

Article

Insights Chinese Primary Mathematics Teachers Gained into their Students' Learning from Using Classroom Assessment Techniques

Xiaoyan Zhao ^{1,2,*} , Marja van den Heuvel-Panhuizen ^{1,3,4,*}  and Michiel Veldhuis ^{3,5,*} ¹ Freudenthal Institute, Faculty of Science, Utrecht University, 3584 CC Utrecht, The Netherlands² School of Teacher Education, Nanjing Normal University, Nanjing 210023, China³ Freudenthal Group, Faculty of Social and Behavioural Sciences, Utrecht University, 3584 CS Utrecht, The Netherlands⁴ Faculty of Education, Art and Culture, Nord University, 8049 Bodø, Norway⁵ iPabo University of Applied Sciences, 1061 AD Amsterdam, The Netherlands

* Correspondence: 69005@nynu.edu.cn (X.Z.); m.vandenheuvel-panhuizen@uu.nl (M.v.d.H.-P.); m.veldhuis@uu.nl (M.V.)

Received: 13 March 2019; Accepted: 11 June 2019; Published: 18 June 2019



Abstract: In this study, we explored the insights that Chinese primary mathematics teachers gained into their students' mathematical understanding from using classroom assessment techniques (CATs). CATs are short teacher-initiated targeted assessment activities proximate to the textbook, which teachers can use in their daily practice to make informed instructional decisions. Twenty-five third-grade teachers participated in a two-week program of implementing eight CATs focusing on the multiplication of two-digit numbers, and filled in feedback forms after using the CATs. When their responses described specific information about their students, emphasized the novelty of the gained information, or referred to a fitting instructional adaptation, and these reactions went together with references to the mathematics content of the CATs, the teachers' responses were considered as evidence of gained insights into their students' mathematics understanding. This was the case for three-quarters of the teachers, but the number of gained insights differed. Five teachers gained insights from five or more CATs, while 14 teachers did so only from three or fewer CATs, and six teachers showed no clear evidence of new insights at all. Despite the differences in levels of gained insights, all the teachers paid more attention to descriptions of students' performance than to possible instructional adaptations.

Keywords: classroom assessment techniques; mathematics education; multiplication; primary school; China

1. Introduction

1.1. Assessment in the Hands of Teachers

Any instructional decision making—and thus any form of teaching—requires in one way or another information about students' learning [1]. The more reliable and valid this information is, the better teachers can find a foothold for these decisions. For generating such information, many approaches are possible, ranging from standardized externally developed tests to teacher-made assignments. Contrary to the low reliability that has been attributed to teachers' judgments of students' performance in the past (see, for example, Parkes [2]), nowadays, teacher assessment that is aimed at gaining insights into their students' progress is highly valued and seen as crucial for adapting the teaching to the students' abilities and needs. Teachers are also seen as being in a good position for collecting

information about their students' learning [3]. Teacher-led assessment activities that are interwoven with instruction and fully integrated in the teachers' daily teaching practice, such as questioning, observing students, and giving quizzes or teacher-made written assignments, can provide insights about students' thinking and about what productive and actionable next instructional steps might be taken [4]. When the assessment focuses on figuring out what students know, or what difficulties students have, for the purpose of making decisions about further instruction, it is considered as formative assessment. Formative assessment in which the teacher has the lead is often referred to as classroom assessment [5–10].

What information can be collected through classroom assessment depends largely on what assessment activities are conducted. Helpful assessment activities are those which offer teachers a window into the students' thinking to uncover their mathematical conceptions and skills [10]. Therefore, much attention has been paid to gaining knowledge about how mathematics teachers can improve their assessment activities to acquire adequate information about their students' development (see, for example, Schoenfeld [11]). Research has shown that using various oral questioning strategies and written tasks, and then analyzing students' responses, offers mathematics teachers opportunities to reveal their students' understanding [11,12]. In particular, challenging students with open-ended problems enables teachers to diagnose students' understanding and reveal their methods of problem solving [12,13]. Other measures that make assessments by teachers more informative are using rubrics [14] or concept maps [15] as frameworks for analyzing students' responses. Both measures were found to assist teachers in identifying gaps in their students' understanding of the particular mathematical topics under investigation.

1.2. Assessment Techniques

Assessment techniques are assessment activities by which mathematics teachers can gauge what their students do and do not know, so that they can adjust their teaching to their students' needs. These assessment techniques can be characterized as short, feasible, and for teachers, often well-known activities, which are fully embedded in teachers' teaching practice [16]. Also, several other researchers and educators [17–21] have investigated such assessment techniques.

William et al. [16,20] investigated a large number of assessment techniques that are used to support primary and secondary teachers' formative assessment practice with the goal of making instructional decisions either for direct use or for later teaching. Not surprisingly, again, asking questions turned out to be very helpful. However, the fact that teachers distinguished different types of questions for different moments in a lesson was a new approach to questioning students. At the beginning of a teaching sequence, range-finding questions were used to find out students' previous knowledge (see, for example, "how many fractions can you find between $1/6$ and $1/7$?" [16] (p. 21)). During a lesson hinge point, questions were used to indicate the direction of the remainder of the lesson (see, for example, six polygons were shown, and the students were asked to indicate how many lines of symmetry each polygon has [20] (p. 103)). Finally, exit pass questions, which are asked before students were leaving the classroom, were meant to make decisions about the next lesson. Furthermore, to allow all the students to answer at the same time, William et al. suggested the use of ABCD cards and mini whiteboards [20]. Then, when a question was asked, all the students could show their answers by holding up a card or writing their answer on the whiteboard.

Similar assessment techniques were also discussed by Keeley and Tobey [18], who considered these techniques as useful to give insight into students' factual, conceptual, and procedural mathematical understanding for a broad range of mathematics teachers from Kindergarten to Grade 12. In Andersson and Palm's [17] study, it came to the fore that the primary mathematics teachers involved paid the most attention to those assessment techniques that helped them best to collect information about their students' knowledge and skills. A similar finding was revealed by Wylie and Lyon [21], who conducted research with mathematics and science teachers in high school and found that the most used assessment techniques were asking questions, organizing classroom discussions, and using written tasks.

Characteristic of the aforementioned studies on assessment techniques is that the techniques are general in nature. They can be applied to any subject and to any mathematical topic. When teachers are provided with such examples of assessment techniques, it can happen, as was found by James and McCormick [22] (p. 976), that some teachers understand the “spirit” behind the assessment techniques, and thus are able to adapt them to their teaching, but that others just catch the “letter” of them and carry them out in a ritualized and mechanistic way. The latter may be the result of providing teachers with assessment techniques that are not directly related to the content the teachers are teaching at that moment. To avoid this, and to have assessment techniques that can generate indications for further instruction, the techniques should be content-dependent.

A study in which this content-dependent approach was adopted is that of Phelan et al. [23]. The aim of their study was supporting teachers to assess students’ learning in pre-algebra. To find out what had to be assessed, an expert panel was organized to map algebra knowledge and its prerequisites. Such a map was used to design the questions and tasks that could provide teachers with the necessary information. This innovative content-dependent approach to assessment, which differs from providing teachers with general assessment guidelines, turned out to be rather successful, and apparently had a positive impact on students’ learning [23].

