EURASIA Journal of Mathematics, Science and Technology Education, 2018, 14(10), em1598
ISSN:1305-8223 (online)
Research Paper https://doi.org/10.29333/ejmste/93420

Opportunity-to-Learn to Solve Context-based Mathematics Tasks and Students' Performance in Solving these Tasks – Lessons from Indonesia

Ariyadi Wijaya [1,2*], Marja Van den Heuvel-Panhuizen [1,3,4], Michiel Doorman [1], Michiel Veldhuis [3]

[1] Freudenthal Institute, Faculty of Science, Utrecht University, THE NETHERLANDS
[2] Yogyakarta State University, INDONESIA
[3] Freudenthal Group, Faculty of Social and Behavioural Sciences, Utrecht University, THE NETHERLANDS
[4] Nord University, NORWAY

Received 5 December 2017 ▪ Revised 6 June 2018 ▪ Accepted 9 June 2018

ABSTRACT

This study investigated whether providing opportunity-to-learn can improve Indonesian students' performance in solving context-based mathematics tasks. On the basis of an inventory of Indonesian students' difficulties with these tasks and an analysis of textbooks and classroom practices, an intervention program for mathematics teachers was developed. This program contained tasks with relevant and essential contexts with missing or superfluous information, but without explicitly given mathematical procedures. The program also comprised guidelines for a consultative teaching approach with metacognitive prompts and questions for discussion to promote reflection in class. A field experiment with a pretest-posttest control-group design was carried out in six junior high schools in Indonesia involving 299 eighth-graders. Students in the experimental group made significantly more progress on solving context-based mathematics tasks than students in the control group. Furthermore, an analysis of students' errors revealed that experimental students made significantly fewer task comprehension errors than control students. These results show that providing opportunity-to-learn, that is offering context-based tasks to students, which require mathematical modeling, and having teachers knowing the characteristics of such tasks and using a consultative teaching approach, can improve students' ability in solving context-based tasks.

Keywords: mathematics education, context-based tasks, Opportunity-to-learn (OTL), students' performance, Indonesia

INTRODUCTION

The broad recognition of the importance of mathematics in coping with the demands of the 21st century has led to an emphasis on developing students' ability to apply mathematics as an essential goal of mathematics education (Eurydice, 2011; Graumann, 2011; National Council of Teachers of Mathematics, NCTM, 2000; Organisation for Economic Co-operation and Development, OECD, 2003; Tomlinson, 2004). However, in the Programme for International Student Assessment (PISA) studies (e.g., OECD, 2010, 2013) it was found that many students cannot solve problems that require mathematical modeling of complex everyday situations.

Similar to other countries, Indonesia also considers the application of mathematics as a relevant aspect of the mathematics curriculum (Pusat Kurikulum, 2003). Inspecting the results of the PISA studies demonstrates that Indonesian students perform weakly (last position of all participating countries) on the mathematics subscales of formulating, employing, and interpreting context-based tasks (OECD, 2013). This situation prompted us to set up the *Context-based Mathematics Tasks Indonesia* (CoMTI) project. In this project, 'context-based tasks' were defined as

✉ a.wijaya@uny.ac.id **(*Correspondence)** ✉ m.vandenheuvel-panhuizen@uu.nl ✉ m.doorman@uu.nl ✉ m.veldhuis@uu.nl

Contribution of this paper to the literature

- This study employed a research-based design method in which on the basis of data about students' errors, content offered in textbooks, and teachers' beliefs and teaching practice, instructional material was developed for teaching students solving context-based mathematics tasks.
- This study elaborated on the concept of opportunity-to-learn by integrating what the textbook is offering with what the teachers can contribute.
- This study showed that offering context-based tasks to students that require mathematical modeling, and having teachers knowing the characteristics of such tasks and being able to use a consultative teaching approach, can improve students' ability in solving context-based tasks.

problems situated in real-world settings which contain elements or provide information that need to be organized and modeled mathematically to reach a solution (Freudenthal, 1983). The contexts of the tasks should refer to situations that the students can imagine and that are truly meaningful to them. This implies that in choosing tasks, students' experiences and reference frameworks have to be carefully considered and taken into account, to ensure that these situations indeed make sense to them and have the potential to engage them in a purposeful mathematical activity (e.g. Ainley, Pratt, & Hansen, 2006; De Lange, 2015). The aim of the CoMTI project was to identify ways to improve Indonesian students' ability to apply mathematics in extra-mathematical situations. The first step of this project was a study on identifying Indonesian students' difficulties in solving context-based tasks in which they have to apply mathematics. An error analysis that was performed in this first study (Wijaya, Van den Heuvel-Panhuizen, Doorman, & Robitzsch, 2014) showed that Indonesian students mainly have difficulties in comprehending real-world problems and in transforming them into mathematical problems. Our second study (Wijaya, Van den Heuvel-Panhuizen, & Doorman, 2015a) revealed a relation between these difficulties and insufficient opportunity-to-learn (OTL) to solve context-based tasks offered in Indonesian textbooks. Indonesian textbooks contain a low number of context-based tasks. Furthermore, these context-based tasks are mostly plain word problems, which use dressed-up contexts, explicitly indicate the mathematical procedures to carry out, and provide precisely the information needed to solve the task. Finally, in the third study of the CoMTI project (Wijaya, Van den Heuvel-Panhuizen, & Doorman, 2015b) classroom observations were carried out to identify what opportunities teachers offer their students to learn to solve context-based tasks. It was found that Indonesian teachers mostly used a directive and teacher-centered teaching approach and did not give students opportunities to get involved in and reflect on the process of solving context-based tasks.

In this study, we synthesized the findings of the earlier studies of the CoMTI project to develop an intervention that offers students OTL to solve context-based tasks. This intervention was put to the test in a field experiment in Indonesia to investigate its effect on student performance with the focus on students' scores and students' errors.

THEORETICAL FRAMEWORK

In this section, we describe the theoretical foundations of the present study. First, we discuss the context in context-based problems, different models of problem solving, and the relevant stages in students' reasoning when solving context-based problems. Then, we elaborate on the concept of OTL, and the different contexts in which such opportunities can be offered to students. Finally, we connect this with what we already know about Indonesian students' difficulties in solving context-based tasks, OTL offered in Indonesian textbooks, and OTL offered in Indonesian teachers' teaching practice.

The Context in Context-based Problems

With the context of a problem we refer to the situation in which the problem itself is embedded (Borasi, 1986). A mathematical problem can be embedded in various ways in a context. De Lange describes a use of contexts to camouflage a mathematical problem in cases where the mathematical procedures are immediately recognizable as in plain word problems (De Lange, 1987). He distinguishes this use of contexts from contexts that create the need to find or develop the relevant mathematics to organize, structure, and solve the problem. In the latter case, contexts are used as a didactical tool to support the learning of mathematics (Gravemeijer & Doorman, 1999; Van den Heuvel-Panhuizen, 1996, 2005). In the present study, the focus is on this use of contexts, and in particular on extra-mathematical contexts, which in PISA are defined as problems presented within a situation that can refer to a real world or fantasy setting, and can include personal, occupational, scientific, and public information (OECD, 2003). Since the aim of our intervention is on teaching Indonesian students to solve context-based tasks, we were interested in contexts for supporting a particular topic as well as providing opportunities to address issues like dealing with superfluous or missing information. We were less interested in context-based tasks that hardly require any modelling or tasks that require much contextual knowledge, like interdisciplinary or authentic vocational problems.

Different Models of Problem Solving

The most well-known model of problem solving is that of Polya. In his seminal work "How to solve it" (Polya, 1945) he distinguished the following stages in the process of problem solving: (1) understanding the problem, (2) devising a plan, (3) carrying out the plan, and (4) looking back. These stages can also be found in other models of problem solving, in which particular elements and characteristics of Polya's model were adapted somewhat. For example, Schoenfeld (1985) used a model with five stages by adding a stage of exploration between understanding the plan and divising a plan. In this exploration stage the problem-solving heuristics come into play. A further difference is that according to Schoenfeld the stage of designing a plan is not an isolated activity but implies a global perspective on what has to be done. The model that Mason, Burton, and Stacey (1982, 2010) have developed for problem solving has a different and rather complex structure consisting of phases (entry, attack, and review) and states (getting started, getting involved, mulling, keeping going, insight, being skeptical, and contemplating). Another model that is more in line with the models of Polya and Schoenfeld is the one that Blum and colleagues have created. Characteristic of their approach is that they focus in particular on applied mathematical problem solving (Blum, 1993; Blum & Niss, 1991). This means that they take a situation in the real world as a starting point for problem solving. Therefore, we decided to choose Blum's model for our study.

Stages in Students' Reasoning when Solving Context-based Tasks

In the applied mathematical problem solving as described by Blum and colleagues (Blum, 1993, 2011, 2015; Blum & Borromeo Ferri, 2009) the modelling aspect plays a crucial role. In total they have identified four stages of mathematical modeling for solving context-based tasks. The first stage relates to the process of comprehending what a context-based task is about. In the second stage, students look for a mathematical concept or procedure required to solve the task. In this stage, the real-world problem is transformed into a mathematical problem. The third stage is carrying out the mathematical procedure to solve the mathematical problem. Finally, in the fourth stage, students interpret and validate the solution in terms of the context of the task in addition they also reflect on the whole modeling process.

