

# Classroom assessment in the eyes of Chinese primary mathematics teachers: A review of teacher-written papers



Xiaoyan Zhao<sup>a,\*</sup>, Marja Van den Heuvel-Panhuizen<sup>a,b</sup>, Michiel Veldhuis<sup>b</sup>

<sup>a</sup> Freudenthal Institute, Faculty of Science, Utrecht University, Utrecht, The Netherlands

<sup>b</sup> Freudenthal Group, Faculty of Social and Behavioural Sciences, Utrecht University, Utrecht, The Netherlands

## ARTICLE INFO

### Article history:

Received 7 October 2015

Accepted 16 December 2016

Available online 13 January 2017

### Keywords:

Classroom assessment

Mathematics education

Primary school

Teachers

Review of teacher-written papers

China

## ABSTRACT

In this paper we report on a review of papers written by teachers aimed at knowing more about teachers' perceptions of the current situation of classroom assessment in primary mathematics education in China. The review is based on 266 papers included in the China National Knowledge Infrastructure database. We found that the teacher-authors reflected various aspects of their classroom assessment practice, including the purpose of assessment, the content of it, the person who is the assessor, the assessment methods that are used, and the feedback that is provided. Most attention was paid to feedback; it seems many teacher-authors considered classroom assessment to be equivalent to feedback. In general, the conceived classroom assessment practice as described in the papers echoed well nearly all aspects that are advocated in the Chinese mathematics curriculum standards. The only aspect that was scarcely discussed in the papers was the use of assessment results to adapt and improve instruction.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

*Classroom assessment*, considered as assessment in the hands of teachers for the purpose of informing teaching and learning, has been recognized and promoted in mathematics education all over the world. This important role of classroom assessment is also reflected in the mathematics curriculum reform and the accompanying assessment reform in China, launched in 2001. After more than a decade of reform, however, it is still unclear how mathematics teachers perform assessment in their classrooms. Gaining more knowledge about this can be achieved in different ways; our approach in this study was conducting a review of papers in teacher journals written by Chinese teachers addressing classroom assessment in primary school mathematics education. By analyzing these teacher-written papers, we aimed at casting light on the activities teachers use in the assessment of their students, and whether the reported practice is related to the assessment guidelines in Chinese curriculum documents.

In the remainder of this introduction, we will elaborate successively on the role of classroom assessment in mathematics education in general, the content of the mathematics curriculum

reform in China, the accompanying assessment guidelines, and finally we will formulate our research questions.

### 1.1. Classroom assessment in mathematics education

In the last decades, many countries have reformed their mathematics education towards a curriculum which no longer solely focuses on knowing facts and carrying out routine skills, but also on understanding and higher-order skills such as reasoning, modelling, and problem solving (see, e.g., [NCTM, 1989, 2000](#)). This reform in mathematics education has also changed the view on assessing students' learning and called for a new approach to assessment corresponding to and serving these changes in curricula ([Leung, 2008](#); [Romberg, Zarinnia, & Collis, 1990](#)). A new approach to assessment is required to make it epistemologically consistent with the didactics of mathematics ([Van den Heuvel-Panhuizen & Becker, 2003](#)). Assessment should correspond to the curriculum that is taught and the learning theory that is adhered to ([Shepard, 2000](#)). This means that in addition to students' knowledge and skills also their ability to solve more complex problems should be assessed, that not only the correctness of students' solutions should be the focus of assessment but also the strategies employed by students, and, finally, that assessment is seen as an on-going process integrated within instruction (e.g., [Berry, 2011](#); [Romberg et al., 1990](#); [Shepard, 2000](#); [Suurtamm, Koch, & Arden, 2010](#); [Van den Heuvel-](#)

\* Corresponding author at: Utrecht University, Heidelberglaan 1, Langeveld Building, Room F2.07, 3584 CS Utrecht, The Netherlands.  
E-mail address: [x.zhao@uu.nl](mailto:x.zhao@uu.nl) (X. Zhao).

Panhuizen, 1996). An important characteristic of this new approach to assessment is the awareness that assessment should not only be assessment *of* learning but also *for* learning, that is formative assessment, meaning that assessment should inform teachers' instructional decision making and students' learning (Assessment Reform Group, 1999; Black & Wiliam, 1998a, 1998b; Stiggins, 2002). Formative assessment 'in the hands of teachers' (Van den Heuvel-Panhuizen & Becker, 2003, p. 698) that is interwoven with instruction and fully integrated in the teachers' daily teaching practice is often called 'classroom assessment' (e.g., Black & Wiliam, 1998b; Brookhart, 2004; De Lange, 1999; Shepard, 2000; Stiggins & Chappuis, 2005; Wiliam, 2007). In general, 'classroom assessment' includes all kinds of formative assessment in which the teacher has the lead. This means that the teacher makes decisions about when, for what purpose, and by which method information about students' learning is gathered with the aim of informing further steps in his/her teaching. To gather this information, teachers can use a variety of methods: ranging from observing students' problem solving, listening to students' answers to questions and quizzes, to examining students' written work and administering tests (Keeley & Tobey, 2011; Wiliam, 2011a). Also, as part of classroom assessment, teachers can offer students opportunities for carrying out self- and peer-assessment (e.g., Wiliam, 2011b), in which teachers need to carefully set up and manage the activity while students play the leading role in collecting and using assessment information for their own learning improvement (Andrade, 2010; Topping, 2010).

In accordance with the worldwide reform of mathematics education towards using assessment to enhance teaching and learning, in the last decades, the relevance of classroom assessment is increasingly acknowledged in many countries (Berry, 2011). In addition, the interest in classroom assessment was particularly triggered by the review done by Black and Wiliam (1998b) in which they revealed that teachers' use of classroom assessment can lead to the improvement of students' mathematics achievement. Although this review and its conclusions were criticized, particularly on the reported effect sizes (e.g., Bennett, 2011), enough evidence remains that teachers' use of classroom assessment is linked to an increase in students' learning (e.g., Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, 2012; Kingston & Nash, 2011; McMillan, Venable, & Varier, 2013; Veldhuis & Van den Heuvel-Panhuizen, 2014). Therefore, policymakers have embraced the use of assessment for learning. For example, the US National Council of Teachers of Mathematics (NCTM, 2013) has strongly endorsed the integration of such assessment in daily instruction. In Hong Kong, the Curriculum Development Council (CDC, 2002) recommended that all schools should review their current assessment practices and put more emphasis on assessment for learning.

## 1.2. Mathematics curriculum reform in China

In mainland China, in 2001, the Ministry of Education of the People's Republic of China (MoE) initiated a curriculum reform with the purpose of better preparing students to meet the challenges of the 21<sup>st</sup> century by publishing a curriculum reform outline (MoE, 2001a). To help teachers, textbook designers, and other stakeholders in the nine-year compulsory education develop a clear view on the implementation of the curriculum reform in mathematics education, the MoE (2001b) also published in that same year the mathematics curriculum standards. In this document, one can find detailed information about (1) fundamental ideas about mathematics and mathematics education, and the structure of the mathematics curriculum standards, (2) the objectives of mathematics education in terms of knowledge and skills, mathematical thinking, problem solving, and mathematical

and learning attitude, (3) mathematical content, and (4) suggestions with examples for instruction, assessment, and the design of mathematics textbooks and other materials.

The document of the mathematics curriculum standards (MoE, 2001b) was initially only used in parts of the country. By Fall 2006 it became compulsory nationwide (Ni, Li, Li, & Zou, 2011). One year later, this was followed by the release of a revised version of the mathematics curriculum standards (MoE, 2007). This revised version was developed by a group of fourteen scholars, researchers, teacher educators, and expert teachers in mathematics education, organized and authorized by the MoE. The mathematics curriculum standards issued by MoE in 2001 were modified based on investigations into its use and the suggestions and critical remarks from mathematicians, experienced mathematics educators, and in-service mathematics teachers from more than ten provinces in China (Shi, Ma, & Liu, 2012). The latest version of the mathematics curriculum standards was published in December 2011 (MoE, 2011). In this version, it is emphasized that students should develop the ability of identifying and posing problems together with the ability of analyzing and solving problems. Moreover, it is stressed that attention should be paid to calculation, modelling, geometric visualization, and creativity, together with number sense, symbol sense, space concept, data analysis, reasoning, and application.