To make the assessment even closer to the teaching at hand, Veldhuis and Van den Heuvel-Panhuizen [19] took the textbook used by the teachers as a starting point. They designed brief and targeted activities, called classroom assessment techniques (CATs), that teachers could use in their daily practice to reveal information about students’ learning of a particular mathematical concept or skill. The ultimate goal of CATs is providing teachers with deep insights into students’ mathematical thinking to make adequate instructional decisions. This requires a skillful way of questioning [24,25], or, in the words of Heritage and Heritage [26] (p. 187), “questioning lies at the epicenter of formative assessment.” For the CATs, this implies that they were designed to serve as an eye-opener for teachers to acquire knowledge about their students’ learning that they did not have before. This goes beyond knowing whether students are able to flawlessly carry out particular calculation procedures. Instead, the CATs are intended to delve deeper and reveal whether and how students understand the underlying concepts of problems and see the relationships between problems, and to what extent they are flexible in solving problems. Therefore, rather than just repeating the tasks in the textbook, CATs present the content to be assessed from a different perspective and in an unfamiliar way. In addition to the content-dependency, what is innovative of the CATs is in particular that teachers are offered a new perspective for looking at students’ understanding. This makes CATs different from the usual ways of assessing students, but at the same time, these new activities are close to the known daily teaching practice. Moreover, to make the CATs manageable for the teachers, they have a format that supports teachers to gather the students’ information efficiently and makes the assessment feasible to carry out. The two main formats that Veldhuis and Van den Heuvel-Panhuizen [19] used for their CATs were red/green cards and worksheets. With the students responding to a question by holding up a red or a green card, the teacher can quickly gather information about the group as a whole. The worksheets, mostly containing a few problems on a specific mathematical concept or skill, are meant to provide teachers with more information on individual students’ strategy use.

1.3. A New Approach to Assessment in Mathematics Education in China

Mathematics education in China has a deeply-rooted examination culture [27]. External examinations at the school level and teacher-made end-of-chapter tests at the classroom level used to be the main aspect of teachers’ assessment activity [28]. In 2001, the Ministry of Education [29] formally launched a new approach to assessment aimed at improving teaching and learning. Since then, mathematics teachers are encouraged to get a comprehensive understanding of students’ learning by employing various approaches: for example, written tests, oral tests, open questions, activity reports, observations, interviews, exercises in and after class, and portfolios [30,31]. However, Cai and Wang [32] found that Chinese mathematics teachers in primary education put much more emphasis on

providing information to students than on getting information from them. Furthermore, when taking action to understand students' thinking, teachers are more likely to do so before lessons than during or after lessons [33,34]. Moreover, when teachers plan their teaching, textbooks serve as the main source, rather than findings from assessing their students' learning [32,35]. Also, the exercises in textbooks have an important role in the decisions that teachers make about assessment activities. Yet, such exercises may be more suitable for summative assessment than for classroom assessment [36]. According to Liu [36], this may lead to teachers focusing on assessing the result of learning, namely what basic knowledge and skills students have acquired, instead of assessing how students developed their mathematical thinking during the learning process. Furthermore, studies have revealed that only very limited attention has been paid to improving teachers' assessment practice to get more information about students' learning [37,38].

Taking into account the promising international findings about the use of assessment techniques, we explored whether this approach to assessment could assist Chinese primary mathematics teachers in their assessment practice. Specifically, as a sequel to the studies carried out in the Netherlands [19] in which classroom assessment techniques (CATs) for primary mathematics education were developed and teachers were supported in using CATs, we investigated the use of CATs in China. Six third-grade mathematics teachers of two primary schools in Nanjing, China, participated in a pilot study [39]. The focus of this pilot study was on assessing the topic of division, in particular three-digit numbers divided by a one-digit number. In line with the Dutch studies [19], the CATs in the Chinese pilot study were also based on a textbook analysis and formulated in such a way that they were not just a repetition of what is in the textbook. In this way, CATs might give teachers access to a deeper level of students' skills and understanding. It was found that teachers recognized that it can be very revealing to challenge students with questions that are not completely prepared by the textbook. Also, they appreciated the use of red/green cards for providing quick information. In general, teachers were positive about the CATs as a way to reveal their students' learning in an effective and efficient way.

1.4. The Present Study

Based on these experiences, we set up a study to investigate whether this positive finding holds for a larger group of teachers, and for a different mathematical topic. In particular, we wanted to explore the insights that Chinese primary school mathematics teachers may gain from carrying out CATs. Precisely, our focus was on whether the teachers, through using CATs, could acquire knowledge about their students' learning that they did not have before. The research question of the present study was: *What new insights can Chinese primary mathematics teachers gain about their students' understanding of mathematics from using CATs?*

Since we already had experiences with the Chinese mathematics curriculum in Grade 3 through the pilot study, we chose to do the present study in this grade, too. However, to extend our knowledge about the use of CATs in Chinese mathematics classrooms, we changed the topic of investigation. We stayed in the domain of number and operations, but instead of on division, the focus in the present study was on multiplication, in particular on what insights teachers can gain from using CATs to assess students' understanding of the multiplication of two-digit numbers.

2. Methods

In the study, Chinese teachers were asked to use a number of CATs in their regular teaching of multiplication during the first two weeks of the second semester of Grade 3. Teachers were informed about the CATs through a teacher guide and two researcher-led meetings. Data on how teachers used the CATs and what insights they got from them were gathered through feedback forms and a teacher-written final report

2.1. Participants

For practical reasons, we decided to set up the study in Nanjing. We contacted three local teaching research offices, which are responsible for inspecting the educational quality of the schools and providing professional development to primary school teachers in their administrative districts. One of these offices volunteered to participate. To include various schools in terms of the school's reputation, educational quality, and location, nine out of 40 primary schools in its district were selected by this local teaching research office. The Grade 3 mathematics teachers and their students of these nine schools took part in the study. Thus, our sample consisted of 25 teachers and their students in 25 classes. In all the classes, the same textbook series was used, namely the *Sujiaoban* (苏教版) textbook [40].

2.2. Multiplication of Two-Digit Numbers

For developing the CATs, we first investigated when and how the topic of multiplication of two-digit numbers was addressed in the *Sujiaoban* textbook. We found that this topic was dealt with in the first chapter of the book meant for the second semester of Grade 3. This chapter covers nine lessons taught over around two weeks, consisting of so-called new lessons and revision lessons. A new lesson mostly starts with a new type of problem, which is presented as a context problem, followed by the corresponding bare number problem. Then, examples are given of how to solve this problem type, and finally, exercises are offered to practice this. A revision lesson generally includes exercises for rehearsing and discussing what the students have learned in earlier lessons. The main content components addressed in this chapter include, among others, multiplication with multiples of 10, the structure of the multiplication algorithm, and the ratio aspect of multiplication.

Multiplication with multiples of 10 is presented in a new lesson and starts with a context problem in which Uncle Li is sending 10 boxes of bell peppers, with 12 peppers in each box. This context problem is followed by the corresponding bare number problem. The students need to find out how many peppers are sent in total. The textbook shows that one method of solving 12×10 is to make use of 12×1 . By seeing both multiplications with their answers, the students become acquainted with the strategy of using an analogous problem, that is, using a problem whose answer is known or easy to calculate to find the result of an analogous problem. After this, the textbook provides three sets of exercises of multiplications with multiples of 10: $16 \times 1 =$, $16 \times 10 =$; $70 \times 6 =$, $70 \times 60 =$; $5 \times 40 =$, $50 \times 40 =$.

In the next new lesson, the structure of the multiplication algorithm is introduced. Here, special attention is paid to how the result of the multiple-of-10 part of the multiplication is notated, namely without writing down a zero and leaving the one-position empty (Figure 1). This structural understanding is further supported by exercises in which the students are provided with an empty structure of the multiplication algorithm that they have to fill in (Figure 2). In addition, the students have to explain what they need to calculate in each step.

$\begin{array}{r} 24 \\ \times 12 \\ \hline 48 \end{array}$ <p>The amount in 2 boxes</p>	$\begin{array}{r} 24 \\ \times 12 \\ \hline 48 \\ 240 \end{array}$ <p>The amount in 10 boxes</p>	$\begin{array}{r} 24 \\ \times 12 \\ \hline 48 \\ 240 \\ \hline 288 \end{array}$ <p>The amount in 12 boxes</p>
--	--	--

Figure 1. Structure of the multiplication algorithm.

$$\begin{array}{r}
 22 \\
 \times 23 \\
 \hline
 \square\square \\
 \square\square \\
 \hline
 \square\square\square
 \end{array}
 \qquad
 \begin{array}{r}
 32 \\
 \times 31 \\
 \hline
 \square\square \\
 \square\square \\
 \hline
 \square\square\square
 \end{array}
 \qquad
 \begin{array}{r}
 43 \\
 \times 21 \\
 \hline
 \square\square \\
 \square\square \\
 \hline
 \square\square\square
 \end{array}$$

Figure 2. Exercises with empty multiplication structure.

