Article

Teachers' Use of Descriptive Assessment in Primary School Mathematics Education in Iran

Marja Van den Heuvel-Panhuizen [1,2,*], **Ali Akbar Sangari** [1] **and Michiel Veldhuis** [1,3]

[1] Utrecht University, 3584 CS Utrecht, The Netherlands; a.a.sangari@uu.nl (A.A.S.); m.veldhuis@ipabo.nl (M.V.)
[2] Nord University, 8026 Bodø, Norway
[3] iPabo University of Applied Sciences, 1061 AD Amsterdam, The Netherlands
[*] Correspondence: m.vandenheuvel-panhuizen@uu.nl or m.vandenheuvel-panhuizen@nord.nl

Abstract: In many countries, assessment and curriculum reforms came into being in recent decades. In Iran, an important educational assessment reform took place called Descriptive Assessment (DA). In this reform, the focus of student assessment was moved from a more summative approach of providing grades and deciding about promotion to the next grade to a more formative approach of providing descriptive feedback aimed at improving student learning. In this study, we evaluated how seven fourth-grade mathematics teachers used the principles of DA. Data were collected by a questionnaire on assessment practices and beliefs, lesson observations, and interviews. Although the teachers varied in how they assess their students, in general their assessment practice is by and large in line with the DA guidelines. Nevertheless, in some respects we found differences. When assessing their students, the teachers essentially do not check the students' strategies and when preparing the report cards, they still use final exams because they do not sufficiently trust the assessments methods suggested by DA. The guideline to use assessment results for adapting instruction is also not genuinely put into action. The article is concluded by discussing dilemmas the teachers may encounter when implementing DA.

Keywords: descriptive assessment; assessment practice; assessment reform; mathematics education; primary school teachers; Iran

1. Introduction

1.1. A Worldwide Change in Assessment in Mathematics Education

In recent decades, a significant change has taken place worldwide in the field of educational assessment. Instead of using assessment to measure whether students acquired a particular skill or understanding at the end of a learning process, which largely defined the 20th-century approach to assessment, the 21st-century international assessment agenda shows signs of increasing recognition of the use of assessment for learning purposes [1]. This change in approach to assessment can be found in many countries in both the Western and Eastern worlds, in which assessment previously was mainly used for the aim of selection and accountability, e.g., [1–3]. At the turn of the century, a general need arose to use assessment as a tool to support learning and improve teaching, and to place greater emphasis on the process by which teachers seek to identify and diagnose student learning difficulties in order to provide students with quality feedback to improve their achievement. Besides this call for assessment to promote successful learning, the pressure to reform assessment also came from the demand for better data to inform educational decision-making and the necessity of assessing a broader range of skills than is possible with traditional assessment approaches [4]. The latter demand was largely related to the major turnaround regarding the what and how of teaching mathematics that took place around the 1990s. A pivotal feature of the then-advocated-for view of mathematics education consisted of broadening the goals to be pursued [2]. As proposed in steering curriculum documents published in the USA [5], the UK [6], and Australia [7,8], this implied that in mathematics

education more attention had to be paid to fostering students' ability to problem solve, carry out investigations, model problem situations, and communicate mathematical ideas. For the sake of this reformed conception of the mathematics curriculum, several countries, including the USA, the UK, Australia, and the Netherlands [2], also started to explore assessing students in ways that were sensitive to these new standards by emphasizing the use of, for example, open-ended tasks and contextualized settings and rejecting time restrictions when students are assessed.

This process of coming to an assessment reform following a curriculum reform—or, in the reverse direction, using reformed assessment as a leverage of a curricular reform [2]—is in line with the conception of assessment as an integral part of the teaching–learning process [9–11]. If the main purpose of assessment is to inform instruction, assessment and instruction should be epistemologically consistent [12]; or, to put it differently, a lack of consistency between both would cause a conflicting situation between new views of instruction and traditional views of assessment [13].

The approach to assessment as an integral part of instruction also means a change in the traditional assessment policy and practice by assigning an increasingly larger role to the teacher [12,14,15]. After all, teachers are in the best position to collect information about the learning of their students [16]. By using assessment as an ongoing process interwoven with instruction, teachers can acquire direct information to make adequate instructional decisions and adapt their teaching to students' needs in order to improve their learning. Several studies have shown that teachers' assessment activities can indeed have a positive effect on student performance, e.g., [17–23].

As a result of these promising findings, teachers' assessment practice has become a key factor in improving mathematics education and has been put on the policy agendas in many countries [1]. Therefore, all over the world, investigations have been conducted to find out mathematics teachers' current assessment practice and beliefs on assessment, e.g., [24–31].

Like in many other countries, a reform of the educational system has also taken place in Iran, accompanied by a reformed approach to assessment that includes a more important role for educational evaluation in the classroom, and has a larger focus on supporting students' learning than on classifying them [32]. However, in English-language educational research literature little can be found about these developments in Iran. Our intention with the present study is to make up for this omission and fill this knowledge gap by investigating how primary school mathematics teachers in Iran perform assessment in their teaching.

1.2. A Change in Assessment in Mathematics Education: The Case of Iran

1.2.1. The Iranian Education System

Iran has a highly centralized education system, in which the Supreme Leader and several autonomous councils within the central government are involved in proposing, approving, and supervising the educational policy. The Supreme Leader of Iran is Ayatollah Ali Khamenei. He is the highest-ranking spiritual leader and the most powerful official in Iran. As supreme leader, he has control over the executive, legislative, and judicial branches of the government, as well as over the military and media. The autonomous councils include the "Supreme Council of Cultural Revolution" (which makes many high-level political decisions; most importantly, it determines the principles of cultural policy and goals of educational policy) and the "Supreme Council of Education" (which is involved in all educational policy decision-making about primary and secondary education; everything that happens in education in Iran should first have been officially approved by this council). The Ministry of Education, with the deputy ministries, organizations, and centers that belong to it, is in charge of controlling and handling all practical aspects of the educational levels from preschool to grade 12. This means that the Ministry of Education is responsible for developing curricula and textbooks, publishing and distributing educational materials, planning and conducting professional development

and education for teachers, and carrying out student assessments and examinations (see http://timssandpirls.bc.edu/timss2015/encyclopedia/countries/iran-islamic-rep-of/; accessed on 7 July 2020).

The education system in Iran is divided into five cycles: preschool, primary education, lower secondary education, higher secondary education, and higher education. Schooling starts with a one-year preschool for five-year-old children to prepare them for entering primary school. There is no examination at the end of this cycle and children proceed automatically to the following cycle. Primary education lasts six years, covering grades 1 to 6, and is meant for children from six to 11 years old. At the end of each academic year students have to take exams on the basis of which it is decided whether they are promoted to the following grade. At the end of grade 6, students take a provincial examination. Students who pass this examination receive an elementary school certificate and are qualified to proceed to lower secondary education, also called the middle or guidance cycle. This cycle covers grades 6 to 8 and is meant for children from 11 to 13 years old. This cycle is concluded with an assessment to determine students' placement in the academic/general or the technical/vocational stream of higher secondary education. This last cycle of formal education covers the grades 9 to 12 and is meant for students from ages 14 to 17 (see http://wenr.wes.org/2017/02/education-in-iran; accessed on 7 July 2020).

1.2.2. The Need for an Educational Change in Iran

Although over time the education system and curricula in Iran have undergone several changes, a real radical change in the mathematics curriculum happened in 2011, including a modification in the chosen approach, the content organization, and the context [33,34]. One of the driving forces of this reform, like in many other countries, was the unexpected low performance of the Iranian students in Trends in International Mathematics and Science Study (TIMSS) of 2007 [35]. To improve students' performance a new pedagogical and curriculum approach was proposed with much attention to having students actively involved in learning mathematics and using "real world" contexts for almost every mathematics concept and skill [34]. This new way of thinking about education also differed largely from the traditional approach to assessment, which was mainly used to make judgments about whether students passed the end-of-grade or -school examination. In this approach, much attention was placed on quantitative assessment and too little on student–teacher interaction, and especially excessive emphasis on final scores was found to cause damage to students' creativity and take away students' opportunity to develop higher cognitive skills [36]. Therefore, to overcome these difficulties new ways of assessment were developed to combat the challenges the educational system faces [32,37].

1.2.3. Toward a New Assessment Approach in Iran

Different from what happened in many other countries, in Iran the thinking about a different way of assessing students did not come after a new curriculum was put into operation. The new approach to assessment was developed concurrently with the National Curriculum that was approved in 2012, and it even foreshadowed this new curriculum in some sense. In 2002, the Supreme Council of Education (SCE) [38,39] officially demanded that the Ministry of Education work on a revision of the existing assessment system, to decrease the stress of negative competition caused by score-oriented assessment and to enhance teacher–student interaction by giving feedback. At the end of 2002, this work led to proposing a project called Descriptive Assessment [40]. Soon after that, the Supreme Council of Education published the "Principles of Assessing Educational Achievement" [41]. In this short document, an explicit connection was made between teaching and assessment by describing the goals of assessment (e.g., improvement of teaching and learning), the place of assessment in the teaching–learning process (e.g., being part of it), what should be assessed (e.g., students' solution processes), what should be taken into account when designing assessments (e.g., students' views, attitudes, and skills), how they should be

assessed (e.g., with a variety of assessment methods), and who is responsible for the assessment (e.g., the teacher and the school).

