practiced the introducing and/or closing the lesson routines.

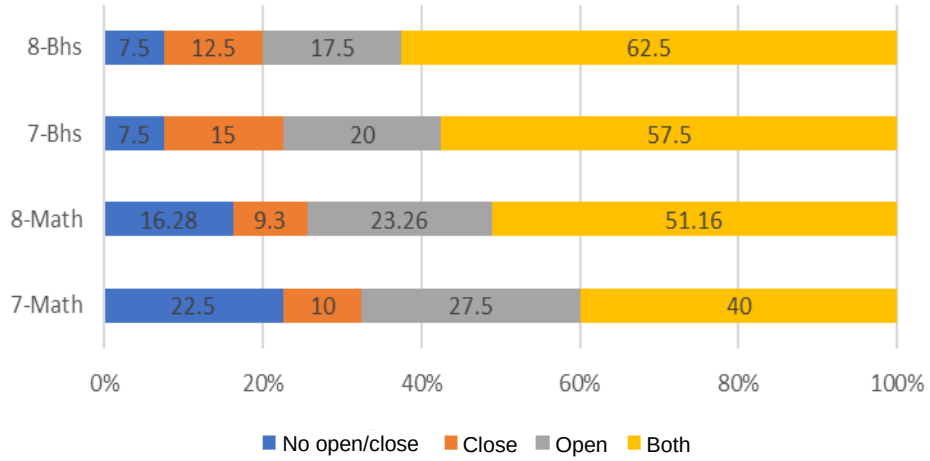

Figure A1. Proportion of teachers introducing and/or closing the lesson

b. Presentation and practice session

Presentation. Teacher presentation includes two activities:

- Expository, which derives from the "teacher explains learning materials" indicator (coded as A1). On any sheet, if the teacher only explains what students should learn without other teaching activities, it is marked as "1".
- Problem-solving, which derives from the "teacher provides example problems and the solution" indicator (coded as A2). On any sheet, if the teacher only provides example and shows the solution without other teaching activities, it is marked as "1". Another category will be included in this activity if A1 and A2 are marked as "1" on any sheet. We include

these activities as problem-solving. Thus, on any sheet, if the teacher gives a problem set, shows the solution, and explains how to solve it, it is marked as "1", which is the activity of problem-solving.

If one or both activities occur and do not coincide with practice session on any sheet, it means the teacher is only doing the presentation, which is marked as "1".

Practice session. The teacher is giving a practice session when activities below are performed:

- Assignment, which derives from the "teacher gives assignments to students in the classroom relating to today's learning materials" indicator (coded as A3a and A3b). If A3a and/or A3b marked as "1" on any sheet, it means the teacher is solely giving assignments during the observation intervals.
- Assessment, which derives from the "teacher checks students' understanding regarding assignment" indicator (coded as E2). If the teacher is doing this activity during each five-minute observation interval, we mark it as "1" for assessment.

If the occurrence of one or both activities do not coincide with the teacher's presentation on any sheet, it means the teacher only carries out practice session, which is marked as "1".

Our observation instrument captures several activities during the five-minute observation interval. Thus, to analyze the lesson structure, we created the following additional categories to teacher's presentation and practice session:

- **Combination**, which is marked as "1" if the teacher carries out both activities, presentation and practice session, on any sheet.
- **Others**, which is marked as "1" if any activity other than presentation and practice session occurs in each five-minute observation interval.

The following graph shows the result of processing the raw observation data based on the lesson structure theme in presenting materials and/or practice session of the lesson. The graph shows that the proportion of instructional time teachers use for practice sessions tends to be larger than presenting the material. The proportion of doing both activities, the combination strategy, tend to be larger in mathematics class. This is common because teachers use sets of problems, explain them, and at the same time ask students to solve them when teaching mathematical concepts.

