

A research program for studying the development and impact of formative assessment

TORULF PALM, CATARINA ANDERSSON,
ERIKA BOSTRÖM AND CHARLOTTA VINGSLE

This paper outlines the research program for the formative assessment group at Umeå Mathematics Education Research Centre. The program was presented in a symposium at the conference, and focuses on the study of the development and impact of formative assessment. The main purpose of the research carried out by the research group is to provide research results that will be used outside the research community for educational decisions on systemic level, or as support for improved teaching and learning at classroom level. The paper outlines the fundamental ideas of the program, current studies, and examples of completed studies.

In 1998 Black and Wiliam published their influential review on the impact of formative assessment. They concluded that large-scale student achievement gains are possible when formative assessment is employed in classroom practice. This sparked a strong upsurge in the interest in formative assessment, and the number of published articles about formative assessment has grown dramatically during the 21st century (Hirsch & Lindberg, 2015). In addition, the significance of formative assessment for educational practice is emphasized by international organisations such as the Organisation for Economic Co-operation and Development (OECD, 2005). The number of studies about formative assessment in Sweden is also growing, but is still quite limited, both in general (Hirsh & Lindberg, 2015), and specifically in mathematics (Ryve et al., 2015).

Formative assessment can be defined in the following way:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about next steps in instruction that are likely to be better, or be better founded, than the decisions they would have taken in the absence of evidence that was elicited.

(Black & Wiliam, 2009, p. 9)

Torulf Palm, Umeå University
Catarina Andersson, Umeå University
Erika Boström, Umeå University
Charlotta Vingsle, Umeå University

This definition affords several foci for formative assessment, and Black and Wiliam's review, which was based on a shorter formulated definition with the same meaning, indeed included studies examining the impact of different strategies for formative assessment. However, the exact meaning of the concept differs between scholars. Some focus on the teacher using tests to gather evidence of student learning, with subsequent adjustment of instruction. Others focus on the feedback from the teachers, on the role students can play to support each other's learning, or on students' participation in the formative assessment process as self-regulated learners. Some scholars researching these strategies of formative assessment use the term formative assessment (or assessment for learning), while others use denotations specifying the specific focus, for example feedback. Research reviews focusing on each of these strategies have confirmed their potential for enhancing student achievement. The reviews include feedback (Hattie & Timperley, 2007), self-regulated learning (Dignath & Büttner, 2008), self-assessment using rubrics (Panadero & Jönsson, 2013), and peer-assisted learning (Rohrbeck et al., 2003). Research reviews focusing mathematics have shown strong relationships between student achievement, and teachers' adjustment of teaching based on collected evidence of student learning (National Mathematics Advisory Panel, 2008; Yeh, 2009) and self-regulated learning (Dignath & Büttner, 2008).

However, a strong research base supporting how to effectively help teachers to implement a high quality formative assessment practice is lacking (Schneider & Randel, 2010; Wiliam, 2010), and many professional development initiatives have been unsuccessful in accomplishing a substantially developed formative assessment practice to the extent that increased student achievement was obtained (Randel et al., 2011; Schneider & Randel, 2010).

The different strategies of formative assessment above share a core of modifying teaching and learning based on identified student learning needs, but focus on different aspects of formative assessment. Thus, a classroom practice that integrates these strategies into a unity could open up extended learning opportunities. However, such a practice would be more complex and provide further difficulties in its implementation. Suggestions for such conceptualisations exist (e.g. Wiliam & Thompson, 2008). However, even though some successful attempts have been made with a random selection of teachers (Wiliam, Lee, Harrison & Black, 2004; Andersson, 2015), studies provide evidence on the difficulty of supporting teachers to developing such a formative assessment practice to the extent that it significantly affects student achievement (Bell et al., 2008; Randel et al., 2011).

The research program

The research group in formative assessment at Umeå University currently includes 10 researchers from Umeå Mathematics Education Research Centre

(UMERC). The main purpose of the research carried out is to provide research results that will be used outside the research community for educational decisions on systemic level, or as support for improved teaching and learning at classroom level. The choices and design of research projects reflects the desire to achieve the goal of being in direct service to the education community.

Therefore, the group has a strategy to engage in combined research and school development projects, in which we collaborate with schools and municipalities for mutual benefit. A main type of research carried out by the group includes designing professional development programs, and studying the significance of characteristics of such programs for outcomes such as teachers' development of a formative classroom practice and student learning. Research also focuses on the relation between characteristics of formative assessment and student outcomes. Current studies conducted by the research group include (1) a three-year combined research and school development project in an upper-secondary school based on professional development in formative assessment conceptualised as a unity of integrated strategies, and (2) a study of the impact of improved teacher support for students' self-regulated learning on students' mathematics learning activities in the early school years.

