International Journal of Learning, Teaching and Educational Research
Vol. 11, No. 1, pp. 53-74, April 2015

A Study of Formative Assessment Strategies in Teachers' School-Based In-Service Training

Eva Nyberg and Mona Holmqvist Olander
Department of Pedagogical, Curricular and Professional Studies
University of Gothenburg
Gothenburg, Sweden

Abstract. The aim of this study was to explore a team of teachers' (*n*=4) use of theoretically based formative assessment strategies within the course of a learning study. The thematic analysis is based on video observations of teachers' discussions during planning meetings, teaching in the classroom and evaluation meetings. The subject-specific content focused on learning about fractions, specifically the concepts of double and half, in three groups of six- to seven-year-old students (*n*=51 in total). An iterative process was used in which the teachers in the study used video recordings as a tool for analyzing their work in the classroom. The thematic analysis shows that the use of a general learning theory – variation theory – strengthens the effect of the teachers' formative assessment. Without explicit use of the assumptions from the theoretical framework, the formative assessment strategies had only a minor impact on students' learning outcomes.

Keywords: Formative assessment, Classroom study, Elementary school, Variation theory; Learning study.

1. Introduction

There have been several attempts to develop teachers' formative assessment by means of in-service training. A range of studies concerning programs for in-service training aiming at developing teachers' competence in performing formative assessment have been carried out, with various outcomes (e.g., Bennet 2011, Phelan, Choi, Vendlinski, Baker & Herman, 2011). It thus seems rather difficult to transform formal training about formative assessment into classroom practice. Wiliam (2006) claims that "tools for formative assessment will only improve formative assessment practices if teachers can integrate them into their regular classroom activities" (p. 287). School-based in-service training could therefore be one way to develop teachers' abilities to use formative assessment to increase the students' learning outcomes. This school-based research project involved a team of teachers who had previously participated in an in-service training course on formative assessment, and the focus was on their use of

formative strategies in their classrooms during an iterative process in which the researchers were partly involved.

Formative assessment can be approached in several different ways. In this study, the in-service training was aimed at encouraging the teachers to base their formative assessment actions on theoretical conjectures. Our hypothesis is that formative assessment strategies such as feedback become more powerful if they are based on a theory of learning, and we therefore explore the effect of applying the principles of variation theory (e.g., Marton, 2014) during planning, implementation and evaluation of teaching and learning, in order to increase student learning.

Research has shown that formative assessment often lacks a theory of action, which makes it difficult to evaluate and understand the mechanisms causing the intended effects:

> Unless we understand the mechanisms responsible for change, we won't know if the effects are due to those mechanisms or to irrelevant factors. We also won't be able to predict the conditions, or population groups, for which the formative assessment is likely to work (Bennet, 2011, p. 14).

Bennet (2011) also argues that a teacher's hypothesis about a students' understanding is dependent on the teacher's cognitive model, which can help the teacher to evaluate a student's understanding. Without a theory of learning, there is a risk that the teachers carrying out the formative assessment will fail to identify the underlying mechanism for how the learning takes place. This is also stressed by Black and Wiliam (2009, 2012) and Wiliam (2009), who argue that to know what feedback to give to students, the teacher needs both a theoretical model of how students learn and the ability to apply this theoretical understanding in a specific context. In light of this, this research project aims to further study the way in which the theoretical assumptions of variation theory can be used as guiding tools for teachers to assess their students formatively.

In formative assessment, the data gathered during assessment is used to inform decisions regarding modifications and adaptations of the teaching to meet the learning needs of the students (e.g., Bennet, 2011; Black & Wiliam, 2009). This includes using evidence about difficulties collected in one group of students "to modify instruction for another group of students at some point in the future" (Wiliam, 2009, p. 26). This means that it is only if the result from the assessment is used to inform teaching and learning that it can be said to have a formative function. Phelan et al. (2011) produced a study pointing out the difficulties in identifying what constitutes effective formative assessment. They studied teachers and students in a formative assessment intervention, aimed at improving student performance in mathematics. Surprisingly, the authors did not find any significant differences between the treatment group and the control group; the hypothesis that the improvement would be greater in the treatment group—in which formative assessment was used—remained unverified. Although the extensive research overview by Black and Wiliam (1998) delivered