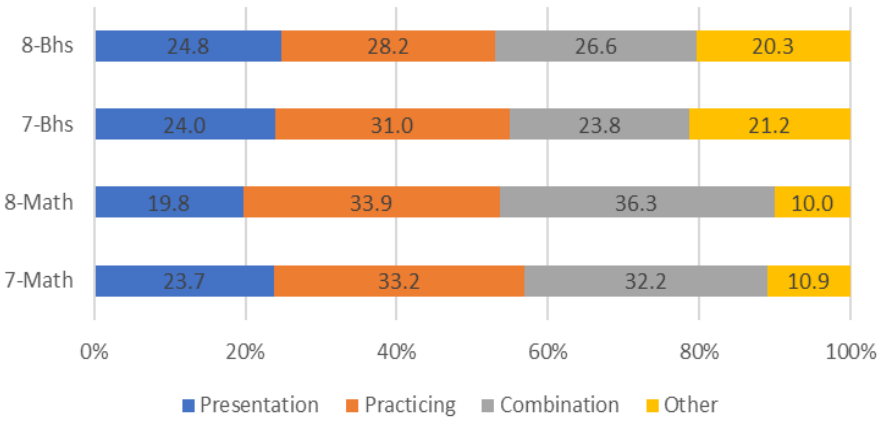

Figure A2. Proportion of teacher's instructional time in presenting and/or practicing

c. *Conventional method*. We also derive a variable to understand whether a teacher's presentation and/or practice session is carried out in a conventional method. The conventional method is marked as "1" on any sheet, which means the teacher does not perform any of these teaching practices:

- Teacher connects today's content with other content that has been taught in the past period (coded as B1). If the teacher does not perform this practice, it is marked as "0".
- Teacher connects today's content with situations or daily activities that are commonly experienced or done by students (coded as B2). If the teacher does not perform this practice on any sheet, it is marked as "0".
- Teacher reiterates learning objectives relating to today's material (coded as B3). If the teacher does not perform this practice on any sheet, it is marked as "0".
- Teacher uses instructional media and involves more than five students when using it (coded as C3 and D1). If the teacher does not perform this practice on any sheet, it is marked as "0".

The following graph shows the result of analyzing the observation data based on the teacher's presentation/practice session and conventional method. There is still a higher proportion of teacher's presentation and practice session carried out without teacher connecting the material with previous content or daily life situation; teacher reminding students of learning objectives; and teacher using learning media that involves more than five students.

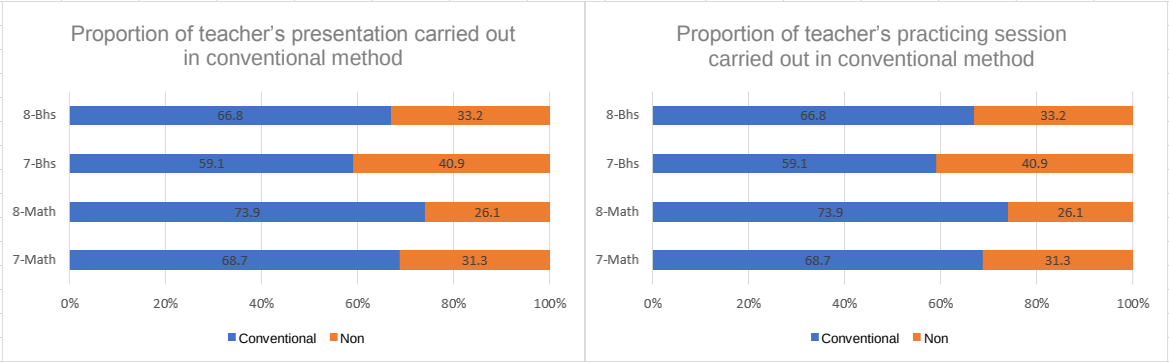

Figure A3. Proportion of teacher's presentation and practice session carried out in a conventional method

(2) The Structure and Types of Teacher-Student Interaction

Following the definition in Indonesia TIMSS Video Study (Ragatz et al., 2015), we use three categories to identify structure and types of interaction between teacher and students in the classroom: group work interaction, seatwork interaction, and whole-class interaction. Our observation instrument includes the "teacher gives assignments to students in the classroom relating to today's learning materials" indicator—whether individual assignment or assignment in pairs or groups. We converted this indicator to show structure of teacher-student interaction in the classroom.