In each of these stages, students face difficulties and can make errors. Research has shown that in the first stage students often misunderstand the meaning of the tasks and misinterpret the terms used in the tasks (Bernardo, 1999; Klymchuk, Zverkova, Gruenwald, & Sauerbier, 2010). In the second stage, students struggle with identifying the mathematical concept or procedure that is needed to solve the tasks (Clements, 1980; Klymchuk et al., 2010). This difficulty relates to students' tendency either to ignore the context and try to apply a routine mathematical procedure without realistic considerations (Blum, 2015; Verschaffel, Greer, & De Corte, 2000; Xin, Lin, Zhang, & Yan, 2007) or to take too much account of the context of the tasks so that no mathematical concept or procedure is used (Boaler, 1994). In the third stage, students can make errors in carrying out mathematical procedures. In the fourth stage, students often have difficulties in interpreting a solution in terms of the context, and give solutions that are not relevant to the context of the tasks (Greer, 1997). Furthermore, validating the results and reflecting on, and exposing, the whole modelling process are mostly not present in students' solutions (Blum, 2015).

The Concept of Opportunity-to-learn

The concept of OTL emerged when researchers and educators started to question the factors that could explain students' unsatisfactory performance. Since the 1960s when OTL was coined by Caroll (1963) when referring to sufficient time for students to learn, the relation between OTL and students' achievement has been documented in many studies (e.g. Grouws & Cebulla, 2000). In particular, researchers in comparative studies became aware that when comparing the achievements of students, students' OTL resulting from curricular differences had to be taken into account (McDonnell, 1995). Therefore, OTL was often used to find an explanation why students from different countries performed differently in international comparative studies.

For example, the First International Mathematics Studies examined "whether or not […] students have had the opportunity to study a particular topic or learn how to solve a particular type of problem" (Husén, 1967, pp. 162-163) to find factors influencing students' performance across countries. In the PISA studies, a slightly different description of OTL is used: "the relative exposure that students of different backgrounds may have to specific content in the classroom […] reflected in the instructional time school systems and teachers allocate to learning a particular subject or content" (OECD, 2016, p. 206). In the present study, we used a broader definition of OTL, in addition to the exposure to mathematics content it also includes the characteristics of instruction and the learning environment that is offered. To measure OTL we followed Brewer and Stasz (1996) and included instructional resources such as textbooks and teachers' instructional strategies.

Opportunity-to-learn offered by textbooks

Textbooks are considered as the main instructional material for teachers (Brewer & Stasz, 1996) mediating between the intended and the implemented curriculum (Valverde, Bianchi, Wolfe, Schmidt, & Houang, 2002). Research in different countries has shown that textbooks strongly influence students' learning (Schmidt, McKnight, Valverde, Houang, & Wiley, 1997; Tornroos, 2005). What is offered in textbooks can be regarded as an important measure for students' OTL. In this respect several aspects of textbooks can be taken into account. A first aspect is the exposure of particular content. Tornroos (2005) found a relation between student achievement on a test and the amount of textbook content related to the test items. A second aspect concerns the support provided by the textbooks to help students understand the content. As found by Xin (2007), students tend to solve word problems by using the solution strategies suggested in their textbooks. A third aspect is the nature of the tasks or exercises that are provided in the textbooks. In this respect, Howson (2013) mentioned the setting of the tasks, the quality of the tasks, the number of the tasks, and the role of the contexts which are used in the tasks. The characteristics of tasks, such as the cognitive demands of tasks and the required types of responses in tasks, are also important attributes (Charalambous, Delaney, Hsu, & Mesa, 2010).

The importance of providing students more context-based tasks was highlighted by Ikeda (2007) who argued that a lack of such tasks in classroom practices contributed to students' low performance. With respect to providing OTL to solve these tasks, Maass (2007) emphasized the importance of giving students tasks that have superfluous and missing information. Such tasks are necessary to direct students to pay attention to the context of the tasks and to teach students to distinguish between relevant and irrelevant information. Furthermore, it is also essential for developing students' modeling competence to provide them with real-world problems that do not provide explicit suggestions about the required procedures to solve the problems (Maass, 2010).

Opportunity-to-learn offered by teachers' instructional strategies

A number of studies (e.g. Eurydice, 2011; Grouws & Cebulla, 2000; Hiebert & Grouws, 2007) highlight that student performance is also affected by the OTL offered by teachers through their instructional strategies. The quality of teachers' use of these strategies, specifically how they teach mathematics and engage their students, influences how well students learn. With respect to the teaching of context-based tasks, several researchers (Antonius, Haines, Jensen, Niss, & Burkhardt, 2007; Blum, 2011; Forman & Steen, 2001) suggested the use of a teaching approach in which teachers take a consultative role and give students opportunities to actively build new knowledge and reflect on their learning process. Another important role of teachers in teaching context-based tasks is offering students more opportunity to reflect on and evaluate their own ideas (Doerr, 2007). A key aspect of consultative teaching is to keep a balance between teacher guidance and students' independence. Both Antonius et al. (2007) and Blum (2011) recommended the use of metacognitive prompts to create this balance and Montague (2007) emphasized that these metacognitive prompts help students to become active learners.

Metacognitive prompts can be provided in the form of self-addressed questions; students are asked to question themselves while solving a problem. Self-addressed questions are an important stimulus to help students regulate and reflect on their solving process (Kramarski, Mevarech, & Arami, 2002; Montague, 2007, 2008). Another kind of metacognitive prompt is giving a verbal prompt or instruction to help students focus attention on particular aspects of the solving process and to assist them in carrying out the solving process (Goldman, 1989; Montague, 2007; Montague, Warger, & Morgan, 2000). For example, the instruction to underline the important information in a task can be used to guide students to focus on identifying relevant information. Asking students to paraphrase a task is also an important prompt. Karbalei and Amoli (2011) and Kletzien (2009) found that students who explain in their own words what the task is about gain a better understanding of the task.

Indonesian Students' Difficulties when Solving Context-based Tasks

To investigate students' difficulties when solving context-based tasks we (Wijaya, Van den Heuvel-Panhuizen, Doorman, & Robitzsch, 2014) analyzed the errors made by students. Related to the four stages of solving context-based tasks (Blum, 2011; Blum & Borromeo Ferri, 2009) four error types were distinguished: comprehension, transformation, mathematical processing, and encoding errors. Comprehension errors refer to students' inability to understand a context-based task, which includes the inability to select relevant information. Transformation errors correspond to students' failure in identifying the mathematical procedure required to solve a context-based task. This error also applies to the inverse situation, when students cannot interpret a (mathematical) model of a situation. Mathematical processing errors are related to mistakes in carrying out mathematical procedures; for example, errors in calculation or solving algebraic formula. Encoding errors refer to answers that are unrealistic and do not match the context of the task.

The error analysis revealed that 38% of Indonesian students' errors when solving context-based tasks were comprehension errors. Students often misunderstood the question of a context-based task. Students also made remarkable errors in selecting relevant information: they tended to use all information given in the text without considering its relevance. Transformation errors were found in 42% of students' errors: many students used the wrong procedure to solve context-based tasks. With respect to mathematical processing errors, it was found that 17% of all errors were of this type. Finally, encoding errors were found only in 3% of students' errors. An example of this error type is students giving 70 meters as the pace length of a human.

Opportunity-to-learn offered in Indonesian textbooks

To find possible explanations for Indonesian students' difficulties when solving context-based tasks, we first investigated the OTL to solve context-based tasks offered in Indonesian textbooks. A textbook analysis revealed correspondences between Indonesian textbooks and the errors made by students (see Wijaya, Van den Heuvel-Panhuizen, & Doorman, 2015a). First of all, it was found that only 10% of the tasks in the textbooks were context-based tasks. With such a low number of context-based tasks, we cannot expect Indonesian students to be good at solving such tasks. When zooming in on the characteristics of the context-based tasks in the textbooks, the correspondences between textbooks and students' errors were even more obvious. Three quarters of context-based tasks in the textbooks used camouflage contexts, which means the contexts of the tasks can be neglected and the procedures are explicitly provided. When solving such tasks students did not have to think about transforming the tasks into mathematical problems. This finding corresponded to the high percentage of students' transformation errors. With respect to the type of information, 85% of the context-based tasks contained matching information, meaning that only the relevant information was provided. This lack of opportunity for students to deal with superfluous and missing information related to the large number of errors Indonesian students made in selecting information.

Opportunity-to-learn offered by teaching practices of Indonesian teachers

A further step to investigate to what degree students are offered OTL to solve context-based tasks was by investigating Indonesian teachers' teaching practices. The data about teachers' teaching practice were collected through a written questionnaire and classroom observations (see Wijaya, Van den Heuvel-Panhuizen, & Doorman, 2015b). The questionnaire focused on the characteristics of context-based tasks used by teachers in their teaching practices. It was found that 67% of the teachers gave context-based tasks with explicit procedures in every lesson or weekly. Regarding context-based tasks with superfluous information, 40% of the teachers reported that they almost never gave such tasks.

Classroom observations were conducted in four classrooms to examine in more detail the teachers' teaching practices. The classroom observations focused on identifying whether teachers used a consultative teaching or a directive teaching approach. As reported in Wijaya, Van den Heuvel-Panhuizen, and Doorman (2015b), the observations revealed that the teachers tended to use directive teaching. Furthermore, the teachers only focused on the correctness of students' mathematical solutions without connecting the solutions to the context of the task. This finding indicates that the teachers did not provide sufficient opportunities for students to actively learn to deal with context-based tasks.

Research Question

The aim of the present study was to test whether students' performance in solving context-based tasks can be improved by offering them OTL solving these tasks. The ingredients for creating this OTL were derived from research literature and from the findings of our three earlier studies. This OTL comprised offering students context-based tasks with particular characteristics, which were found to be lacking in Indonesian textbooks, and using a consultative teaching approach, which teachers were found to rarely use. To provide students with as much OTL as possible in a coherent way we developed an intervention in which both were integrated. To investigate whether this is an effective way to improve students' ability to solve context-based tasks we focused on students' scores and on students' errors.