In the subsequent new lesson, the students are prompted to further strengthen their understanding of the structure of the multiplication algorithm. To achieve this, the textbook offers only the start of the algorithm for 24×53 (Figure 3). The students have to complete the remaining steps of the algorithm. Right after this, the textbook provides a description in words of the steps to be taken when carrying out the multiplication algorithm of two-digit numbers. The students are told to first choose the digit in the ones place of the lower number to multiply the upper number, and then do the same for the digit in the tens place of the lower number. After this, for every calculated product, they have to write the last digit of the product in the same column as the digit chosen from the lower number. Finally, the students need to add the two products.

$$\begin{array}{r}
 24 \\
 \times 53 \\
 \hline
 72
 \end{array}$$

Figure 3. Multiplication algorithm for which the start is given.

The ratio aspect of multiplication is dealt with in revision lessons halfway and at the end of the chapter. In the first problem offered to the students, they have to find out how many pencils there are in 10, 20, 40, and 80 boxes when in one box, there are 10 pencils. The problem is presented in a ratio table (Figure 4). The first column shows that in five boxes, there are 50 pencils in total. The students have to fill in the remaining empty cells. In the end, they have to explain what they can discover from the ratio table. The focus in this problem is on the external ratio between the number of boxes and of pencils, or in other words, on the functional relationship between them. This is even clearer in the next ratio-table-like problem (Figure 5), in which the students are explicitly asked to multiply two given numbers. Also, this functional relationship is emphasized by the notation of the following accompanying exercises $15 \xrightarrow{\times 20} \square$ and $15 \xrightarrow{\times 20+20} \square$, in which the students are required to fill in the empty frames.

Number of boxes	5	10	20	40	80
Number of pencils	50				

Figure 4. Ratio table problem.

Number	5	10	20	40	80
Number	20	20	20	20	20
Product					

Figure 5. Multiplying two numbers.

2.3. CATs for Assessing Multiplication of Two-Digit Numbers

To provide teachers with a tool for getting insights in their students’ learning and actionable clues for their next steps in teaching, we developed eight CATs (see Appendix A): five CATs in the format of the red/green cards, and three in a worksheet format. Each CAT contains two similar tasks

for doing two assessments, if necessary. Here, exemplarily, we explain four CATs in detail. Three are meant for assessing multiplication with multiples of 10 (CAT-1), the structure of the multiplication algorithm (CAT-3), and the ratio aspect of multiplication (CAT-4). Finally, near the end of the chapter, when students have learned the multiplication algorithm for two-digit numbers, it is assessed whether the students' understanding goes beyond mechanically carrying out the algorithm (CAT-8). To show the possible ways of collecting information with the CATs, we chose two CATs of each format: CAT-1 and CAT-4, which had a red/green card format, and CAT-3 and CAT-8, which had the individual worksheet format. In addition, CAT-1, CAT-3/CAT-4, and CAT-8 were meant to be used at the beginning, in the middle, and by the end of teaching with this chapter, respectively. A further reason for discussing these CATs anticipates on our finding that the CATs differed in the degree to which they gave teachers insights. Choosing these CATs gave us the opportunity to provide a fair picture of what CATs can bring about. CAT-1 and CAT-4 were CATs that helped fewer teachers gain insights than CAT-3 and CAT-8.

2.3.1. CAT-1: Family Problems

Multiplications with multiples of 10 are often considered as rather easy problems. Solving 12×10 by thinking of the analogous problem 12×1 and adding a zero is not hard. However, understanding why this simple adding of a zero works is something else. To really grasp the content component of multiplication with multiples of 10, a deeper understanding of the 10-based number system is necessary. Just being able to put a zero at the end of a result, in the case of problems within the number range of two-digit numbers (e.g., using 70×6 to find 70×60), does not guarantee that the students comprehend this content component of multiplication. Therefore, using the exercises in the textbook in which the numbers are below 100 has limited value for assessing whether students truly understand multiplication with multiples of 10. Students have learned to add one zero in the chapter, and in the assessment based on these problems, they have to add one zero, too. Students can pass this test by carrying out mechanically what they have practiced. To learn more about students' understanding, we developed a CAT in which the scope went beyond the two-digit number range. If students understand the 10-based number system, they can use the analogy strategy also for a broader collection of problems.

CAT-1 (Figure 6) has the red/green card format and starts with the multiplication 97×8 , to which the answer is given. Then, several related multiplication problems follow, such as 970×8000 . These problems are at first sight not easy to solve by mental calculation. In the CAT, students are not asked to solve these problems, but only whether they think they are able to solve them. They show their answer by raising the green ("Yes") or the red card ("No"). By tallying the green and red cards, the teacher gets an immediate overview of students' responses. In this way, he/she observes whether students' understanding of multiplication with multiples of 10 goes beyond mechanically adding one zero and whether they see the analogy and think they can make use of it.

It is known that 97×8 equals 776.
 Do you think you can solve the following problems?
 (Yes- Green card ; No- Red card)
 a) 97×80 b) 97×800 c) 97×8000 d) 970×8000

Figure 6. CAT-1: Family problems.

2.3.2. CAT-3: Breaking down a Multiplication

Knowing how an algorithm is built up can help for using it. Therefore, in the chapter, much attention is paid to the structure of the multiplication algorithm. Students were explained how the results of multiplications with multiples of 10 are notated, they had to fill in an empty structure of the multiplication algorithm, and they were taught how to carry out the algorithm step-by-step. However, being able to write down the algorithm perfectly and even arriving at the correct answer

does not necessarily mean that students understand what they are doing and understand the structure of multiplications with two-digit numbers.

CAT-3 (Figure 7) has a worksheet format and is meant to give teachers an extra opportunity to assess whether their students can identify the components of a multiplication and how they understand what is behind the algorithm. In this CAT, the same numbers are used as in the textbook, namely 24×53 . However, now students have to unravel this multiplication instead of carrying it out. By using distributive and associative properties, this can lead to four sub-multiplications, namely 3×4 , 3×20 , 50×4 , and 50×20 , or in any other order. The teacher hands out the worksheet and checks students' responses after class and uses the gained information for decisions about further instruction in the next lessons.

24 × 53 means that you have to calculate

a) 3×4 and 3×20 and $__ \times 4$ and $__ \times 20$

b) $4 \times __$ and $4 \times __$ and $__ \times 3$ and 20×50

c) 20×50 and $20 \times __$ and 4×50 and $__ \times 3$

Figure 7. CAT-3: Breaking down a multiplication.

2.3.3. CAT-4: Completing the Ratio Table

The ratio table problems provided in the textbook, in which students just have to multiply the numbers in the top row by 10 or 20, may not really reveal how students understand the relationships between the numbers in the ratio table. The focus is on vertical reasoning and is based on knowing how many are in one unit. In fact, a ratio table is not necessary for solving these problems. It seems to be just a format to notate multiplication problems with multiples of 10. CAT-4 (Figure 8) was developed for assessing the ratio aspect of multiplication and has a broader operationalization of ratio and the use of the ratio table.

You are NOT allowed to calculate how many pencils there are in one box. Can you still find out how many pencils there are in the following boxes? (Yes- **Green card** ; No- **Red card**)

Number of boxes	6	11	12	13	17	18	22	23
Number of pencils	72	132						

Figure 8. CAT-4: Completing the ratio table.

