



Improving classroom assessment in primary mathematics education in the Netherlands

Michiel Veldhuis, Marja Van den Heuvel-Panhuizen

► **To cite this version:**

Michiel Veldhuis, Marja Van den Heuvel-Panhuizen. Improving classroom assessment in primary mathematics education in the Netherlands. Konrad Krainer; Naďa Vondrová. CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Feb 2015, Prague, Czech Republic. pp.375-381, Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education. <hal-01286850>

HAL Id: hal-01286850

<https://hal.archives-ouvertes.fr/hal-01286850>

Submitted on 11 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Improving classroom assessment in primary mathematics education in the Netherlands

Michiel Veldhuis¹ and Marja van den Heuvel-Panhuizen^{1,2}

1 Utrecht University, Freudenthal Institute for Science and Mathematics Education, Utrecht, Netherlands, m.veldhuis1@uu.nl

2 Faculty of Social and Behavioural Sciences, Utrecht University, Utrecht, Netherlands

We report on the results of the Improving Classroom Assessment (ICA) project in the Netherlands that was aimed at improving primary students' mathematics achievement through improving their teachers' classroom assessment. Towards this end we first investigated primary teachers' assessment practice in a large-scale survey study. After having described and profiled teachers' current assessment practice, we designed a number of classroom assessment techniques, which were tested for feasibility in two small-scale sub-studies. Finally, we evaluated the effectiveness of the use of these classroom assessment techniques in a large-scale evaluation. Results indicate that students generally benefit from their teachers' improved use of classroom assessment techniques in mathematics.

Keywords: Classroom assessment; primary school; assessment technique; achievement.

BACKGROUND AND FOCUS OF THE PROJECT

Developing and keeping track of students' mathematical abilities are important parts of every primary mathematics teacher's daily practice. In order for teachers to gauge their students' learning, classroom assessment plays a pivotal role (Cizek, 2010). By using classroom assessment teachers can gather information about their students' mathematical skills and level of understanding. Collecting information on students' learning is primordial for at least two reasons: to find out whether the instruction has had its desired effect and to generate ideas for how to proceed in the subsequent lessons. Based on assessment information teachers can align their teaching to their students' needs, which in turn can result into adapting their teaching, but can of course also mean not changing anything and continuing with what was planned before.

Many of the characteristics of classroom assessment appear to be part of merely good teaching practice, as Ginsburg (2009) wrote in the context of mathematics education:

Good teaching [...] sometimes involves the same activities as those comprising formative assessment: understanding the mathematics, the trajectories, the child's mind, the obstacles, and using general principles of instruction to inform the teaching of a child or a group of children (p. 126).

Classroom assessment is broader: it comprises all activities that permit teachers to find out where their students are at a particular moment in terms of comprehension of the subject and to give information on what is going right and wrong. Policymakers as well as influential researchers have urged the educational community, and in particular teachers, to embrace (formative) classroom assessment in their practice. For instance, the U.S. National Council of Teachers of Mathematics (NCTM, 2013) recently took the following position on formative assessment in mathematics education:

The use of formative assessment has been shown to result in higher achievement. The National Council of Teachers of Mathematics strongly endorses the integration of formative assessment strategies into daily instruction (p. 1).

Teachers are the only ones that can actively integrate these formative assessment strategies into their practice. Advocating positions such as these were mainly inspired by the influential review study by Black and Wiliam (1998) that reported the different practical expressions of classroom assessment to be the most effective interventions for teachers to improve student learning. Recently, several researchers have

questioned the size of the effectiveness of (formative) assessment on student learning through reviews or meta-analyses of existing studies (e.g., McMillan, Venable, & Varier, 2013). Common to these critical examinations, although their specificities differ, is that they do not contest the positive effect formative assessment is purported to have on student achievement, but only the size of this effect.

Why then is classroom assessment by teachers supposed to lead to improved student learning in mathematics? In order to answer this question researchers have drawn parallels between the concepts and practices of formative assessment, self-regulated learning, feedback, and scaffolding (see for an overview, among many others, Clark, 2012). An intuitive way of saying it would be: if teachers are better aware of their students' mathematical abilities and understanding, then they can undoubtedly better adapt their teaching to the needs of the students. In doing this and providing explicit and implicit feedback students also become more aware of their own functioning, and the circle is complete: students and teacher simultaneously advance. This does have its limits, because "the teacher *must actually use the assessment data* to inform some change in the conduct of instruction" (Erickson, 2007, p. 189, original emphasis). In order for teachers to be willing to use classroom assessment techniques, these have to provide them with valuable and easily usable information about students' understanding of mathematics in a timely manner, otherwise it would not contribute to better teaching and, in the end, to better student achievement.