This led to the following research question: *Does providing students with an opportunity-to-learn to solve context-based tasks contribute to students' performance in solving these tasks?* In particular we investigate what the effect is on the students' success rate when solving context-based tasks and on the number and types of errors.

Based on the studies, described in Section 2.2, that investigated the effect of OTL on students' mathematics achievement, our first hypothesis was that students who got the OTL solving context-based tasks will show more improvement of their achievement in solving these tasks than students who did not get this OTL. Furthermore, our second hypothesis was that the OTL will help students to reduce the number of errors when solving context-based tasks.

METHOD

Design of the Study

To answer the research question, we carried out a field experiment with a pretest-posttest control-group design. In the experimental group, the teachers used an intervention program consisting of five lessons that offer students OTL to solve context-based tasks. In the control group, the students followed a teaching program that was developed on the basis of the textbook that they regularly use and that did not include the components of the intervention program. For the study we collected two types of data. To investigate whether the intervention had an effect on the students' performance in solving context-based tasks, student data were collected by means of a pretest and posttest. In addition, classroom data were collected both to check the fidelity of the intervention and to have more detailed information of how the OTL worked in classroom. The classroom data consisted of video-recordings and field notes, made by the first author when observing the experimental and control classes. The video-recordings were done with two cameras; i.e. a static camera to capture whole-class activities and a dynamic camera to capture specific students' activities. These video-recordings were transcribed and translated in English, to allow all authors to understand what students and teachers were saying. Also the teachers were asked to keep logs of their lessons. More details about the measurement of the students' performance and the content of the intervention program follow in the next sections.

Setting and Participants

The study took place in Grade 8 in six junior high schools located in the province of Yogyakarta, Indonesia. The schools were selected in this province for reasons of convenience. The location of the schools made it possible to prepare the teachers, to secure the data collection, and to perform observations. In each of the six schools there was an experimental class and a control class. In four schools (coded as PR, S, M, and G) Grade 8 was taught by only one teacher; therefore in these schools this teacher taught both the experimental and the control class. In the other two schools (coded as PL and B) the experimental class and the control class were taught by different teachers. The decision about which teacher was in which condition was left to the school principal in order to afford a benevolent collaboration within the study. All eight participating teachers had a bachelor degree in mathematics education and considerable teaching experience, ranging from 5 to 32 years (M = 18.9 years; SD = 9.9 years). In each school, several textbooks were in use, but all schools also had one textbook in common, which was *Matematika* (Textbook for Junior High School, Grade VIII: 2A & 2B).

In total, 311 eight graders (M = 13.8 years; SD = 0.5 year) participated in the study, involving 146 students in the experimental group and 165 in the control group. The data analysis included only the 299 students (M = 13.7 years; SD = 0.5 year) who were present during both pretest and posttest. Of these students, 144 students were in the experimental group and 155 students in the control group.

Intervention Program

The purpose of the intervention was to offer students OTL to solve context-based tasks. The intervention program was designed based on the findings of our earlier studies and comprised two components: a set of context-based tasks with particular characteristics and a consultative teaching approach with metacognitive prompts (see **Appendix 1**).

Context-based tasks

To compensate what was lacking in Indonesian textbooks, the intervention program offered students context-based tasks. We developed context-based tasks by considering three characteristics. The first characteristic was that the context-based tasks should have *relevant and essential contexts* that require students to connect the contextual problem to a mathematical strategy and to identify the relevant information for performing that strategy. The tasks should be a real problem for students. For example, in the Internet task (see **Figure 1**) students encounter a problem that is rather authentic to them. Figuring out what internet fee Doni has to pay is a meaningful context for students because there is really something at stake. Students are stimulated to take the context into account when solving the task instead of just using the numbers involved to carry out a particular calculation.

An internet provider Inter-NET offers two different programs. Program Smile charges customers 30.500 IDR as monthly fee and 40 IDR/1 Megabyte (MB). Program Shine charges customers 20.000 IDR as monthly fee and 52 IDR/1 MB. The registration fees including the price of modem for both programs are the same, namely 300.000 IDR. In January Doni subscribed to the program Shine. In May Doni used 550 MB of internet data. How much money did he pay in May?

a. *Underline all information given in the task and circle only the information we need to answer the question.*

b. *Reformulate the task with less words by leaving out unnecessary information.*

c. How much money did Doni pay in May?

Doni's internet usage is increasing. Now, he has a problem to decide whether it will be wise to change the internet program.

d. *What strategy can we use to solve Doni's problem?*

e. When Doni's internet usage is increasing, is it better for him to change the internet program? If so, when should he do it?

Figure 1. Internet task (translated from Bahasa Indonesia)

The second characteristic concerned the type of information provided in the context-based tasks. To offer students OTL selecting relevant information, context-based tasks were designed with *missing or superfluous information*, i.e. tasks that have less or more information than needed to find the solution. For example, to decide how much Doni has to pay in May, students do not need to use the monthly fee for the program Smile (30,500 IDR), the fee per 1 MB (40 IDR), or the registration fee (300,000 IDR).

The third characteristic required for the context-based tasks was that they do *not contain apparent indications about the procedures* that can be used to solve the problem. If explicit information is given about what procedure to apply, then students are not offered OTL to decide what would be a suitable mathematical procedure. For example, the Internet task does not directly ask students to use a particular mathematical procedure, such as doing a calculation or drawing a graph. Instead, students are asked to decide whether it is better for Doni to change his internet program and, if yes, when it is better to change. This means that students have to come up with a procedure by themselves.

Consultative teaching approach

The second component of the intervention program was a consultative teaching approach. To conduct this approach, we provided the teachers with suggestions to give students metacognitive prompts and stimulate discussions that promote reflection for all the stages of solving context-based tasks.

The *metacognitive prompts* were meant to point students to important aspects of the tasks and the solving process. A first metacognitive prompt was asking students to underline all the information included in a context-based task and to discuss the included information. For example, in the Internet task, this prompt was given in the first assignment. As a second metacognitive prompt, students were asked to use their own words to explain the Internet task. This paraphrasing strategy was aimed to help students get a better understanding of the problem. Finally, a third type of metacognitive prompt was to elicit self-questioning. For example, in the Internet task, students were stimulated to ask questions, such as "What strategy can we use to solve Doni's problem?"

To make the metacognitive prompts an integral part of the intervention program they were embedded in the context-based tasks. Metacognitive prompts were printed in italics to make students (and teachers) aware that the questions differed from the mathematical questions. In the first two lessons, metacognitive prompts were provided for every question in the context-based tasks, but in the later lessons the prompts gradually faded out.

The *suggestions for discussions to promote reflection* were not embedded in the tasks, but were only included in the teacher guide. This means that suggestions were provided for additional instructions or guiding questions. For example, when students had difficulties with only circling the relevant information in the Internet task, the teacher could suggest students to focus on the question they had to answer ("How much money did he pay in May?") and identify the keywords of this question. Regarding the latter, different students can come up with different keywords and discussion is triggered when they have to explain why they think a particular word could be considered a keyword.

Outline of the intervention program

The complete intervention program consisted of five 80-minutes lessons, each consisting of an explanation section and an exercise section. The actual intervention took place in the exercise sections, which each lasted about 20 to 30 minutes. The topic that was being taught during the intervention period was graphical representations of linear equations; therefore we chose this topic in designing the intervention program. In total, the program included nine context-based tasks; in every lesson two and in the last lesson one. Moreover, in order not to lessen too much the experience of students in solving mathematics tasks without a real-world context, the lesson series also contained five of such tasks.

Teacher training

To help the teachers to implement the intervention program, we provided them with a teacher guide. In the teacher guide, information was given about the mathematics topics related to the context-based tasks included in the intervention program. Furthermore, the teacher guide informed the teachers about the learning goals, i.e. the mathematical competences or skills to be developed by working on the task. Also, suggestions were given about how the teacher could help students to achieve the learning goals. Finally, the teacher guide gave examples of possible students' answers and how to respond to these answers in order to support the students' understanding.

As an addition to informing the teachers by means of the teacher guide, two weeks prior to the intervention the experimental teachers were trained to conduct the intervention. The training consisted of two meetings of 90 to 120 minutes. During these meetings, the overall goal of the intervention program, the structure of the program, and the materials used in this program were explained. After this, attention was paid to the importance of the competence of solving context-based tasks and the characteristics of context-based tasks. Finally, examples were discussed of how teachers could help students to deal with superfluous or missing information and select relevant information, how teachers could give metacognitive prompts, and how to stimulate students to discuss and reflect on the solving process.

Fidelity of the implementation of the intervention program

To monitor whether the teachers conducted the intervention in the intended way, classroom observations were carried out by the first author. Checking the implementation fidelity was especially important with respect to the teachers that taught both an experimental class and a control class. In this way we could ensure that in both classes the appropriate treatment was given. However, not all lessons could be observed because the intervention in the six schools took place over the same period of time. In general, at least two lessons were observed in each classroom. After observing a lesson, a discussion was held with the teacher to reflect on how the intervention was carried out and prepare for the intervention in the next lesson. In this discussion, the first author also regularly reminded the teachers, especially those who taught both groups, to strictly follow the program for each class. For the lessons that were not observed the teachers were asked to keep a log. In this log they were asked to report how many tasks they discussed in class.