A few years later, the Supreme Council of Education published the *Academic Assessment Regulation for Primary School* [42] and stipulated that from 2009 on teachers in primary school had to use Descriptive Assessment (DA). Characteristic of this qualitative, process-oriented approach to assessment is that teachers collect and document evidence of student learning by a variety of methods, on the basis of which they have to provide students and their parents with rich descriptions of students' strengths and weaknesses accompanied by concrete advice about the areas that need improvement. (The term "descriptive assessment" seems not to be unique for the assessment reform that is developed in Iran. Ndoro et al. [43] also coined this term, but in their research, it had another meaning. Here descriptive assessment is an approach to strengthen desirable behavior in children. The aim of descriptive assessment as they used it was to improve the effectiveness of adult–child interaction by using directive prompting to achieve the children's compliance with instruction. Nevertheless, from this guidance and steering perspective there is certainly a commonality with the emphasis on giving descriptive feedback in the descriptive assessment as developed in Iran.)

In the *Teacher's Guidance for Descriptive Assessment* [44] the aforementioned assessment goals are further operationalized. As the first goal, it is stated that the assessment has to enrich the teaching–learning process by considering assessment an essential and integral component of the teaching–learning process. The assessment should head for supporting the students' learning rather than only classifying students and giving grades. Moreover, the focus should be on getting students to achieve the curriculum goals instead of merely following the textbook. The second goal is that the assessment should identify and develop students' talents and should focus more on the assessment of students' developmental trends and pay attention to assessing non-cognitive domains of learning, such as creativity. The third goal is in line with this and emphasizes that the students' multidimensional development should be facilitated by adapting the teaching–learning environment to encourage students to develop in this way. The fourth goal underlines the importance of raising parental awareness and involvement in the students' development. The fifth goal aims to motivate students by involving them in the assessment process, for example by self-assessment and encouraging their creativity. The sixth goal is about using the descriptive developmental information about students over the school year to decide about their attainment of the educational goals and making the decision about grade promotion. The final and seventh goal is to use the gathered student data not only to provide students, teachers, and other interested parties with useful information about the students' development but also to use this information to reconsider the adequacy of teaching methods and find ways to improve and enrich them.

Apart from the goals, this *Teacher's Guidance for Descriptive Assessment* [44] also explains how teachers can assess students' achievement by collecting and documenting evidence regarding their learning and performance. To do so, different methods can be used, including observations of students during the classes, interviews, student portfolios, performance tasks, and various kinds of tests. Characteristic of this new approach to assessment is that much attention has to be paid to giving feedback to the students. This can be provided orally throughout the academic year and two times a year it has to be done in writing by filling in two types of a descriptive report card. The A form of this card is meant to discuss the student's progress with the parents in a parent–teacher meeting. On this form the student's performance is described in one of the four-categories of descriptive feedback: "very good," "good," "acceptable," and "more work is needed." At the end of the grade this card also contains the decision about the student's grade promotion. The B form is used by the teachers to explain in much more detail how students' performance develops and is explicitly used to give the teacher of the next grade a clear picture of a student's knowledge and skill level. This form is put in the student portfolio at school, in which the teacher keeps all the results of tests and any other documents that show student's learning

process. Every student has their own portfolio. Students' parents can see and check their child's activity when they are at the parent–teacher meetings.

The goal of the current study was to explore whether and how this new approach to assessment is used in the classroom practice. More specifically our research question was, *In what ways do primary mathematics teachers in Iran follow the guidelines of Descriptive Assessment?* To answer this research question, we related teachers' actual assessment practice to the ideas about DA that are articulated in the official documents in Iran. We carried out a multiple-case study in which the beliefs, attitudes, and classroom assessment practice of seven fourth-grade teachers were investigated by means of three research methods: a questionnaire survey, semi-structured interviews, and classroom observations.

2. Materials and Methods

2.1. Participants

This multiple-case study was based on seven fourth-grade female mathematics teachers (age $M = 46$, {min, max} = {38, 54}) from six schools in Tehran (see Table 1 for background information on these teachers). In the last three years, they had all attended the professional development course "New Assessment Approach," which was meant to inform teachers about DA. The teachers were part of a larger group of Iranian teachers who participated in TIMSS 2015. We sent these teachers a written questionnaire with questions about their beliefs on and practices of assessment in mathematics education. Of the 38 teachers in Tehran to whom we sent the questionnaire, 17 filled it in and returned it. These 17 teachers were contacted by phone with the request to participate in interviews and classroom observations. Two teachers could not participate because they were not teaching in Grade 4 anymore and eight teachers could not participate due to the Iranian regulations for doing research in schools. In the end, seven teachers participated in the interviews and classroom observations, held between April and June 2016.

Table 1. Background of the teachers.

Teacher	Years of Teaching Experience	Academic Degree	Teaching Certification for Primary School	Number of Students
❶ [1]	4	Bachelor's degree in Chemistry	No	28
❷	26	Elementary Education	Yes	32
❸	30	Elementary Education	Yes	30
❹	26	Elementary Education	Yes	33
❺	20	Bachelor's degree in Arabic Literature	No	35
❻	28	Elementary Education	Yes	36
❼	30	Elementary Education	Yes	34

[1] The seven dots refer to the seven teachers.

The study was conducted in accordance with the Declaration of Helsinki. To contact schools and collect data we obtained approval from the Tehran General Education Department. Representatives from participating schools and the teachers signed letters of consent and they collected signed permission forms from the parents to conduct observations and making videos of the students in class. Based on this it can be stated that all subjects gave their informed consent for inclusion before they participated in the study.

2.2. Data Collection

The data we used for this case study came from four sources. In spring 2016, the data collection started with administering a questionnaire on teachers' assessment practices in mathematics education. The second source of data was the pre-observation interview with teachers about their assessment plans for the lesson to be observed and their beliefs and attitudes about assessment in mathematics classes. The third source of data was the classroom observation of teachers during a mathematics class to see how they used descriptive assessment in teaching mathematics. The fourth source of data was the post-observation interview to examine what teachers learned from the assessment activities.

2.2.1. Assessment Questionnaire

The teachers responded to an Iranian version of a questionnaire on teachers' assessment practices in mathematics education. The questionnaire was based on a Dutch questionnaire [27] also used in a study in China [45]. The questionnaire consisted of 40 questions related to background information of the teachers, including their age, experience, and education, as well as questions about their assessment practice, the frequency of assessment, the purpose of assessment, the methods they used, their assessment beliefs, their views regarding the relevance of assessment methods, and whether they considered assessment an interruption of their teaching or a means to improve their teaching.

2.2.2. Pre-Observation Interview

In this semi-structured and videotaped interview the teachers were questioned about their lesson plans and the assessment activities (or activity) that would be carried out in the lesson observed by the researcher. The teachers were also asked whether these assessment activities were typical for this class and what the purpose was of the assessment activities. The focus in the questions in the pre-observation interview was on revealing what kind of assessment the teachers had planned for the upcoming lesson and what the reasons behind it were.

2.2.3. Content of the Observed Lesson

The lesson of all seven teachers was an introduction to probability from Chapter 7, Section 2, of the textbook *Mathematics in Grade Four* (ریاضی چهارم دبستان), which is the official mathematics textbook in Iran. We chose to observe the same lesson for all teachers in order to get insight into their assessment choices while keeping the mathematical content identical. In this lesson, students are taught to judge the probability of the occurrence of particular events, using five qualitative sentences, namely, "It does not happen at all," "There is little chance it happens," "There is equal chance it happens," "There is a high chance it happens," and "It surely happens." In the textbook, two main exercises are proposed to students. In the first exercise, they have to decide which of the probability sentences applies most to the probability of picking a marble of a particular color from bags that are filled with three or four marbles of different colors (see Figure 1). In the second exercise, students have to judge the probability of particular outcomes of spinners (see Figure 2). In the *Teacher's Guidance for Descriptive Assessment* the purpose of this chapter is described as understanding the concept of probability and gaining skills in determining the expectancy of possible outcomes of an event.
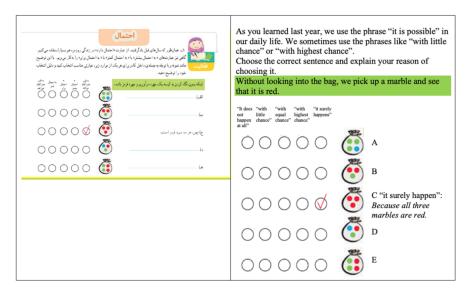

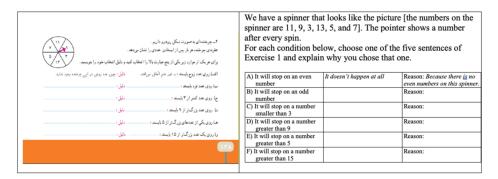**Figure 1.** Exercise 1 of Chapter 7 of the grade 4 textbook.

Figure 2. Exercise 2 of Chapter 7 of the Grade 4 textbook.

2.2.4. Observation of Classroom Activities

The previously described lesson was observed and videotaped. The aim of the observations was to collect occurrences of assessment activities carried out by the teachers during the lesson. For example, we tried to discern what kinds of assessment methods the teachers used during the lesson, the purpose of using them, and their plan for using the collected evidence and gathered information. Finally, it was important to find out to what degree teachers' assessment methods corresponded with the DA guidelines.

2.2.5. Post-Observation Interview

Following each classroom observation, a semi-structured interview with the teacher was conducted and videotaped to reflect on their taught lesson. The focus was to find out whether the teacher was able to collect the information she was aiming for, what she would do with the information to inform and improve her instruction, and whether the assessment activity went as planned. In other words, we asked the teachers to explain what they learned from the assessment activities and how they were going to use the information gained from the assessment activities.

2.3. Document Analysis

To compare the teachers' use of assessment with the use as it is intended in the DA, approach, we developed a framework for analyzing the collected data. To develop this framework, we determined and collected all the important official documents related to DA. To do this we contacted different governmental organizations that are responsible for primary education (Ministry of Education, Department of Primary School) and educational assessment (Ministry of Education, Department of Primary School, Unit of Assessment), which provided us with the official regulations and transcripts of meetings of the Supreme Council of Education. In this way, we were able to base our framework on the guidelines for assessment that are given in the *Teacher's Guidance for Descriptive Assessment* [46] and the *Academic Assessment Regulation for Primary School* [42].