### 1.3. Assessment as described in the mathematics curriculum standards in China

China has a long history of examination-oriented education (Berry, 2011), which has been changed remarkably when in 2001, together with the curriculum reform, a new approach to assessment was promoted. In the curriculum reform outline (MoE, 2001a) it is mentioned that the assessment reform can be characterized by reducing the overemphasis on using assessment for differentiation and selection purposes, and using assessment to facilitate students' development, teachers' enhancement, and the improvement of the teaching and learning practice. The latest version of the mathematics curriculum standards (MoE, 2011) released some ten years later contains specific information about how assessment is conceptualized within the Chinese curriculum reform movement. To better support teachers' assessment practice, the mathematics curriculum standards document gives guidelines for the following aspects of assessment: (1) main purposes of assessment, (2) the content of assessment, (3) who can be an assessor, (4) the methods that can be used for assessment, and (5) suitable ways of reporting and using assessment results.

#### 1.3.1. Purpose of assessment

In contrast to the use of assessment for differentiating and selecting students – which was common practice before the reform – the new approach to assessment is aimed at contributing to the teaching-learning process. In line with this, three purposes are mentioned in the mathematics curriculum standards: assessment should be used to get a comprehensive understanding of students' learning, to help students to enhance their learning, and to facilitate teachers to improve their instruction. However, the description of the purposes is very brief, and no further explanations or examples are given.

#### 1.3.2. Content of assessment

For the content of assessment it is stipulated that it should address what mathematics students have to learn and what mathematical competences they have to develop. Table 1 shows the assessment guidelines and the two examples given for the competence domains in the mathematics curriculum standards of the nine-year compulsory education.

**Table 1**  
Assessment guidelines and examples for the competence domains included in the mathematics curriculum standards of the nine-year compulsory education in China (derived from MoE, 2011).

Competence domain	Assessment guideline	Example provided
Knowledge and skills	Assessment of mastery of basic knowledge and skills, which also includes assessing flexible application of these skills and students' understanding of mathematical concepts. For students who do not reach the objectives, providing extra instruction and more time is suggested.	By the end of Grade 3, students should be able to calculate within one minute 8 to 10 addition or subtraction problems up to 20, or 8 to 10 multiplication or division problems within the multiplication table.
Mathematical thinking and problem solving	Assessment of mathematical thinking and problem solving should be carried out by multiple methods during the whole process of mathematics learning, especially by using context problems.	In the second stage (Grade 4–6), teachers can provide the following task: Here is a 50 centimeter long string. Now use it to make a rectangle with sides of integers; and find when the rectangle has the biggest area. Teachers can provide questions with different difficulty levels to assess their students.
Mathematical and learning attitude	Assessment of students' mathematical and learning attitude should be carried out during the daily teaching, mainly using observations and activity reports. Also interviews can be carried out, to understand students' situation, their initiatives, interest, confidence, courage, and their cooperation and communication with others.	

### 1.3.3. Assessor

Regarding the person who is doing the assessment, it is underlined in the mathematics curriculum standards that both the teacher, students themselves, their peers, and parents can participate as assessors. By establishing such a multi-actor system of assessment, both teachers' teaching and students' learning can be assessed. For example, by the end of a chapter the teacher can ask students to make a summary about their learning gains and difficulties. In this way the teacher can assess whether students have a good understanding of what is taught in the chapter. Additionally, with such a summary the students can reflect on what they have learned, the problems they have encountered, and how they could make improvements. Moreover, they can share and discuss their findings regarding their own learning and difficulties with their peers, which can make students learn from each other's experiences. If possible, parents are also welcome to join such assessment practice.

### 1.3.4. Method of assessment

Regarding the method of assessment, the document of the mathematics curriculum standards advises various methods including written tests, oral tests, open questions, activity reports, observations, interviews, exercises in and after class, and portfolios. Teachers also should understand the characteristics of all these different methods and choose an appropriate method that fits both the content to be assessed and their students' learning situations. The importance of written tests to assess students is explicitly emphasized, which also applies to the primary school grades. In connection with this, suggestions are given about how to design and conduct written tests.

### 1.3.5. Report and use of assessment results

The main guideline for reporting assessment results is that the feedback should contribute to the enhancement of students' confidence and their learning interests, help them to develop good learning habits and facilitate their learning. In the feedback to students, the assessment results should be provided descriptively combined with a grade or mark, with the focus on what the students learned, the progress they made, their potential, and where they need to improve. This feedback can be given orally or on paper. It is also emphasized that the process of getting feedback can be an emotional experience for students. Furthermore, the feedback should not only provide students with success

experiences and boost their confidence, but should also let them know their weaknesses and where to improve.

About the use of assessment results it is only briefly described in the mathematics curriculum standards how teachers can benefit from the findings from an assessment. Based on the information about the students' learning level and their learning difficulties, it is suggested that teachers can adapt and improve their instruction. Finally, although the assessment promoted here is generally in alignment with the purposes and the characteristics of classroom assessment, the term 'classroom assessment' is not employed in the mathematics curriculum standards.

### 1.4. Research questions

Over a decade ago the mathematics curriculum reform was launched in China. Whether and how the fundamental ideas of the reform were implemented in classrooms is an issue that lately has been receiving attention in the Chinese academic community. For example, Ni et al. (2011) investigated the impact of the curriculum reform on classroom teaching and learning, and found positive changes in teachers' beliefs, the cognitive level of the learning tasks provided to students, and the balanced development in students' mathematics achievement. Concerning the types of assessment teachers used, Ni et al. (2011) discovered that the teachers in the reform group were more able to employ reform-related types of assessment. However, other researchers have stated that assessment is a 'bottleneck' for teachers (Zhu, 2012) and that teachers need to improve their classroom assessment (Zhong, 2012). Despite these findings about assessment, it is still largely unclear how mathematics teachers perform assessment in their classrooms, especially in primary school. To fill this knowledge gap, we set up the current study. We aimed to gain more knowledge, by examining teacher-written papers, on the activities teachers use in the assessment of their students, and whether this reported practice is related to the assessment guidelines in the curriculum documents. Even though there may be a profound gap between what teachers reported in their papers and what they factually did regarding to their assessment activity, what can at least be extracted from the teachers' publications are what assessment activities and views they have in their mind, which can be considered as a first requirement to use them in practice. Moreover, Herse (1979) stated that one's manner of presenting mathematics is an indication of what one believes to be most

essential in it. Similarly, teachers' descriptions and views about what classroom assessment is and should be in their practices, can be a good resource to reveal what they believe to be of most importance in classroom assessment. Following Thompson (1984, 1992) who defines conception as a general mental structure encompassing beliefs, views, preferences, and the like, the research questions of our study were:

1. What do teacher-written papers reveal about Chinese primary school mathematics teachers' conceptions of classroom assessment? More specifically, for what purpose do teachers use assessment, what content is assessed, who acts as an assessor, which assessment methods are used and how are the assessment results reported and used?
2. How are the teachers' conceptions of classroom assessment, as reflected in teacher-written papers, related to the assessment guidelines as included in the mathematics curriculum standards released in 2011?

## 2. Method

### 2.1. Selection of teacher-written papers

To select papers we used the China National Knowledge Infrastructure (CNKI) database, which is the most comprehensive online resource for accessing China's intellectual output and includes journal papers as well as master and doctor theses, papers in proceedings, newspaper articles, and yearbooks. The selection of the papers (see Table 2) was carried out in July 2013.

Our goal was to collect teacher-written papers that cast light on the activities conducted by primary school mathematics teachers in the assessment of their students. Since a direct search for teacher-written papers is not possible in this database, we started with searching for papers in the Education category of CNKI related to "assessment", in Chinese *Pingjia* (评价), which is the literal translation of assessment. This first selection step resulted in 209,492 papers having *Pingjia* (评价) in the title, keywords, or abstract. The earliest one was published in 1949, concerning evaluating the price of materials in military system (Announcement, 1949).