This CAT has a red/green card format and is called 'Completing the ratio table'. It is meant to challenge students in their work with the ratio table and give teachers extensive information of students' understanding of the ratio aspect of multiplication. The difficulty for the students in this CAT is that the number of pencils in one box is not given. Moreover, they are not allowed to calculate this number. The numbers in the ratio table have been chosen in such a way that students are prompted to find other methods to fill in the empty cells. For example, if in six boxes there are 72 pencils, then you also know how many there are in 12 boxes. Similarly, if in six boxes there are 72 pencils and in 11 boxes there are 132, then you can also know directly how many pencils there are in 17 boxes. Reasoning and calculating similar to this means that the ratio table is not only used vertically, but also horizontally. To a certain degree, the textbook also gives opportunities to elicit this richer method of using the ratio table, when the students are asked to reflect on what they discovered in the ratio table. In particular, this is the case when problems such as $15 \xrightarrow{\times 20} \square$ are followed by $15 \xrightarrow{\times 20 + 20} \square$, and there is an opportunity to discuss that this is equal to 16×20 . This way of reasoning about the ratio table is in

any case explicitly promoted in CAT-4 and can provide teachers with extra information about their students' understanding of the ratio aspect of multiplication.

2.3.4. CAT-8: Solving Problems without Algorithm

The exercises provided in the chapter are mainly on solving multiplication problems, with or without context, by using the algorithm. By the end of the chapter, the expectation is that most students are quite able to correctly perform the algorithm. However, after lots of practice, it could happen that students carry out every step of algorithm perfectly, but merely in a mechanical way. CAT-8 (Figure 9) has a worksheet format and allows the teacher to assess whether students really understand what a multiplication means and thus also what the algorithmic procedure actually implies.

Here you see a multiplication problem. You have to solve it WITHOUT using the algorithm. Write down how you solved it.

	In this way I solved the problem	Answer
59×62		_____

Figure 9. CAT-8: Solving problems without algorithm.

In this CAT, students have to solve multiplication problems without using the algorithm. The main idea is that when students cannot solve a multiplication problem without using the algorithm, they will probably not have a sufficient understanding of multiplication, which might put them in trouble when learning to solve more complicated multiplication problems with, for example, three-digit numbers or decimal numbers. Examining the worksheets after class offers the teacher clues about whether and what instructional supports students need before finishing this chapter.

2.4. Teacher Training

To familiarize the teachers with the CATs and assist them in using the CATs in class, two two-hour meetings were organized during the first two weeks of the second semester of Grade 3. The first meeting took place before the teaching of the chapter on multiplication started. Each teacher received a package including a teacher guide with the material (PowerPoint slides, red/green cards, worksheets) needed for carrying out the CATs in their teaching of this chapter. During the meeting, some general information about classroom assessment was presented, and the CATs that could be used in the coming week were discussed. The second meeting began with sharing the teachers' experiences in using the CATs. After that, the CATs for the second week were discussed.

2.5. Data Collection

During the two-week intervention, each of the 25 teachers was observed at least twice regarding their use of the CATs. As a result, we could see that the teachers basically conducted the CATs as expected. To know what insights teachers might gain into students' learning, they were asked to fill in feedback forms and write a final report. Teachers filled in the feedback forms every time they used a CAT, and wrote the final report after finishing the chapter. Specifically, the teachers commented on the feedback form on whether using the CAT helped them gain new information about their students' learning, whether they adapted their further instruction, and what the new information and the instructional adaption looked like. In the final report, among others, the teachers were asked to suggest two CATs to be included in the textbook as assessment exercises and explain why they had chosen these CATs.

2.6. Data Analysis

To get an overall picture of the teachers' experiences with the CATs, we first scanned all the filled-in feedback forms. The responses of the 25 teachers on the feedback forms and final reports were included in the final analysis. To answer the research question regarding teachers' insights into students' mathematics learning, we had to identify which teachers got insights when using the CATs. Therefore, all the teachers' responses were gathered and translated into English. First, all the responses were scrutinized by the three authors separately. Each author identified the information the teachers gained about their students and had to decide whether there was evidence of gaining insights. In order to award criteria for having gained insights, the authors had to specify why they thought so. Next, for each CAT, the authors' decisions on whether the teachers' responses showed indications of insights, and the reasons for making these decisions were compared and discussed. In some cases, it was easy to agree on whether there was evidence of teachers gaining insights. When the decisions differed, they were discussed until 100% agreement was reached. Then, we checked all of our decisions again, which led to some changes of our earlier decisions. During this checking process, we also finalized the formulation of the criteria for indicating gaining insights.

In the end, this resulted in four unique criteria:

1. *Referring to the mathematical content the CAT is supposed to assess.* For this, teachers can use their own words or give a clear description of the purpose of the CAT by using (partly) the wording that appeared in the teacher guide. However, this criterion is not met when teachers only refer to the CAT in general terms not mentioning the mathematics assessed.
2. *Providing specific information about students.* This includes mentioning the proportion of students showing a particular performance on the assessed content or describing the difficulties students encountered with this content.
3. *Describing the novelty of the gained information about students.* This means that teachers learn something "new", "unexpected", "surprising", or "that was not known before" about students' understanding of the assessed content.
4. *Explaining an instructional adaptation matching the findings from the CAT.* Such an instructional adaptation has to correspond to the information about the assessed content as revealed by using the CAT; general phrases such as "providing additional exercises" or "give extra instruction" are not sufficient.

Showing that one has learned something from doing an assessment is a multifaceted phenomenon. It can be expressed in different ways. Teachers can say something about the performance of their students, can emphasize that they discovered new information in the students' performance, or can discuss their decisions about further teaching. All these responses can indicate that teachers have learned something from assessing their students with the CATs. Yet, to fully classify a teacher's response as having gained insights from a CAT, a first requirement is that the teacher refers to the mathematical content the CAT is supposed to assess. Just talking about students' performance in general terms is not sufficient. So, our final decision rule to qualify a teacher's response as having gained insights is that it should meet Criterion I and at least one from Criterion II, III, and IV. Based on this decision rule, a final round of checking was carried out by the first author. This resulted in qualifying 57 teachers' responses out of the total 200 possible responses (25 teachers x eight CATs) as having gained insights. Table 1 provides examples of the qualifications of teachers' responses about CAT-1.

Table 1. Examples of the teachers' responses about CAT-1 Family problems, and whether they were qualified as showing having gained insights. CAT: classroom assessment technique.

Example of Teacher's Response	Criterion Met	Having Gained Insights
<p><i>Example 1: Teacher S14</i> "In my expectation, the vast majority of the students can find the correct answer without being disturbed by the increasing number of zeroes. However, the reality is that students felt difficult when the numbers became bigger and bigger [II, III]. Therefore, it illustrated that the students were not flexible enough when solving by mental calculation the problem with multiples of 10 [I]. Also, it means that the students lack the ability of reasoning and generalization [I]."</p>	I, II, III	Yes
<p><i>Example 2: Teacher S09</i> "The students could not recognize such analogous problems very well [I]. Most students showed their cards based on their 'gut feeling' [II]. Only a small proportion of the students could find the rule [III]. As the number of zeroes increased, the accuracy went down [II]."</p>	I, II	Yes
<p><i>Example 3: Teacher S04</i> "Most of the students could give the answer quickly, but some individual students need help [II]."</p>	II	No
<p><i>Example 4: Teacher S10</i> "This CAT helps the students to extend the boundary of knowledge. The difficulty level is appropriate. The students showed strong interests when doing the CAT. In general, the CAT is helpful to extend students' knowledge and to improve their initiatives."</p>	[no criterion applicable]	No

3. Results

3.1. Teachers Gaining Insights from Using the CATs

Table 2 gives an overview of the CATs used by the teachers (see the white cells) and whether the teachers gained insights into their students' understanding of multiplication of two-digit numbers (see the white cells containing a ✓). Of the 25 teachers involved in the study, 22 used all the eight CATs, one (H02) used seven CATs, and two (H04, N05) used only five CATs. In total, 193 responses of using the CATs were collected, of which 57 responses (30%) clearly showed that the teachers gained insights. Five teachers gained insights from five or more CATs (the *High Insight* group), 14 teachers did this only from three or fewer CATs (the *Some Insight* group), and six teachers did not show clear evidence of insights, no matter how many CATs they used (the *No Insight* group). CAT-3 and CAT-8 appeared to be the most informative CATs for teachers; for these respective CATs, there were nine and 11 teachers who showed evidence of gaining insights. Gaining insights here means that clear indications could be identified in a teacher's response to a particular CAT: that the teacher referred to specific information about his/her students, or emphasized the novelty of the gained information, or referred to a fitting instructional adaptation, and these reactions went together with references to the mathematics content of the CATs.