The experiences during the classroom observations, the information gained from the video data, and teachers' logs showed that on average the teachers taught only six out of nine context-based tasks. With respect to the consultative teaching, we found that in general the teachers carried out the intervention as planned. In particular, they frequently emphasized the importance of answering the written questions in the tasks that were related to the metacognitive prompts. Excerpt 1 illustrates how a teacher helped students to comprehend the Internet task (see **Figure 1**). Before the students in her class started to work on the task, the teacher reminded the students to read the instruction and questions thoroughly (see Line 2-3 and Line 7-8 in Excerpt 1). The teacher also read the task aloud (see Line 3-6 in Excerpt 1), which seemed to be a kind of directive teaching. However, the teacher did this only for the first part of the task and she asked the students to continue reading the task by themselves. Another guidance given by the teacher was in the form of emphasizing keywords that are related to metacognitive strategies. This kind of guidance can lead students to become aware of the important words or instruction.

Excerpt 1. Teacher's guidance in the comprehension stage

Teacher :	Now, it is the time for you to solve Task 3 [*the Internet task*]. [*The teacher gave students time to take a look at the task.*] Firstly, you have to thoroughly read the instruction of the task and also the questions. "An internet provider Inter-NET offers two different programs. Program Smile charges customers 30,500 IDR as monthly fee and 40 IDR/1 Megabyte (MB)." [*The teacher reads Task 3 aloud*]. Okay, finish reading the task by yourself. Once again, read the task thoroughly and try to comprehend the task and the question before you start solving it.	[1] [2] [3] [4] [5] [6] [7]
Students:	[*The students continue reading the task.*]	[8]
Teacher :	Now, let's see Question a. What is being emphazed in this question? *[Silence for one minute.*] Underline all information provided in the task and circle ONLY the information which is NEEDED to solve the task. [*Note: The capitalizing means that the teacher put an emphasis in his tone when mentioning these two words.*]	[9] [10] [11] [12]

However, discussions of students' answers and tentative ideas were less visible in the classroom practices. Excerpt 2 illustrates an example of the teachers' focus on the correctness of students' answer instead of presenting and comparing strategies. This excerpt discusses a task in which the students have to determine the price of a taxi ride for a given particular distance. The task contains a distance-price graph. Presenting the graph in the task had the purpose of creating opportunities for students to have more strategies at hand and to be able to select an appropriate strategy when confronted with non-standard context-based tasks. However, not all teachers picked up on this idea. The teacher in Excerpt 2 (see Line 12-14) was satisfied with a student's answer that the result was found by doing a calculation. When the correct answer was given, the underlying reasoning of this answer was not asked for and other students were also not challenged to provide alternative answers or strategies.

Excerpt 2. Teacher's focus on the correctness of students' answer

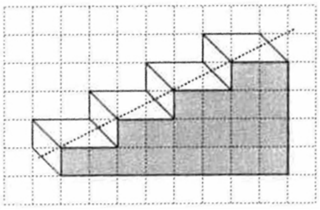
	[The question to be answered was: How much does taxi Sentosa charge for a 20 km journey?]	[1] [2]
Student 1 :	Sir, here is my answer [*The student is showing his answer IDR 37,500.]*	[3] [4]
Teacher:	(…) OK, that is correct. Now, you solve the other questions. Wait a minute, did you use a calculation or read the answer from the graph?	[5] [6] [7]
Student 2 :	[The teacher is *walking to the front of the classroom and posing the question to all students*]. How much should a customer pay if he travels with taxi Sentosa for 20 km? How much?	[8] [9] [10]
Students:	IDR 37,500	[11]
Teacher :	By calculation … or … ?	[12]
Students:	Calculation	[13]
Teacher:	OK, that is good. For the next question…	[14]

Furthermore, it was observed that, especially in unexpected situations, teachers had difficulties in fully implementing the intervention. It was observed that teachers frequently looked at the teacher guide when students' responses were far from the provided examples. Also, the teachers occasionally showed their impatience in waiting for students' answers.

Regular program

Students in the control classes were taught as they were usually taught on the basis of the textbook that the six schools had in common. This means that the teachers used a teacher-centered approach in which they mainly explained and demonstrated how to solve tasks. To make the mathematics content the same in all the control classes, the control teachers were asked to follow a program consisting of 19 mathematics tasks without context and three context-based tasks. These tasks were taken from the textbook that was used by all control teachers. Consequently, the context-based tasks all had a camouflage context and explicitly mentioned the mathematical concepts related to the task. For example, the two tasks shown in **Figure 2** include a staircase and a ski slope, which

1. The following figure is a side view of a staircase. Calculate the gradient of the staircase.



2. Determine the gradient of the following ski slope.

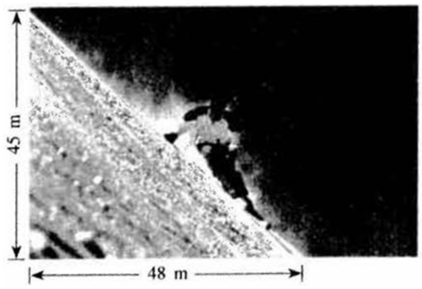


Figure 2. Context-based tasks included in the regular program

are real-world contexts, yet students do not have to think about a staircase and a ski slope in reality. They can just do the calculation that is asked for, based on the figure of the staircase and the numbers provided in the picture of the mountainside (e.g. slope = $\frac{45}{48}$ = 0.9). Furthermore, these two tasks explicitly mention the mathematical concept gradient, which is relevant to solve these tasks.

Measurement of Students' Performance in Solving Context-based Tasks

Composition of the test

The test items used for measuring students' performance in solving context-based tasks were selected from released PISA items (OECD, 2009b). In the selection, we took the three characteristics of the context-based tasks used in the intervention program into account, i.e. tasks that have a relevant context, have superfluous or missing information, and do not contain explicit indications about the required solution strategies. The test comprised 17 PISA mathematics items (which are called 'units') comprising 30 questions. The 17 mathematics items were divided over two booklets. Booklet A contained eight mathematics items (comprising 15 questions) and Booklet B contained nine mathematics items (comprising 15 questions). The questions were equally distributed over the two booklets according to: (1) the cognitive demand of the questions, including reproduction, connection, and reflection, (2) the difficulty level of the questions as indicated by the percentage of correct answers found in the 2003 PISA study (OECD, 2009a), and (3) the mathematical topics involved in the questions. The aim of the study was not to develop students' content-specific skills, but students' generic skills in solving context-based tasks. Therefore, the items which were used in the test addressed various mathematics topics. Only five out of 30 questions were related to the topic of the intervention, i.e. graphical representations of linear equations.

To avoid a re-test effect due to administering the same items twice, the group of students in each class was randomly split in half, leading to two groups of which one got Booklet A as a pretest and Booklet B as a posttest, with the other group getting these booklets in the reverse order.

Scoring the correctness of students' answers

To score students' responses the scoring scheme of the PISA studies (OECD, 2009b) was used. Of the total of 30 questions, 24 questions were scored as correct (1), incorrect (0) or no answer (9). The other questions had a partial credit scoring, including five questions that were coded as correct (2), partially correct (1), incorrect (0) or no answer (9) and one item that was coded as correct (3), partially correct level 2 (2), partially correct level 1 (1), incorrect (0) or no answer (9). The maximum score for Booklet A was 18 and the maximum score for Booklet B was 19.

Psychometric properties of the test

To check the reliability of the test, we calculated Cronbach's alphas. For the complete test, i.e. Booklet A and B combined, this gave a good α of .80. For the booklets separately we found α = .75 for Booklet A, and α = .69 for Booklet B. As the complete sample was split into two groups and each group got a different booklet as pretest and posttest, we also checked whether the reliability estimates per booklet changed for whether it was used as a pretest or as a posttest. Cronbach's alpha for Booklet A as pretest was α = .64 and as a posttest α = .79, for Booklet B as a pretest α = .68 and as a posttest α = .69. These changes are relatively small and all alphas indicate acceptable to good reliabilities.

Table 1. Means and standard deviations of the raw test scores in the pretest and posttest in both conditions for both booklet orders

Experimental group	Pretest		Posttest		
	M	*SD*	*M*	*SD*	*Gain score*
Booklet order A-B	6.36	3.29	7.82	3.39	+ 1.47
Booklet order B-A	6.60	3.19	8.64	4.18	+ 2.04
Control group	**Pretest**		**Posttest**		
	M	*SD*	*M*	*SD*	*Gain score*
Booklet order A-B	6.67	2.85	7.66	2.83	+ 0.99
Booklet order B-A	6.46	2.95	7.04	4.03	+ 0.48

Data Analysis

Analysis of students' errors

To analyze the errors students made when solving the test items we used the framework developed in our earlier study (Wijaya, Van den Heuvel-Panhuizen, Doorman, & Robitzsch, 2014) (see **Appendix 3**). The error analysis was done for incorrect, or partially incorrect, responses. Students could make more than one type of error; therefore more than one code could be given to a student's work. Additionally, we looked at the number of students that made particular types of errors.

The coding of students' errors was carried out by the first author. The interrater reliability of the coding was checked through an extra coding by a mathematics teacher who was not part of this study. The extra coding was done on 12% of students' incorrect responses which were randomly selected. A Cohen's Kappa of .78 indicated a substantial agreement between the first author and the second coder (Landis & Koch, 1977).

Statistical analysis of the effect of the intervention

We first report descriptive statistics (*M*, *SD*, and gain scores) on students' performance on the pretest and posttest. To investigate whether students in the experimental condition improved more than those in the control condition we used an analysis of co-variance (ANCOVA) with the standardized posttest score as dependent variable and the standardized pretest score as covariate. Standardized scores were used in this analysis, because students could obtain different total scores on the two booklets (A and B) they got as pretest or posttest. We standardized the scores per booklet and per order of presentation of the booklets. So the mean of all students that got Booklet A as a pretest was used to standardize the scores of the students who were in this group, and for the students who got Booklet B as a pretest the mean of all students who were in this group was also used for standardizing the scores. This same procedure was repeated for the scores on the posttest.