The assessment guidelines provided in these two documents cover the entire assessment process, from gathering information to following up on the assessment information. In total, we identified five key categories of guidelines: directions for (A) the collection of student data, (B) the actors involved in the assessment, (C) the focus of the assessment, (D) the documentation of the assessment findings, and (E) the follow-up of the assessment. In the following, we describe what the documents suggest about these key categories.

2.3.1. Collection of Student Data (A)

The first key category concerns the assessment methods the teachers are suggested to use. The focus is on the way student data should be collected. In total five possible methods are described: observing students in class (A.1), asking questions in class (A.2), administering written tests (A.3), providing performance tasks (A.4), and having students working on a project (A.5).

For observing students in class (A.1) both structured and unstructured ways are mentioned. When teachers use checklists, rating scales, or a specific way of logging student work, this is considered a structured observation. In an unstructured observation, teachers observe students when working individually or on group assignments such as a project. The purposes of the observation are checking students' answers, collecting information about the quality of their answers, and giving them hints for improvement. Asking oral questions in class (A.2) may be done at the beginning and the end of a lesson. These questions may be posed to individual students or to all students in class. The questioning can be done randomly as well as purposely to collect some specific information form particular students. Regarding the use of written tests (A.3), the guidelines suggest that they can be designed by one or several teachers. The tests should fit to the assessed grade level. Administering written tests may be done daily, weekly, or monthly or when moving from one chapter to the next. The questions in the test should cover the topic of the chapter involved. The teachers can use the textbook as a framework for designing written tests. The evaluation of the students' work on the written tests should be done by writing descriptive feedback on the students' papers. In addition, teachers may also review and check students' answers by asking some students to solve questions on the blackboard. Providing performance tasks (A.4) is a method of collecting student data in which students are asked to work on practical tasks. This method is suggested because while students are working on these tasks, teachers can see and assess the process and products of students' learning and gather information on their reasoning processes. Having students working on a project (A.5) is a method suggested in the guidelines particularly meant for assessing whether students have internalized and can apply what they have learned so far. Projects give students the opportunity to present both results and procedures. The difference between performance tasks and written tests is that these student projects are more complicated and specifically include a set of goal-oriented activities.

2.3.2. Actors Involved in the Assessment (B)

The second key category concerns the actors who are performing the assessment or who are being assessed. The teacher can assess individual students (B.1) and groups of students (B.2), and students can assess peers (B.3) and themselves (B.4).

Assessing individual students (B.1) is considered to be important for teachers to know the academic performance of the students. Tracking individual students can help teachers to accurately and reliably fill in individual students' report cards. Several ways are suggested for assessing individual students. The teachers can call on students to directly answer a question or to evaluate the response a classmate provides to a question or assignment. For assessing groups of students (B.2) the teachers may give them a group activity based on the objectives of a lesson. It is recommended that the group have a leader and that the students work together and try to learn from each other during the group work. The results of the activity may be reported to other groups. Both the teachers and the other students can assess students' work and provide descriptive feedback. The teacher can use the results of such group activities to fill in students' descriptive report cards. Peer assessment (B.3) is suggested to be used by teachers to create chances for criticizing and assessing other students' performance by their peers. When students assess their peers, students answer other students' oral questions, report the results of an activity, or work on assignments on the blackboard in front of other students. Finally, the guidelines mention students assessing themselves (B.4) as a mean to increase students' responsibility for their own learning. Thinking about their answers and activities may affect the learning process positively. For example, students can evaluate their own work by answering questions such as, "Why do you think your answer or solution is correct?" In this way, the students can assess what they have learned and correct the potential inferential errors personally.

2.3.3. Focus of the Assessment (C)

The third key category concerns the teachers' focus when checking students' responses to assignments. The assessment guidelines emphasize that the teacher leads the assessment during the academic year and is responsible for the quality of information and data. The teacher has to observe the students' performances and outcomes during and at the end of tasks and has to record the essential points. The recording of data should be in agreement with the educational goals to be interpretable for the students and the parents. Teachers are recommended to take notes and describe students' learning process comprehensively and write notes that involve the analysis of students' performance on tasks and tests. Two general ways of checking responses are suggested: checking answers (C.1) and checking strategies (C.2).

When teachers focus on checking answers (C.1), students' written answers are the main source of information about students' learning. The teacher guide lists acceptable answers and how these answers should be obtained. When teachers focus on checking strategies (C.2), students' written or verbal solution strategies are the main source of information. In this way teachers can also value alternative solutions to assignments (different from the prescribed textbook solution) if they want to.

2.3.4. Documentation of the Assessment Findings (D)

The fourth key category concerns how teachers document their findings from the assessments of their students' learning. Three possible ways of documentation are suggested: reporting the performance in a notebook (D.1), keeping a portfolio (D.2), and filling in the report cards (D.3).

Reporting students' performance in the teachers' notebook (D.1) implies the traditional way of recording and bookkeeping students' performance on written tests, teacher observations, and notes related to teaching and students' other assessments. The guidelines require teachers to describe a student's performance in a sentence instead of a grade between 0 and 20. Teachers have to use this information when they fill in the students' report cards. Keeping a portfolio (D.2) is also suggested. This means that they have to collect students' performance on written tests in a folder to be discussed when meeting students' parents. Parents are supposed to visit school at least two times in every semester. Teachers can also ask students to bring the portfolio home to show their work to their parents. In this way teachers can involve parents in the students' learning processes.

The report cards (D.3) have two different versions. Version "A" is designed for topics addressed in a particular chapter of the textbook. This version comes from the *Teacher's Guidance for Descriptive Assessment* [46]. For example, for mathematics in Chapter 7 of the Grade 4 textbook (see Section 2.2.3), the card contains the categories "understanding the concept of probability" and "acquiring skills in expectancy of possible outcomes of an event." Version "B" of the report cards is more general and meant to be used in all subjects and in every grade. It should be used to report students' achievements at least twice in an academic year for every subject, using the descriptive feedback scale "very good," "good," "acceptable," and "needs more training and attempts."

2.3.5. Assessment Follow-Up (E)

The fifth key category concerns the teacher's actions after having used assessment methods and gathered information on students' learning. Three possible teacher actions are suggested: giving descriptive feedback (E.1), discussing descriptive report cards (E.2), and adapting further instruction (E.3).

Giving descriptive feedback (E.1) implies that the teachers are supposed to give clear information to students about their performance after they have been assessed by written tests, observation, performance tasks, student presentations, or any other assessment method. Descriptive feedback has to consist of a sentence that shows students' learning difficulties and contains some hints on how to improve and work towards their specific learning goals. When teachers discuss the descriptive report cards (E.2), they should

provide the students with further background information about their learning. Moreover, the descriptive report cards should also be discussed with the parents. Adapting further instruction following assessment (E.3) is one of the main goals of descriptive assessment. This entails that teachers should modify, improve, and enhance the teaching–learning process in the classroom by encouraging effective teacher–student interactions. Teachers can adapt further instruction by interpreting the data and information they collect.

The aforementioned five key categories of the DA guidelines were transformed into key questions and resulted in the following framework (see Figure 3) that was used to analyze the collected data.

```
Teacher's classroom assessment practice
A. How are student data collected?
        A.1 By observing students in class
        A.2 By asking questions in class
        A.3 By administering written tests
        A.4 By providing performance tasks
        A.5 By having students working on a project
B. Who is assessed or assessing?
        B.1 Individual students are assessed
        B.2 Groups of students are assessed
        B.3 Students assess peers
        B.4 Students assess themselves
C. What is the focus on in assessment?
        C.1 On checking students' answers
        C.2 On checking students' strategies
D. How is the assessment documented?
        D.1 By reporting students' performance in the notebook
        D.2 By keeping a portfolio
        D.3 By filling in the descriptive report cards
E. How is the assessment followed-up?
        E.1 By giving descriptive feedback
        E.2 By discussing descriptive report cards
        E.3 By adapting further instruction
```

Figure 3. Framework for analyzing the collected data.

2.4. Analysis of the Collected Questionnaire, Interview, and Observation Data

First, transcriptions were made of the videotaped interviews and lessons. Since these data and the responses to the questionnaire are in Farsi, they were all translated into English. The correctness of the transcription and translation was checked by a mathematics educator from Iran. Based on the feedback we got from her we adapted the English data accordingly.

The analysis of the data started with making for each teacher a separate report of the responses given in the questionnaire, the notes taken during the observations, and the transcripts of the interviews. Hereafter, we organized the data per category and subcategory. This means that the information we gained from the seven teachers were put together and structured in frequency tables, to which we added examples of the teachers' assessment practices and how they thought about assessment. During the process of quantitative and qualitative data analysis, our findings were checked over and over again and several rounds of discussion took place to come to the following results.

3. Results

3.1. How Are Student Data Collected? (A)

In their answers to the questionnaire Teachers ❶❷❸❹❺❻❼ indicated that they are **observing students in class (A.1)** several times a week. Teachers ❶❷❸❺ reported that observing students while they are doing exercises provides very relevant information about their students' mathematics skills and knowledge. Neither during the observation of the teachers' lessons nor during the interviews were indications found that the teachers made observations in a structured way. None of the teachers used checklists, rating scales, or logs to structure their observations. Observing students in class as an assessment method was also mentioned in the pre-observation interview. This was done by Teachers ❶❻❼.