Table 2. Overview of whether clear evidence of teachers (n = 25) gaining insights into their students’ understanding of multiplication by using the CATs was identified ¹.

Insight	Teacher	CAT								Total of ✓
		1	4	7	2	5	6	3	8	
High Insight	H01	✓	✓	✓		✓	✓	✓	✓	7
	H02		✓	✓	✓	■	✓	✓	✓	6
	H03	✓			✓	✓		✓		5
	H04	■	■	✓	■	✓	✓	✓	✓	5
	H05			✓	✓	✓	✓		✓	5
Some Insight	S01			✓	✓				✓	3
	S02				✓		✓		✓	3
	S03		✓			✓		✓		3
	S04		✓					✓	✓	3
	S05					✓	✓		✓	3
	S06	✓				✓			✓	3
	S07		✓						✓	2
	S08					✓	✓			2
	S09	✓						✓		2
	S10								✓	1
	S11							✓		1
	S12							✓		1
	S13				✓					1
	S14	✓								1
No Insight	N01									0
	N02									0
	N03									0
	N04									0
	N05		■			■		■		0
	N06									0
Total of ✓		5	5	5	6	8	8	9	11	57

¹ Black cell means the teacher did not use the CAT; empty white cell means the teacher used the CAT but no clear evidence of the teacher gaining insights from using the CAT was identified; white cell with “✓” means the teacher used the CAT and clear evidence of the teacher gaining insights from using the CAT was identified.

To give more information about the specific insights the teachers appeared to gain into their students’ mathematical understanding, we now focus on the teachers’ responses to the four earlier described CATs (1, 3, 4, and 8). Based on the four criteria for gained insights, in the next sections, we discuss for each of the CATs why we considered particular responses of the teachers as indications for having gained insights.

3.2. Four Examples of Teachers Gaining Insights from Using CATs

3.2.1. Insights from Using CAT-1: Family Problems.

All teachers, except one (H04), returned a response for CAT-1. In five teachers’ responses, clear evidence was found that they gained insights about whether, and to what extent, their students understood multiplication with multiples of 10 beyond the two-digit number range. Four of the teachers who gained insights directly referred to the mathematical content that CAT-1 is supposed to assess. For example, they reported about students’ analogous thinking (H03, S09) or about students’ flexible use of known rules of multiplication with multiples of 10 (H01, S14). Teacher S06 described that CAT-1 aims to assess “a method” that “makes use of the given problems” and that shows its “advantage when the number of zeroes increases”. In our view, this indicates that the teacher recognized what the task is about, resulting in a good lens for diagnosing what students do or do not understand. In their responses, the five teachers dealt with their students’ performance when the numbers in the problems

contained more than two digits. For three teachers (H01, S06, S14), this came down to reporting that most students were able to deal with this and that only a few students “could only solve [until] 97×80 and 97×800 ” (H01). Three teachers (H03, S09, S14) reported that fewer students provided correct answers when the numbers became bigger, and that “only a minority of the students could determine the rule” (S09). This came as a surprise to one teacher (S14), who explained that “in my expectation, the vast majority of students would find the correct answers without being disturbed by the increase in the number of zeroes”. Only two teachers (H03, S06) mentioned how they would adapt their instruction; they were going to include analogous problems for students to practice with.

In 19 teachers’ responses, the mathematical content assessed by CAT-1 was not mentioned. When teachers described students’ performance, it was in very general terms. Some teachers reported to be satisfied with their students’ performance; for example, students were able to calculate “according to the given characteristics” (S05), other teachers pointed out their students’ shortcomings in understanding “the rule” (H02, H05, S01). Similarly, when teachers wrote about their instructional adaptations, they often used general terms such as “more exercises” (S07) or “extra instruction” (N05). Interestingly, two teachers (S11, N04) decided not to adjust their further teaching because they considered the content of CAT-1 to be too similar to what is in the textbook. In contrast, another teacher (N01) provided as a reason for not adapting her instruction that “there is no such type of exercise in the textbook”.

3.2.2. Insights from Using CAT-3: Breaking down a Multiplication

Twenty-four teachers provided a response for CAT-3. Nine teachers were found to have gained insights into whether their students could identify the components of a multiplication of two-digit numbers and how their students understand what is “behind” the algorithm. Only one of these teachers (S11) reflected on what is assessed in CAT-3 in her own words: “breaking down the multiplication problem into four components is factually the same as showing how the algorithm works to calculate the multiplication of two-digit numbers”. The remaining teachers who gained insights referred to the mathematical content assessed by CAT-3 in terms of “the meaning of multiplication of two-digit numbers” (H03) or “the meaning behind the algorithm” (H01). When describing students’ performance, most teachers (H01, H02, S03, S04, S09, S11, S12) made a clear distinction between students’ understanding of the structure of the multiplication algorithm and students’ ability to apply the procedures. For example, one teacher (S12) reported that most of her students “master the procedure of calculating, but their understanding about how it works is not good enough”. Another teacher (S11) found that “25 out of the 39 students could answer all the blanks correctly”. In contrast, the remaining teachers reported that at least half of the students were unable to break down 24×53 , although they were able to find its result. Two teachers (H01, H04) said this was not what they expected. Another teacher (S03) also expressed her surprise, as she “thought the students would not even understand the question”, but “the situation in fact was a bit better”. For further teaching, this teacher wanted to pay more attention to help the students “understand the meaning of each step of performing the algorithm”.

The remaining 15 teachers did not refer to the mathematical content assessed in CAT-3. Instead, they mainly focused on describing students’ performance. Most teachers reported that their students had difficulties in breaking down the problem—for example, “students could only break down 24×53 into $24 \times 50 + 24 \times 3$ ” (S06)—or difficulties in understanding the question—for example, “students had never been trained to break down a multiplication problem into four parts” (S05). Overall, students in these 15 classes did not perform well on CAT-3. However, about half of the teachers would not adjust their further teaching. The main reason they gave was that CAT-3 was too different from what they taught in class about the multiplication of two-digit numbers, and it therefore could “disturb students’ thinking” (N01) or “make students confused” (N02).

3.2.3. Insights from Using CAT-4: Completing the Ratio Table

Twenty-three teachers provided a response about their use of CAT-4. Also for this CAT, five teachers' responses reflected gained insight into their students' understanding. Four of these teachers' responses (H01, H02, S03, S07) reflected the aim of finding out how students use the different relationships between the provided numbers in the ratio table. One teacher (S04) only referred to vertical reasoning in the ratio table: "most students still needed to calculate how many pencils there are in one box in order to find the total number". In the other four classes (H01, H02, S03, and S07), the teachers reported that many of their students could solve the problems by observing the provided numbers of boxes and establishing relationships. One teacher (H02) realized that "vertical and horizontal reasoning can be considered at the same time" and she planned "in future teaching, to encourage students to think and solve problems from multiple perspectives".