the clear message that more formative assessments in school can lead to great improvements in student learning, the above-mentioned study by Phelan et al. (2011) emphasizes that it is not simply the case that any formative assessment tool leads to an improvement in student performance. The teachers' competence in performing the formative assessment is most likely crucial for the outcome, and therefore teacher variables need to be more closely analyzed. What matters seems to be how the formative assessment is carried out, i.e., *what* is noted by the teacher and *the way in which* this is addressed during the instruction. This means that effective formative assessment is constituted by both knowledge of the content in question and knowledge of what it takes to learn this specific content, in line with Bennet (2011), as well as Black and Wiliam (2009). This is the reason why we have added a theoretical framework to strengthen the teachers' knowledge about learning during the in-service training.

In this article, we have taken into consideration the fact that to be formative, the instruction needs to be specific at a micro-level in relation to what is elicited through the assessment. This involves determining in what way the content is offered in relation to how it is experienced by the students and doing this according to theoretically based assumptions about what it takes to learn. The teachers in our study were guided by the variation theory of learning (Lo & Marton, 2012; Marton & Booth, 1997; Marton, 2014; Lo, 2013; Holmqvist, Gustavsson & Wernberg, 2008; Holmqvist, 2011), to which they were introduced by the researchers participating in the interventions during the introduction to the in-service project. The concept of variation in variation theory does not refer to varying methods, but rather to varying the features of the content that have not been previously discerned by the students. The assumptions are that to discern new aspects of the object of learning, these have to vary against an invariant background consisting of the aspects already known. Variation theory will be more thoroughly described below.

2. Aim and research questions

The aim of this study was to explore a team of teachers' (n=4) use of theoretically based formative assessment/feedback strategies during a school-based in-service training in a primary school. The subject-specific content focused on learning about fractions, specifically the concepts of double and half, in a group of six- to seven-year-old students. The research questions were:

> 1. In what way do the teachers take advantage of the variation theory of learning in their formative feedback to the students?
>
> 2. What impact does the theoretically based formative assessment have on the teachers' way of constructing lessons in new groups of students?

3. The iterative in-service training process

In this section we describe the tools used by the teachers in their in-service training: variation theory (theoretical framework) and the learning study process (an iterative process including video recordings).

3.1 Variation theory as a guiding principle

Educational learning theories often describe learning in terms of conditions for learning. For example, a theory might state that learning takes place in an interaction between people, or as an internal construction by the individual, or through a combination of both. However, these theories do not usually offer explicit guidance on how to develop domain-specific learning or how to use the theories in practice. Variation theory is built on the research field of phenomenography, which originated in the 1970s and is concerned with studying qualitatively different ways of experiencing the same phenomena (Marton, 1981; Marton & Booth, 1997; Marton, 2014). Variation theory is thus useful in providing a way to analyze what it takes to learn. It is also used as a guiding principle in lesson planning (Holmqvist, Gustavsson & Wernberg, 2008; Kullberg, 2010; Lo & Marton, 2012). Its point of departure is non-dualistic, meaning that the content to be learned cannot be separated from the learner's understanding of it (Lo & Marton, 2012).

Discernment, simultaneity and variation are cornerstones in variation theory. To learn, one has to discern what one has previously not been aware of and relate it to what one already knows. By varying aspects of the content that have not been previously discerned by the learner against an invariant background of aspects that are already known, new aspects become discernible. These patterns of variation are taken into consideration during planning (the intended object of learning) and implementation of the lesson (the enacted object of learning), and when analyzing the learning outcomes shown by the students' test results (the lived object of learning). One important point is that the pattern of variation will differ depending on the students' previous knowledge; it is based not only on what it takes to learn in relation to specific content but also on what it takes to learn specific content in relation to what is already known by the learner (Pang & Marton, 2013). The differences in improvement are thus not due solely to the design, method or other strategies used – for example formative assessment techniques as such (Phelan et al., 2011) – as the students' perspective has to be addressed in an adequate way if the full power of formative assessment is to be exploited. The design of the instruction aims to make visible those aspects that are necessary for further learning, called critical aspects. In variation theory, this refers to aspects of the content that have not yet been discerned by the learner but have to be discerned in order to develop further knowledge. As soon as the aspect is discerned, it stops being critical. If a lesson offers only aspects of the content that are not critical for learning, for example aspects that the learner has already discerned, then no learning will take place.