In the ICA-project we strived to improve primary students' mathematics performance through improving their teachers' use of classroom assessment in Grade 3. As a start we investigated the current classroom assessment practice of primary mathematics teachers in the Netherlands (Study 1). Secondly we identified a meaningful profile characterization of these teachers' mathematics assessment practice (Study 2). Now that the current practice was known we could test the feasibility of classroom assessment techniques that were designed to match the mathematics curriculum of the second half of Grade 3 in the Netherlands and provide valuable information to the teachers (Study 3). Finally we evaluate the effectiveness of the use of these classroom assessment techniques in a large-scale experimental study (Study 4).

STUDY 1: PRIMARY TEACHERS' USE OF CLASSROOM ASSESSMENT IN MATHEMATICS

We conducted a survey of the classroom assessment practices of Dutch primary school teachers in mathematics education (Veldhuis, van den Heuvel-Panhuizen, Vermeulen, & Eggen, 2013). International studies have shown that teachers use a wide range of methods to collect information about their students' learning (e.g., Suurtamm, Koch, & Arden, 2010). To find out students' skills and comprehension level, teachers can use methods ranging from standardized tests and tests that come with a textbook, to asking questions and observing students while they are working. The assessment methods teachers choose to reveal their students' learning processes depend on several factors. A first factor that has been found to affect this choice is teachers' beliefs concerning classroom assessment (Dixon, Hawe, & Parr, 2011). A second factor in choosing a particular assessment method, beside beliefs, concerns the assessment purpose teachers have in mind (Suurtamm et al., 2010), for instance a formative or summative purpose. A further determining factor of using particular assessment methods is the view on education in which the assessment takes place. The methods used for assessment often correspond to the approach to education as reflected in the adhered learning theory and the curriculum that is taught (Shepard, 2000).

Method

We investigated, using an online questionnaire, how primary teachers in the Netherlands collect information on their students' progress in mathematics and how teachers' assessment methods, purposes, and beliefs about the usefulness of assessment are related. This questionnaire contained 40 items, pertaining to the teachers' (i) background characteristics, (ii) mathematics teaching practice, (iii) assessment practice, and (iv) perceived usefulness of assessment. Questions with different formats were included: fixed-response and items with a rating scale, but also some open-ended items. The sample of participating teachers was obtained through an open invitation by e-mail, which was sent successfully to 5094 primary schools for regular education in the Netherlands. Teachers who were willing to respond to the online questionnaire were promised a set of digital mathematical exercise material as a reward.

Results and discussion

In total 960 teachers at 557 Dutch primary schools responded to the questionnaire. Observation-based assessment methods of questioning, observing, and correcting written work, were the most frequently – that is weekly – applied methods, whereas instrument-based methods, particularly using textbook tests and student monitoring tests were employed several times a year (see Figure 1).

Teachers used assessment mainly for formative purposes and they considered the assessment methods they used themselves as most relevant. We found that teachers in primary mathematics education in the Netherlands use a variety of assessment methods, use instrument-based and observation-based assessment methods on average just as frequently and find assessment generally useful. This perceived usefulness is shown by the overall very positive reactions teachers gave on the different uses of assessment. The two main instrument-based assessment methods, textbook tests and tests from a student monitoring system, are reported as the most relevant, with asking questions and observing students the most relevant of the observation-based assessment methods. Furthermore, the teachers' responses to the questionnaire revealed that they used assessment both for formative and summative purposes. The results of our survey indicate that teachers do use assessment information for various purposes, from giving feedback via adapting instruction to stimulating thinking.

STUDY 2: PRIMARY TEACHER PROFILES IN MATHEMATICS ASSESSMENT

After this general overview of the current assessment practice we were interested in finding out more about individual teachers. The second study was aimed at gaining knowledge of how the assessment practices of individual teachers could be characterized within the universe of assessment skills and activities. In fact, we wanted to understand assessment from the conglomerate of choices a single teacher is making when collecting information about his or her students' learning process. To achieve this we performed a secondary analysis of the earlier gathered questionnaire data to identify a profile characterization of every teacher's assessment practice (Veldhuis & van den Heuvel-Panhuizen, 2014). The rationale for distinguishing assessment profiles of teachers is that these can contribute to our theoretical understanding of assessment as it teachers carry it out. In addition, knowledge about these assessment profiles can help us in a practical sense with designing tailor-made courses for professional development that fit the teachers' needs.