RESULTS

Descriptive Statistics

Before the intervention, students of the experimental and control group turned out to be quite comparable with respect to their ability to solve context-based tasks, as evidenced by their almost identical mean scores on the pretest (see **Table 1**). On average the students scored low on the test. The mean scores were around 7 while the maximum score was 18 or 19, depending on the booklet. The gains in the mean scores after the intervention were also low. Yet, in both conditions, the students scored generally higher on the posttest than on the pretest.

Effect of Opportunity-to-Learn on Students' Success Rate

To investigate the effect of the intervention program we carried out an ANCOVA with posttest score as dependent variable, intervention as independent variable, and pretest score as covariate. As the total scores of the two booklets differed, from here on standardized scores are reported. The standardized scores on the posttest were significantly higher for the students in the experimental group (M_{exp} = 0.11, SD_{exp} = 1.05) than in the control group (M_{con} = -0.10, SD_{con} = 0.93) as such showing a small effect of the intervention ($F(1, 274) = 4.092$, $p = .044$, $\eta_p^2 = .015$). There was also a significant main effect for the school the students were in ($F(5, 274) = 32.516$, $p < .001$, $\eta_p^2 = .372$). In **Figure 3** the different gain scores (i.e. standardized posttest score minus standardized pretest score) for the students in the control and experimental condition in the six schools are displayed.

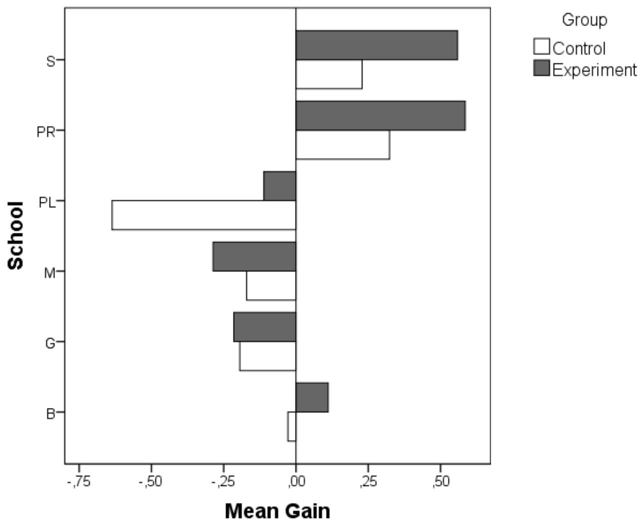


Figure 3. Mean gain scores for students in the control (white) and experimental (gray) condition in the six schools

Table 2. The number and types of error made by the students in the pretest and the posttest

Types of error	Sub-types of error	Group	Number of errors		Percentage of Change
			Pre-test	Post-test	
Comprehension	Errors in understanding instruction	Exp.	68	51	- 25% [a]
		Control	84	93	11%
	Errors in understanding a keyword	Exp.	22	36	64%
		Control	14	39	179%
	Errors in selecting information	Exp.	125	86	- 31%
		Control	125	127	2%
	Total	Exp.	215	173	- 20%
		Control	223	259	16%
Transformation	Procedural tendency	Exp.	7	15	114%
		Control	2	9	350%
	Taking too much account of context	Exp.	11	8	- 27%
		Control	20	10	- 50%
	Wrong mathematical procedure	Exp.	487	376	- 23%
		Control	582	442	- 24%
	Treating a graph as a picture	Exp.	68	59	- 13%
		Control	69	87	26%
	Total	Exp.	573	458	- 20%
		Control	673	548	- 19%
Mathematical processing[b]		Exp.	195	165	- 15%
		Control	230	239	4%
Encoding[c]		Exp.	32	33	3%
		Control	58	53	- 9%
Total[d]		Exp.	1015	829	- 18%
		Control	1184	1099	- 7%
	Overall total errors[d]		2199	1928	

[a] A negative value means a decrease
[b] The sub-types of mathematical processing error depend on the mathematics topics addressed in the test items
[c] No sub-type for encoding error
[d] Because of multiple coding, the number of errors was greater than the number of incorrect responses.

Effect of Opportunity-to-Learn on Students' Number and Types of Errors

In total, we found 1942 incorrect responses in the pretest and 1705 incorrect responses in the posttest. In the experimental group students gave fewer incorrect responses (pretest: 892, posttest: 744) than in the control group (pretest: 1050, posttest: 961). With Fisher's exact test we found that this difference in number of incorrect responses between the two groups was not significant (p = .171). An error analysis of the incorrect responses revealed that the students in the pretest made 2199 errors (1015 in the experimental group and 1184 in the control group) and in the posttest 1928 errors (829 in the experimental group and 1099 in the control group) (**Table 2**).

The overall decline in the number of errors was in the experimental group about ten percent points larger than in the control group. This difference was found to be significant (χ^2 (1, n = 4127) = 4.149, p = .042, r = .032). The type of error students in each condition made most frequently, on both the pretest and the posttest, was the error of selecting the wrong mathematical procedure (>20% of all errors), a sub-type of Transformation errors. In choosing the mathematical procedure students made many more errors than in carrying out the calculation (<12% of all errors were mathematical processing errors).

A further analysis with respect to the types of errors showed that in the experimental group the number of comprehension errors decreased by 20%, whereas in the control group the occurrence of these errors increased by 16%. However, the numbers of students for which we found these changes were rather small. In the experimental group, 103 students made comprehension errors in the pretest while in the posttest this was done by 95 students. In the control group these numbers were 113 and 119 respectively. Furthermore, the number of errors in selecting information reduced with 31% in the experimental group, while in the control group this number increased with 2%.

This apparent positive relation between the intervention and the reduction of this type of error was supported by drawings in the posttest work of the students in the experimental group, which contained clear signs of identifying and selecting relevant information for the mathematical processes that were needed to solve the problem. This was, for example, the case in the Skateboard task (**Figure 4**) in which students were asked to calculate the minimum and the maximum price for self-assembled skateboards.

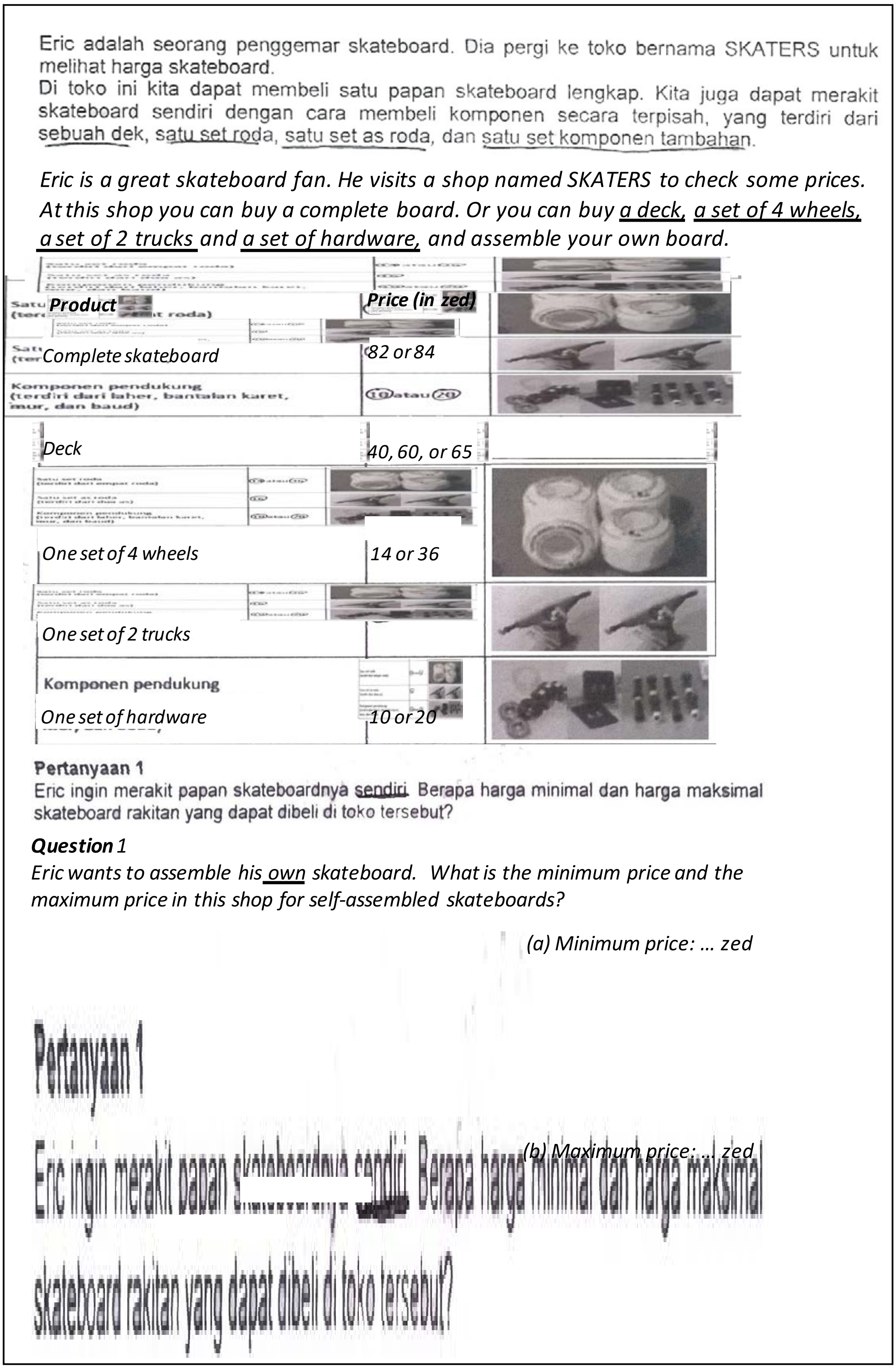
This task provided a price list that included irrelevant information, i.e. the prices of complete skateboards. In the posttest we found traces of the metacognitive prompt 'underlining information'. In the stem of the task, students in the experimental group underlined the names of components that are needed to assemble a skateboard (e.g., see **Figure 4**). Underlining information indicates the students' awareness of and actions for identifying the keywords of the task.