3.2 Learning study as a school developmental model

The learning study is a model developed by researchers from Hong Kong and Sweden (e.g., Lo, Pong & Chik, 2005; Holmqvist, 2010; Marton, 2003; Marton & Tsui, 2004), inspired by the Japanese lesson study (Lewis, Perry & Murata, 2006). It entails a systematic and cyclical process consisting of a number of stages, starting with the choice and study of the object of learning to be addressed and taught. In this iterative model, the design of new lessons is based on the analysis of earlier lessons, with the aim of further developing students' previous learning outcomes. The object of learning is initially chosen by the teachers and defined

in detail after a screening, an interview or test to assess students' pre-knowledge, which identifies the aspects that may be critical for learning. A pretest is thereafter constructed based on the results of the initial screening and guided by the chosen theory of learning, in this case, variation theory. The teachers and researchers collaborate in planning the first lesson, which is taught by one of the teachers in the group. The lesson is video recorded. During the subsequent meeting the lesson is analyzed by the teachers and the researchers using both the video recording and the results of the pretest and posttest. The focus in these meetings is the learning outcomes, which provide the basis for the development of a new lesson to be taught to another group of students. If the learning study also has a research aim, an analysis of the whole process is made to report the results of the study. Learning studies, which involve teachers and researchers in collaboration (Marton & Tsui, 2004), have been implemented in 120 schools in Hong Kong in a project called 'Variation for the improvement of teaching and learning project' (VITAL). The project has been evaluated (Elliott & Yu, 2008) showing the benefits for both student and teacher learning (Elliott & Yu, 2013). On the other hand, research on the collaboration between teachers and researchers shows tensions of several different kinds: "outsiders' versus insiders' perspectives; academic versus grounded knowledge; unclear hierarchical statuses; and diverse and conflicting agenda" (Adamson & Walker, 2011). Empirical studies on the implementation of learning studies in other parts of the world, in this case Sweden, have also found them to be effective both for students and for professional development among teachers (Holmqvist, 2006, 2011). The first research project where learning studies were implemented in Sweden was funded by the Swedish Research Council (Holmqvist, 2002), and 18 learning studies in three different school subjects were studied (Gustavsson, 2008; Wernberg, 2009; Kullberg, 2010). The results show an increased learning outcome both in the short-term and the long-term (Holmqvist, 2011). In a similar model for school development, the lesson study (Stigler & Hiebert, 2009), teaching is developed through a similar cyclical process. The main difference is that learning studies are guided both by a theory of learning and a focus on content-specific research into students' understandings of the specific object of learning (Holmqvist, Gustavsson & Wernberg, 2007, p. 189). The focus in learning studies is on the learners' understanding of the content, while the focus in lesson studies is mainly on the improvement of the lesson itself or on other issues needing to be improved in the classroom. The results of an evaluation of the Swedish learning studies show that the teachers see learning studies as time-consuming and difficult to work with continuously, despite the high student learning outcomes (Olteanu & Lennerstad, 2011). Lewis (2015) points out the lesson study as a so-called improvement science, which theorizes the need for two different types of knowledge sciences: a system of basic knowledge from the discipline of education and a system of profound knowledge. The learning study process, in which a combination of theoretically based conjectures and the teachers' deep knowledge about teaching is needed, also seems to fulfill the requirements for being labeled as improvement science.

3.3 Formative assessment in the learning study model

In a learning study, the aim of the assessment is to make in-depth microanalyses of the students' learning in order to identify what the critical aspects of the content in question are in relation to the learners, and how these aspects can be made discernible in a powerful way. Having identifying the critical aspects, the teachers use and try out this information in their formative assessment—that is, they use it to guide their teaching in the classroom. Our hypothesis, therefore, is that formative assessment informed by variation theory enables teachers to pinpoint the aspects that should be focused on in the learning situation and to change this focus depending on the learners' needs. James and Pedder (2006), when discussing results from a project including a survey of 1000 teachers, assert that programs of professional development "should be focused on classrooms and classroom practice" (p. 39). One of their conclusions is that both individual and social processes are "important conditions for the promotion of assessment for learning in classrooms" (James & Pedder 2006, p. 39). They put forward the concept of "research lessons" as described by Stigler and Hiebert (1999) as one possible approach to developing assessment practices in the classroom. Pedder and James (2012) further stress the importance of collaborative, classroom-based professional learning "for fostering effective assessment for learning" (p. 41). We agree with these claims, and consider the learning study as a school-development process in line with the findings about powerful professional development. By the use of a theoretical perspective on learning, such as variation theory, the formative assessment in the learning study helps the teachers to understand what needs to be changed in instruction or which feedback is relevant in order to increase students' learning, using the theory as a guiding tool.