Because *Pingjia* (评价) is a very broad term, which includes not only assessment at classroom level, but also all kinds of external evaluations – such as the evaluation of the quality of a school – a further selection was necessary to exclude forms of assessment that were not of interest in our review. In the second selection step a total of thirteen search words were used. These words were all related in some way to assessment conducted in the classroom such as 'classroom assessment', 'classroom teaching and learning assessment', 'assessment for learning', 'student assessment' (including both self- and peer-assessment), 'learning assessment', and 'formative assessment'. In this way we got 30,826 papers that address internal assessment, that is, assessment at the classroom level. To collect only the papers about mathematics education the third selection step was searching for papers with 'mathematics' in the title, keywords, or abstract. The result was 2750 remaining papers. The fourth selection step was meant to only keep the papers about assessment in primary school. To optimize this search we did it in two steps: first the search option of the CNKI was used to search for 'primary school' in the full papers and then the

**Table 2**  
Steps in selecting teacher-written papers on classroom assessment in primary mathematics education in China.

Papers in CNKI database belonging to the category Education			
Step	Selection focus		Resulting papers
1	Term '评价' (assessment and/or evaluation) <i>Search in title, abstract, and keywords</i>	209,492	Papers addressing: -all kinds of assessment
2	Terms '课堂评价' (classroom assessment), '课堂教学评价' (classroom teaching and learning assessment or classroom teaching assessment), '形成性评价' (formative assessment), '为了学的评价' (assessment for learning), '教师评价' (teacher assessment), '学生评价' (student assessment) <sup>a</sup> , '作业评价' (assignment assessment), '即时评价' (immediate assessment), '延时评价' (delayed assessment), '过程评价' (assessment on process), '多元评价' (multiple assessment), '评价与提问' (assessment and questioning) and '评价与反馈' (assessment and feedback) <i>Search in title, abstract, and keywords</i>	30,826	Papers addressing: -internal assessment
3	Term '数学' (mathematics) <i>Search in title, abstract, and keywords</i>	2750	Papers addressing: -internal assessment -mathematics
4	Term '小学' (primary school) <i>Search in full paper</i>	904	Papers addressing: -internal assessment -mathematics -primary school
904 papers transferred to EndNote			
5	Publication year 2011 or 2012 <i>Search in the EndNote column "Publication year"</i>	360	Papers published in 2011 or 2012 addressing: -internal assessment -mathematics -primary school
6	Paper contains formative assessment conducted by the teacher, that is, contains classroom assessment <i>Search in full paper</i>	283	Papers published in 2011 or 2012 addressing: -mathematics -primary school -classroom assessment
7	Paper is written by a teacher <i>Search in author's affiliation</i>	266	Teacher-written papers published in 2011 or 2012 addressing: -mathematics -primary school -classroom assessment

<sup>a</sup> Student assessment includes self- and peer-assessment.

resulting papers were checked by a quick read. The remaining collection consisted of 904 papers that were downloaded from the CNKI database and put in EndNote.

As shown in Fig. 1, the earliest paper referring to assessment at the classroom level in primary mathematics education was published in 1985. It is about assessing the quality of students' learning (Rong, 1985). In general, before 2001 very few papers were published about assessment at the classroom level in primary school mathematics education. In 2002, one year after the publication of the curriculum reform outline (MoE, 2001a) and the mathematics curriculum standards (MoE, 2001b), a notable increase in the number of published papers on assessment can be seen. A further boost took place in 2007 when a revised version of the mathematics curriculum standards was published (MoE, 2007). Since then the number of published papers showed a steady increase.

Since these papers still involved all kinds of assessment taking place in classrooms, a further selection was necessary. However, just searching for 'classroom assessment' would not have given the intended selection of papers. As in the rest of the world, in China there are many interpretations of the term 'classroom assessment'. According to Zhong (2012), classroom assessment covers both assessment of teachers' teaching and students' learning, whereas Wang (2011) considers classroom assessment as only appraising or correcting students' performance. "Classroom teaching and learning assessment", one of the search words used in the second selection step, is also defined differently between authors, ranging from assessing teachers, to assessing students, to assessing the process and effectiveness of teaching, and to assessing teaching and learning as a whole (Cao, Li, & Qing, 2011). So, to guarantee that the papers in the review were about activities conducted by teachers to assess their students' learning, the papers had to be read more closely. However, doing this for 904 papers was not feasible. Therefore, we decided to include only the papers published in the two latest full years, 2011 and 2012, in our collection. This means that the papers in the review appeared one decade after the mathematics curriculum reform was launched in 2001. This fifth selection step brought us to 360 papers. Then, in the sixth selection step, the full papers were read to identify those that discuss classroom assessment, that is, formative assessment conducted by the teacher. This resulted in 283 papers. Finally, a seventh selection step was carried out in which the information about the authors' affiliations was used to identify the papers that were written by teachers. In fact, except for thirteen papers from

researchers and four papers from consultants, the majority papers, 266 out of 283, were written by teachers, which were used for the review. A numbered list of the references of these papers is in Appendix A (see Supplementary material).

## 2.2. Coding framework

The development of a coding framework (see Table 3) for carrying out the review in a systematic way started with reading the first fifty papers published in 2011 to find out whether the aspects of assessment for which guidelines are provided in the mathematics curriculum standards could function as suitable categories to be included in the coding framework. It turned out that most of the aspects were indeed discussed in the papers, including the purpose of assessment, the content of assessment, the person who is the assessor, and the method of assessment. With respect to reporting and using of the assessment results, it was found that most of the papers addressed mainly the aspect of reporting, that is, the provision of feedback; whether and how the assessment results were used for adapting and improving teaching was rarely described in the papers.

After the main categories were established, the sub-categories were determined. The suggested assessment guidelines in the mathematics curriculum standards were the first source for deciding possible subcategories. The other source was what was discussed by the teacher-authors. For example, according to the mathematics curriculum standards, three main purposes of assessment can be distinguished: checking students' understanding, stimulating students to learn, and informing teachers' instructional decision-making. Moreover, in the teacher-written papers, some teachers also talked about using assessment to create a harmonious classroom environment or to promote their students' confidence. So, these two were added as subcategories of the purpose of assessment in our framework.

Table 3 shows the subcategories that were used for the purpose of assessment, the content of assessment, the person who is the assessor, the methods of assessment, and the feedback given to the students. For the category of teaching adaptation no sub-category was provided, since it was only briefly described in the mathematics curriculum standards and was rarely discussed in the teacher-written papers. Examples for each subcategory and the category of teaching adaptation can be found in Appendix B (see Supplementary material). Almost all the categories and the subcategories mentioned in Table 3 are not exclusive. This means

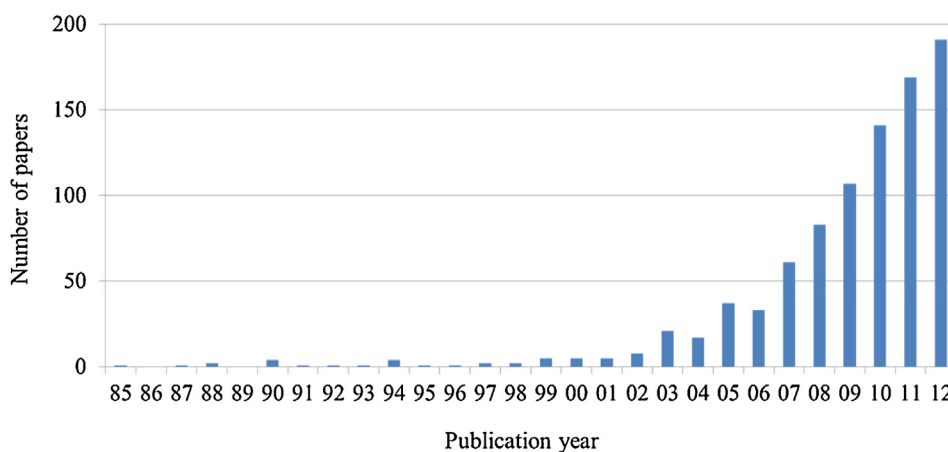


Fig. 1. Frequency distribution of papers on assessment at the classroom level in primary mathematics education found in the CNKI database during 1985–2012.