Regarding the transformation errors, we found that in general there was no difference between the experimental group and the control group. In both groups there was on average a decrease of 20% in the number of errors. The changes in the numbers of students for which we found these transformation errors were also rather small. In the experimental group 143 students made transformation errors in the pretest while in the posttest this was done by 139 students. In the control group these numbers were 153 and 145 respectively. At item level, however, the change in errors differed between the two groups. For tasks addressing an interpretation of a graph – i.e. the topic that was taught during the intervention – we found that for the students in the experimental group, in contrast with the students in the control group, the number of transformation errors in the posttest was smaller than the number of errors in the pretest. For example, this was the case in the Speed of Racing Car task (see **Figure 5**) in which students were asked to choose from five alternative race tracks along which a racing car could have produced the speed graph that was previously shown.

A usual transformation error for this task was treating a graph as a literal picture of a situation. Students tended to choose Track E because it resembled the shape of the speed graph. For this particular task, the occurrence of this error in the experimental group decreased by 17% in the posttest, whereas in the control group the number of errors did not change. This result can be explained by the fact that the intervention program included tasks on interpreting graphs in real-world situations and did not provide explicit procedures for solving these tasks. In contrast, the regular program used tasks that were mostly tasks without context by which the students were immediately directed to the procedure of using linear equations to create tables and use the numbers in the tables to draw graphs as a final result (with x horizontally and y vertically). Students in the experimental group had more opportunities to learn to interpret and reason with graphs in context-based tasks.

Similar to comprehension errors, for mathematical processing errors the number of errors decreased in the experimental group while it increased in the control group. When looking at the number of students who made these errors, in both groups a similar decrease is visible (experimental group, pretest: 111 students, posttest: 96 students; control group, pretest: 112 students, posttest: 101 students). When taking a closer look at the sub-types of mathematical processing errors, we found the largest difference between the two groups in arithmetical errors: in the posttest the number of arithmetical errors in the experimental group barely changed, whereas in the control group it increased by 70%. This seems to be quite surprising, since the regular program had a strong focus on processing skills. An explanation could be that processing skills were strongly connected to specific tasks. Consequently, when students are not trained to use them in unfamiliar situations these skills might become less useful.

Lastly, regarding encoding errors, we found only a small number of errors and in both groups the number of these errors hardly changed between pretest and posttest, which was also reflected in the number of students that made these errors (experimental group, pretest: 32 students, posttest: 33 students; control group, pretest: 58 students, posttest: 53 students).



Eric adalah seorang penggemar skateboard. Dia pergi ke toko bernama SKATERS untuk melihat harga skateboard.
Di toko ini kita dapat membeli satu papan skateboard lengkap. Kita juga dapat merakit skateboard sendiri dengan cara membeli komponen secara terpisah, yang terdiri dari sebuah dek, satu set roda, satu set as roda, dan satu set komponen tambahan.

Eric is a great skateboard fan. He visits a shop named SKATERS to check some prices. At this shop you can buy a complete board. Or you can buy a deck, a set of 4 wheels, a set of 2 trucks and a set of hardware, and assemble your own board.

Product	**Price (in zed)**	
Complete skateboard	*82 or 84*	
Komponen pendukung (terdiri dari laher, bantalan karet, mur, dan baud)	10 atau 20	
Deck	*40, 60, or 65*	
One set of 4 wheels	*14 or 36*	
One set of 2 trucks		
Komponen pendukung *One set of hardware*	*10 or 20*	

Pertanyaan 1
Eric ingin merakit papan skateboardnya sendiri. Berapa harga minimal dan harga maksimal skateboard rakitan yang dapat dibeli di toko tersebut?

***Question** 1*
Eric wants to assemble his own skateboard. What is the minimum price and the maximum price in this shop for self-assembled skateboards?

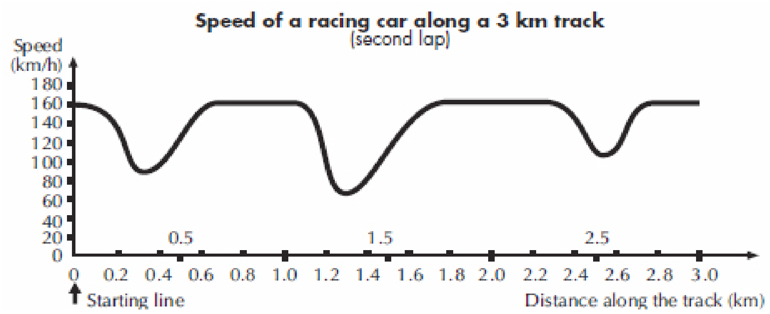
(a) Minimum price: … zed

Pertanyaan 1
Eric ingin merakit papan skateboardnya sendiri. Berapa harga minimal dan harga maksimal skateboard rakitan yang dapat dibeli di toko tersebut?

(b) Maximum price: … zed

Figure 4. Skateboard task with student work

This graph shows how the speed of a racing car varies along a flat track 3 kilometre track during its second lap.



Question

Here are pictures of five tracks. Along which one of these tracks was the car driven to produce the speed graph shown earlier?

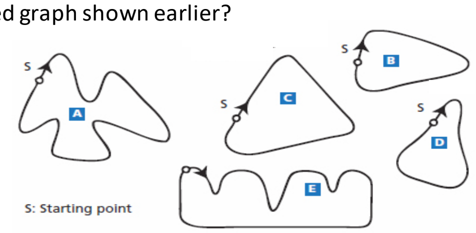


Figure 5. Speed of Racing Car task

DISCUSSION AND CONCLUSION

The purpose of the field experiment in this study was to investigate whether Indonesian students' ability to solve context-based tasks could be improved by offering them an intervention program in which tasks were offered with relevant and essential contexts, missing or superfluous information, and without explicitly given mathematical procedures. In addition, teachers were asked to apply a consultative teaching approach by giving students metacognitive prompts and questions for discussion to promote reflection. Contrary to our expectations, we did not find strong evidence for the effect of this intervention on the students' achievement in solving context-based tasks in terms of their improvement in test score. Students in the experimental group made more progress than students in the control group. This difference in improvement between the two groups was significant but the effect of the intervention was rather small. Furthermore, we found a significant difference in the decline of errors in advantage of the experimental group. This was specifically the case for comprehension errors; within this category there was a substantial decrease in the number of errors in selecting information. Taken together, the findings about students' errors signify the potential benefit of the metacognitive prompts as implemented in the tasks and in the consultative teaching strategies. Asking students to circle only relevant information and using their own words to explain a task, seem to be effective in guiding students to looking thoroughly at the information provided in a task.

Yet, these promising results did not occur for all types of errors. For the errors in transforming a real-world problem into a mathematical problem no intervention effect was found. Here, in general, there was no difference between the experimental and the control group. A possible explanation for this finding might be that during the transformation stage teachers often used a directive teaching approach in which they told students which mathematics concepts or strategies are related to a particular task. Also, often they did not let their students explore and discuss different strategies. In this respect, the teachers did not implement the intervention fully as intended. Instead of consultative teaching their teaching clearly showed directive characteristics, which, as is shown by other studies (Maulana, Opdenakker, Den Brok, & Bosker, 2012; World Bank, 2010), is the main way of teaching in Indonesia. Because of this tradition the teachers involved in our study may have had a lack of experience with metacognitive instruction and therefore, as is also emphasized by Kramarski et al. (2002), had difficulties in using consultative teaching. This came especially to the fore when the teachers were faced with unexpected situations. Evidently, a professional development of two meetings of 90 to 120 minutes and a teacher guide were not enough to fully apply a consultative teaching approach.

Another factor that might not have been helpful for lowering the number of transformation errors was the focus of the intervention on one specific mathematical topic: graphical representations of linear equations. A consequence of this was that during the intervention, by design, it might have been clear for the students into what mathematical problem the real-world problems had to be transformed – namely, into a linear equation – and therefore they might not have had enough occasions to practice identifying the needed procedure. It might have been better to also include different topics in our intervention, which is in agreement with the situatedness of learning, and with Howson (2013), who pointed out that including context-based tasks in a particular chapter of a textbook will discourage students from thinking about the required procedure before they solve the tasks.

A further shortcoming of our experiment was that due to limited financial resources – and thus restricted research time – we could only include a small number of schools. This made it impossible to compose two groups of matched schools followed by random allocation to the experimental or control group. Instead, in our design, all schools had both conditions and, consequently, the danger of a spillover effect. Moreover, because most schools had only one teacher who taught Grade 8, we had to accept that experimental students and control students were taught by the same teacher. This is not an ideal situation, because what the teachers did in the experimental group might have had an influence on their teaching practices in the control group. Although, based on our observations, it is not likely that this was the case, it is known from research (Kunter et al., 2013) that teachers' professional knowledge affects students' performance. So, it might be possible that the knowledge about solving context-based tasks that the teachers learned in their role as teacher of an experimental class may have spilled over in the control class. An additional weakness in our design was the lack of a measure of students' general mathematics achievement level. The reason for this was that the different districts where the schools were located each had their own tests for Grade 8, which made the scores of the students in different schools not comparable.