4. Procedure of the Study

4.1 Method

The analysis of the teachers' strategic use of variation theory is qualitative and based on video recorded meetings and interventions as well as observations. A thematic analysis was made (Boyatzis, 1998) based on several readings of the material (which was transcribed verbatim), as well as watching and re-watching of the video recorded meetings and lessons. The students' summative assessments were used as a triangulation to strengthen the observations and find out whether the teachers' use of theoretical assumptions was reflected in students' learning outcomes.

4.2 Context

The study took place during an in-service training project, which was conducted in a school district in a rural area close to a small town in the south of Sweden. In the Swedish school system, all classes are mixed with regard to both gender and abilities. Students spend 9 years in compulsory school, from age 7 to 16. The first time the children receive grades is at the age of 12, i.e., in grade 6. All classes include children of the same age and during the first 6 years, the classes normally consist of groups of 15-25 children.

The content of the lessons in the Swedish school system is governed by a national curriculum (Swedish National Agency for Education, 2011), but the design of instruction and methods used is left to the teachers. The education to become a primary school teacher consists of four years of academic teaching studies and vocational training at university level.

4.3 Participants and implementation

Two researchers and four teachers from two different primary schools participated in the study, which took place over one semester. 51 seven-year-old students (24 girls and 27 boys) were involved in the study. The students' learning process and learning outcomes are not, however, in focus in this paper; the findings about the students form another part of the project (Holmqvist & Nyberg, 2014). Three of the teachers were each in charge of one lesson, including pretest, research lesson and posttest. One teacher, who was responsible for developing mathematics instruction in the school district, participated only in the planning meetings. Before the first meeting, as a part of the overall in-service project in the school district, the teachers were introduced to variation theory and the learning study model in a lecture by one of the researchers. The teachers were also given a book about the theoretical background and concept of the learning study (Holmqvist et al., 2006). The researchers' deeper knowledge of variation theory, however, was important throughout the discussions. The object of learning chosen by this group of teachers was the ability to double and halve numbers. This topic was chosen based on the teachers' previous observations that children had difficulty learning how to perform these operations with numbers, but not with concrete objects. The teachers had previously been introduced to formative assessment.

The lessons and the planning meetings were video recorded. Before each new lesson, the recording of the previous lesson was analyzed, and experiences from that lesson were evaluated and discussed, including the results of the pretest and posttest. Three lessons were conducted, meaning that there were four meetings between the teachers and researchers (see Fig. 1). During the first meeting, the pretest was constructed and the first lesson was jointly planned. During the second and third meetings, the previous lessons were evaluated with respect to learning outcomes, and the coming lessons were planned on the basis of this evaluation. During the fourth meeting, the last lesson and the results from the tests were discussed, as was the outcome of the learning study as a whole.

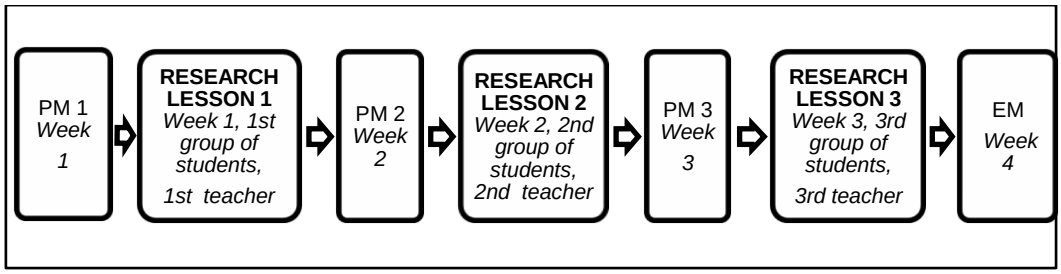

Figure 1: The four-week time-line of the learning study. PM=Planning meeting, EM=Evaluation meeting.