a. Group work interaction

If the teacher gives classroom assignment in pairs or groups (coded as A3a), it is marked "1" on any sheet (observation intervals) indicating "yes" or "occurred". We acknowledge this indicator

as a group work interaction. To identify the type of interaction between teacher and students in this group work setting, we use the "teacher checks students' understanding regarding assignment" indicator. The indicator is marked "1" on any sheet if there is a group work setting with teacher involvement. The indicator is marked "0" if there is a group work setting without teacher involvement, i.e., student only.

b. Seatwork interaction

If the teacher asks students to do classroom assignments individually (coded as A3b), it is marked "1" on any sheet (observation intervals) indicating "yes" or "occurred". We acknowledge this indicator as a seatwork interaction. Similar with the type of interaction in group work setting, we use the "teacher checks students' understanding regarding assignment" indicator (coded as E2). The indicator is marked "1" on any sheet if there is a group work setting with teacher involvement. The indicator is marked "0" if there is a seatwork time without teacher involvement, i.e., student only.

c. Whole-class interaction

On any sheet (observation intervals), if the teacher does not give group or individual classroom assignment, we regard this classroom setting as a whole-class interaction, which is marked as "1". We use "students' involvement in using learning materials provided by the teacher" (coded as D1) and "active participation of the student" (coded as D2) indicators as evidence of teacher-student interaction. If one or both indicators occur in a whole-class setting, with the participation of more than five students, it is marked as "1", which means whole-class time with teacher-student interaction.

Using structure and type of interaction, we can show the proportion of "time" allocated by the teacher in a whole-class activity compared to group work and seatwork. The Indonesia TIMSS Video Study shows that teacher-student interaction in a whole-class time had a positive relationship with student learning, whereas whole-class setting without evidence of student engagement (teacher-only time) had a negative relationship (Ragatz et al., 2015). Thus, analyzing the structure and type of interaction can be used to portray good teaching practices.

The following graph shows the result of processing the raw observation data based on the structure and types of teacher-student interaction theme. The proportion of instructional time is largely carried out in whole-class setting, whether it is mathematics or Indonesian classroom.

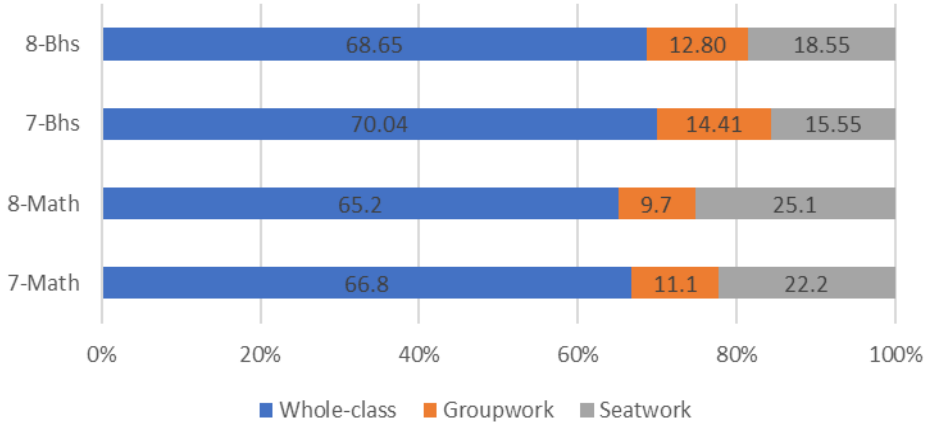

Figure A4. Proportion of instruction time by structure of interaction

Using the structure and types of teacher-student interaction theme, we can also show the type of interaction between teacher and students during each class setting. During the group work setting, teacher-student interaction occurs when the teacher does not merely assign group assignment, but also follows-up on students' understanding about the assignment. The graph below shows that teacher-student interaction is relatively low during group work time.