Although our study had some flaws that lowered the robustness of our findings – which is almost inevitable when doing research in the real setting of educational practice – it did show us the necessary ingredients of the teaching that can foster students' ability in solving context-based tasks requiring mathematical modeling (e.g. transforming the contextual problem to a mathematical problem). The OTL that is needed for this, implies, in the first place, a classroom practice in which students are offered tasks in which they can really apply mathematics. Secondly, to make this happen, it is necessary that teachers have knowledge of the characteristics of context-based tasks, including their cognitive demands and the types of contexts. A third requirement is that teachers are offered possibilities for developing the ability to apply a consultative teaching approach in which they can give students metacognitive prompts and can elicit discussion and reflection. These three ingredients are not new, in the sense that they have already been demonstrated in earlier research, but the experiences from our study have shown how this can work in practice in a project where they are all combined. However, as these ingredients of OTL were all blended in the intervention, our study does not provide us with the opportunity to differentiate between the contribution of each of them.

REFERENCES

Ainley, J., Pratt, D., & Hansen, A. (2006). Connecting engagement and focus in pedagogic task design. *British Educational Research Journal, 32*(1), 23-38. https://doi.org/10.1080/01411920500401971

Antonius, S., Haines, C., Jensen, T. H., Niss, M., & Burkhardt, H. (2007). Classroom activities and the teacher. In W. Blum, P. L. Galbraith, H.-W. Henn & M. Niss (Eds.), *Modelling and Applications in Mathematics Education: The 14th ICMI Study* (pp. 295–308). New York: Springer. https://doi.org/10.1007/978-0-387-29822-1_31

Bernardo, A. B. I. (1999). Overcoming obstacles to understanding and solving word problems in mathematics. *Educational Psychology, 19*(2), 149–163. https://doi.org/10.1080/0144341990190203

Blum, W. (1993). Mathematical modelling in mathematics education and instruction In T. Breiteig, I. Huntley, & G. Kaiser-Messmer (Eds.), *Teaching and learning mathematics in context* (pp. 3-14). Chichester, UK: Ellis Horwood Limited.

Blum, W. (2011). Can modelling be taught and learnt? Some answers from empirical research. In G. Kaiser, W. Blum, R. B. Borromeo Ferri & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling* (pp. 15–30). New York: Springer. https://doi.org/10.1007/978-94-007-0910-2_3

Blum, W. (2015). Quality Teaching of Mathematical Modelling: What Do We Know, What Can We Do? In: S.J. Cho (Ed.), *The Proceedings of the 12th International Congress on Mathematical Education - Intellectual and Attitudinal Challenges* (pp 73-96). New York: Springer.

Blum, W., & Borromeo Ferri, R. B. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and Application, 1*(1), 45–58.

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects - state, trends and issues in mathematics instruction. *Educational Studies in Mathematics, 22*, 37–68. https://doi.org/10.1007/BF00302716

Boaler, J. (1994). When girls prefer football to fashion? An analysis of female under achievement in relation to realistic mathematics contexts. *British Educational Research Journal, 20*(5), 551–564. https://doi.org/10.1080/0141192940200504

Borasi, R. (1986). On the nature of problems. *Educational Studies in Mathematics 17*(2), 125-141. https://doi.org/10.1007/BF00311517

Brewer, D. J., & Stasz, C. (1996). *Enhancing opportunity to learn measures in NCES data*. Santa Monica, CA: RAND.

Carroll, J. (1963). A model of school learning. *Teachers College Record, 64*, 723-733.

Charalambous, C. Y., Delaney, S., Hsu, H-Y., & Mesa, V. (2010). A comparative analysis of the addition and subtraction of fractions in textbooks from three countries. *Mathematical Thinking and Learning, 12*(2), 117–151. https://doi.org/10.1080/10986060903460070

Clements, M. A. (1980). Analyzing children's errors on written mathematical task. *Educational Studies in Mathematics, 11*(1), 1–21. https://doi.org/10.1007/BF00369157

De Lange, J. (1987). *Mathematics, Insight and Meaning*. Utrecht, the Netherlands: OW&OC.

De Lange, J. (2015). There is, probably, no need for this presentation. In A. Watson & M. Ohtani (Eds.), *Task design in mathematics education – ICMI Study 22* (pp. 287-308). Cham, Heidelberg, New York, Dordrecht, London: Springer. https://doi.org/10.1007/978-3-319-09629-2_10

Doerr, H. (2007). What knowledge do teachers need for teaching mathematics through applications and modelling? In W. Blum et al. (Eds.), *Modelling and applications in mathematics education* (pp. 69–78). New York: Springer. https://doi.org/10.1007/978-0-387-29822-1_5

Eurydice (2011). *Mathematics education in Europe: Common challenges and national policies*. Brussels: Education, Audiovisual and Culture Executive Agency.

Forman, S. L., & Steen, L. A. (2001). Why math? Applications in science, engineering, and technological programs. *Research Brief, American Association of Community Colleges.*

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures.* Dordrecht, the Netherlands: Reidel.

Goldman, S. R. (1989). Strategy instruction in mathematics. *Learning Disability Quarterly, 12*(1), 43–55. https://doi.org/10.2307/1510251

Graumann, G. (2011). Mathematics for problem in the everyday world. In J. Maasz & J. O'Donoghue (Eds.), *Real-world problems for secondary school mathematics students: Case studies* (pp. 113–122). Rotterdam: Sense Publishers. https://doi.org/10.1155/2011/526167

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics, 39*, 111-129.

Greer, B. (1997). Modeling reality in mathematics classrooms: The case of word problems. *Learning and Instruction, 7*(4), 293–307. https://doi.org/10.1016/S0959-4752(97)00006-6

Grouws, D. A., & Cebulla, K. J. (2000). *Improving student achievement in mathematics*. Brussels: International Academy of Education.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte: Information Age Publishing.

Howson, G. (2013). The development of mathematics textbooks: historical reflections from personal perspective. *ZDM – The International Journal on Mathematics Education, 45*(5), 647–658. https://doi.org/10.1007/s11858-013-0511-9

Husén, T. (Ed). (1967). *International study of achievement in mathematics: A comparison of twelve countries (Vol. II)*. New York: John Wiley & Sons.

Ikeda, T. (2007). Possibilities for, and obstacles to teaching applications and modelling in the lower secondary levels. In W. Blum et al. (Eds.), *Modelling and applications in mathematics education* (pp. 457–462). New York: Springer. https://doi.org/10.1007/978-0-387-29822-1_51

Karbalei, A., & Amoli, F. A. (2011). The effect of paraphrasing strategy training on the reading comprehension of college students at the undergraduate level. *Asian EFL Journal, 13*(3), 229–344.

Kletzien, S. B. (2009). Paraphrasing: An effective comprehension strategy. *The Reading Teacher, 63*(1), 73–77. https://doi.org/10.1598/RT.63.1.7

Klymchuk, S., Zverkova, T., Gruenwald, N., & Sauerbier, G. (2010). University students' difficulties in solving application problems in calculus: Student perspectives. *Mathematics Education Research Journal, 22*(1), 81–91. https://doi.org/10.1007/BF03217567

Kramarski, B., Mevarech, Z. R., & Arami, M. (2002). The effects of metacognitive instruction on solving mathematical authentic tasks. *Educational Studies in Mathematics, 49*(2), 225–250. https://doi.org/10.1023/A:1016282811724

Kunter, M., Baumert, J., Voss, T., Klusmann, U., Richter, D., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology, 105*(3), 805–820. https://doi.org/10.1037/a0032583

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Maass, K. (2007). Modelling tasks for low achieving students – first results of an empirical study. In D. Pitta-Pantazi & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education CERME 5* (pp. 2120–2129). Larnaca, Cyprus.

Maass, K. (2010). Classification scheme for modelling tasks. *Journal fur Mathematik-Didaktik, 31*(2), 285–311. https://doi.org/10.1007/s13138-010-0010-2

Mason, J., Burton, L., & Stacey, K. (1982 / 2010). *Thinking mathematically*. Dorchester, England: Pearson Education Limited. Second Edition.

Maulana, R., Opdenakker, MC., Den Brok, P., & Bosker, R. J. (2012). Teacher-student interpersonal behavior in secondary mathematics classes in Indonesia. *International Journal of Science and Mathematics Education, 10*(1), 21–47. https://doi.org/10.1007/s10763-011-9276-1

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis, 17*(3), 305-322. https://doi.org/10.2307/1164509

Montague, M. (2007). Self-regulation and mathematics instruction. *Learning Disabilities Research and Practice, 22*(1), 75–83. https://doi.org/10.1111/j.1540-5826.2007.00232.x

Montague, M. (2008). Self-regulation strategies to improve mathematical problem solving for students with learning disabilities. *Learning Disability Quarterly, 31*(1), 37–44.

Montague, M., Warger, C., & Morgan, T. H. (2000). Solve it! Strategy instruction to improve mathematical problem solving. *Learning Disabilities Research and Practices, 15*(2), 110–116. https://doi.org/10.1207/SLDRP1502_7

NCTM (2000). *Principles and Standard for School Mathematics*. Reston: Author.

OECD (2003). *The PISA 2003 Assessment Framework - Mathematics, Reading, Science, and Problem Solving Knowledge and Skills*. Paris: Author.

OECD (2009a). *Learning Mathematics for Life. A View Perspective from PISA*. Paris: Author.

OECD (2009b). *Take the Test. Sample Questions from OECD's PISA Assessments*. Paris: Author.