As a complement to such developmental research, we are also engaged in another type of cooperation with a municipality. In this project we study the impact of a professional development program (PDP) in formative assessment on mathematics teachers' practice. This PDP is organised by the municipality itself and is carried out in all their schools at compulsory level. Such research is conducted to gain research insights about implementations made with the intent to improve teaching and learning, and the results are intended to be used in subsequent professional development initiatives to improve the support to teachers. Another kind of research we carry out, as a complement to developmental research in collaboration with schools and municipalities, is laboratory studies about the impact of different types of reasoning on student achievement and how these types can be supported by formative feedback. These studies are made in collaboration with the UMERC research group on mathematical reasoning. The results of these studies are intended to be used in upcoming school developmental projects. For the same reason, the writing of research reviews is another complement to the developmental research that is the main focus of the group's research activities. A review of the impact of different approaches for formative assessment on student mathematics achievement is completed and currently under review, and a review on the impact of formative feedback on different types of mathematical reasoning is in progress.

The focus of our previous studies has been on mathematics, and this subject will continue to be of special importance in future studies. In addition, current studies also include other subjects, as we now conduct research projects involving all teachers and subjects in whole schools. Some studies focus on feedback or self-regulated learning, which aim at a specific aspect of formative

assessment, but in most of our studies the content includes strategies for several aspects of formative assessment.

Examples of completed studies in the research program

In the following we describe some recently completed studies in a research project about the effects of a teacher professional development program in formative assessment we developed. In the first of these studies the teaching practice of a random selection of mathematics teachers was analysed before they entered the program. The specific aim was to investigate how the teachers used formative assessment. This is of importance since little is known about Swedish mathematics teachers' current use of formative assessment (Ryve et al., 2015), and thus about the possible value of, and specific content to include in, professional development programs in formative assessment.

The same teachers were then freed from teaching for 20% during one term for participating in the professional development program (PDP). The following school year they went back to normal teaching loads again. We examined the impact of the PDP on the teachers' practice and their students' achievement in mathematics, as well as the reasons for the type of changes the teachers made in their classroom practice due to the PDP.

In a follow-up study an in-depth analysis is provided of the knowledge and skills used by one of the year 4-teachers when applying formative assessment principles. At the heart of definitions of formative assessment lies the idea of collecting evidence of students' learning, and using this information to modify teaching and learning to better meet students' learning needs. Such regulation of learning processes would require skills to elicit the thinking underlying students' oral and written responses, and the capacity to make suitable instructional decisions based on this thinking. Sufficient knowledge about the character and use of mathematics teachers' knowledge and skills when practicing formative assessment is lacking (Heritage, Kim, Vendlinski & Herman, 2009). The aim of this study is to characterize the knowledge and skills that the teacher uses in her formative assessment practice during whole-class sessions.

Methods

A framework for operationalization of formative assessment conceptualised as a unity of integrated strategies by Wiliam and Thompson (2008) was used both for the development of the PDP, and for the analysis of teachers' practice. The framework comprises a big idea of using assessment to identify student learning needs and modifying teaching to meet these needs. As a complement it includes five key strategies involving the teacher and students in the processes of identifying the learning goals, the students' learning, and how to take the next step in the learning towards the goals. The key strategies are (1) clarifying

learning intentions and criteria for success, (2) engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding, (3) providing feedback that moves learners forward, (4) activating students as instructional resources for one another, and (5) activating students as the owners of their own learning.

To examine teachers' formative assessment practice before the PDP two random samples of mathematics teachers from a mid-sized Swedish municipality were analysed (21 teachers teaching school year 4, and 17 teachers teaching year 7). The teachers were interviewed and their classroom practices were observed twice. The interviews and observations were semi-structured and interview guides and observation schemes directed the data collection. The analysis was guided by the Wiliam and Thompson framework. The purpose with the analysis was to identify the activities of the teachers' classroom practice that were used regularly and could be regarded as formative assessment.

To investigate the impact of the program on the teachers' practice interviews with the teachers about their changes in practice were made in the end of the school year that followed the PDP. In addition, complementary data were collected through unannounced classroom observations during this school year. Data about the teachers' practice before the PDP was already available from the first study. Using the framework by Wiliam and Thompson the analysis of the data was carried out with the purpose of identifying the characteristics of the teachers' changes in their formative assessment practice. To collect additional data for the study about the reasons for the type of changes the teachers made in their practice, teacher questionnaires were administered immediately after the PDP and in the end of the school year following the PDP.

To study the impact of the professional development on student achievement, control groups were used. For both school year 4 and 7, all teachers not randomly selected to participate in the PDP constituted the control groups. To compare the increase in achievement both the students to the teachers in the intervention groups and the students to the teachers in the control groups took a mathematics pretest in the beginning of the school year after the PDP, and a posttest in the end of the same school year.