The length of intervention for each group of students was one lesson comprising approximately 15 minutes of pretest, 30 minutes of instruction, and 15 minutes of posttest. The lessons were planned and evaluated on the basis of variation theory.

4.4 Data

The data collected during the school-based study consists of video recorded meetings (3), video recorded lessons (3) and one final observed meeting.
Participating observations: Four planning meetings with the teachers were conducted during the four weeks of the learning study. Each meeting took place at the schools where the teachers worked. The meetings lasted 2 hours.
Videotaped lessons: Three lessons were videotaped. Each lesson lasted 1 hour. Group 1 (Cycle 1) consisted of 24 students, group 2 (Cycle 2) consisted of 13 students and group 3 (Cycle 3) consisted of 14 students.

4.5 Analyses

The video recordings from the lessons (n=3) and planning meetings (n=4) were transcribed verbatim. The data was analyzed as an exploratory single case study (Stake, 2006; Yin, 2009) at a fine-grained level (Phelan et al., 2011). For the qualitative analysis of the teachers' use of theoretically based formative assessment strategies and reasoning, a thematic analysis was used as a first step (Boyatzis, 1998). The analysis was based upon how the core concepts of variation theory – contrast, simultaneity and variation (Holmqvist, 2011; Kullberg, 2010; Marton, 2014; Lo, 2012) - were used by the teachers, and in what respects they were used in the planning, implementation and evaluation of each lesson. Thereafter a more detailed and specific analysis was made across the three lessons including the planning and evaluation meetings, during which the different data sources were compared and analyzed in parallel.

4.6 Ethical considerations

The study followed the ethical considerations described by the Swedish Research Council (Hermerén, 2011). All participants, including parents, teachers, school leaders and children, were continuously informed about the aims of the study, the use of the data and their rights to confidentiality and to withdraw from participation. The parents also signed an agreement to let their children participate in the study and to enable the researchers to use the data.

5. Results

The analysis describes the different ways the teachers' formative assessment strategies are expressed during the study. The analysis ends up with three themes regarding the teachers' development of formative assessment strategies guided by the theoretical assumptions. The first theme describes how the teachers, through their formative assessment, gradually developed insight about what critical aspects are and how they can be used to increase the students' understanding of halving and doubling (Theme A). The second theme highlights another example of how the teachers' formative assessment led to increased evidence-based insight into the students' understanding during the course of the

learning study (Theme B), by taking into consideration the connection between the object of learning and the learners pre-knowledge (non-dualism). The third theme addresses the teachers' feedback during their respective lessons, focusing mainly on how the teachers used variation theory (Theme C) in their feedback to the students.

5.1 Theme A: Critical aspects of the content

The knowledge about what critical aspects are and how they can be dealt within instruction is the main focus in this theme. The teachers found that to understand the concept of doubling, the students had to discern variation in amounts at the same time as discerning the invariant aspect of doubling—that is, taking the same amount "one more time" and adding it to the original amount. This section presents the analysis of how this was addressed and handled during the course of the learning study.

5.1.1. The first meeting

During the first planning meeting, the test questions and the design of the first lesson were discussed and constructed. The design was based on the results from the screening carried out by the teachers themselves, the teachers' previous experiences from teaching this content, the researchers' knowledge of common obstacles for students, and the researchers' extensive experience of using variation theory for test construction and lesson planning. The various aspects the object of learning were analyzed, to determine which aspects the students had already discerned and which ones they needed to discern in order to develop their knowledge.

During the screening, when one of the students was asked by the teacher to double the number four, the child answered five (i.e., four plus one). This student evidently interpreted "add one more time" as "add one (the number 1)". During the planning meeting before the first lesson, one of the teachers confirmed that this was a critical aspect of the object of learning.

> Excerpt 1, first planning meeting (T1 = Teacher 1):
>
> T1: Yes, we have many children who cling to that; they don't get doubling, for them it is … plus one. We tried previously when we taught this particular topic not to start with one, because we thought that was what made them stick to this, but there are those who still hold on to "plus one" even if we start by asking them to double three.

The teachers and the researchers consequently set up the hypothesis that one important part of the instruction might be to avoid the expression "one more time", which the students could interpret to mean "plus one". Instead, the expression "the same amount one more time" was to be used during the first research lesson to make it possible for the students to discern that "one more time" is not the same as "one more". For the students who had not already noticed the difference, this might be a critical aspect.