**Table 3**  
Coding framework for teacher-authors' conceptions of classroom assessment.

Category	Subcategory	
Purpose	Checking students' understanding Stimulating students to learn Informing teachers' instructional decision-making Establishing a harmonious classroom environment Promoting students' confidence	
Content	Basic knowledge and skills Mathematical thinking and problem solving Mathematical and learning attitude	
Assessor	Teachers Students (including assessing themselves and peers) Parents	
Method	Observation Questioning Classroom discussion In-class assignment After-class assignment Presentation Portfolio Quiz/written test Mathematics diary	
Feedback	Focus	Task-related Process-related Person-related
	Nature	Only positive Balanced (positive and what needs to be improved)
	Mode	Verbal Written Body language Material incentives
	Timing	Immediate (during class) Delayed (after class)
Teaching adaptation		

that multiple coding of the papers was possible. The subcategories were exclusive only with respect to the nature of the feedback (the feedback was either positive or balanced).

In order to code a paper with a “yes” for a particular category or a subcategory, two requirements from the coding protocol had to be met at the same time. The first requirement was that there must be a clear message from the teacher-author in which it is expressed, for example, in what way classroom assessment *should be* conducted. Such a message could be conveyed by the teachers' own general statements, by quotations from the mathematics curriculum standards, or by concrete examples from the classroom. However, negative statements or unfavorable examples alone were not coded, since saying what was *not* appropriate did not directly indicate what was advocated by the teacher. The other requirement was that a category or a subcategory should really be addressed by the author, rather than casually mentioned. Yet, the amount of attention that was required to consider a category or a subcategory as being addressed differed between categories. For example, how to use a mathematics diary to assess students could be treated in the whole paper, while the purpose of classroom assessment could be dealt with by only referring to it once at the beginning of the paper. Because feedback was mentioned in almost every paper and many of them provided no further explanation, coding these papers as dealing with feedback would not be informative. Therefore, the feedback in a paper was only coded if at least a quarter of the paper discussed feedback.

The coding was done by the first author of this paper. After the coding was finished, 10% of the papers (of each publication year, the

first thirteen papers in the list of the teacher-written papers in Appendix A) were reviewed by a second coder not involved in this study. The interrater reliability was found to be substantial, with  $\kappa = 0.73$ . In the 26 papers, 86% of the codes assigned by the external coder were identical to the codes of the first author.

### 3. Results

#### 3.1. Characteristics of the papers included in the review

Of the 266 teacher-written papers about classroom assessment in primary mathematics education that were analyzed, 123 were published in 2011 and 143 were published in 2012. The authors were from 26 out of the 31 provinces, municipalities, and autonomous regions in mainland China. More than half of the papers (58%) were written by teachers from the provinces Jiangsu, Zhejiang, and Fujian, which are the three provinces with the highest overall level of educational development (Wang, Yuan, Tian, & Zhang, 2013). Surprisingly no papers were found from authors from Beijing or Tianjin, which are the municipalities with the second and fourth highest overall level of educational development in mainland China (Wang et al., 2013).

The teacher-written papers mainly involve teachers' reflections on their own practice of classroom assessment or tips learned from others. The length of the papers ranges from about half a page to four pages and most papers have one or two pages, containing about 2000 to 3000 Chinese characters. These papers usually start with a statement on the importance of classroom assessment, a quotation from the mathematics curriculum standards, or an example of unfavorable assessment practice the teacher-authors encountered. Then, in the main part of the papers the teacher-authors mostly gave their own suggestions for conducting classroom assessment, illustrated by concrete examples. For example, feedback should pay attention to the strategies students used, and it is better to correct low-achievers' work face to face. In the conclusion part of the papers, the teacher-authors generally provided a short conclusion to highlight the main idea of the papers. In addition to quoting from the mathematics curriculum standards, many teachers quoted educators, psychologists, or celebrities both national – such as Xingzhi Tao – and international – such as Howard Earl Gardner or Abraham Lincoln – to support their ideas. The teacher-authors also used a variety of metaphors to describe classroom assessment, like the “GPS to guide students' learning in class”, the “catalyst of teaching and learning”, and the “booster to facilitate learning”. In 60 papers we found that classroom assessment was associated with beauty, art, or positive emotional feelings. Specifically, in 7 papers the teacher-authors talked about the beauty of the classroom assessment, like in #10 (the paper numbered 10 in Appendix A) it was said that “assessment is the most gorgeous flower in classroom teaching”, and in #101 the teacher-author said that doing classroom assessment “turned out to be a beautiful story.” In 14 papers it was mentioned that doing classroom assessment is a kind of art. Finally, in 39 other papers adjectives were used to describe emotion-related effect of classroom assessment, such as that classroom assessment makes the mathematics teaching or the class beautiful, glorious, energetic, poetic, charming, warm or sweet.

#### 3.2. Teacher-authors' conceptions of classroom assessment

Table 4 shows that of the 266 teacher-written papers about classroom assessment in primary mathematics education, 131 papers (49%) mention the purpose of classroom assessment, 187 papers (70%) address the content to be assessed, all 266 papers (100%) discuss who the assessor is, 208 papers (78%) describe the

**Table 4**  
Number of papers referring to the different aspects and its subcategories of classroom assessment (N = 266).

Category	Subcategory	Number of papers			
Purpose	Stimulating students to learn	131	106 (81%)		
	Checking students' understanding		90 (69%)		
	Informing teachers' instructional decision-making		78 (60%)		
	Establishing a harmonious classroom environment		4 (3%)		
	Promoting students' confidence		4 (3%)		
Content	Basic knowledge and skills	187	129 (69%)		
	Mathematical thinking and problem solving		105 (56%)		
	Mathematical and learning attitude		106 (57%)		
Assessor	Teachers	266	265 (100%)		
	Students (including students themselves and their peers)		139 (52%)		
	Parents		30 (11%)		
Method	Exercises in class	208	129 (62%)		
	Homework		93 (45%)		
	Questioning		83 (40%)		
	Observation		29 (14%)		
	Discussion		26 (13%)		
	Quiz/written test		25 (12%)		
	Portfolio		23 (11%)		
	Presentation		11 (5%)		
	Mathematics diary		11 (5%)		
Feedback	Focus	198	146		
				Task-related	115 (58%)
				Process-related	80 (40%)
	Person-related		72 (36%)		
	Nature		Balanced (positive and what needs to be improved)	134 (68%)	
			Only positive	64 (32%)	
	Mode		196	Verbal	153 (77%)
				Written	76 (38%)
				Body language	44 (22%)
				Material incentives	26 (13%)
Timing	194	Immediate (during class)	147 (74%)		
		Delayed (after class)	79 (40%)		
Teaching adaptation		9			

used method(s), and in 198 papers (74%) at least a quarter of the paper refers to feedback. However, in only 9 papers (3%) it was found that the teacher-authors explicated making teaching adaptation based on the assessment information.

### 3.2.1. Classroom assessment purpose

In total, we found 131 papers referring to one or more purposes of classroom assessment. Stimulating students to learn is mentioned most often (in 81% of the papers). This purpose is followed by checking students' understanding (69%) and informing teachers' instructional decision-making (60%).

In addition, a few papers refer to using classroom assessment to establish a harmonious classroom environment and to promote students' confidence. These purposes of doing assessment are both mentioned in 3% of the papers. For example, in paper #9 (see Appendix A) the teacher-author wrote that she was "using immediate assessment to create positive emotion [in the class] and to create a harmonious and democratic classroom learning environment." Similarly, in #10 the teacher-author expressed it in the following way: "The purpose is to find students' strengths

during their learning process, to give encouragement, and to create a harmonious environment for teaching and learning." In #192 the teacher-author was very clear that the purpose of assessment was to contribute to students' confidence. He even titled his paper as "Three approaches to using assessment to build students' confidence." According to him, "teachers are required to [ . . . ] help students to recognize themselves and build confidence."