When the remaining 18 teachers referred to what CAT-4 aims to assess, it was in very general terms; they mentioned "the relation between the numbers" (S02), "multiple perspectives" (S05), or "other strategies" (S09). Regarding students' performance on CAT-4, five teachers reported their satisfaction with students' reactions. In the other classes, the majority of students "was limited in their thinking" (S10) or "just made a guess" (N03). Based on CAT-4, 13 teachers wrote they would provide "more exercises". Two of them (S05, S13) in fact considered reasoning horizontally in a ratio table not to be "something every student needs to master". The teachers who indicated not making instructional adaptations gave different explanations. For example, "it is not the content in the textbook" (N01) or "it is not necessary to explain such difficult content in Grade 3. It is of course good if students can understand; it is also fine if they do not understand" (S09).

3.2.4. Insights from Using CAT-8: Solving Problems without Algorithm

All teachers returned a response for CAT-8. In 11 teachers' responses, clear evidence was identified of gained insights into students' capability of solving multiplication problems of two-digit numbers without using the algorithm. In their responses, these teachers referred to the mathematical content assessed by CAT-8, for example about their students' understanding of multiplication, students' flexible use of different solutions, or students having a 'mindset' about solving multiplication problems. In particular, these teachers described their students' performance of providing different solutions to solve 59×62 . For example, one teacher (H02) found that some students understood the connection between *Lattice multiplication* and breaking down the multiplication. Another teacher (S01) wrote that "a small part of my students could solve the problem by using the distributive property". Contrastingly, another teacher (S05) found that "part of the students thought 'without using the algorithm' meant 'no accurate answer being required', [and] therefore they only made an estimation of the product". Furthermore, two teachers (H01, S10) reported that their students were used to writing down the algorithm when given a multiplication problem, and they did not know how to start now. Only one teacher (H05) really described the novelty of her gained insights. She had expected her students to not be able to solve 59×62 without using the algorithm, but in fact many of them used the method of *Lattice multiplication*. Regarding instructional adaptation, two teachers (H02, H04) valued the "openness and flexibility" (H02) of CAT-8 to "give students more space to think and imagine freely" (H04) and would use it in future teaching.

Of the 14 teachers with no evidence for insights, most only reported briefly that their students could not solve problems without using the algorithm. Some teachers did refer to CAT-8 as aiming to "develop students' divergent thinking" (S09), to "extend students' learning" (S12), or to "remind students to solve problems in different ways" (S11). Three teachers (S03, S13, N06) seemed not to understand what CAT-8 aims to assess. They thought that the assessed mathematical content was students' ability to apply the properties of multiplication; for example "it is difficult for students to understand the distributive property" (S13). Another teacher (S03) made it clear that "students are going to systematically learn the properties [in Grade 4]". Seven teachers used general phrases to describe their instructional adaptation; for example, "providing additional instruction for those who

are able to learn more” (N03) or “providing extra exercises to revise this content” (N01). The other seven teachers would not make any instructional adaptation. Two of them (S11, N04) did this because their students did not have problems in solving the problems, and the remaining five teachers felt using the algorithm was more suitable to solve these problems.

4. Conclusions and Discussion

In this study, we investigated the insights into students’ understanding that Chinese primary mathematics teachers gained by using CATs. CATs are meant to give teachers access to their students’ understanding of a particular mathematical concept or skill, by posing questions that purposely present the mathematical content from a different perspective and in an unfamiliar way. In this way, CATs may provide teachers with new lenses to observe and understand their students’ learning. In addition, the gained insights may offer teachers clues to answer the question of what to do next in their teaching. With respect to the latter, all the teachers in our study, no matter the extent of gained insights, paid more attention to describing what they found out about their students’ performance than to indicating possible instructional adaptations. This echoes the finding of Heritage et al. [41] that teachers are better at drawing inferences about students’ understanding from assessment information than at deciding the next instructional steps.

Although using CATs to assess students’ learning was quite new for the participating Chinese mathematics teachers, most of them did gain insights into students’ mathematics understanding of multiplication of two-digit numbers. Gaining insights was evidenced by either describing specific information about their students, emphasizing the novelty of the gained information, or giving suggestions for a fitting instructional adaptation, while at the same time referring to the mathematics content of the CATs. Three-quarters of the teachers gave signs of having gained insights. However, another finding was that the teachers differed greatly in the extent to which they gained these insights. Five teachers acquired insights from five or more CATs, while 14 teachers did so only from three or fewer CATs, and six teachers did not show clear evidence of insights, no matter how many CATs they used. These teachers were respectively assigned to the *High Insight*, *Some Insight*, or *No Insight* group. The teachers belonging to the *High Insight* group could clearly recognize that offering students problems that differ slightly from the way they are presented in the textbook can be an eye-opener for the teachers and can offer them the opportunity to acquire knowledge about their students’ learning that they did not have before. In line with what was found by James and McCormick [22] for the ‘Assessment for Learning’ practice, these teachers seemed to have adopted the CATs as a means to assess students according to the ‘spirit’ of the CATs, whereas the *No Insight* teachers seemed to have implemented the CATs in a more mechanical way, more according to the ‘letter’ of the CATs, following the prescribed procedure of the CATs and carrying them out in class without getting a better understanding of students’ mathematics learning. In contrast with the teachers belonging to the *High Insight* group, who clearly favored the revelatory capacity of assessment tasks that differ in a specific way from the teaching tasks, the teachers in the *No Insight* group did not value them as such, and did not appear to notice the difference. Several of these teachers considered the CATs to be (too) similar to the textbook, and did not want to repeat what they had already taught. Yet, others emphasized that the CATs were too different or too difficult compared to their regular teaching. The latter group of teachers appeared to hold a strict view of what should and could be assessed, which might lead to them not perceiving the CATs as helpful for obtaining more insight into students’ understanding regarding the topic that is at hand in the textbook and their teaching. As a result, they might not have used the CATs and the elicited information optimally, and thus did not show evidence of gained insights from using the CATs. A possible reason why teachers did not gain insights may have to do with the short duration of the intervention. A two-week support is a very limited time for teachers to become accustomed to using the CATs and recognizing the valuable information revealed by using them. Perhaps more insights would have been observed if a longer intervention had taken place. Moreover, the teachers, including those in *High Insight* and *Some Insight* groups, were not involved in designing

the CATs, and this could also be a reason that they did not notice or value the information about students' learning. When teachers can take part in the development of the CATs and are supported regarding seeing the connections between the CATs and the textbook content, this might lead to gaining more insights into students' understanding. In addition, that the local teaching office, rather than the individual teachers, decided to participate in this study may have influenced the teachers' willingness regarding using CATs and filling in feedback forms. Also, the felt need of teachers to finish their planned regular teaching may have been a factor that teachers considered when deciding whether to use the CATs, and might have led to not always getting the most out of the CATs.

Having only a short intervention and not having the teachers involved in designing the CATs are shortcomings of our study that should be kept in mind when interpreting our findings. Also, it is important to take into account that our conclusions about whether the teachers obtained insights from using CATs were based on their self-reported data. Further data collection, such as observing teachers using CATs in class and directly asking them about their gained insights, could shed new light on possibly gathered new information about their students' mathematics learning. Furthermore, examining students' responses and worksheets could be included in order to triangulate teachers' reported insights. Another limitation of our study is that the CATs used in the context of Chinese primary mathematics education so far, including the pilot study [39], were designed based on one particular textbook series and only involved teachers from the city of Nanjing. Whether Chinese primary teachers who use different mathematics textbook series or who are from different regions can get new insights from implementing CATs remains unclear. Further studies are necessary in this respect.

A strong recommendation for additional studies is in any case to have a design in which teachers are included in developing CATs. Moreover, future research may investigate why some teachers gain less insights than others. To know more about this, it could be worthwhile to explore whether teachers with different assessment profiles [38,42] benefit differently from using CATs. For example, recently, it has been shown that significant differences exist between Chinese expert teachers in primary school mathematics education and their non-expert colleagues in their perception and reported behavior of understanding their students' mathematics thinking [34]. Future research may also focus on how to use the gained insights for making instructional decisions, because this is an issue to which the teachers in our study did not pay much attention. In line with this, new research may be conducted to examine whether and how teachers' gained insights affects students' performance.