One of the concepts in the theoretical framework used here is about how to make aspects of the content discernible for the students. In this case the concept of simultaneity was used. The first lesson was therefore planned to allow the students to apply doubling and halving at the same time, but without having the original number of objects on their desks. The decision not to let the students see the objects representing the original number was based on previous experience of students making a mechanical visual doubling, without really understanding the concept of doubling, e.g., by adding the same amount once again, which results in an understanding of double as "the same as"; doubling four then gives four instead of eight, as the student added four to the original amount.

5.1.2 Second meeting

The first lesson was evaluated in the second meeting. The analysis during the meeting of the pretest and posttest focused on the theoretical concept of simultaneity. However, in their analysis of the test results, the teachers found no significant improvement regarding mean points. The students (n=24) scored 3.9 (out of 10 maximum points) in the pretest and 4.7 in the posttest (p=0.073). The analysis of the test revealed that after instruction, the students still had problems discerning what doubling means. In one of the test questions the students were asked to draw double or half of a particular number of objects, having chosen the original number themselves. This was as difficult for them after instruction as before instruction. It seemed that the students had not yet learned to separate the original number from the concept of doubling or understood that doubling and halving was not related to a specific number but involved the same general processes regardless of what number used. To allow the students to separate the rules for doubling and halving from the object being processed, the group of teachers and researchers decided that simultaneity was to be used. By contrasting halving and doubling, the aim was to enable the student to separate the original amount from half the amount as well as double the amount. The aim of looking at halving and doubling with different numbers was to help the students to generalize and discern the general idea of halving and doubling instead of e.g., adding one to the original amount in all cases. The review of the video recording of the lesson during the meeting revealed another way of understanding doubling that the group had not previously encountered, in which doubling an amount was taken to mean adding two, instead of taking the original amount once and once again. The information "take the original amount twice" was understood as +2. The contrast between the original amount and the doubled amount had thus not been explicit enough and therefore the difficulty remained.

> Excerpt 2, second planning meeting (R1 = Researcher 1):
>
> R1: We said it had to do with "one more time", adding or subtracting one, but here they seem to come to the conclusion to add two because they understand "take the same amount twice" to mean "add two". You responded to that well, but we will take this to the next lesson because we didn't know it before.

The review during the meeting of the video recording of the first lesson clearly showed that the teacher had followed the agreed plan based on the micro-analysis, that is, using the expression "the same thing one more time" rather than "the same thing twice":

> Excerpt 3, first lesson (T1 = Teacher 1):
>
> T1: Now, you are to add twice as many as six. I put six pieces here. I have now added six one time. If we want to have double the amount, we must have this six and then I have to add another six pieces.

One conclusion after this discussion was that it was very important to be even more explicit and use the phrase "the same amount as the first and then the same amount one more time" instead of "the same amount twice".

Another discussion concerned how to design the task to help the children distinguish between the original amount and the new amount. While watching the video recording, the teacher who had taught the first lesson reflected on whether one reason for the students' confusion might be that the students did not have the initial number of objects in front of them while working with their tasks.

> Excerpt 4, second planning meeting (T1 = Teacher 1, R1 = Researcher 1):
>
> T1: So, the question is whether it is best to put the objects away or whether it would have been better to let them have the original number of objects in front of them and let them do it again, so to speak … do it next to … so that both were there to make it possible for them to compare …
>
> R1: Maybe it would be even clearer to have a borderline in between … so: "This is what we had from the beginning, and now we are going to put down double that amount, so this is what should be there now" … We would then not get the problem of their just adding the same amount … to be clearer that they should not just list … but to understand doubling, you must understand that you have a new amount that is twice as large as the other.

Here, the teacher who had given the first lesson wondered whether it might have been better for the students' learning if they had had the original amount that they were asked to double or halve in front of them so that they would be able to contrast this with the new amount. This discussion concerned how the children could be helped to separate the original amount from the solution, avoiding just adding the same amount as the initial and adding this to the 'doubled' amount (double of four experienced as plus four). In the end, the group agreed upon increasing the contrast between the amounts, and the plan for the second lesson was therefore to work with the original number of objects and its double or half simultaneously but separated from each other. To make

this relation clear, it was decided that the original amount should be placed on one side of a border and the doubled or halved amount should be placed on the other. In this way, the original number of objects and the halved or doubled amounts could be contrasted and discernible simultaneously. This should be done by both the teacher and the students while they solved tasks during the lesson, and the initial number of objects should be explicitly contrasted with the new amount, in accordance with the assumptions of variation theory.