Teacher ❶ said, "*I will observe students when they work, answer their questions, and correct them individually.*"

During the classroom observations, it was noted that Teachers ❶❷❸❺❼ observed students while they worked on textbook assignments. Teacher ❶ told her students, "*Do not forget to think! I will come to you and check your answer when you work on your assignment.*" While observing her students, she noticed a group of students who answered the first exercise (see Figure 1) in the wrong way. The students tried to solve the assignment by actually using the marbles (instead of analyzing the situation and determining the theoretical probability that was aimed for), which led them to obtain different results. Therefore, Teacher ❶ asked them to put the marbles away and try to find a true result based on the a priori information. She tried to explain the concept of theoretical probability; after this, the students changed their approach and corrected their answer. Teacher ❷ also informed students while observing their work on the second exercise (see Figure 2) and she tried to motivate her students and told them, "*I would like to see who is the first one in finding the answer to the assignment.*" During the observation, she encouraged students by saying to the class, "*Well done! Some answers are very well.*" Checking students' answers when observing them was for some teachers of great importance. This was, for example, the case for Teacher ❺. When she saw a student who did not write the correct answer in his textbook, she went over to him and asked him to write the correct answer. In addition, Teacher ❼ observed students while they answered the written test questions and gave them tips, such as, "*Read this question carefully.*" She allowed one student to read the written test question, "*We threw a coin 5 times, if we throw it one more time what is the result and why?*" to all students because he did not understand the meaning of the question. The student answered, "*We should throw a coin in this situation!*" The teacher replied, "*No, just write your expectation.*"

In the post-observation interview Teachers ❶❷❸❺❼ clarified that the main purpose of observing students is checking students' answers.

The teachers' responses to the questionnaire told us that **asking questions (A.2)** is done in two different ways: either to all students or to individual students. Teachers ❶❷❺❼ ask individual students questions several times a week and Teachers ❸❹❻ weekly. Asking the whole class is done by Teachers ❷❹❻❼ several times a week, by Teachers ❶❺ weekly, and by Teacher ❸ monthly. Teachers ❷❸❼ indicated that posing questions in class to all or individual students provides very relevant information about students' mathematics skills and knowledge.

During the classroom observation it was noticed that Teachers ❶❷❸❹❺❻❼ started teaching by using oral questions. Teacher ❶ started her teaching by asking all the students general questions, such as, "*Who knows the meaning of probability?*" Then all students provided the answer to this question at the same time. Teacher ❹ asked specific students questions, trying to make a connection between students' prior knowledge and the new subject. She asked questions such as, "*What do you know about probability? What did you learn about probability last year? Could you give me an example of an event with low probability?*"

In the post-observation interview, Teacher ❶ explained her reasons for asking individual students questions in class. She clarified it as follows: "*I try to ask and involve all the students poor and strong, in particular, the noisy students!*" Teacher ❷ also gave details about her reasons for asking questions. She pointed out, "*In this way, students join the lesson, and I can assess their knowledge. I like to involve students during the teaching. I start the lesson with some questions to students. It helps me to explain more.*" Teacher ❻ cleared up the difficulty of asking questions in class and her solution to overcome this:

> *I will not try to ask a question to all students, because it makes noise. I try to ask questions individually; I will start from left to right of class to ask questions or randomly based on my notebook. In other words, I do it to keep the students concentrated.*

Teacher ❼ brought to the fore that she uses oral questioning to design written tests. Her explanation was:

When I ask oral questions, the students who know the answer raise their hand and this tells me how many students understood the question. If a few students answer the question I try to include questions like that in the exam.

Regarding teachers' practice and beliefs about **administering written tests (A.3),** information was brought together about the frequency of administering written tests, the perceived relevance of using written tests, the sources that are used to design them, the purpose for administering them, and the format of the test problems. Teachers ❶❷❸❹❺❻❼ reported in the questionnaire that they use written tests weekly, covering the lessons of that week. Teachers ❶❸❻❼ indicated that they see them as very relevant for providing information about students' learning. In the pre-observation interview Teachers ❶❸❺❻❼ explained that in addition to these weekly tests they give monthly a more overall test or an end-of-chapter test monthly. Moreover, Teachers ❶❷❸❹❺❻❼ mentioned that they administer a final exam at the end of each semester.

In the pre-observation interview it came to the fore that Teachers ❶❷❸❹❺❻❼ design written tests by themselves. Teachers ❶❷❸❹❻❼ use student textbooks and supplementary material, which they slightly adapt for this. Teacher ❼ expressed this as follows: "*I tried to design the tests based on student's textbook. For example, the textbook talks about colored beads, but I replaced it with numbers.*" Teachers ❷❺❻❼ made it clear that they discuss the design of written tests with a colleague. Teacher ❷ explained, "*When I want to design a written test, I talk sometimes with my colleagues in the same grade. [. . .] Sometimes she writes and designs an examination and the other times me.*" Other sources that were mentioned for designing tests were the use of test item banks (Teacher ❷), making use of students' mistakes (Teacher ❼), and making use of a teachers' own experience (Teachers ❺❻). In general, the teachers use at least two sources to design written tests, but Teacher ❼ uses four sources. She said in the pre-observation interview, "*I am famous for taking too many exams from students between my colleagues,*" and, "*I believe that my students are more successful because of it.*"

In the pre-observation interview Teachers ❶❷❸❹❺❻❼ indicated that the main purpose of administering written tests is to understand students' learning. Teachers ❶❹❺ mentioned motivating students, and teacher ❹ made it clear that for her a main purpose was moving to the next lesson. She expressed this as follows:

When I teach if I ask students "did you learn it?" they say "yes." During an assessment, they understand that they have a lot of questions because they are really involved in the problems. For example, in the last session, we had an examination about bar charts, students asked me questions like "from which number do we have to start," "how do we have to draw it," or "which number do we have to write where." I did not hear these kind of questions when teaching!

With respect to the format of the test problems, the questionnaire and pre-observation data revealed that Teachers ❶❷❸❹❺❻❼ use both open-answer and closed-answer formats when administering written tests. They do this in a balanced way. Teacher ❷ indicated being very in favor of using open-answer questions and being negative about closed-answer questions:

I don't like multiple-choice questions because I am not sure about students' responses. But if I use open-answer questions, I can understand if it is his reply or he used answers of his classmate. I always say to my students to be yourself during the exam. But they are young! They do not listen to me! Sometimes their answer is by chance. If I use open answer questions it helps me to know in which part they need help or in which part they understand.

On the issue of **providing performance tasks (A.4),** data was gathered about the types of performance tasks used (using a tool, giving a presentation, designing something, writing a text), the purpose of using performance tasks, and the frequency and relevance of using them. In the questionnaire it was revealed that Teachers ❶❷❸❹❺❻❼ use performance tasks, but their frequency of using them differs a lot. Teacher ❹ reported that she uses performance tasks only several times a year, whereas Teachers ❶❷❸❻ reported doing this

monthly, Teacher ❺ weekly, and teacher ❼ several times a week. Teachers ❺❼ as well as teachers ❶❹ explained that they consider providing performance tasks a very relevant way of assessing students' learning. However, Teachers ❶❷ stated that for them providing performance tasks is a time-consuming activity, therefore they only offer performance tasks when they have enough time, otherwise they prefer to use written tests. The mentioned purposes for using performance tasks were improvement of students' learning (Teachers ❶❷❸), motivating students (Teachers ❶❼), and having students involved in the learning process (Teachers ❷❹❻).

About the types of performance tasks used, the pre-observation interview brought to the fore that one of the most mentioned types of performance task was students asking to use a tool to solve a problem. This was mentioned by Teachers ❶❷❸❹❻❼. Asking students to write a text was brought up by teachers ❸❹❺❻. Giving a presentation was signaled by Teachers ❶❷❹❺❼. Teacher ❹ mentioned that she uses giving a presentation for discussing homework: "*If students do homework, especially when they are on holiday, they have to present their answer in class as a presentation.*" Teacher ❺ said, "*After having corrected students' written test, I ask a student to come in front of the blackboard and solve the written test question. Other students can check their answers and correct it if they made a mistake.*" Performance tasks in which students have to design something were brought up by Teachers ❶❸❹❺. For example, Teacher ❸ explained that she is asking her students to design a game, whereas Teacher ❶ mentioned that she asks, as homework, for her students to think of a probability sentence for a problem from the textbook.

In the observed lesson, the use of performance tasks in which the students have to work with a tool was seen in the lessons of Teachers ❶❸❹❻❼ when the students were asked to use physical tools such as coins, spinners, and marbles to solve the probability tasks in Figures 1 and 2. These teachers asked students randomly to come in front of class and to use a big bag of marbles to solve the teachers' questions. Teacher ❸ divided her students into two-person groups and gave a coin to every group and asked them to toss it six times. Beforehand she asked students to write down their expectations about the number of thrown heads and tails. In the end, the students reported their results one by one and discussed their findings in class. Giving a presentation as a performance task was observed in the lessons of Teachers ❶❷❹❺❼. For example, Teacher ❶ designed a group activity in which every group had to present the results of their group to another group. Moreover, the students had to answer questions the teacher posed during the presentation.

Assessment by **having students working on a project (A.5)** was not explicitly asked in the questionnaire nor in the interviews. This method was also not mentioned when the teachers were requested to add further methods to assess their students. Furthermore, this method was not witnessed in the observed lessons.

3.2. Who Is Assessed or Assessing? (B)

The data showed that **assessing individual students (B.1)** is a clear part of a teacher's classroom activity. In the questionnaire and during the pre- and post-observation interviews, Teachers ❶❷❸❹❺❻❼ mentioned that they assess individual students at least several times a week.

For **assessing groups of students (B.2)** the questionnaire data revealed that Teachers ❶❸❹❻❼ do this at least monthly. Teachers ❷❺ do this several times a year. During the observation, group assessment was only seen in the lessons of Teachers ❶❸.

In the questionnaire it came to the fore that Teachers ❶❷❸❹❺❻❼ are **using peer assessment (B.3)**. Teachers ❶❸❹❻ indicated using this method several times a year or yearly, whereas Teachers ❷❺❼ mentioned that they use it at least weekly. Teachers ❹❺ selected peer assessment as a very relevant assessment method. During the observation peer assessment was only seen in the lesson of Teacher ❶.