Despite the aforementioned limitations and the questions that still have to be answered, we think we can conclude that our study provides evidence that Chinese primary school mathematics teachers may gain insights into their students' understanding of mathematics from using assessment techniques such as CATs. Yet, for the majority of the teachers in our study, it seems to be necessary to offer them more time and support to get acquainted with this assessment approach. Also, more opportunities could be provided to support teachers to see the connections between CATs and the textbook. Using the CATs implies a strong formative approach to assessment. For Chinese primary school mathematics teachers, who often put more emphasis on providing information to students than on getting information about students [32], this may mean a change of perspective. Our CATs may help them develop a more formative approach to assess students' learning. More research is certainly necessary, especially studies that investigate how teachers' culturally-based beliefs about teaching affects their formative use of assessment and that examine how to support teachers to become independent users of formative assessment.

Author Contributions: Conceptualization, Methodology, Analyses, Writing: X.Z., M.v.d.H.-P., M.V.; Funding Acquisition, X.Z. (China Scholarship Council) and M.v.d.H.-P. (Netherlands Organization for Scientific Research).

Funding: This research was funded by China Scholarship Council, grant number 201206860002; and the Netherlands Organization for Scientific Research, grant number NWO MaGW/PROO: Project 411-10-750.

Acknowledgments: The authors thank Lianhua Ning from Nanjing Normal University, China, for helping to contact schools, and all the teachers involved in this study for their cooperation and contribution.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. The eight CATs for multiplication of two-digit numbers.

Title	Format	Example	Purpose																		
Family problems (CAT-1)	Red/Green cards	It is known that 97×8 equals 776. Do you think you can solve the following problems? (Yes- Green card ; No- Red card) a) 97×80 b) 97×800 c) 97×8000 d) 970×8000	Assessing whether students can recognize analogous problems and are aware of the relationship among the results of these problems																		
How many approximately? (CAT-2)	Red/Green cards	Here you see five bags with a different weight. How many kilograms approximately in one bag? Do you think the answer is 30, 40 or 35? (Yes- Green card ; No- Red card) <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>Bag 1</td> <td>Bag 2</td> <td>Bag 3</td> <td>Bag 4</td> <td>Bag 5</td> </tr> <tr> <td>33 kg</td> <td>34 kg</td> <td>37 kg</td> <td>37 kg</td> <td>36 kg</td> </tr> </table>	Bag 1	Bag 2	Bag 3	Bag 4	Bag 5	33 kg	34 kg	37 kg	37 kg	36 kg	Assessing whether students can find the approximate average of a series of numbers								
Bag 1	Bag 2	Bag 3	Bag 4	Bag 5																	
33 kg	34 kg	37 kg	37 kg	36 kg																	
Breaking down a multiplication (CAT-3)	Worksheet	Here you see a multiplication problem. You do not have to calculate the result. You only have to fill in the blanks. 24×53 means that you have to calculate a) 3×4 and 3×20 and $_ \times 4$ and $_ \times 20$ b) $4 \times _$ and $4 \times _$ and $_ \times 3$ and 20×50 c) 20×50 and $20 \times _$ and 4×50 and $_ \times 3$	Assessing whether students understand how the multiplication algorithm works																		
Completing the ratio table (CAT-4)	Red/Green cards	You are NOT allowed to calculate how many pencils there are in one box. Can you still find out how many pencils there are in the following boxes? (Yes- Green card ; No- Red card) <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>Number of boxes</td> <td>6</td> <td>11</td> <td>12</td> <td>13</td> <td>17</td> <td>18</td> <td>22</td> <td>23</td> </tr> <tr> <td>Number of pencils</td> <td>72</td> <td>132</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>	Number of boxes	6	11	12	13	17	18	22	23	Number of pencils	72	132							Assessing whether students can make use of the ratio table by reasoning horizontally
Number of boxes	6	11	12	13	17	18	22	23													
Number of pencils	72	132																			
Bigger or smaller? (CAT-5)	Red/Green cards	Are the statements correct? (Yes- Green card ; No- Red card) a) The product of 24×45 is smaller than 1200. b) The product of 18×32 is bigger than 600.	Assessing whether students can estimate the product by reasoning																		
Quick check of answers (CAT-6)	Red/Green cards	Here you see multiplication problems with their answers. Do you see a mistake? (Yes- Green card ; No- Red card) $50 \times 200 = 1000$ $37 \times 34 = 1250$ $23 \times 38 = 874$	Assessing whether students can quickly check the correctness of the result of multiplication problems without performing the algorithm																		
Fruit language (CAT-7)	Worksheet	On this worksheet you have to rewrite the problems in fruit language. Use the given info and look at the examples. <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td></td> <td>Problem</td> <td>In fruit language</td> </tr> <tr> <td rowspan="2">Given info</td> <td>7×18</td> <td></td> </tr> <tr> <td>5×18</td> <td></td> </tr> <tr> <td rowspan="4">Examples</td> <td>14×18</td> <td></td> </tr> <tr> <td>12×18</td> <td></td> </tr> <tr> <td>18×20</td> <td>_____</td> </tr> <tr> <td>24×18</td> <td>_____</td> </tr> </table>		Problem	In fruit language	Given info	7×18		5×18		Examples	14×18		12×18		18×20	_____	24×18	_____	Assessing whether students can use the associative and distributive property of multiplication to restructure a multiplication problem	
	Problem	In fruit language																			
Given info	7×18																				
	5×18																				
Examples	14×18																				
	12×18																				
	18×20	_____																			
	24×18	_____																			
Solving problems without using the algorithm (CAT-8)	Worksheet	Here you see a multiplication problem. You have to solve it WITHOUT using the algorithm. Write down how you solved it. <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td></td> <td>In this way I solved the problem</td> <td>Answer</td> </tr> <tr> <td>59×62</td> <td></td> <td>_____</td> </tr> </table>		In this way I solved the problem	Answer	59×62		_____	Assessing whether students have a deep understanding of the multiplication operation and whether they have, instead of the algorithm, other strategies available to solve multiplication problems												
	In this way I solved the problem	Answer																			
59×62		_____																			

References

- Shavelson, R.J. *The Basic Teaching Skill: Decision Making (R & D Memorandum No.104)*; Stanford University: Stanford, CA, USA, 1973.
- Parkes, J. Reliability in classroom assessment. In *Sage Handbook of Research on Classroom Assessment*; McMillan, J.H., Ed.; Corwin Press: Thousand Oaks, CA, USA, 2013; pp. 107–123. Available online: <http://sci-hub.tw/10.4135/9781452218649.n7> (accessed on 11 March 2019).