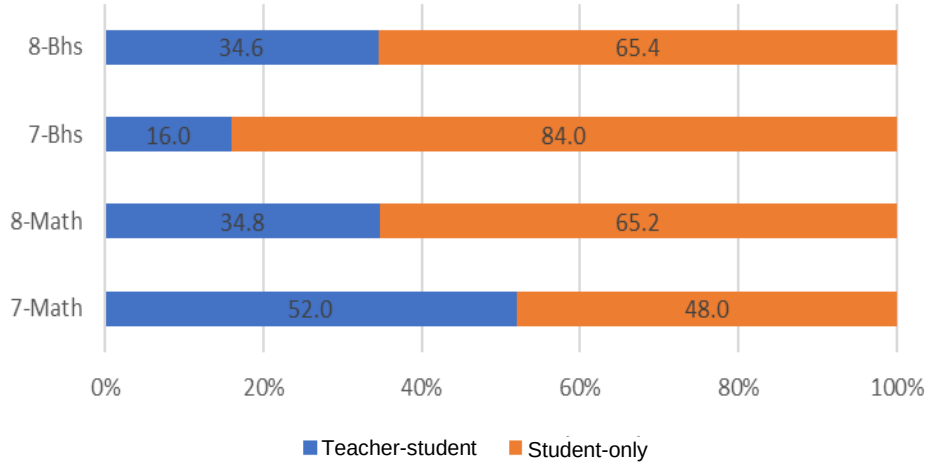

Figure A5. Proportion of groupwork time by interaction type

Teacher-student interaction occurs more during individual or seatwork setting than during group work time. The interaction is also likely to occur during mathematics lesson than during Indonesian lesson.

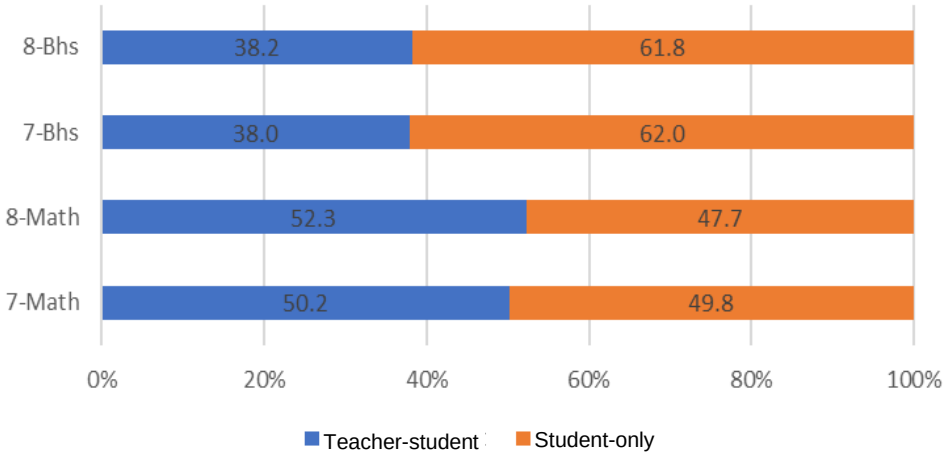

Figure A6. Proportion of individual/seatwork time by interaction type

During the whole-class time, the proportion of teacher-student interaction is almost a fifty-fifty split with teacher-only interaction type. It shows that the teacher is able to involve students' participation in about 50% of their whole-class teaching time.