In the questionnaire, **using self-assessment (B.4)** was reported by Teachers ❶❷❸❹❺❻❼. Part of them, Teachers ❷❸❺❻❼, use self-assessment weekly or several times a week, Teacher ❹ use it only monthly, and Teacher ❶ only yearly. Teachers ❸❻ selected self-assessment as a

very relevant assessment method. The same was also done by Teacher ❶. When interviewing the teachers, it turned out that Teachers ❶❷❸❹❺❻❼ actually had difficulties trusting the results of both peer- and self-assessment. They explained that they find these methods helpful for the students, but that students in general are not capable of coming up with a trustworthy evaluation.

3.3. What Is the Focus on in Assessment? (C)

In the pre-observation interview it became apparent that **checking students' answers (C.1)** is done by Teachers ❶❷❸❹❺❻❼. In general, they do this by asking students to come the blackboard. The students have to solve the problems by following the steps suggested in the textbook. Alternative ways are not permitted. If an answer and the solution method are correct, the other students have to write them in their notebook. To check the students' answers given in the written tests, the teachers provide descriptive feedback, but this is mostly limited to just writing "correct" or "incorrect" (see also E.1), which actually means that the focus is just on students' answers and not on their strategies.

Only Teacher ❶ explained in the pre-observation interview that she is also **checking students' strategies (C.2)**. This checking was observed in her lesson as well. The teacher was examining her students when they worked on their assignments. After the students answered the oral questions, the teacher reacted by saying, "Why?" and, "Explain more." In the case the students' answers were not correct, she gave them support to reach a correct answer by providing them with a practical example. For example, she asked a student to put a colored bead in the bag. In the post-observation interview the teacher mentioned that she learned something new from TIMSS: "*In the TIMSS, I saw that some questions that students can answer with different strategies. After seeing this I decided to ask the students to explain their answers.*" The teacher indicated that she was eager to know students' reasoning and strategies.

3.4. How Is the Assessment Documented? (D)

Teachers ❶❷❸❹❺❻❼ answered in the questionnaire that they are **reporting students' performance in the notebook (D.1)** to use later when preparing the descriptive report cards. The notebook they use is designed and suggested by the *Teacher's Guidance for Descriptive Assessment* [46]. In the pre-observation interview Teachers ❶❷❸❹❺❻❼ indicated that according to them, this notebook does not have enough space to report the students' performance in a descriptive way.

Based on the questionnaire it was found that **keeping a portfolio (D.2)** is done by Teachers ❶❷❸❹❺❻❼. To build up such a portfolio these teachers ask students to save a sample of their work, such as a written test, monthly or at least several times in a year. In the post-observation interview, Teachers ❶❷❸❹❺❻❼ added to this that they use and check the students' portfolios when they are filling in the descriptive report cards. Only Teachers ❺❻ thought that portfolios provide very relevant information about the students' learning. Teachers ❶❷❸❹❼ explained that they do not find keeping a portfolio very relevant. They indicated that in most of the cases they just collect the student work to put it in the portfolio without checking it, because they do not have enough time to do this.

The questionnaire data revealed that documenting the students' learning results by **filling in the descriptive report cards (D.3)** is done by Teachers ❶❷❸❹❺❻❼. These cards have to be filled in for each student at the end of each semester. The teachers use all the information they collected during the semester for this. In the pre-observation interview it was found that Teachers ❶❷❸❹❺❻❼ in addition also administer a final exam to get extra information. This final exam largely determines what is on the descriptive report cards. However, during the interview it also turned out that the teachers think that the descriptive report cards cannot only be based on the results of this final exam. All teachers made it clear that the use of a final exam is an approach that comes from the old assessment regulation. The new DA guidelines do suggest that teachers not use a final exam when preparing the report cards at the end of the semester.

3.5. How Is the Assessment Followed Up on? (E)

In the answers to the questionnaire it was shown that for Teachers ❶❷❸❹❺❻❼ the assessment is followed up on by **giving descriptive feedback (E.1)**. The frequency and the kind of feedback that is given differed between teachers. Oral feedback is given several times a week by Teachers ❶❷❸❹❺❻❼. Written feedback is given several times a week by Teacher ❼, weekly by Teachers ❶❷❸❺❻, and monthly by Teacher ❹. Feedback to individual students is given several times a week by Teachers ❶❸❹❻❼ and weekly by Teacher ❷❺. Feedback to the class as a whole is given several times a week by Teachers ❷❸❺❼, weekly by Teacher ❻, and monthly by Teachers ❶❹.

In addition, differences in the content of the given feedback were found. Feedback to explain why an answer is wrong is given by Teachers ❶❷❸❹❺❻❼. This is done several times a week by Teachers ❶❺, weekly by Teachers ❷❹❻❼, and monthly by Teacher ❸. Giving hints was mentioned by Teachers ❶❷❸❹❺❻❼. They are mostly given by Teachers ❶❷❺❼, who do this several times a week. Teachers ❸❻ give weekly hints and Teacher ❹ monthly. Giving feedback with an extra assignment for the students to practice is done weekly by Teachers ❶❷❸❹❺❻❼.

In the pre-observation interview Teachers ❶❷❸❹❺❻❼ made it known that they consider giving descriptive feedback useful for the students, because in this way they can explain to their students why an answer is wrong. Yet at the same these teachers added that giving descriptive feedback implies a lot of work. Teachers ❷❸❹ mentioned that therefore they just write "correct" or "incorrect." Teacher ❶ expressed her difficulties with descriptive feedback as follows: "*I write a sentence as descriptive feedback on students' work as descriptive feedback. I know, descriptive feedback is useful for students, but it is extra work for me as teacher.*"

With respect to **discussing descriptive report cards (E.2)** the pre-observation interviews brought to light that Teachers ❶❷❸❹❺❻❼ discuss the results on the descriptive report cards with the students and with the parents. The latter is done by sending them a copy of the card or by inviting them to school. Similarly, the students have to take the portfolios home to show them to their parents. In this way parents are provided the opportunity to monitor the achievement of their children.

The questionnaire data revealed that using assessment results for **adapting further instruction (E.3)** is done monthly by Teachers ❸❹❺❻ and several times a year by Teachers ❶❷❼. Regarding the way they adapted their teaching, Teachers ❶❷❸❹❺❻❼ indicated that they use assessment data several times a year to change the speed of their lessons by spending more time on a subject and having the students solving more similar problems. As is described with respect to giving descriptive feedback (E.1), all teachers use assessment data to explain why answers are wrong, give hints, and give an extra assignment for practice. These measures can also be seen as adapting further instruction. In the post-observation interview Teachers ❶❷❸❹❺❻❼ explained that the assessment provides practical information for adapting instruction, but they added that based on the plan to finish the student textbook completely by the end of the academic year, they find it difficult to adapt their instruction. However, when the teachers were asked what they learned from the assessment, Teachers ❶❷❸❹❺❻❼ made it clear that they believe that assessment helps the students, not the teacher. Changing the instruction in the sense of teaching the subject again in another way in the next lesson seemed not to be their interpretation of adapting further instruction. Teacher ❶ expressed this as follows: "*I will change my teaching in this content when I am a teacher in this grade next year.*"

3.6. Teachers' Experiences with DA and the Use of DA

Besides providing answers to the aforementioned five questions, the collected data also brought more information about DA and the teachers' thoughts, experiences, and concerns to light. When the teachers were asked in the questionnaire about their agreement with the statement that DA has changed their teaching, Teachers ❶❹❺❻❼ agreed and Teachers ❷❸ agreed completely. In the pre-observation interview Teachers ❶❻ indicated

that they believe that DA reduces competition and hence stress. Teacher ❶ added to this that she thinks that lower stress is good for students. Teacher ❻, however, made the point that low competition causes lower achievement.

Teachers ❷❹❺ explicitly voiced some concerns about the help that is provided to utilize DA. Teacher ❷ found it difficult that DA is understood differently. She experienced that her colleague received different instruction for documenting evidence of learning. She was told in the course that the notebook is very important to be filled in, but her colleague received the information that keeping a portfolio is more important than recording the assessment results in the notebook. According to Teacher ❺ it was not helpful that the content of the course was very general. Therefore, she underlined that it cannot be used for mathematics lessons. Similarly, Teacher ❹ mentioned that the teacher guide for DA does not help mathematics teachers, because it does not provide applicable information, a framework, or an example to show what DA is in a mathematics class.

3.7. Overview of the Teachers' Classroom Assessment Practices

To give an overview of the findings of the assessment practices of the seven teachers, Table A1 in the Appendix A shows what we found about the five key categories of the DA guidelines that we detected in the *Teacher's Guidance for Descriptive Assessment* [46] and the *Academic Assessment Regulation for Primary School* [42]. For each teacher (A) how they collected student data, (B) who was involved in the assessment, (C) what the focus of the assessment was, (D) how the assessment findings were documented, and (E) how the assessment was followed up on is displayed. Next, Table A2 in the Appendix A clarifies for how many teachers particular classroom assessment practices were in line or not in line with the DA guidelines.

In the category of collecting student data (A), the teachers used all the suggested DA methods but one. They did carry out observation in class, ask questions in class, administer written tests, and provide students with performance tasks, but they did not assess students by having them work on a project. Another deviation from the DA guidelines is that the teachers, when they did observations, did not use a systematic approach such as using checklists or rating scales. Furthermore, another significant difference from the DA guidelines is that the teachers at the end of the semester still administered a final exam to fill in the descriptive report cards.

Regarding the actors involved in assessment (B) we found that the teachers mostly assessed individual students. Assessing groups of students was less frequently done. With varying frequencies, all teachers indicated that in line with DA guidelines, they did use peer- and self-assessment. However, they did not see these methods as very relevant for gaining information about their students' mathematics skills and knowledge. They thought students were not capable of coming up with trustworthy evaluations.