Analyses

We analyzed the survey data in two steps. To identify the latent structure of what was measured by the questionnaire and be able to construct assessment profiles of teachers we used a combination of latent variable modeling techniques. To explore the underlying structure of the items measuring teachers' math-

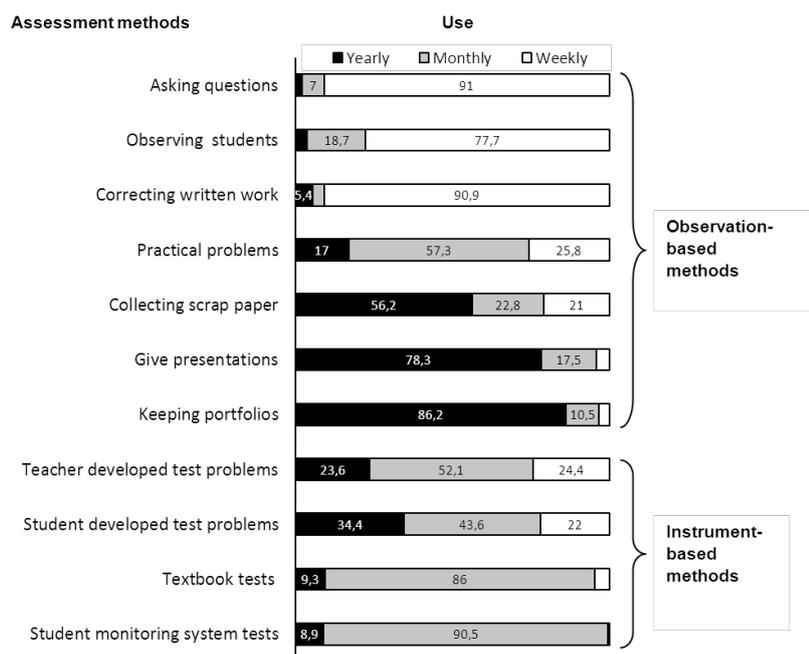


Figure 1: Percentages of frequencies of assessment methods' use (ns > 940)

ematics assessment practice, we performed exploratory factor analyses. To investigate whether these latent factors could also be used to interpret classes of teachers, we performed a latent class analysis. This is a statistical technique permitting the identification of underlying classes of individuals based on differences in their responses on items in a questionnaire or test. The teachers in our sample were assigned to the different latent classes—that we will call assessment profiles—through modal assignment, i.e. they were assigned to the latent class to which they had the highest probability of belonging.

Results and discussion

After comparing one- to seven-factor solutions and eliminating items with cross loadings over |0.4|, an exploratory factor analysis delivered a five-factor solution that had a good enough fit ($\chi^2(1076, N = 960) = 5494.1, p < .001, RMSEA = .064, CFI = .961$). We named the five factors based on the items they contained: *Goal centeredness of assessment* (items on teachers’ purposes of assessment), *Authentic nature of assessment* (items on authentic assessment methods), *Perceived usefulness of assessment* (statements on usefulness), *Diversity of assessment problem format* (items on problem formats), and *Allocated importance of assessing skills and knowledge* (items on the importance of assessing particular skills and knowledge). To be able to characterize teachers’ assessment practice and assign them to different assessment profiles we performed a

latent class analysis using all variable scores as input. As such we were able to check whether we would be able to show differences between the latent classes of teachers on the five factors we found in the separately performed factor analysis. Four latent classes provided the best fitting solution. To find out whether teachers thus assigned to the four latent classes differed on the five factors of assessment practice identified before, we performed several analyses of variance. The results showed that teachers from one latent class to another differed significantly from each other (see Figure 2 for the size and the direction of these differences). These differences suggest that teachers with particular assessment profiles have qualitatively different assessment practices.

The assessment profile to which most teachers (35.5%) in our sample belonged was the *mainstream assessors* profile. In this profile most teachers regularly used different types of assessment, test-based and observation-based, for both summative and formative purposes. On all factors teachers with this profile scored around the mean.