OECD (2010). *PISA 2009 Results: What students know and can do. Student performance in reading, mathematics, and science* (Vol. I). Paris: Author.

OECD (2013). *PISA 2012 Results: What students know and can do. Student performance in mathematics, reading and science*. Paris: Author.

OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. PISA, OECD Publishing, Paris. https://doi.org/10.1787/9789264266490-en

Pólya, G. (1945). *How to solve it*. Princeton, NJ: Princeton University Press.

Pusat Kurikulum (2003). *Kurikulum 2004: Standar kompetensi mata pelajaran matematika Sekolah Menengah Pertama dan Madrasah Tsanawiyah*. Jakarta: Departemen Pendidikan nasional.

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-5786-5

Schoenfeld, A. H. (1985). Mathematical problem solving. Orlando, FL: Academic Press Inc. https://doi.org/10.1016/B978-0-12-628870-4.50009-8

Tomlinson, M. (2004). *14-19 Curriculum and Qualifications Reform. Final report of the Working Group on 14-19 Reform*. London: Department for Education and Skills.

Tornroos, J. (2005). Mathematics textbooks, opportunity to learn and student achievement. *Studies in Educational Evaluation, 31*(4), 315–327. https://doi.org/10.1016/j.stueduc.2005.11.005

Valverde. G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book. Using TIMSS to investigate the translation of policy into practice through the world of textbooks.* Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-007-0844-0

Van den Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*. Utrecht, the Netherlands: CD-β Press/Freudenthal Institute, Utrecht University.

Van den Heuvel-Panhuizen, M. (2005). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics, 25*(2), 2-9 + 23.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Lisse: Swets & Zeitlinger.

Wijaya, A., Van den Heuvel-Panhuizen, M., & Doorman, M. (2015a). Opportunity-to-learn context-based tasks provided by mathematics textbooks. *Educational Studies in Mathematics, 89*(1), 41–65. https://doi.org/10.1007/s10649-015-9595-1

Wijaya, A., Van den Heuvel-Panhuizen, M., & Doorman, M. (2015b). Teachers' teaching practices and beliefs regarding context-based tasks and their relation with students' difficulties in solving these tasks. *Mathematics Education Research Journal, 27*(4), 637–662. https://doi.org/10.1007/s13394-015-0157-8

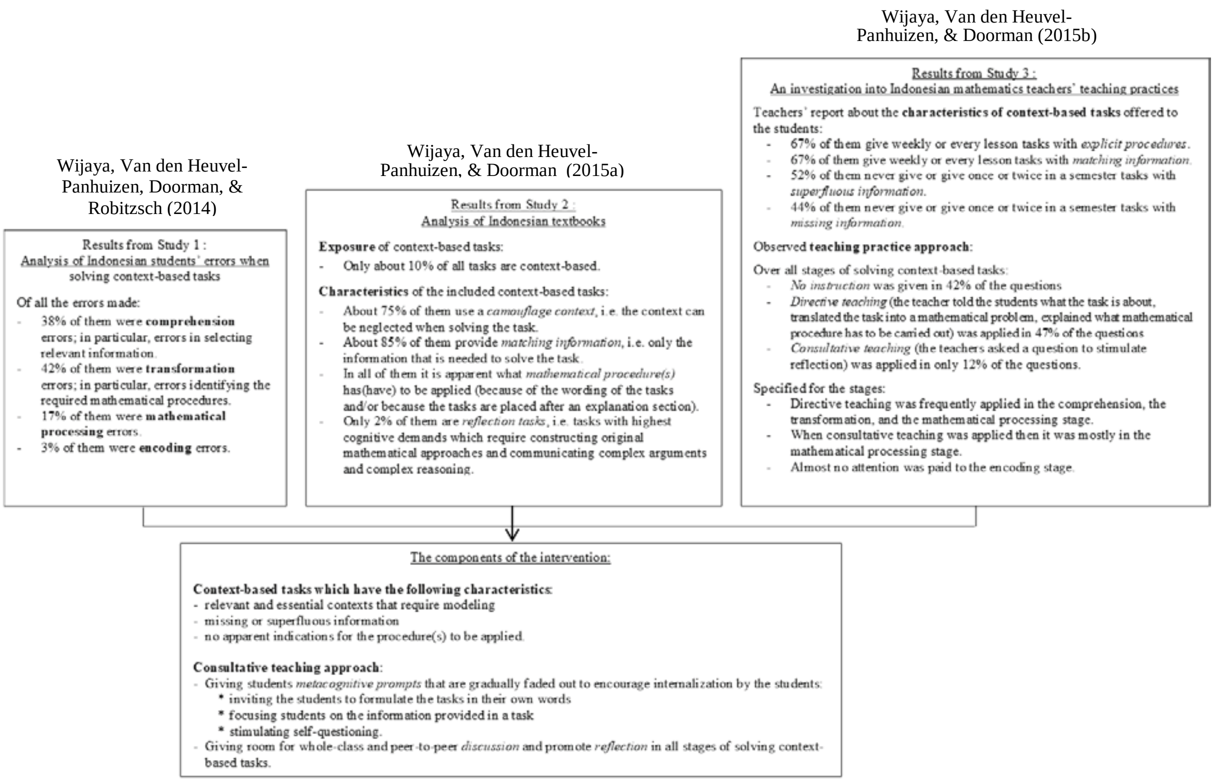
Wijaya, A., Van den Heuvel-Panhuizen, M., Doorman, M., & Robitzsch, A. (2014). Difficulties in solving context-based PISA mathematics tasks: An analysis of students' errors. *The Mathematics Enthusiast, 11*(3), 555–584.

World Bank (2010). *Inside Indonesia's mathematics classrooms: A TIMSS Video Study of teaching practices and student achievement*. Jakarta: Author.

Xin, Y. P. (2007). Word problem solving tasks in textbooks and their relation to student performance. *Journal of Educational Research, 100*(6), 347–359. https://doi.org/10.3200/JOER.100.6.347-360

APPENDIX 1

Intervention Program Informed by the Findings from our Previous Studies



APPENDIX 2

Outline of the Intervention Program

Lesson	Mathematical sub-topic	Goal related to mathematical content	Goal related to solving context-based tasks	Context-based task	Bare mathematics task
1	Sketching graph of straight line	- Making table of ordered pairs of linear equation - Sketching graph of equation on Cartesian coordinate by plotting ordered pairs	- Identifying and completing missing information. - Selecting relevant information. - Paraphrasing the tasks. - Identifying appropriate mathematics procedure.	- *Taxi task*: Comparing taxi tariffs - *Internet task*: Comparing internet fees	1
2	Gradient (1)	- Understanding definition of gradient - Determining gradient of a straight line through two points	- Identifying and completing missing information. - Selecting relevant information. -Identifying appropriate mathematics procedure.	- *Experiment task*: Determining the fastest speed of bike simulation - *Price task*: Determining the increase rate of oil price	1
3	Gradient (2)	- Determining gradient of parallel lines - Determining gradient of perpendicular lines	- Identifying and completing missing information. - Identifying appropriate mathematics procedure. - Interpreting the solution in terms of the problem situation.	- *Bus task*: Comparing bus speed - *Water pump task (1)*: Visualizing the rates of filling water tanks	1
4	Equation of straight line (1)	Determining equation and sketching graph of: - A line with a gradient of *m* and through point (x_1,y_1) - A line through points (x_1,y_1) and (x_2,y_2)	- Selecting relevant information	- *Olympic task*: Predicting the number of future participants - *Hospital task*: Estimating drug dosage	-
5	Equation of straight line (2)	Determining equation and sketching graph of: - A line through point (x_1,y_1) and parallel to another line - A line through point (x_1,y_1) and perpendicular to another line	- Selecting relevant information - Interpreting the solution in terms of the problem situation	- *Water pump task (2)*: Estimating filling rates	2

APPENDIX 3

Coding Scheme for Error Types when Solving Context-based Mathematics Tasks

Error type	Sub-type	Explanation
Comprehension	Misunderstanding the instruction	Student incorrectly interprets what (s)he is asked to do.
	Misunderstanding a keyword	Student misunderstands a keyword, which is usually a mathematical term.
	Error in selecting information	Student is unable to distinguish between relevant and irrelevant information (e.g. using all information provided in a task or neglecting relevant information) or is unable to gather required information which is not provided in the task.
Transformation	Procedural tendency	Student tends to directly use a mathematical procedure such as formula or algorithm without analyzing whether or not it is needed.
	Taking too much account of the context	Student's answer only refers to the context or real-world situation without taking the perspective of the mathematics.
	Wrong mathematical procedures/concepts	Student uses mathematical procedures or concepts which are not relevant to the tasks.
	Treating a graph as a picture	Student treats a graph as a literal picture of a situation. (S)he interprets and focuses on the shape of the graph, instead of on the properties of the graph.
Mathematical Processing	Algebraic error	Error in solving algebraic expression or function.
	Arithmetical error	Error in calculation.
	Error in mathematical interpretation of graph:	
	- Point-interval confusion	Student mistakenly focuses on a single point rather than on an interval.
	- Slope-height confusion	Student does not use the slope of the graph but only focuses on the vertical distance.
	Measurement error	Student cannot convert between standard units (e.g. from m/minute to km/h) or from a non-standard unit to a standard unit (e.g. from step/minute to m/minute).
	Improper use of scale	Student cannot select and use the scale of a map properly.
	Unfinished answer	Student uses a correct procedure, but (s)he does not finish it.
Encoding		Student is unable to correctly interpret and validate the mathematical solution in terms of the real-world problem. This error is reflected by an impossible or not realistic answer.