The focus on assessment practice (C) was clearly on checking the correctness of the answers. Only one teacher indicated that she also checked the strategies her students used to solve the problems. In this respect the assessment practice of the group of teachers involved in this study was clearly inconsistent with the DA guidelines.

When it comes to the documentation of the assessment results (D) all teachers followed the DA guidelines. They used the notebook, kept portfolios, and filled in the descriptive report cards. About the notebook that is included in the *Teacher's Guidance for Descriptive Assessment* [46], all teachers commented that there was not enough space for taking notes. Keeping portfolios as a method was not highly valued. Five of the teachers explained that they did not find keeping a portfolio very relevant for acquiring knowledge about their students' learning.

The last step in the assessment cycle is making use of the assessment information. According to the DA guidelines, the assessment should be followed up on (E) by giving students descriptive feedback, discussing the descriptive report cards with the students and their parents, and using the information gained from the assessment to adapt further instruction to tailor it to the needs and possibilities of the students. All teachers found

giving descriptive feedback very helpful. It was clearly a regular part of their assessment practice. They provided the feedback both orally and in writing, and to individual students as well as to the class as a whole. As far as the type of given feedback is concerned, they explained students' wrong answers, gave them hints, and offered them extra assignments to practice. The suggestion in the DA guidelines to discuss the descriptive report cards with students and parents was also fully carried out. Relating to adapting further instruction, the teachers did indicate that they used assessment information for that. However, doing this monthly or only several times a year did not give the impression that they applied the assessment information according to the intention of the DA guidelines. These guidelines articulate strongly that modifying, improving, and enhancing the teaching–learning process in the classroom is one of the main goals of DA. The finding that the teachers only scarcely used assessment information to adapt their instruction, which was corroborated by two other statements by them, namely, that they found it difficult to use assessment information for this and that they thought that assessment was helpful for students, but not for teachers.

4. Discussion

4.1. Alignment with the DA Guidelines

The various sources we used to reveal the ways DA is used in primary school mathematics education in Iran led us to conclude that in general the teachers' assessment practices are in line with the DA guidelines. They apply the suggested methods for collecting student data, they are not the only assessor but also use peer- and self-assessment, they use the suggested forms of documentation, and applied to some degree assessment results for making decisions about follow-up actions. However, at the same time, at several points the way they assess their students also differs from the assessment as intended in the DA guidelines. Although the teachers use observations, ask questions, and administer written tests and performance tests, they are not fully convinced of the relevance of these methods for gaining information about their students' mathematics skills and knowledge. Similarly, they find peer- and self-assessment not trustworthy. In addition, when preparing the report cards at the end of the semester they apparently do not have sufficient confidence in the results obtained by DA-suggested assessments methods and still use a final exam. Nonetheless, while doing this, they made it clear that they are quite aware that this is an approach that belongs to the old assessment guidelines.

Other aspects of the teachers' assessment practices that differ from the ideas of DA are that the teachers, when assessing their students, mainly focus on the correctness of the given answers and barely on the solution strategies used. This also makes it understandable that the teachers do not really see that the assessment can be helpful for them and can be a means to improve their teaching. Paradoxically, the latter was precisely one of the main reasons to reform the assessment and to introduce DA in Iran. An explanation for this contradiction became evident during the interviews when the teachers expressed that they are very bound to the timetable. Their ultimate concern is that at the end of the year they have completed the book. This requirement gives them little room to change the teaching program.

4.2. Dilemmas the Teachers May Face

Basically, this responsibility to complete the book in time implies that the teachers have to cope with two conflicting systems in which assessment and instruction are not epistemologically consistent [12]. The DA asks to adjust the teaching to the needs of the students in order to support the improvement of their learning, whereas the envisioned curriculum and the accompanying teaching approach require teaching strictly according to the book in a fix order and speed.

For Iranian teachers these conflicting systems can make it so that they are faced with several dilemmas, similar to what Suurtamm and Koch [47] revealed when they investigated Canadian mathematics teachers who were in a process of transforming their assessment practices. To analyze the assessment experiences of these teachers, they used a

framework adapted from Windschitl [48], which brought them to the identification of four dilemmas with which the teachers were confronted: a conceptual, a pedagogical, a cultural, and a political dilemma. These dilemmas may also have arisen in the circumstances of the Iranian teachers. Similar to the examples given by Suurtamm et al. [49], the Iranian teachers may have encountered a conceptual dilemma when they had to change from seeing assessment as an event at the end of a unit or semester to assessment as an ongoing act. As was shown in a study of an assessment reform in Namibia, this requires teachers to make a paradigm shift in attitudes from the past assessment approach, grounded in a teacher-centered education view of education, to a learner-centered model [50]. A pedagogical dilemma may have emerged when they were not sure how to use new assessment methods to collect and document student data. This came to the fore when, during the interviews, teachers questioned the appropriateness of the guidelines in the *Teacher's Guidance for Descriptive Assessment* [44]. In addition, the teachers' responses gave food for hypothesizing that a lack of proper theoretical and practical knowledge of DA may have influenced its accomplishment. A cultural dilemma may arise because a new assessment practice could challenge the established classroom, school, or general culture. Finally, a political dilemma may become manifest when the teachers have to cope with particular national, district, or school policies with respect to education, such as having the book finished at the end of a semester. As Suurtamm et al. [49] emphasized, it is important to be aware of the dilemmas teachers are confronted with when they are reforming their classroom assessment, because different types of dilemmas might require different types of professional support and training.

4.3. Our Findings Considered in the Light of Other Studies of DA

The results found in our study about the seven teachers' assessment practices can be put in broader perspective when giving consideration to the findings of two other studies in Iran in which DA was investigated. Like our study, the first study by Ostad-Ali et al. [51] took place in Tehran. In the survey that was carried out, 30 primary school mathematics teachers were involved from schools in the city's Zone 1. The authors reported that the teachers had a positive view towards DA, whereas quantitative evaluation was experienced to have a negative impact on the mental health of families and students. Notwithstanding these positive results for DA, the study also revealed various problems with it. The teachers mentioned that they did not receive adequate training in or sufficient information about DA. They also found DA not effective enough at improving the students' attitudes toward learning mathematics. Finally, they articulated that the mathematics textbook and DA do not match. The content of the mathematics book is not in alignment with the time that is necessary to implement DA and does not offer guidance for DA. Therefore, they proposed teaching teachers the implementation of DA by offering them workshops by the authors of the mathematics textbook and also using accompanying educational technology for the workshops.

The second study by Baluchinejad et al. [52] was conducted in Saravan, which is in Iranian Balochistan in the southeast of the country. The researchers found a significant positive relationship between the use of descriptive evaluation tools by teachers and the self-efficacy and academic achievement of students in Grade 4. Unfortunately, the published paper did not provide information about how the tools were used or how the teachers valued the reformed approach to assessment.

This was precisely the focus of the study of Choi [53], who investigated teachers' responses in South Korea to the national large-scale assessment reform that was enacted in that country. The study examined South Korean primary school teachers' implementation of the assessment reform and how implementation related to their capacity for and willingness to change. A large-scale questionnaire survey involving 700 teachers followed by a small case study of four teachers revealed that teachers' assessment practices were shifting away from a traditional way of assessment, but that their responses to the reform varied. The implementation did not progress for all teachers at the same speed. Moreover,

although teachers easily implemented superficial aspects of the reform, they struggled with implementing substantial aspects of it. An important conclusion of this study, which was also relevant for Iran, was that to make the "move from policy changes toward real changes in practice, teachers should have sufficient and high-quality learning opportunities" [53] (p. 596). Teachers need knowledge of the principles of a reform and need to know why they should implement it, otherwise they will not bring about fundamental changes. Therefore, Choi [53] suggested that policymakers should consider the development of a "scaffolding instrument" to complement the mandate of the reform and to reduce the discrepancies between policy (the intended reform) and practice (the enacted reform).

4.4. Limitations, Suggestions for Further Research, and a Recommendation

Although our study gave insight into the differences between the DA guidelines and what is happening in teachers' assessment practice, our findings should be considered with prudence. Due to the small number of teachers involved, we cannot come to a general conclusion about how DA is implemented in Iran. To make more robust and comprehensive statements about this, further research is necessary with a larger and more representative sample. As was also suggested by others [51], DA should also be investigated in other parts of Tehran and even other parts of the country with other cultural, economic, and geographical circumstances. Another limitation for which ameliorations are needed is that we did not investigate the teachers' assessment literacy and what they learned in the professional development courses they attended. With respect to the latter, we did not have the opportunity to investigate such a course. Furthermore, because a change in assessment practice also is pertinent to students and parents, it would have been better if their perspective also had been included in our study.

Despite the aforementioned shortcomings of our study, on the basis of the findings that have emerged, one recommendation can be given as a start. The struggle of the teachers with the "how" of the DA, their low confidence in the results of certain DA methods, and the pressure they feel to finish the book, indicate the need for perspicuous, domain-specific guidelines, and workshops where teachers can discuss their assessment practice and can enrich their assessment to become a tool to improve teaching and support students' learning.

Appendix A

Table A1. Overview of the teachers for whom we found particular classroom assessment practices as suggested by the Descriptive Assessment (DA) guidelines.