The next biggest group (28.5%) were the *enthusiastic assessors*. Teachers with this profile were very aware of the different possibilities assessment offers them and used them likewise. On all components these teachers scored above the mean, with a peak on *Goal centeredness of assessment*. An almost equally large

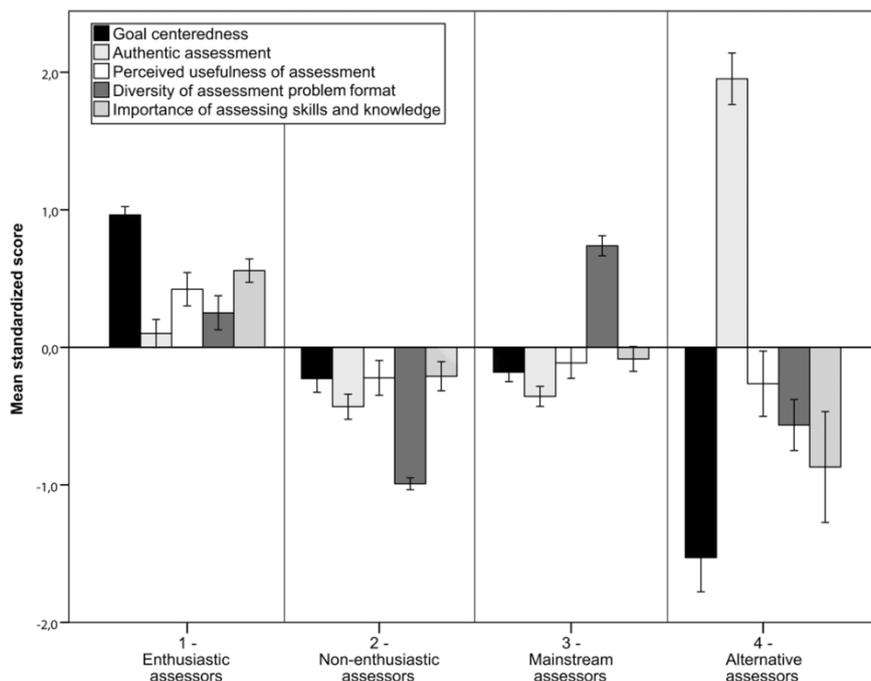


Figure 2: Mean standardized scores on factors for teachers in the four latent classes

group of teachers (25.8%) were the *non-enthusiastic assessors*. These teachers viewed assessment more often in a negative way and used it accordingly less. On all factors, teachers with this profile scored below average. Finally, there were the *alternative assessors* (10.3%). Teachers with this profile had an ambiguous view of assessment. Although they reported a lot of own input in assessment and devised their own tasks and tests, they did not find assessment important or necessary.

Through this profile characterization of teachers' assessment practice we were able to select some of the skills and activities from the universe of assessment skills and activities of teachers. In this way we brought structure to the many possible characterizations of assessment practice and gained a more clear idea of what to expect from teachers prior to the introduction of the assessment techniques in our consecutive studies.

STUDY 3: EXPLORING THE FEASIBILITY AND EFFECTIVENESS OF CLASSROOM ASSESSMENT TECHNIQUES IN MATHEMATICS

In this small-scale study the focus switched from the current practice of teachers to how this practice could be improved. Many types of formative or classroom assessment techniques have been proposed and used in international research (e.g., Black, Harrison, Lee, Marshall, & Wiliam, 2004) or in more practice-oriented work (e.g., Keeley & Tobey, 2011). In mathematics education there exist many different forms of these classroom assessment techniques, through our survey of current practice we could determine that these techniques were not very often used by primary teachers in the Netherlands. The Dutch Inspectorate (Inspectie van het onderwijs, 2013) also pointed out that many primary (40%) and secondary schools (33%) do not systematically use assessments to monitor their students' progress. In any case, from available findings it becomes clear that there is a need for investigation of the use of classroom assessment techniques in mathematics education.

The purpose of this study was to investigate the feasibility and effectiveness of classroom assessment techniques for mathematics in primary school. We wanted to find out whether teachers and students were prone to use assessment techniques and whether the use of an ensemble of these techniques would be related to an increase in achievement.

Method

Ten teachers (with 214 students; 14 to 29 students per class) participated in monthly workshops in the second semester of Grade 3 in two consecutive sub-studies (four teachers in the first; six teachers in the second). In the workshops, consisting of three or four teachers and the first author, classroom assessment techniques were presented, discussed, and evaluated. The teachers were approached by e-mail and volunteered to participate. The schools were all situated in urbanized areas with highly mixed student populations, and the teachers used four different textbooks that were all based on realistic mathematics education principles as is common in the Netherlands.