The teacher chosen for in-depth analysis was one of the year 4-teachers who had made significant changes in her teaching towards a more formative assessment practice. Mathematics lessons were observed and audio-recorded for 2 months. A number of episodes involving formative assessment were analysed. The analysis was carried out in three steps. First, the interactions between teacher and students were assessed as formative if existence of the three phases; eliciting, interpreting, and use of information were identified. Second, the teacher's actions during the phases were described. Finally, the knowledge and skills the teacher used were characterized. The definition of formative assessment above by Black and Wiliam (2009) was used as an analytic

tool to identify the formative practice. A framework based on Shulman (1986), and Ball, Thames and Phelps (2008) was used to characterize knowledge and skills used by the teacher.

Results

The results of the study about the characteristics of the teachers' formative assessment practice before the PDP show that all teachers used formative assessment to some extent in their classrooms. Together they performed activities within all five key strategies and had different ways of adjusting instruction based on the information they collected about student learning. The study also identifies the characteristics of this practice, and the results indicate that there were relatively small differences in the classroom practice of year 4 and year 7 concerning formative assessment. However, it is clear from the study that there is much room for improvement in both quality and quantity of the formative assessment practice, and the study points to potential areas of development.

The results of the studies of the impact of the professional development program on the year 4 teachers and their students show that the PDP motivated the teachers to make large changes in their teaching. They added new formative assessment activities into their classroom practice to a level that had significant impact on student achievement in mathematics. The classes taught by the teachers who had participated in the PDP improved their achievement more than the classes in the control group, and this difference was statistically significant.

All teachers had implemented some of the formative assessment activities presented in the PDP, modifications of these or modifications of previously used activities. The teachers' changes span from complementing previous teaching with new activities that enhance the big idea in formative assessment to a classroom practice that is radically developed in its very foundation. None of the teachers seem to have only implemented an instrumental use of new formative assessment activities, which have been reported in several other studies (e.g. James & McCormick, 2009). Based on Wiliam and Thompson's framework (2008) further analysis shows that the teachers had developed their formative assessment practice in three dimensions: (1) the processes in teaching and learning of identifying the learning goals, the students' learning, and how to take the next step towards the goals, (2) agents in the classroom, and (3) the time from assessment to modification of teaching and learning. This three-dimensional development may have afforded new opportunities for student learning. First, the integration of the three key processes of teaching and learning may enhance student learning. Strengthening one of the processes improve the combined value of using them together. The second dimension indicates that further learning opportunities may occur by involving all agents (teacher, student, and peers) in the assessment process. The teacher and students work together to support

learning through interaction during all three learning processes and the quality of students' support to each other and students' self-regulated learning can be improved. Lastly, shortened time between assessment and modification makes formative assessment more time efficient. Less time is spent on activities less optimal for the learning and less time is spent on waiting for help from the teacher, since the students are less dependent on the teacher.

Results also show that the reasons for the teachers' implementation of formative assessment activities were well explained by the expectancy-value theory of achievement motivation (Wigfield & Eccles, 2000). The teachers developed high value beliefs for the outcome of formative classroom practice as well as high expectancies of success to be able to carry out this kind of teaching. The value beliefs included for example high experienced utility value for both themselves and for the students, and only moderate costs in terms of time and effort. According to expectancy-value theory these variables are decisive for the motivation of action. Identified important aspects of the professional development program that motivated the teachers were: (1) A formative and process-oriented character, (2) activities directly useable in classrooms, (3) positive experience of using formative assessment activities, (4) connection between theory and practice, (5) time, and (6) knowledgeable support.

Similar to the studies about the year 4-teachers, preliminary results show that after the PDP all year 7-teachers were also highly motivated to develop their practice. They also did so, but in different ways and to different degrees. The most common and frequent change was that the teachers more often, and in a structured way, elicited evidence of all students' learning with the purpose of adjusting their instruction (Key strategy 2), which led to more well-founded and more frequent adjustments of their teaching. Another common change was that they used more effective activities to engage and create thinking among all students during whole-class sessions. Only small or moderate changes were related to Key strategy 4 (peer-assisted learning) and Key strategy 5 (students as self-regulated learners). Thus, much of the responsibility for the formative classroom practice was still on the teachers. The analysis of the impact of the changes in teaching on student achievement has not yet been completed.