Key Category Detected in the DA Guidelines	Teacher							*N*
	1	**2**	**3**	**4**	**5**	**6**	**7**	
A. How Are Student Data Collected?								
A.1 By observing students in class								
Observe students several times a week	x [1]	x	x	x	x	x	x	7
Observe students in a structured way								0
Observing students is observed during the lesson	x	x	x			x	x	5
Find observing students very relevant	x	x	x		x			4
A.2 By asking questions in class								
Ask individual students questions (several times a week: 4 teachers; weekly: 3 teachers)	x	x	x	x	x	x	x	7
Ask the whole class questions (several times a week: 4 teachers; weekly: 2 teachers; monthly: 1 teacher)	x	x	x	x	x	x	x	7
Asking questions is observed during the lesson	x	x	x	x	x	x	x	7
Find asking students very relevant		x	x				x	3
A.3 By administering written tests								
Administer written tests weekly to cover the lessons	x	x	x	x	x	x	x	7
Use additional written tests	x		x		x	x	x	5
Use a final exam at end of semester	x	x	x	x	x	x	x	7
Design written tests	x	x	x	x	x	x	x	7
Use closed- and open-answer problem formats	x	x	x	x	x	x	x	7
Find administering written tests very relevant	x		x			x	x	4
A.4 By providing performance tasks								
Use performance tasks monthly or several times a year	x	x	x	x		x		5
Use performance tasks weekly or several times a week					x		x	2
Find performance tasks time consuming	x	x						2
Type of performance task: using a tool	x	x			x	x	x	6
Type of performance task: writing a text			x	x	x	x		4
Type of performance task: giving a presentation	x	x		x	x		x	5
Type of performance task: designing something	x		x	x	x			4
Find performance tasks very relevant	x			x	x		x	4
A.5 By having students working on a project								0
B. Who is assessed or assessing?								
B.1 Assessing individual students (several times a week)	x	x	x	x	x	x	x	7
B.2 Assessing groups of students (at least monthly: 5 teachers; several times a year: 2 teachers)	x	x	x	x	x	x	x	7
B.3 Using peer assessment								
Use peer assessment weekly or several times a week		x			x		x	3
Use peer assessment yearly or several times a year	x		x	x		x		4
Peer assessment is observed during the lesson	x							1
Have difficulties with peer assessment (not trustworthy)	x	x	x	x	x	x	x	7
Find peer assessment very relevant				x	x			2
B.4 Using self-assessment								
Use self-assessment weekly or several times a week		x	x		x	x	x	5
Use self-assessment yearly or several times a year	x			x				2
Have difficulties with self-assessment (not trustworthy)	x	x	x	x	x	x	x	7
Find self-assessment very relevant	x		x			x		3
C. What is the focus on in assessment?								
C.1 Checking students' answers	x	x	x	x	x	x	x	7
C.2 Checking students' strategies	x							1
D. How is the assessment documented?								
D.1 Reporting students' performance in the notebook	x	x	x	x	x	x	x	7
Find that there is not enough space in the notebook	x	x	x	x	x	x	x	7
D.2 Keeping a portfolio	x	x	x	x	x	x	x	7
Find keeping a portfolio very relevant					x	x		2
Find keeping a portfolio not very relevant	x	x	x	x			x	5
D.3 Filling in the descriptive report cards	x	x	x	x	x	x	x	7
E. How is the assessment followed up on?								
E.1 Giving descriptive feedback								
Give oral feedback several times a week	x	x	x	x	x	x	x	7
Give written feedback weekly or several times a week	x	x	x		x	x	x	6
Give written feedback monthly				x				1
Give individual feedback several times a week or weekly	x	x	x	x	x	x	x	7
Give feedback to class several times a week or weekly		x	x		x	x	x	5
Give feedback to class monthly	x			x				2
Type: Explaining wrong answer	x	x	x	x	x	x	x	7
Type: Giving hints	x	x	x	x	x	x	x	7
Type: Feedback with extra assignment	x	x	x	x	x	x	x	7
Find giving descriptive feedback useful	x	x	x	x	x	x	x	7
Find that giving descriptive feedback implies a lot of work	x	x	x	x	x	x	x	7
E.2 Discussing descriptive report cards								
Discuss descriptive report cards with students	x	x	x	x	x	x	x	7
Discuss descriptive report cards with parents	x	x	x	x	x	x	x	7
E.3 Adapting further instruction								
Adapt further instruction monthly			x	x	x	x		4
Adapt further instruction several times a year	x	x					x	3
Adapt the speed of lessons	x	x	x	x	x	x	x	7
Find it difficult to adapt instruction	x	x	x	x	x	x	x	7
Find assessment more helpful for students than for teachers	x	x	x	x	x	x	x	7

[1] The x indicates that for this teacher a particular assessment practice was found.

Table A2. Number of teachers for whom we found classroom assessment practices in line or not in line with DA guidelines (*N* = 7).

Teachers' Classroom Assessment Practices			
In line with descriptive assessment	*N*	Not in line with descriptive assessment	*N*
A. How are student data collected?			
A.1 By observing students in class			
Observe students several times a week	7	Do not observe students in a structured way	7
Find observing student very relevant	4		
A.2 By asking questions in class			
Ask individual students questions (several times a week: 4 teachers; weekly: 3 teachers)	7		
Ask the whole class questions (several times a week: 4 teachers; weekly: 2 teachers; monthly: 1 teacher)	7		
Find asking questions very relevant	3		
A.3 By administering written tests			
Administer written tests weekly that cover lessons	7		
Use additional written tests	5	Use a final exam at end of semester	7
Design written tests	7		
Use closed- and open-answer problem formats	7		
Find administering written tests very relevant	4		
A.4 By providing performance tasks			
Use performance tasks (monthly or several times a year: 5 teachers; weekly or several times a week: 2 teachers)	7		
Find performance tasks time consuming	2		
Type of performance task: using a tool	6		
Type of performance task: writing a text	4		
Type of performance task: giving a presentation	5		
Type of performance task: designing something	4		
Find performance tasks very relevant	4		
A.5 By having students working on a project		Do not use projects for assessing students	7
B. Who is assessed or assessing?			
B.1 Assess individual students (several times a week)	7		
B.2 Assess groups of students (at least monthly: 5 teachers; several times a year: 2 teachers)	7		
B.3 Using peer assessment			
Use peer assessment (weekly or several times a week: 3 teachers; yearly or several times a year: 4 teachers)	7		
Find peer assessment very relevant	2	Find peer assessment not trustworthy	7
B.4 Using self-assessment			
Use self-assessment (weekly or several times a week: 5 teachers; yearly or several times a year: 2 teachers)	7		
Find self-assessment very relevant	3	Find self-assessment not trustworthy	7
C. What is the focus on in assessment?			
C.1 Checking students' answers	7		
C.2 Checking students' strategies	1	Do not check students' strategies	6
D. How is the assessment documented?			
D.1 Reporting students' performance in notebook	7		
Find that there is not enough space in the notebook	7		
D.2 Keeping a portfolio	7		
Find keeping a portfolio very relevant	2	Find keeping a portfolio not very relevant	5
D.3 Filling in the descriptive report cards	7		
E. How is the assessment followed-up?			
E.1 Giving descriptive feedback			
Give oral feedback (several times a week)	7		
Give written feedback (weekly or several times a week: 6 teachers; monthly: 1 teacher)	7		
Give individual feedback (several times a week or weekly)	7		
Give feedback to class (several times a week or weekly: 5 teachers; monthly: 2 teachers)	7		
Type: Explaining wrong answer	7		
Type: Giving hints	7		
Type: Feedback with extra assignment	7		
Find giving descriptive feedback useful	7		
Find that descriptive feedback implies a lot of work	7		
E.2 Discussing descriptive report cards			
Discuss descriptive report cards with students	7		
Discuss descriptive report cards with parents	7		
E.3 Adapting further instruction			
Adapt further instruction (monthly: 4 teachers; several times a year: 3 teachers)	7		
Type: Adapting speed of lessons	7		
Find it difficult to adapt instruction	7		
		Do not really find assessment helpful for the teacher	7