The feasibility of the classroom assessment techniques was investigated by conducting regular classroom observations of every teacher in between workshops. These observations were intertwined with short informal interviews. To investigate the effectiveness of the use of classroom assessment techniques we used a pre-/post-test evaluation of students' mathematics achievement. The pre-test data consisted of the results from the midyear student-monitoring test for Grade 3 (Cito LOVS M5) and the results from the end of year student-monitoring test for Grade 3 (Cito LOVS E5) served as post-test data (Janssen, Verhelst, Engelen, & Scheltens, 2010). These biannual student-monitoring tests are used in virtually all primary schools in the Netherlands to monitor students' development in mathematical ability over the years. The teachers administered the tests in their own classes as is common in educational practice in the Netherlands. The scores on these tests are mathematical ability scores calculated through item response theory models.

We proposed a collection of classroom assessment techniques consisting of short activities of less than 10 minutes to the teachers. The techniques were supposed to help teachers to quickly find out something about their students' mathematical skills and understanding, provide teachers with indications for further instruction, and focus on some of the mathematics content of the second semester of Grade 3. Most techniques were centred on the assessment of number knowledge, mainly in the context of addition and subtraction, but they could also be used to assess multiplication and division tables. In Figure 3 we provide an example of a technique called Red/Green cards.

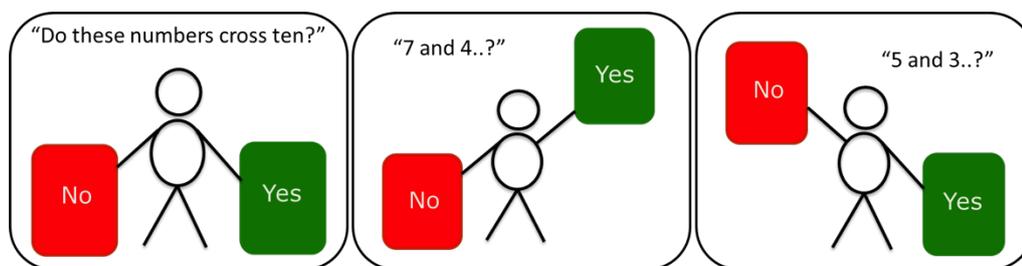


Figure 3: An example of the Red/Green cards. The focus is on number sense: the comprehension that two numbers together can be more or less than 10, 100, or 1000

The teacher asks all students a series of questions that can be answered quickly with Yes (green) or No (red). By inspecting the waving red and green cards the teacher gets an immediate overview of all students' responses. Especially when students have to carry out mental additions and subtractions with two-digit numbers it is crucial that they can instantaneously identify whether two numbers cross ten or not, because this has consequences for the strategy to be applied. This technique provides the teacher quickly with information on particular number sense knowledge of the students.

Results and discussion

Teachers and students reported enjoying the techniques and finding them useful in the sense that they provided them with valuable information that supported their teaching and learning. Teachers also mentioned that the techniques were easy to apply in their classrooms. In terms of mathematics achievement, results indicate students improving considerably ($M_{\text{gain substudy 1}} = +9.7$; $M_{\text{gain substudy 2}} = +7.6$). It could of course be expected that students advance in their mathematical ability, whether teachers perform specific assessment activities or not; the scores of the national norm sample also showed this direction ($M_{\text{gain norm}} = +5.1$).

Even though the treatment group was relatively small and there was no control group in this study, these results do provide an indication for the feasibility and effectiveness of the use of the classroom assessment techniques in mathematics: teachers use the techniques and students appear to advance more from the midyear to the end of the year testing than expected.

STUDY 4: TEACHERS' USE OF CLASSROOM ASSESSMENT IMPROVING STUDENTS' MATHEMATICS ACHIEVEMENT

To verify whether the students' achievement improvement we found in Study 3 was really due to the teachers' use of the classroom assessment techniques in mathematics and not just to an attention (also Hawthorne) effect, we replicated this investigation in a large-scale experiment.

Method

The same pre-/post-test design was used as in Study 3, but now with a control group and an extra manipulation. Thirty teachers (and their 616 students) participated and were randomly distributed over the three experimental conditions and one control condition. In the experimental conditions teachers participated in the same type of workshops as in Study 3. These experimental conditions differed on the intensity of the professional development teachers received: in Experimental I there was one workshop, in Experimental II, two workshops, and in Experimental III, three workshops, and in the control condition teachers did not have any workshops. In these workshops the same classroom assessment techniques were discussed; there was more time for every technique if teachers had more workshops.