### 3.2.2. Content of classroom assessment

In 187 papers the aspect of what has to be assessed is addressed. Table 4 shows that most attention is paid to assessing students' basic knowledge and skills (69%), followed by assessing mathematical thinking and problem solving (56%) and assessing students' mathematical and learning attitude (57%). In addition to the different content that is assessed, in 82 papers (44%) the teacher-authors mentioned explicitly that the various aspects of students' competences are interrelated and therefore these competences should not be assessed in isolation but together. For example, the teacher-author of #195 stated that classroom assessment "should not only focus on the understanding and

application of knowledge and skills, both also on emotion and attitude to facilitate students' overall, harmonious and sustainable development."

**3.2.2.1. Basic knowledge and skills.** We found that many teacher-authors wrote about the necessity and importance of assessing students' basic knowledge and skills, but only a few of them described explicitly how to assess this content; however, where papers contain more detailed information about how to assess basic knowledge and skills, the given examples mainly consist of bare number problems. Another finding was that some teacher-authors reflected a broad interpretation of basic knowledge and skills, meaning that they did not only focus on the correctness of the problems, but also tried to get information about how students solved them and whether they really understood them. For example, the teacher-author of paper #16 underlined that "the principle standard [to assess] is not to see whether they learned by heart a formula or can use the formula to calculate the right answers, but to see whether they know how the formula is developed" and the author of #98 "asked her students to describe how they solved the problem of 45-9".

**3.2.2.2. Mathematical thinking and problem solving.** The importance of assessing mathematical thinking and problem solving is recognized by many teachers, like the teacher-author of #44, who said: "The ultimate goal of students' learning is application of knowledge and skills they learned in real contexts. Therefore, after students have learned some new knowledge, it should be assessed not only whether they command the knowledge, but also whether they can use the knowledge to solve problems." Also, we found suggestions for how to assess this content. According to the teacher-author of #16, the teachers should design problems within a real life situation, like how much to pay when buying one hundred notebooks with discount. Providing open-ended problems was used in #98, in which the teacher asked her students to design a rectangular garden with a particular area. Furthermore, in paper #70, the teacher required his students to think of their own questions based on a frequency table and then solve their own problems posed on these data.

**3.2.2.3. Mathematical and learning attitude.** We found that many teacher-authors noticed that teachers need to keep an eye on students' learning interest, their initiatives and engagement in learning activities, and their communication and cooperation with others. To assess students' attitude several indicators were mentioned, such as how many students raised their hands to answer the teachers' questions, whether they showed self-confidence when answering questions and whether they wrote their homework in a clean and neat way. The teacher-authors of #70 and #205 also paid attention to whether their students behaved well, like whether they started to interrupt while others were still talking. And the teacher-author of #137 mentioned that he/she assessed whether the students were sitting upright in class.

### 3.2.3. Assessor in classroom assessment

In all 266 papers except for one, it was found that the teacher is the person who conducted the classroom assessment. Nevertheless, in 139 papers (52%) students also play a role as assessor. Taking paper #6 as an example, the teacher-authors described that they "organized the students to correct their exercises together in a group to develop their sense of doing assessment." In #5, the teacher-author emphasized that "by using self- and peer-assessment, the function of classroom assessment will be largely magnified. The students are not just the passive assesseees

anymore; they take part in assessment actively and experience the joy of doing assessment."

Also, in 30 papers (11%), it was found that parents are encouraged to be involved as assessors. The teacher-author of #81 wrote: "[W]ithout the participation of the parents, the assessment of students is not complete. By asking parents to observe how their children learn mathematics at home and sending the assessment results back to the teachers, the teachers can get a better understanding of the students' learning and adjust their instruction instantly." The teacher-author of paper #145 wrote that an assessment form had been designed for the parents to assess their children's performance of doing homework.

### 3.2.4. Method of classroom assessment

Table 4 shows that there are 208 papers discussing the method of classroom assessment in general and the methods that were found in these papers. The most frequently discussed method is doing exercises in class (62%). This method is followed by giving homework (45%) and questioning (40%). Less attention is paid to carrying out observations (14%), organizing a discussion (13%), doing quizzes (12%) and keeping a portfolio (11%). Rarely applied methods are asking students to do a presentation (5%) and write a mathematics diary (5%).

**3.2.4.1. Exercises in class.** Giving exercises as an assessment method can be done at the beginning of a lesson to check whether the students are ready for the new learning material. The teacher-author of paper #11 described that she "prepared ten problems of mental calculation to ask her students to solve before teaching calculation in a smart way." Also, exercises were used after the instruction of a certain topic. For example, the teacher-author of #60 used exercises in class "to know immediately to what extent the students have mastered the most important and most difficult part of the topic in this lesson." Similarly, in paper #160, after the students learned the decimals, the teacher gave the exercise: "Which '0' can be left out [without changing] the following numbers? (3.09, 0.300, 1.800, 500, 5.780 and 0.040)." When using exercises in class to assess students' learning a few teachers also provided open-ended exercises. For example, the teacher-author in #88 asked students to think of their own questions related to percentages and solve them given the situation that "there are 20 apples and 25 pears".

**3.2.4.2. Homework.** Written homework gives teachers several opportunities to get information about their students' learning. For example, in paper #127, the teacher-author described that based on the homework he found "one student made the same mistake when solving a series of problems:  $46 \times 54 + 46 \times 46 = 46 \times 3 + 54 = 192$ ,  $25 \times 99 + 25 = 25 \times 2 + 99 = 149$ , and  $99 \times 99 + 99 = 99 \times 3 = 297$ ." In #77, it was emphasized that "what students' homework looks like also could reflect students' learning habit and attitude." Furthermore, instead of assigning simple written homework, the teachers tended to provide various types of homework. Like the teacher-author of #3 asked the students to go to a supermarket and collect the prices of goods, and the teacher-author of #147 asked her students to estimate the distance between school and their homes.

**3.2.4.3. Questioning.** By asking students to answer specific questions, teachers can quickly find out where students are. Questions were usually posed to be answered by any student in the class, but only one or a few students were selected to give their answers. For example, in paper #66, the teacher provided a true-or-false question: "Cutting a round piece of paper into two parts, then one of the two parts will be  $\frac{1}{2}$  [of the piece of the paper], true



or false?” She found that some students said “true” and some said “false”. Also, a few teachers raised questions to be answered by the whole class at the same time. The teacher-authors of #214 described an example that “[s]tudents with a prime [student number] are required to raise their right hands; while those with a [student] number that has a divisor other than 1 and itself raise their left hands.” In addition, teachers also mentioned that only asking questions that can be answered by a simple “yes” or “no” is not enough. For example, the author of paper #69 reported that teachers should use more “why” questions to look closer at students’ thinking. Furthermore, the teachers noticed that the questions should be clear to the students, and sufficient time should be given to guarantee students can think carefully.

**3.2.4.4. Observation.** In their writings the teachers made it clear that the observation took place when students were involved in various activities, such as doing exercises, answering a question, and discussing with their neighbors. In these observations an eye could be kept on the students’ mathematics learning by focusing on how many students solved a problem, what strategies were used, and what mistakes were made. The author of #83 emphasized that such observations require that teachers “observe their students through the lens of mathematics for finding information about students’ mathematics learning.” In addition, the observation was used to throw light on the students’ attitude and social behavior. For example, by observing how a student raised her hand, the author of #64 noticed that the student was lacking in confidence to give her opinion. Furthermore, in a few papers, for example in #10 and #116, the teacher-authors made observation forms to guide their classroom assessment practice.

**3.2.4.5. Discussion.** The method of organizing a discussion was chosen because in this method, according to the teacher-authors, more students could be involved and engaged in the assessment. Moreover, discussions are considered to have the potential to reveal students’ deeper thinking. In the papers we found both discussion in small groups and in the whole class. The topics of discussion mainly involve the key concept of a lesson, different strategies of solving a problem and students’ mistakes. For example, the author of paper #8, after the instruction, assessed his students by letting them discuss the relationship between rectangle and square in groups. In paper #38, it is described how an experienced teacher organized an assessment by asking his students to discuss three different answers to the problem of factorizing 36.