A main conclusion from the in-depth analysis of one of the year 4-teachers is that the formative assessment practice is a very complex, demanding and difficult task for the teacher in several ways. The analysis identified 13 activities the teacher used in the formative practice. Six of those formed the base of the teacher's formative assessment. These activities included the use of all-response systems, random selection of students to answer questions and the use of extended time to think. The teacher also engaged the students in taking an active part in the formative assessment practice. For instance, the students were asked to give examples of how to write fractions equal to $\frac{3}{2}$. They gave their answers on their miniwhiteboards (an all-response system) so the teacher

could receive information about all students' understanding. The teacher noted that not all of the suggestions were correct and decided to write the students' suggestions on her own big whiteboard: $15/100$, $15/10$ and $1\frac{1}{2}$. Then the students were given time to pair-wise assess which of the suggested fractions were actually equal to $3/2$. The teacher then randomly selected students to argue for why a certain fraction is equal to $3/2$. The other students listened to the arguments and were then given the possibility to agree with the arguments or not, and to provide their own arguments or counter-arguments.

In the minute-by-minute formative assessment practice the teacher handled unpredictable situations and made decisions about teaching and learning in a matter of seconds. Even though the teacher had some thinking time between eliciting information and using information, unexpected questions or answers occurred which put the teacher in situations where flexibility and decisions were required instantly. Knowledge of how students learn mathematics was the most frequent type of teacher knowledge used during the activities and was for example used to understand different kinds of student misconceptions.

Final remarks

Together the studies show the feasibility of supporting teachers to develop their formative assessment practice in a way that improves student achievement. But, it can be expected that teachers would need substantial time and support.

There are different advantages with different ways of organising research. The description of the research program outlined in this article points to some of the benefits of a group working together and coordinating research endeavours. The results and experiences drawn from each of the above mentioned individual studies inform the design and understanding of the other studies. For example, experiences from the impact of the PDP on teachers' practice and their students' achievement, in combination with the study on the reasons for teachers' change and the in-depth analysis of the knowledge and skills used by one of the teachers' formative assessment practice, are currently used in a new combined research and school development project. A group of researchers with a common research agenda can more quickly and efficiently gather valuable experiences to be used in a specific context. In addition, the combined results from many related studies can provide a broader picture of a phenomenon under study. This may be a particularly valuable characteristic to be able to offer for a research group interested in engaging in collaboration with schools or municipalities.

References

- Andersson, C. (2015). *Professional development in formative assessment: effects on teacher classroom practice and student achievement* (Doctoral dissertation). Department of science and mathematics education, Umeå University.
- Ball, D. L., Thames, M. H. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Bell, C., Steinberg, J., Wiliam, D. & Wylie, C. (2008, March). *Formative assessment and teacher achievement: two years of implementation of the Keeping learning on track program*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
doi:10.1080/0969595980050102
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
doi:10.1007/s11092-008-9068-5
- Dignath, C. & Büttner, G. (2008). Components of fostering self-regulated learning among students. A Meta-analysis on Intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231–264. doi:10.1007/s11409-008-9029-x
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Heritage, M., Kim, J., Vendlinski, T. & Herman, J. (2009). From evidence to action: a seamless process in formative assessment? *Educational Measurement: Issues and Practice*, 28(3), 24–31.
- Hirsh, Å. & Lindberg, V. (2015). *Formativ bedömning på 2000-talet*. Stockholm: Swedish Research Council.
- James, M. & McCormick, R. (2009). Teachers learning how to learn. *Teaching and Teacher Education*, 25(7), 973–982. doi:10.1016/j.tate.2009.02.023
- OECD (2005). *Formative assessment: improving learning in secondary classrooms*. Paris: OECD.
- Panadero, E. & Jönsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: a review. *Educational Research Review*, 9, 129–144. doi:10.1016/j.edurev.2013.01.002
- Randel, B., Beesley, A. D., Apthorp, H., Clark, T. F., Wang, X., et al. (2011). *Classroom assessment for student learning: impact on elementary school mathematics in the central region*. Washington: National Center for Education Evaluation and Regional Assistance.
- Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W. & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school studies: a meta-analytic review. *Journal of Educational Psychology*, 95(2), 240–257.

- Ryve, A., Nilsson, P., Palm, T., Van Steenbrugge, H., Andersson, C., et al. (2015). *Kartläggning av forskning om formativ bedömning, klassrumsundervisning och läromedel i matematik*. Stockholm: Swedish Research Council
- Schneider, M. C. & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 251–276). Abingdon: Routledge.
- Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, 15 (2), 4–14.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25 (1), 68–81.
doi:10.1006/ceps.1999.1015
- Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). Abingdon: Routledge.
- Wiliam, D., Lee, C., Harrison, C. & Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education Principles Policy and Practice*, 11 (1), 49–65. doi:10.1080/0969594042000208994
- Wiliam, D. & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: shaping teaching and learning* (pp. 53–82). Mahwah: Lawrence Erlbaum Associates.
- Yeh, S. S. (2009). Class size reduction or rapid formative assessment? A comparison of cost-effectiveness. *Educational Research Review*, 4 (1), 7–15.
doi:10.1016/j.edurev.2008.09.001