Results and discussion

Teachers' use of classroom assessment techniques in mathematics was associated with students' improved mathematics achievement. More specifically, when teachers participated in three workshops and, as such, developed more ownership of the techniques, their students showed more improvement in terms of mathematics achievement than in the other conditions ($M_{\text{gain Experimental III}} = +8.1$; $M_{\text{gain other}} = +5.9$; cf. Figure 4 with the results of an ANCOVA, correcting posttest scores for pretest differences). Supporting teachers in the use of classroom assessment techniques for math-

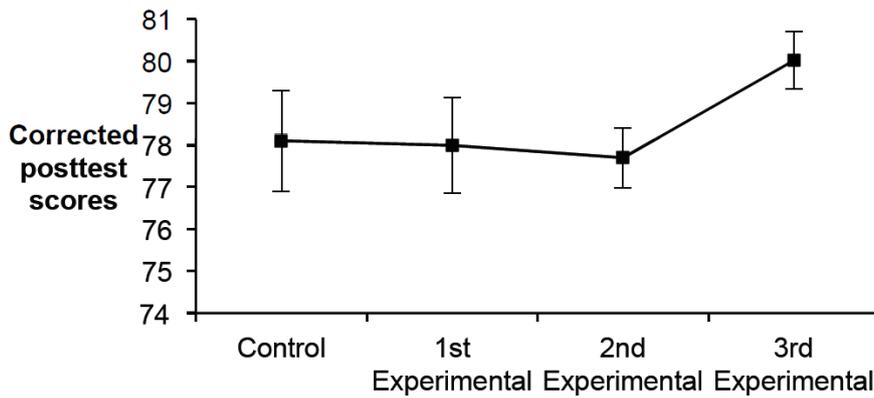


Figure 4: Posttest scores corrected for pretest differences for the four conditions (with error bars)

ematics in three workshops clearly benefits students' mathematics achievement.

ACKNOWLEDGEMENT

The ICA (Improving Classroom Assessment) project was supported by a grant from the Netherlands Organization for Scientific Research (NWO MaGW/PROO: Project 411-10-750).

REFERENCES

- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education: Principles, Policy & Practice*, 5(1), 7–74.
- Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. In H.L. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment* (pp. 3–17). Abingdon, UK: Routledge.
- Clark, I. (2012). Formative assessment: assessment is for self-regulated learning. *Educational Psychology Review*, 24, 205–249.
- Dixon, H. R., Hawe, E., & Parr, J. (2011). Enacting assessment for learning: the beliefs practice nexus. *Assessment in Education: Principles, Policy, & Practice*, 18(4), 365–379. doi: 10.1080/0969594X.2010.526587
- Erickson, F. (2007). Some thoughts on "proximal" formative assessment of student learning. *Yearbook of the National Society for the Study of Education*, 106(1), 186–216.
- Ginsburg, H. P. (2009). The challenge of formative assessment in mathematics education: Children's minds, teachers' minds. *Human Development*, 52, 109–128.
- Inspectie van het onderwijs (2013). *De staat van het onderwijs* [The current status of education]. Utrecht, The Netherlands: Author.
- Janssen, J., Verhelst, N., Engelen, R., & Scheltens, F. (2010). *Wetenschappelijke verantwoording van de toetsen LOVS Rekenen-Wiskunde voor groep 3 tot en met 8*. Arnhem, The Netherlands: Cito.
- Keeley, P., & Tobey, C.R. (2011). *Mathematics Formative Assessment: 75 Practical Strategies for Linking Assessment, Instruction, and Learning*. Thousand Oaks, CA: Corwin.
- McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative assessment on student achievement: so much more is needed. *Practical Assessment, Research & Evaluation*, 18(2), 1–15.
- National Council of Teachers of Mathematics (2013). *Formative Assessment: a position of the National Council of Teachers of Mathematics*. Position document, NCTM. http://www.nctm.org/uploadedFiles/About_NCTM/Position_Statements/Formative%20Assessment1.pdf.
- Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Suurtamm, C., Koch, M., & Arden, A. (2010) Teachers' assessment practices in mathematics: Classrooms in the context of reform. *Assessment in education: Principles, Policy & Practice*, 17(4), 399–417.
- Veldhuis, M., & Van den Heuvel-Panhuizen, M. (2014). Primary school teachers' assessment profiles in mathematics education. *PLoS ONE*, 9(1), e86817.
- Veldhuis, M., Van den Heuvel-Panhuizen, M., Vermeulen J. A., & Eggen, T. J. H. M. (2013). Teachers' use of classroom assessment in primary school mathematics education in the Netherlands. *CADMO*, 21(2), 35–53.