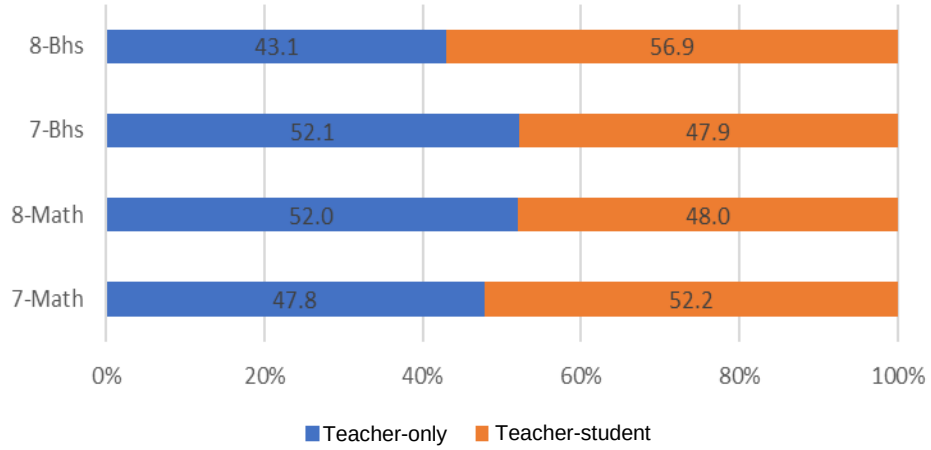

Figure A7. Proportion of whole-class time by interaction type

(3) Questioning

During the lesson, we observe whether the teacher asks open or closed questions. Based on the snapshot approach in our observation instrument, we observe the teacher's effort in asking questions to students in each observation sheet. When the teacher asks one question, either open-ended (coded as E4) or close-ended (coded as E3), or several questions within five minutes of one observation sheet, we mark it as "1" in any sheet. Open-ended questions tend to extend and construct students' thinking. Effective teacher uses such form of questions to encourage higher order thinking (Ragatz et al., 2015). Thus, we measure the teacher's effort in asking one or more open-ended questions as one of the good teaching practices.

a. *Teacher's effort in asking one or more open-ended questions during the lesson*

   Teacher's effort in allocating her or his instructional "time" to ask open-ended questions during the lesson is marked as "1".

b. *Percentage of open-ended to total questions in a lesson*

   The total number of the teacher's effort in asking questions is the sum of closed-ended question indicator (coded as E3), marked as "1", and the sum of open-ended question indicator (coded as E4) marked as "1", in all observation sheets (observation intervals) during the lesson. We use the percentage of open-ended questions to the total questions in one lesson to assess teacher's effort in asking open form questions.

   There is limitation in analyzing the aspect of questioning as a good teaching practice when using the snapshot approach. We record teacher's effort to ask open-ended questions using our observation instrument. However, the snapshot approach is unable to capture whether the questions gauge student response or raise classroom discussion. To understand the way the teacher asks the open-ended questions and how the students' response, one can use the videotape for more in-depth analysis.

   The following graph shows the result of processing the raw observation data based on the questioning theme. The graph shows that most teachers make an effort of asking open-ended questions at least one time during the lesson.

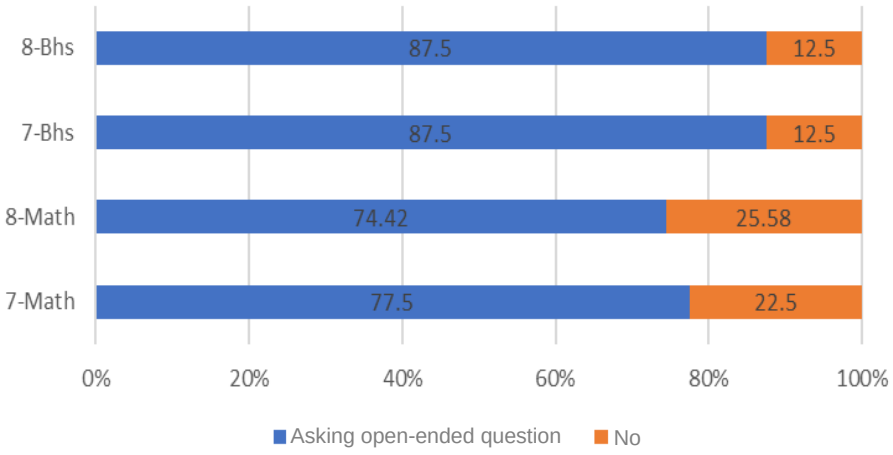

Figure A8. Proportion of teachers asking open-ended question

Previous graph shows that most teachers asking open-ended questions. However, the number of open-ended questions compared to total questions asked in the classroom is relatively low. It means that most of the time, teachers are asking closed-ended questions. The occurrence of open-ended questions in Indonesian class is more frequent than in mathematics class.