3. Harlen, W. *Assessment of Learning*; Sage: London, UK, 2007; Available online: <http://sci-hub.tw/10.4135/9781446214695> (accessed on 11 March 2019).
4. Shepard, L.A.; Penuel, W.R.; Pellegrino, J.W. Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educ. Meas. Issues Pract.* **2018**, *37*(1), 21–34. [[CrossRef](#)]
5. Black, P.; Wiliam, D. Assessment and classroom learning. *Assess. Educ. Princ. Policy Pract.* **1998**, *5*, 7–74. [[CrossRef](#)]
6. Brookhart, S.M. Classroom assessment: Tensions and intersections in theory and practice. *Teach. Coll. Rec.* **2004**, *106*, 429–458. [[CrossRef](#)]
7. De Lange, J. *Framework for Classroom Assessment in Mathematics*; NICLA; WCER: Madison, WI, USA, 1999.
8. Shepard, L.A. The role of assessment in a learning culture. *Educ. Res.* **2000**, *29*(7), 4–14. [[CrossRef](#)]
9. Stiggins, R.; Chappuis, J. Using student-involved classroom assessment to close achievement gaps. *Theory Pract.* **2005**, *44*(1), 11–18. [[CrossRef](#)]
10. Wiliam, D. Keeping learning on track: Classroom assessment and the regulation of learning. In *Second Handbook of Research on Mathematics Teaching and Learning*; Lester, F.K., Ed.; Information Age Publishing: Greenwich, UK, 2007; pp. 1053–1098.
11. Schoenfeld, A.H. Summative and formative assessments in mathematics supporting the goals of the common core standards. *Theory Pract.* **2015**, *54*, 183–194. [[CrossRef](#)]
12. Lin, P.J. Conceptualizing teachers' understanding of students' mathematical learning by using assessment tasks. *Int. J. Sci. Math. Educ.* **2006**, *4*, 545–580. [[CrossRef](#)]
13. Panizzon, D.; Pegg, J. Assessment practices: Empowering mathematics and science teachers in rural secondary schools to enhance student learning. *Int. J. Sci. Math. Educ.* **2008**, *6*, 417–436. [[CrossRef](#)]
14. Gallego-Arrufat, M.J.; Dandis, M. Rubrics in a secondary mathematics class. *Int. Electron. J. Math. Educ.* **2014**, *9*, 73–82.
15. Jin, H.; Wong, K.Y. Mapping conceptual understanding of algebraic concepts: An exploratory investigation involving grade 8 Chinese students. *Int. J. Sci. Math. Educ.* **2015**, *13*, 683–703. [[CrossRef](#)]
16. Leahy, S.; Lyon, C.; Thompson, M.; Wiliam, D. Classroom assessment: Minute-by minute and day by day. *Educ. Leadersh.* **2005**, *63*(3), 18–24.
17. Andersson, C.; Palm, T. The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learn. Instr.* **2017**, *49*, 92–102. [[CrossRef](#)]
18. Keeley, P.; Tobey, C.R. *Mathematics Formative Assessment: 75 Practical Strategies for Linking Assessment, Instruction, and Learning*; Corwin Press: Thousand Oaks, CA, USA, 2011.
19. Veldhuis, M.; Van den Heuvel-Panhuizen, M. Supporting primary school teachers' classroom assessment in mathematics education: Effects on student achievement. *Math. Educ. Res. J.* **2019**. Available online: <https://link.springer.com/article/10.1007/s13394-019-00270-5> (accessed on 3 June 2019).
20. Wiliam, D. *Embedded Formative Assessment*; Solution Tree: Bloomington, IN, USA, 2011.
21. Wylie, E.C.; Lyon, C.J. The fidelity of formative assessment implementation: Issues of breadth and quality. *Assess. Educ. Princ. Policy Pract.* **2015**, *22*, 140–160. [[CrossRef](#)]
22. James, M.; McCormick, R. Teachers learning how to learn. *Teach. Teach. Educ.* **2009**, *25*, 973–982. [[CrossRef](#)]
23. Phelan, J.; Choi, K.; Niemi, D.N.; Vendlinski, T.; Baker, E.L.; Herman, J. The effects of POWERSOURCE[©] assessments on middle-school students' math performance. *Assess. Educ. Princ. Policy Pract.* **2012**, *19*, 211–230. [[CrossRef](#)]
24. Davidson, A.; Herbert, S.; Bragg, L.A. Supporting elementary teachers' planning and assessing of mathematical reasoning. *Int. J. Sci. Math. Educ.* **2018**. [[CrossRef](#)]
25. Martino, A.M.; Maher, C.A. Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us. *J. Math. Behav.* **1999**, *18*, 53–78. [[CrossRef](#)]
26. Heritage, M.; Heritage, J. Teacher questioning: The epicenter of instruction and assessment. *Appl. Meas. Educ.* **2013**, *26*, 176–190. [[CrossRef](#)]
27. Zhang, D.; Lee, P.Y. Examination culture and mathematics teaching. In *Proceedings of the ICMI–China Regional Conference of Mathematical Education*, Beijing, China, 5–8 August 1991.
28. Li, Y. Purpose, function and types of mathematics assessment in China. *ZDM* **2000**, *32*, 192–200. [[CrossRef](#)]

29. Ministry of Education of the People's Republic of China (MoE). Jichu jiaoyu kecheng Gaige Gangyao (Shixing) [Curriculum Reform Outline of Basic Education (Trial Version)]. Available online: <http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/167343.html> (accessed on 11 March 2019).
30. Ministry of Education of the People's Republic of China (MoE). *Quanrizhi Yiwu Jiaoyu Shuxue Kecheng Biao Zhun (Shiyangao)*; [Mathematics Curriculum Standards of Nine-Year Compulsory Education (Trial Version)]; Beijing Normal University Press: Beijing, China, 2001.
31. Ministry of Education of the People's Republic of China (MoE). *Yiwu jiaoyu shuxue kecheng biao zhun (2011 nian ban)* [Mathematics Curriculum Standards of Nine-Year Compulsory Education (2011 Version)]. Available online: http://ncct.moe.edu.cn/2014/curriculumPolicy_1115/3175.html (accessed on 11 March 2019).
32. Cai, J.; Wang, T. Conceptions of effective mathematics teaching within a cultural context: Perspectives of teachers from China and the United States. *J. Math. Teach. Educ.* **2010**, *13*, 265–287. [[CrossRef](#)]
33. Cai, J.; Ding, M.; Wang, T. How do exemplary Chinese and U.S. mathematics teachers view instructional coherence? *Educ. Stud. Math.* **2014**, *85*, 265–280. [[CrossRef](#)]
34. Zhu, Y.; Yu, W.; Cai, J. Understanding students' mathematical thinking for effective teaching: A comparison between expert and nonexpert Chinese elementary mathematics teachers. *Eurasia J. Math. Sci. Technol. Educ.* **2018**, *14*, 213–224. [[CrossRef](#)]
35. Li, Y.; Chen, X.; Kulm, G. Mathematics teachers' practices and thinking in lesson plan development: A case of teaching fraction division. *ZDM* **2009**, *41*, 717–731. [[CrossRef](#)]
36. Liu, J. *Xiaoxue Dinianji Shuxue Jiaokeshu Zhong De Xiti Yanjiu—Jiyu Zhongguo Xinjiapo Jiaokeshu Bijiao De Shijiao*; [Study in Mathematics Textbook Tasks for Lower Grades in Primary School—From the Perspective of Comparison Between Textbooks in China and that in Singapore]. Master's Thesis, Northeast Normal University, Changchun, China, 2012.
37. Gu, F.; Gu, L. Characterizing mathematics teaching research specialists' mentoring in the context of Chinese lesson study. *ZDM* **2016**, *48*, 441–454. [[CrossRef](#)]
38. Zhao, X.; Van den Heuvel-Panhuizen, M.; Veldhuis, M. Chinese primary school mathematics teachers' assessment profiles: Findings from a large-scale questionnaire survey. *Int. J. Sci. Math. Educ.* **2018**, *16*, 1387–1407. [[CrossRef](#)]
39. Zhao, X.; Van den Heuvel-Panhuizen, M.; Veldhuis, M. Teachers' use of classroom assessment techniques in primary mathematics education—an explorative study with six Chinese teachers. *Int. J. STEM Educ.* **2016**, *3*, 19. [[CrossRef](#)]
40. Jiangsu Phoenix Education Publishing House. *Sujiaoban Jiaokeshu*; (Xiaoxue Shuxue Sannianji Xiace) [Sujiaoban Textbook (Mathematics Textbook for Grade 3 in Primary Education, Volume 2)]; Jiangsu Phoenix Education Publishing House: Nanjing, China, 2014.
41. Heritage, M.; Kim, J.; Vendlinski, T.; Herman, J. From evidence to action: A seamless process in formative assessment? *Educ. Meas. Issues Pract.* **2009**, *28*(3), 24–31. [[CrossRef](#)]
42. Veldhuis, M.; van den Heuvel-Panhuizen, M. Primary school teachers' assessment profiles in mathematics 718 education. *PLoS ONE* **2014**, *9*, e86817. [[CrossRef](#)]