**3.2.4.6. Quiz/written test.** From the papers it can be derived that quizzes are normally held after a period of learning, for example at the end of a chapter. Quizzes mostly have a broad coverage of assessed content and generally take about one lesson. Quizzes, contrast with the method of giving exercises which is often used immediately after the instruction, is mostly focused only on the topic taught in that lesson, and needs only a few minutes. In the papers, the teacher-authors agreed that doing quizzes is a powerful way to elicit information about students’ mastery of basic knowledge and skills. However, they can also be used to assess students’ deep understanding and ability to solve problems. For example, the authors of #201 and #232 suggested to reduce problems which demand rote memory and include well-designed context problems or open questions, while papers #16 and #61 recommended to provide quizzes with different difficulty levels and offer the students the opportunity to choose a quiz that fits their own situation.

**3.2.4.7. Portfolio.** The references made in the papers to keeping a portfolio as a method to assess students’ learning make it clear that both teachers and students can decide what to put in the portfolios and that a wide variety of files can be collected, for example, a student’s best homework, a creative solution to a problem, a summary of what a student learned and some typical mistakes a student made. In the papers we found two main reasons for using portfolios to assess students. According to the teacher-author of paper #132, using portfolios could help teachers to “track individual students’ longitudinal development.” Furthermore, the teacher-authors of paper #165 valued the advantage of using portfolios to improve students’ self-assessment, because “when students are responsible for making decisions about what to put into their portfolio, they get the opportunity to reflect on their learning and assess themselves.”

**3.2.4.8. Presentation.** Only a few papers mention asking the students to give a presentation as a way of assessing them. Yet teacher-authors who referred to this method emphasized that one can get deep insight into students’ understanding in this way. In paper #213, the students were asked to do a self-study about a new topic and give a short presentation about what they have learned. This let the teacher find out whether this new topic was easy or difficult for the students. Another example is from the teacher-author in #220. He asked his students to give a presentation about their self-designed problems and their solutions.

**3.2.4.9. Mathematics diary.** The few teacher-authors who mentioned the mathematics diary are quite positive about this assessment method. Asking students to keep a mathematics diary is considered as a rich resource to reflect all aspects of students’ mathematics learning. In their mathematics diaries, students can ask questions or explain their confusion, report mistakes they made and their corrections, describe phenomena they found in daily life which are related to mathematics, or tell a story about mathematicians. The teacher-author of #252 included part of a student’s mathematics diary in the paper:

After we learned calculation in a smart way, Teacher Chen reminded us to summarize where it is easily to make a mistake. I noticed I need to be more careful in two situations [namely, when removing parentheses or adding parentheses without changing the original problem]. At that moment, I thought how it was possible for me to make such simple mistakes. However, it turned out I did make an error in my homework. [ . . . ] it was just because I was careless. How ashamed I am!

The possibility to have a wide scope in this assessment method is reflected in paper #68. Here, the teacher-author made it clear that “based on students’ mathematics diaries, teachers can assess students’ learning about knowledge and skills. More importantly, teachers are able to discern students’ learning interest and attitude as well.”

In addition, in 36 papers we found that the teacher-authors mentioned designing tasks which included exercises in class and homework, mainly for students to practice more and to learn from these tasks. Only in 17 of these papers did the teacher-authors mention the design of questions and quiz items to assess their students. For both purposes they emphasized that more context problems and open-ended problems should be offered to students. Moreover, some teacher-authors pointed out that it is important to offer “fun” problems to students and to offer different problems to individual students.

### 3.2.5. Feedback

The aspect of assessment that is mostly addressed in the reviewed collection of teacher-written papers on classroom

assessment is providing feedback to students. Although we only counted a paper as referring to feedback when substantial attention is paid to this aspect of assessment, this still was the case in 198 (74%) of those papers. Actually, in 89 of them nearly the whole paper is dedicated to the issue of giving students feedback.

A further finding was that in 64 papers the teacher-authors seemed to consider classroom assessment as equivalent to feedback. For example, according to the teacher-author of #144, assessment is “providing students with feedback about their learning situation.” The author of #158 made it clear that: “classroom assessment is, during the process of teaching and learning, the positive or negative judgements made by assessors based on assessee’s performance.” In paper #111 the teacher-author used interchangeably the terms “课堂评价 (classroom assessment)” and “课堂评价语言 (classroom assessment language)”. She gave as an example of her ‘assessment’: “Cao is very shy in the discussion and she did not dare to express her thinking.” Also we found that when the teacher-authors reflected on their practice of classroom assessment, they sometimes did not give any information about the concrete questions they asked their students, but only mentioned their reactions after the students gave their answer. Like the author of #166 who summarized: “[W]hen students reply with a wrong answer, I can say ‘you are not far from the right answer’ or ‘I know you have thought it over, shall we first listen to other students?’” The same was done by the author of #119. When she described her assessment practice, she just summed up the possible reactions she could give, such as: “[T]o students who are careless, I can say ‘I hope you can say goodbye to carelessness and make friend with carefulness’ or ‘your handwriting is very beautiful. It is much better if you can be careful [while solving problem].’”

With respect to the different aspects of feedback found in the papers, Table 4 shows that in the 146 papers that discuss the focus of feedback, the feedback is mostly task-related (58%), followed by process-related (40%) and person-related (36%) feedback. As for the nature, in the 198 papers that refer to this aspect of feedback we found fewer papers in which the teacher-authors only thought of giving positive feedback (32%), that is praising students, and more papers in which the teacher-authors reflected a balanced way of providing feedback (68%). The latter means that the feedback is not completely positive, but also contains clear information about what needs to be improved.

The mode of feedback was found to be addressed in 196 papers with verbal feedback (77%) as the dominant mode, followed by written feedback (38%) and feedback given by body language (22%). Using material incentives as feedback (13%) is mentioned only in some papers. In the 194 papers that give information about the timing of providing feedback we found that much more feedback was given during class (74%) than after class (40%).

### 3.2.6. Teaching adaptation

Only in 9 papers, the teacher-authors mentioned how results of classroom assessment were used for adapting and improving their instruction. From the examples described in #38 and #161, we found that the teachers gave supplementary exercises immediately after finding their students could not answer the questions correctly. Teacher-authors of #86 and #219 described that teachers are required to analyze students’ results on written tests and students’ mistakes in order to adapt their further teaching. Similarly, in paper #234, the teacher-author recommended to analyze students’ mistakes in their homework to make instruction meet students’ needs. In the remaining papers, the teacher-authors mentioned that they thought of using or had used assessment results for instructional decision making.

### 3.3. Relation between assessment conceptions in the papers and the assessment guidelines

The papers revealed that the teacher-authors took the assessment guidelines in the mathematics curriculum standards as a source for their classroom assessment practice. Evidence for this could be found in 142 papers. The teacher-authors explicitly stated that their classroom assessment activities are in line with what is advocated in the curriculum reform. For example, the teacher-author of #94 wrote: “[T]he idea of assessment in the [2001] curriculum reform is student-development-oriented; I explored how to assess my students [based on this new idea] in my classroom teaching.” Other teacher-authors clearly paraphrased the guidelines or even exactly cited them. In total, in 118 papers literal quotations were found.

#### 3.3.1. Purpose of assessment

Regarding the purpose of assessment, the document of the mathematics curriculum standards emphasizes that the main purpose of assessment is to (1) “get the whole picture of process and outcomes of students’ mathematics learning”, (2) “stimulate students to learn”, and (3) “improve teachers’ instruction” (see MoE, 2011, p. 33). All three purposes are also mentioned by the teacher-authors, with stimulating students to learn discussed most often. Furthermore, we found in 41 papers that the teacher-authors used the entire and exact wording of the assessment guidelines. In addition, the teachers also mentioned two other purposes, namely establishing a harmonious classroom environment and promoting students’ confidence. However, these were only found in a very few papers.