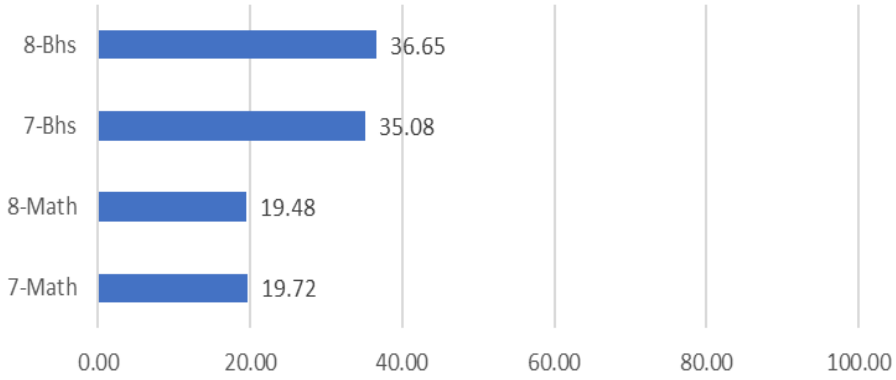

Figure A9. Percentage of open-ended question to total question

(4) Supportive Classroom Environment

Using the observable aspect in our instrument, we define supportive environment for learning in two criteria: classroom condition and teacher's behavior.

Classroom condition is marked as "1" if all the indicators below are marked as follows:

- Good lighting (coded as OK7), marked as "1".
- Good air circulation (coded as OK8), marked as "1".
- All students are able to see the writing on the board clearly from their sitting position (coded as OK11), marked as "1".
- The number of tables and chairs are enough for the number of students (coded as OK12), marked as "1".

Teacher's behavior is marked as "1" if all the indicators below are marked as follows:

- Teacher calls each student by their respective names (coded as OK1), marked as "1".

- Teacher does not leave the classroom during the lesson (coded as OK3), marked as “0”.
- Teacher does not receive interruption, such as visit from other teachers (coded as OK5), marked as “0”.
- Teacher praises students or gives positive responses (coded as F1), marked as “1”.
- Teacher circulates the classroom during the lesson (coded as F2), marked as “1”.

The following graph shows the result of processing the raw observation data based on the supportive classroom environment theme. Most teachers is provided with a good classroom condition rather than supportive teacher’s behavior. Approximately 74% of the teachers, across subject knowledge and grade, is teaching in a class that meets the criteria of a classroom condition that supports learning.

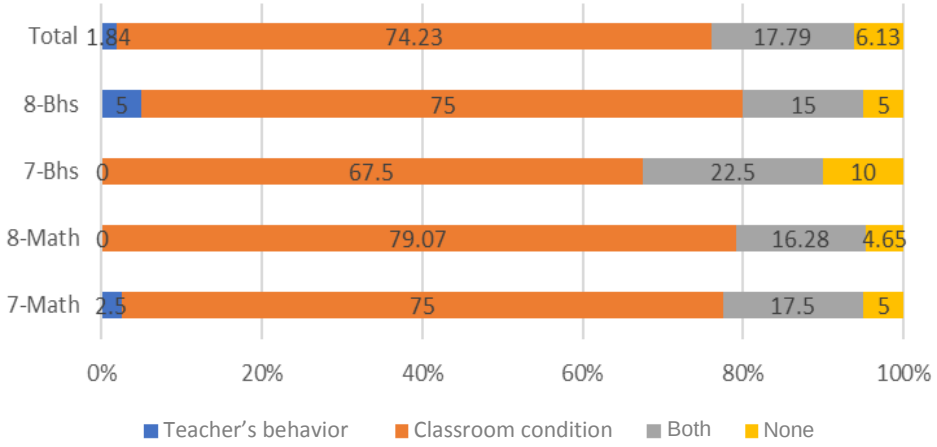

Figure A10. Proportion of teachers with the criteria of supportive classroom environment

In assessing whether the absence of this expected practices would indicate poor student learning, we suggest including other data, such as student learning gain, student characteristics, as well as school and its staff condition. One should notice that the teaching themes we present may not completely satisfy individual interest in specific teaching practices. However, the observable variables in our classroom observation instrument allows for different analysis. One alternative is if there is a clearly defined intervention based on individual/organization interest and good control of counterfactual, then the observation instrument can be used to evaluate the intervention and improve teaching practices.

The SMERU Research Institute

Telephone : +62 21 3193 6336
Fax : +62 21 3193 0850
E-mail : smeru@smeru.or.id
Website : www. smeru. or.id
Facebook : @SMERUInstitute
Twitter : @SMERUInstitute
YouTube : The SMERU Research Institute

9 786237 492375
9 786237 492382