References

1. Berry, R. Assessment reforms around the world. In *Assessment Reform in Education: Policy and Practice*; Berry, R., Adamson, B., Eds.; Springer Science & Business Media: Dordrecht, The Netherlands, 2011; pp. 89–102.
2. Barnes, M.; Clarke, D.; Stephens, M. Assessment: The engine of systemic curricular reform? *J. Curric. Stud.* **2000**, *32*, 623–650. [CrossRef]
3. Fiske, E.B. *Education for All: Status and Trends. Assessing Learning Achievement*; UNESCO: Paris, France, 2000. Available online: https://unesdoc.unesco.org/ark:/48223/pf0000119823 (accessed on 6 July 2020).
4. Masters, G.N. *Reforming Educational Assessment: Imperatives, Principles and Challenges*; ACER: Camberwell, Victoria, Australia, 2013.
5. NCTM (National Council of Teachers of Mathematics). *Curriculum and Evaluation Standards for School Mathematics*; NCTM: Reston, VA, USA, 1989.
6. DES/WO (Department of Education and Science and the Welsh Office). *Mathematics in the National Curriculum*; Her Majesty's Stationary Office: London, UK, 1989.
7. AEC (Australian Education Council). *A National Statement on Mathematics for Australian Schools*; Curriculum Corporation: Carlton, Victoria, Australia, 1991.
8. VCAB (Victorian Curriculum and Assessment Board). *Mathematics Study Design*; VCAB: Melbourne, Australia, 1990.
9. Heritage, M. Formative assessment: What do teachers need to know and do? *PDK* **2007**, *89*, 140–145. [CrossRef]
10. Rothman, R. *Measuring Deeper Learning: New Directions in Formative Assessment*; Jobs for the Future: Boston, MA, USA, 2018.
11. Wiliam, D.; Thompson, M. Integrating assessment with learning: What will it take to make it work? In *The Future of Assessment. Shaping Teaching and Learning*; Dwyer, C.A., Ed.; Routledge: New York, NY, USA, 2008; pp. 53–82.
12. Van den Heuvel-Panhuizen, M.; Becker, J. Towards a didactic model for assessment design in mathematics education. In *Second International Handbook of Mathematics Education*; Bishop, A.J., Clements, M.A., Keitel, C., Kilpatrick, J., Leung, F.K.S., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; pp. 689–716.
13. Shepard, L.A. The role of assessment in a learning culture. *Educ. Res.* **2000**, *29*, 4–14. [CrossRef]
14. Bartlett, D.; Black, P.; Daugherty, R.; Ecclestone, K.; Harlen, W.; English, J.; Gardner, J.; James, M.; Hutchinson, C.; Newton, P.; et al. *The Role of Teachers in the Assessment of Learning*; Nuffield Foundation and Institute of Education: London, UK, 2006. Available online: http://www.nuffieldfoundation.org/sites/default/files/files/The-role-of-teachers-in-the-assessment-of-learning.pdf (accessed on 6 July 2020).
15. Looney, A.; Cumming, J.; Van der Kleij, F.; Harris, K. Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assess. Educ.* **2018**, *25*, 442–467. [CrossRef]
16. Harlen, W. Improving assessment of learning and for learning. *Education* **2009**, *37*, 247–257. [CrossRef]
17. Anderson, 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. Briggs, D.C.; Ruiz-Primo, M.A.; Furtak, E.M.; Shepard, L.A.; Yin, Y. Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educ. Meas.* **2012**, *31*, 13–17. [CrossRef]
19. Cauley, K.M.; McMillan, J.H. Formative assessment techniques to support student motivation and achievement. *Pract. Assess. Res. Eval.* **2010**, *83*, 1–6. [CrossRef]
20. Kingston, N.; Nash, B. Formative assessment: A meta-analysis and a call for research. *Educ. Meas.* **2011**, *30*, 28–37. [CrossRef]
21. Phelan, J.; Choi, K.; Vendlinski, T.; Baker, E.; Herman, J. Differential improvement in student understanding of mathematical principles following formative assessment intervention. *J. Educ. Res.* **2011**, *104*, 330–339. [CrossRef]
22. Veldhuis, M.; Van den Heuvel-Panhuizen, M. Supporting primary school teachers' classroom assessment in mathematics education: Effects on student achievement. *Math. Educ. Res. J.* **2020**, *32*, 449–471. [CrossRef]
23. Wiliam, D.; Lee, C.; Harrison, C.; Black, P.J. Teachers developing assessment for learning: Impact on student achievement. *Assess. Educ.* **2004**, *11*, 49–65. [CrossRef]
24. Buabeng, I.; Atingane, A.B.; Amoako, I. Practices, challenges and perceived influence of classroom assessment on mathematics instruction. *Int. J. Assess. Tools Educ.* **2019**, *6*, 476–486. [CrossRef]
25. McNair, S.; Bhargava, A.; Adams, L.; Edgerton, S.; Kypros, B. Teachers speak out on assessment practice. *Early Child Educ. J.* **2003**, *31*, 23–31. [CrossRef]
26. Suurtamm, C.; Koch, M.; Arden, A. Teachers' assessment practices in mathematics: Classrooms in the context of reform. *Assess. Educ.* **2010**, *17*, 399–417. [CrossRef]
27. Veldhuis, M.; Van den Heuvel-Panhuizen, M.; Vermeulen, J.A.; Eggen, T.J. Teachers' use of classroom assessment in primary school mathematics education in the Netherlands. *Cadmo* **2013**, *21*, 35–53.
28. Yu, W.M.; Kennedy, K.J.; Fok, P.K.; Chan, K.S.J. Assessment reform in basic education in Hong Kong: The emergence of assessment for learning. In Proceedings of the 32nd Annual Conference of International Association for Educational Assessment: Assessment in an Era of Rapid Change: Innovations and Best Practices, Singapore, 21–26 May 2006. Available online: https://repository.eduhk.hk/en/publications/assessment-reform-in-basic-education-in-hong-kong-the-emergence-o-3 (accessed on 6 July 2020).
29. Zhang, Z.; Burry-Stock, J.A. Classroom assessment practices and teachers' self-perceived assessment skills. *Appl. Meas. Educ.* **2003**, *16*, 323–342. [CrossRef]

30. 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]

31. Zhao, X.; Van den Heuvel-Panhuizen, M.; Veldhuis, M. Insights Chinese primary mathematics teachers gained into their students' learning from using classroom assessment techniques. *Educ. Sci.* **2019**, *9*, 150. [CrossRef]

32. Sharief, M.; Naderi, M.; Hiedari, M.S.; Roodbari, O.; Jalilvand, M.R. A study of strengths and weaknesses of descriptive assessment from principals', teachers' and experts' points of view in Chaharmahal and Bakhteyari primary schools. *Int. Educ. Stud.* **2012**, *5*. Available online: http://www.ccsenet.org/journal/index.php/ies/article/view/18075 (accessed on 6 July 2020). [CrossRef]

33. Gooya, Z. A critical investigation of the mathematics curriculum of the National Curriculum document in Iran. *Q. J. Curric. Stud.* **2010**, *5*, 147–164. (In Persian)

34. Gooya, Z.; Gholamazad, S. An overview of changes in school mathematics curriculum in Iran. In *ICMI Study 24 Conference Proceedings. School Mathematics Curriculum Reforms: Challenges, Changes and Opportunities*; Shimizu, Y., Vithal, R., Eds.; University of Tsukuba/ICMI: Tsukuba, Japan, 2018; pp. 125–132. Available online: https://www.mathunion.org/fileadmin/ICMI/ICMI%20studies/ICMI%20Study%2024/ICMI%20Study%2024%20Proceedings.pdf (accessed on 6 July 2020).

35. Rafiepour, F.; Gooya, Z. The necessity and directions of changes in school mathematics curriculum in Iran from the perspective of teachers. *Q. J. Educ. Innov.* **2010**, *33*, 92–120. (In Persian)

36. Arani, A.M.; Kakia, L.; Karimi, V. Assessment in education in Iran. *SA-eDUC J.* **2012**, *9*, 1–11. Available online: www.nwu.ac.za/webfm_send/58395 (accessed on 6 July 2020).

37. Hasani, M.; Ahmadi, H. *Descriptive Evaluation: A New Model in Educational Evaluation*; MoE & OERP: Tehran, Iran, 2005.

38. SCE (Supreme Council of Education). *Negotiations of the Supreme Council of Education-N.674*; SCE: Tehran, Iran, 2002. (In Persian)

39. SCE (Supreme Council of Education). *Negotiations of the Supreme Council of Education-N.679*; SCE: Tehran, Iran, 2002. (In Persian)

40. Hasani, M.; Kazemi, Y. *Descriptive Assessment Project (Goals, Fundamentals, and Solutions)*; Asare Moaser: Tehran, Iran, 2003. (In Persian)

41. SCE (Supreme Council of Education). *Principles of Assessing Educational Achievement*; SCE: Tehran, Iran, 2004. (In Persian)

42. SCE (Supreme Council of Education). *Academic Assessment Regulation for Primary School*; SCE: Tehran, Iran, 2009. (In Persian)

43. Ndoro, V.W.; Hanley, G.P.; Tiger, J.H.; Heal, N.A. A descriptive assessment of instruction-based interactions in the preschool classroom. *J. Appl. Behav. Anal.* **2006**, *39*, 79–90. [CrossRef]

44. MoE (Ministry of Education); OERP (Organization for Educational Research and Planning). *Teacher's Guidance for Descriptive Assessment*; MoE & OERP: Tehran, Iran, 2009. (In Persian)

45. Zhao, Y.; 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*, 1–18. [CrossRef]

46. MoE (Ministry of Education); OERP (Organization for Educational Research and Planning). *Teacher's Guidance for Descriptive Assessment*; Revised version; MoE & OERP: Tehran, Iran, 2011. (In Persian)

47. Suurtamm, C.; Koch, M.J. Navigating dilemmas in transforming assessment practices: Experiences of mathematics teachers in Ontario, Canada. *Educ. Assess. Eval. Acc.* **2014**, *26*, 263–287. [CrossRef]

48. Windschitl, M. Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Rev. Educ. Res.* **2002**, *72*, 131–175. [CrossRef]

49. Suurtamm, C.; Thompson, D.R.; Kim, R.Y.; Moreno, L.D.; Sayac, N.; Schukajlow, S.; Silver, E.; Ufer, S.; Vos, P. *Assessment in Mathematics Education—Large-Scale Assessment and Classroom Assessment*; Springer Nature: Cham, Switzerland, 2016. Available online: https://link.springer.com/book/10.1007/978-3-319-32394-7 (accessed on 6 July 2020).

50. Iipinge, S.M.; Kasanda, C.D. Challenges associated with curriculum alignment, change and assessment reforms in Namibia. *Assess. Educ.* **2013**, *20*, 424–441. [CrossRef]

51. Ostad-Ali, F.; Behzadi, M.H.; Shahvarani, A. Descriptive qualitative method of evaluation from the viewpoint of math teachers and its comparison with the quantitative evaluation (giving scores) method (A case study on the primary schools for girls in Zone 1 of Tehran City). *Math. Educ. Trends Res.* **2015**, *1*, 50–56. [CrossRef]

52. Baluchinejad, T.; Al-Nisa Baluchinejad, J.; Baluchinejad, M.; Dehvari, S.; Baluchinejad, M. Investigating the relationship between descriptive evaluation and self efficacy of students of fourth grade elementary school. *Electron. J. Biol.* **2016**, *12*, 417–421. Available online: https://ejbio.imedpub.com/investigating-the-relationship-between-descriptive-evaluationand-self-efficacy-of-students-of-fourth-grade-elementaryschool.php?aid=11127 (accessed on 8 February 2021).

53. Choi, J. Understanding elementary teachers' different responses to reform: The case of implementation of an assessment reform in South Korea. *Int. Electron. J. Elem. Educ.* **2017**, *9*, 581–598. Available online: https://www.iejee.com/index.php/IEJEE/article/view/177 (accessed on 8 February 2020).