#### 3.3.2. Content of assessment

The descriptions in the papers about the content that is assessed reflect that the teacher-authors’ conceptions are quite in line with the assessment guidelines. For the content, in 55 papers exact quotations from the guidelines were found. For example, the teacher-author of paper #54 wrote: “[A]ccording to the mathematics curriculum standards, ‘as for assessing students’ mathematics learning, students’ understanding and using mathematics knowledge and skills should be focused on. Students’ development of emotion and attitude also need more attention.’” Although all the competence domains – basic knowledge and skills, mathematical thinking and problem solving, and mathematical and learning attitude – are mentioned in the papers, basic knowledge and skills received most attention from the teacher-authors.

#### 3.3.3. Assessor in assessment

Concerning the assessor, the teachers’ conceptions reflected in the papers adhere to the multi-actor assessment system that is promoted in the assessment guidelines. But in the papers, teachers play a dominant role in conducting classroom assessment. Nevertheless, evidence and examples of self- and peer-assessment and parents as assessors can also be found. In paper #64, the teacher-author described: “[T]he [2001] mathematics curriculum standards document points out when assessing students’ learning, self- and peer-assessment should be organized.”

#### 3.3.4. Method of assessment

Most of the assessment methods recommended in the assessment guidelines in the mathematics curriculum standards were also found in the papers. This particularly applies to the exercises in class and after class (“homework” in the papers) and to the oral tests (“questioning” in the papers). To a lesser degree the papers mentioned doing observations, keeping portfolios, and administering written tests (“quizzes” in the papers). However, while the importance of written tests is emphasized in the

mathematics curriculum standards, its use was only found in 25 papers.

### 3.3.5. Report and use of assessment results

The only aspect of assessment for which we found a real difference between what the mathematics curriculum standards document intends and what is in the papers is the report and use of assessment results. Although in the papers much attention is paid to the report of assessment results, that is, to the provision of feedback, which echoes the guidelines, few discussions focus on the use of classroom assessment results for adapting and improving teachers' further instruction to meet the students' needs.

## 4. Conclusion and discussion

### 4.1. Teachers' conceptions of classroom assessment

Through their papers the teacher-authors gave a rich picture about their conceptions of classroom assessment (Research question 1). They described various aspects of their classroom assessment practice, including the purpose, content, assessor, method and feedback, and illustrated these aspects with examples. Moreover, they gave many reflections on how they assess their students. However, though many teacher-authors explicitly stated that improving their instruction is one of the purposes of conducting classroom assessment, in only a few papers discussions were found about instructional decision-making based on the information gained by their classroom assessment activities. One possible reason for this might be that teachers find it difficult to use assessment-based information for adapting further instruction, as it was found by, for example, [Heritage, Kim, Vendlinski, and Herman \(2009\)](#). A second possible reason could be related to the teacher-authors' narrow scope of classroom assessment. Many of them seem to consider classroom assessment equivalent to providing feedback. This view on assessment is, for example, revealed in some papers in which the teacher-authors either defined classroom assessment as providing students feedback or explained their classroom assessment practice by just reflecting on how they reacted or will react when encountering certain student performances or facing students with certain characteristics. Giving such explanations indicates that the teacher-authors thought their practice of classroom assessment started only after their students' performance. In fact, for some of these teachers, classroom assessment is what a teacher says or writes, in such a way that it helps their students to improve. This is also evidenced by their use of the terms "classroom assessment" and "classroom assessment language" interchangeably, implying that for these teachers the language they use when assessing actually is the assessment.

Undoubtedly, providing feedback to facilitate students to move forward is one of the key strategies for effectively implementing classroom assessment ([Black & Wiliam, 2009](#)), and the type of feedback and the way it is given matter its effectiveness of enhancing students' learning ([Hattie & Timperley, 2007](#)). From this perspective, it is encouraging to find that how to give feedback was widely discussed by the teacher-authors and that the feedback they gave was mainly task- and process-related, balanced, and timely. Nevertheless, we were surprised by the huge amount of attention that is paid to feedback, especially when comparing this with the attention that is paid to teaching adaptation. Teachers' competence to adapt their instruction based on evidence is critical for effective teaching. If teachers cannot make instructional decisions according to the assessment information, the promise of classroom assessment to improve students' learning will be impaired ([Heritage et al., 2009](#)).

In addition, it is interesting that some teacher-authors referred to the beauty aspect of classroom assessment. In fact, in the body of assessment literature known to us and mostly based on studies carried out in the Western world, we never came across references to the aesthetic appeal of assessment. It could be that this approach is typical for how teachers in China conceive education and assessment, but we did not find any substantial evidence for this in other studies. Another possible reason for emphasizing the art and beauty aspect of assessment might be that the teacher-authors would give their paper a good reception by making the topic of assessment more attractive for readers.

### 4.2. Relation between teachers' conceptions and the assessment guidelines

Regarding the relation between the classroom assessment as reflected in the teacher-written papers and the assessment guidelines in the mathematics curriculum standards (Research question 2), it is evident that the latter has exerted a great influence on inviting primary school mathematics teachers to think about, to discuss, and to share their use of classroom assessment. From 2002 on, one year after the curriculum reform outline ([MoE, 2001a](#)) and the mathematics curriculum standards ([MoE, 2001b](#)) were published, a steady increase in the number of published papers on classroom assessment was found. In 2011 and 2012, a decade after the advent of the assessment reform, a great number of 266 published teacher-written papers were found to reflect on different aspects of classroom assessment included in the assessment guidelines. By analyzing the papers, we found that the teacher-authors clearly considered these guidelines as a source for conducting their classroom assessment. The teachers often used literal quotations from the documents or paraphrased the guidelines when discussing the purpose of classroom assessment, the content, the assessor, the method and the report of assessment results, that is, giving feedback. The only mismatch with the assessment guidelines lies in the teachers' use of the assessment results. While it is suggested in the mathematics curriculum standards that the assessment results can be used for adapting and improving instruction, this aspect of assessment was hardly addressed in the teacher-written papers. Maybe the reason for this is that the document of the mathematics curriculum standards only briefly describes the use of assessment results.

A further difference between the classroom assessment discussed in the papers and what is suggested in the mathematics curriculum standards relates to the use of written tests or quizzes. Although the guidelines in the latest version of the mathematics curriculum standards ([MoE, 2011](#)) state that written tests are important and give detailed suggestions about designing and conducting these tests, only 25 papers mentioned this method. A possible explanation for this discrepancy is that the 2001 version ([MoE, 2001b](#)) was rather reluctant in emphasizing the use of written tests in primary education. Taking into account that the reviewed papers were published in 2011 or 2012, it is understandable that the reevaluation of written tests in the latest version ([MoE, 2011](#)) is not already reflected in the teachers' papers.

When looking back at our analysis and results, indeed we found that the teachers' conceptions of classroom assessment reported in their papers generally were in agreement with the assessment guidelines. However, together with this positive finding, in retrospect we have noticed that the assessment guidelines also may have affected the teachers' conception of classroom assessment in another way. Although the structure of the assessment guidelines in the Chinese mathematics curriculum standards is very clear by addressing all the key aspects of performing classroom assessment specifically and providing suggestions or examples for each aspect, this structured presentation also can

have a disadvantage. By delivering the message about the new approach to assessment aspect-by-aspect, it becomes maybe not so clear for teachers how these aspects of classroom assessment function as a whole and how to embed them in their daily teaching practice. Therefore, this structure may hinder teachers to get a holistic picture of how classroom assessment works – which is reflected in the teachers' papers – and this might jeopardize the implementation of classroom assessment.

Possible improvements regarding to the presentation of the assessment guidelines can be put on the agenda of the policy makers of the Chinese mathematics curriculum standards. For example, detailed suggestions about how teachers can make use of the assessment results to adapt and improve their teaching are needed. Moreover, after addressing all the key aspects of assessment separately, explanations about how these key aspects work as a whole and how to integrate assessment into teachers' practice should be provided. Finally, it may be helpful to give teachers concrete examples of conducting assessment, which illustrate, in the context of assessing specific mathematics content, for what purposes certain assessment methods are chosen to use, possible assessment results teachers and students may receive, diverse feedback that teachers can provide to students, and various instructional adaptations teachers can do for their further teaching.

#### 4.3. Limitations and recommendations for future research

As our search in the CNKI database has shown, assessment at the classroom level in primary mathematics education is a rather recent educational phenomenon in China and is rapidly gaining ground. How this reform in assessment took place and is still going on, and what were or are the important change agents in getting this reform implemented in educational practice would be very relevant information for further improving education. However, due to the limited time for carrying out this review study we could only include the papers published in 2011 and 2012. To gain a complete picture of the change in classroom assessment, in future research more years could be included. Another extension that is recommended is to have a closer look at the papers written by researchers and consultants, whose views were left out in our review due to the very small number of such papers that were found in the database.

Finally, although teacher-written papers are a valuable source for getting more knowledge about teachers' conceptions of classroom assessment, the findings from the 266 teacher-written papers need to be interpreted with prudence, since only teachers who had their papers on classroom assessment published were included in this study and the teachers' reports for publication may have been affected by social desirability and publication bias. Another concern is that “there are profound gaps between what people know, what they think they know, what they say, and what they do.”<sup>1</sup> Therefore, it is unsure whether what is written in the teacher-written papers can be considered fully equivalent to what the teachers really think of assessment. Further sources could include directly asking teachers about their classroom assessment or carrying out classroom observations to describe what appears to be happening in classrooms. Conducting a survey to further investigate Chinese primary mathematics teachers' use of assessment will be the next step in our research.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.stueduc.2016.12.002>.

## References

- Andrade, H. L. (2010). Students as the definitive source of formative assessment. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Abingdon, UK: Routledge.
- Announcement about reporting and evaluating the price of materials instantly (1949). *Governmental Journal of Jiangxi*, 2, 55–60 (In Chinese).
- Assessment Reform Group (1999). *Assessment for learning: Beyond the black box*. Cambridge: University of Cambridge School of Education.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25.
- Berry, R. (2011). Educational assessment in mainland China, Hong Kong and Taiwan. In R. Berry, & B. Adamson (Eds.), *Assessment reform in education* (pp. 49–61). Netherlands: Springer.
- Black, P., & William, D. (1998a). Inside the black box. Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.
- Black, P., & William, D. (1998b). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
- Briggs, D. C., Ruiz-Primo, M. A., Furtak, E. M., Shepard, L. A., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, 31(4), 13–17.
- Brookhart, S. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, 106(3), 429–458.
- Cao, Y., Li, J., & Qing, H. (2011). A review of classroom instruction and learning assessment in mathematics education in China. *Mathematics Bulletin*, 8, 1–5 (In Chinese).
- Curriculum Development Council (CDC) (2002). *Basic education curriculum guide: Building on strengths (Primary 1/Secondary 3)*. Hong Kong: Author.
- De Lange, J. (1999). *Framework for classroom assessment in mathematics*. Madison, WI: NICLA/WCER.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Heritage, M., Kim, J., Vendilinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice*, 28(3), 24–31.
- Herse, R. (1979). Some proposals for reviving the philosophy of mathematics. *Advances in Mathematics*, 31(1), 31–50.
- Keeley, P., & Tobey, C. R. (2011). *Mathematics formative assessment: 75 practical strategies for linking assessment, instruction, and learning*. Thousand Oaks, CA: Corwin.
- Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37.
- Leung, F. K. (2008). In the books there are golden houses: Mathematics assessment in East Asia. *The International Journal on Mathematics Education*, 40(6), 983–992.
- McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative assessment on student achievement: So much more is needed. *Practical Assessment, Research & Evaluation* 18(2) . . (Retrieved from) <http://pareonline.net/getvn.asp?v=18&n=2>.
- Ministry of Education of the People's Republic of China (MoE) (2001a). *Curriculum reform outline of basic education (trial version)*. . . (In Chinese Retrieved from) <http://moe.edu.cn/publicfiles/business/htmlfiles/moe/s8001/201404/167343.html>.
- Ministry of Education of the People's Republic of China (MoE) (2001b). *Mathematics curriculum standards of nine-year compulsory education (trial version)*. Beijing: Beijing Normal University Press (In Chinese).
- Ministry of Education of the People's Republic of China (MoE) (2007). *Mathematics curriculum standards of nine-year compulsory education (2011 version)*. . . (In Chinese. Retrieved from) <http://wenku.baidu.com/view/3b7a694233687e21af45a93e.html>.
- Ministry of Education of the People's Republic of China (MoE) (2011). *Mathematics curriculum standards of nine-year compulsory education (2011 version)*. . . (In Chinese, Retrieved from) <http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/167340.html>.
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics (2013). *Formative assessment: A position of the National Council of Teachers of Mathematics*. National Council of Teachers of Mathematics. . (Retrieved from) <http://www.nctm.org/about/content.aspx?id=37990>.
- Ni, Y. J., Li, Q., Li, X. Q., & Zou, J. (2011). China's new millennium curriculum reform in mathematics and its impact on classroom teaching and learning. In T. D. Huang, & A. W. Wiseman (Eds.), *The impact and transformation of education policy in China* (pp. 99–124). Bingley: Emerald.

<sup>1</sup> We thank an anonymous reviewer for suggesting this phrasing.



- Romberg, T. A., Zarinnia, E. A., & Collis, K. F. (1990). A new world view of assessment in mathematics. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 21–38). Washington, DC: AAAS.
- Rong, Z. (1985). The design and use the worksheet for analyzing the quality of learning: Management of teaching and learning in primary mathematics education. *Theory and Practice of Education*, 5, 23–26 (In Chinese).
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Shi, N., Ma, Y., & Liu, X. (2012). The revision process and the content revised in the mathematics curriculum standards of nine-year compulsory education. *Curriculum, Teaching Material and Method*, 32(3), 50–56 (In Chinese).
- Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory Into Practice*, 44(1), 11–18.
- Stiggins, R. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan*, 83(10), 758–765.
- Suurtamm, C., Koch, M., & Arden, A. (2010). Teachers' assessment practices in mathematics: Classrooms in the context of reform. *Assessment in Education: Principles, Policy & Practice*, 17(4), 399–417.
- Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, 15(2), 105–127.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York, US: Macmillan.
- Topping, K. J. (2010). Peers as a source of formative assessment. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 61–74). Abingdon, UK: Routledge.
- Van den Heuvel-Panhuizen, M. (1996). *Assessment and Realistic Mathematics Education*. Utrecht, the Netherlands: CD Beta Press.
- Van den Heuvel-Panhuizen, M., & Becker, J. (2003). Towards a didactical model for assessment design in mathematics education. In A. J. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Second International Handbook of Mathematics Education* (pp. 689–716). Dordrecht: Kluwer Academic Publishers.
- Veldhuis, M., & Van den Heuvel-Panhuizen, M. (2014). Exploring the feasibility and effectiveness of assessment techniques to improve student learning in primary mathematics education. In C. Nicol, S. Oesterle, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education*: (5. pp. 329–336).
- Wang, S., Yuan, L., Tian, Z., & Zhang, X. (2013). The Comparative Research on the level of regional development of education. *Research in Education*, 6, 29–41 (In Chinese).
- Wang, X. (2011). Classroom assessment should focus on students' situation. *Education of Liaoning*, 11, 45–46 (In Chinese).
- Wiliam, D. (2007). Keeping Learning on Track: Classroom assessment and the regulation of learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053–1098). Greenwich, CT: Information Age Publishing.
- Wiliam (2011a). *Embedded formative assessment*. Bloomington, IN: Solution Tree.
- Wiliam, D. (2011b). What is assessment for learning? *Studies in Educational Evaluation*, 37, 3–14.
- Zhong, Q. (2012). The challenges of classroom evaluation. *Global Education*, 1(10), 10–16 (In Chinese).
- Zhu, L. (2012). What has the mathematics curriculum standard for compulsory education (2011 version) revised. *Journal of Mathematics Education*, 21(3), 7–10 (In Chinese).