5.1.3 Third planning meeting

After the second lesson, during the third planning meeting, the analysis of the pretest and posttest indicated that the second group of students had understood the content better than the first group. The analysis of the pretest and posttest for this group of students (n=13) showed significant differences between pretest and posttest. Mean scores increased from 6.5 to 8.2 ($p=0.015$), thus indicating increased learning, in contrast to the first group of students, where, as mentioned above, no significant differences were found.

The increased mean scores indicated that the design chosen for this lesson had been successful. However, the tasks in the test that allowed the original number of objects to be chosen by the students themselves were still problematic in this second group of students. This third meeting therefore included discussions about various ways to improve the students' abilities to solve these tasks.

The teacher who taught the second lesson (T2), the teacher who would teach the third lesson (T3), and one of the researchers (R1) discussed the issue of simultaneously contrasting different amounts when working with the concepts of doubling and halving in class.

> Excerpt 5, third planning meeting (T2 = Teacher 2, T3 = Teacher 3, R1 = Researcher 1):
>
> R1: We talked last time about having the original amount, and then half and twice that amount. This change, I think, would be interesting because you would have the example of both half and twice the same original amount. … You put the original amount there at the same time as you tell them to put down half and double that amount.
>
> T3: However, why have you put the ruler along there? (Points at R1's paper, on which a ruler has been laid down to divide the paper into two areas.)
>
> T2: To make them see the point with … the original amount.
>
> R1: It would have been good … if you had … It may be a paper or something like that which you just copy … and then they put down the same original amount as you as the teacher have, in the middle. … So, if you make copies of a handout you can mark HALF and DOUBLE … there … if you want to emphasize that.

> T2: Then, you go on with the simultaneity the whole time … in everything you do, so you don't … it is quicker too.

During this third meeting, the group decided that simultaneity, not only regarding the initial amount and half or double but all three, should be used in the subsequent lesson for both the original numbers and their doubles and halves. The students would work with double and half of the original amount at the same time, meaning that even more simultaneity would be involved in the teacher's instruction. The original amount, half the amount, and double the amount would be contrasted with each other simultaneously. Another important conclusion was that the original amount should remain unchanged in the middle of the children's worksheet, and the teacher should say "the same thing again and once more" to explain the doubled amount. This would be emphasized by the use of one more borderline than in the second lesson—that is, the paper used for the exercise would be divided into three fields, instead of two, with the initial amount placed in the middle, and the doubled and halved amounts on either side, separated by lines.

The pretest and posttest associated with the third and last research lesson did not show improvements over the second lesson. The mean score in this group of students (n=14) was 6.1 in the pretest and 7.4 in the posttest, a difference that was not significant (p=0.162). This can be compared to the previous research lesson (lesson 2), after which the mean posttest result had increased significantly (from 6.5 to 8.2 points). The tasks in which the students themselves chose the original number to double or halve were still problematic, as were other tasks dealing with halving and doubling a predetermined amount.

This result illustrates how small changes in instruction make differences in what the students can learn. It also shows how, even if the teachers in their joint planning of the lesson had found a way to adjust the instruction to the students' knowledge, these plans were not always actually understood and implemented by the teacher giving the lesson. To understand the result, the teachers and the researchers reviewed the videotape of the lesson. This showed that the exercise where the intention had been to use simultaneity to clarify the difference between the original amount and the doubled and halved amounts, was performed mostly by the teacher, but not by the students during their work in the classroom. The change planned for this lesson, to use simultaneous contrast between all the amounts to make the difference even more explicit, was thus only partly carried out by the teacher in charge of the third lesson. The formative assessment used during the lesson was therefore not informed by the theoretical assumptions discussed during the planning meeting. One example was the decreased number of examples of amounts used during the teacher-led instruction, which only included one number (6). Another issue was when the teacher, demonstrating the examples, twice placed the objects in the wrong area on the overhead projector (placing the doubled amount in the area for half). She corrected the error when it was pointed out by a student, but did not explain to the class what had been wrong or why she changed the placement of the objects.

This oversight might give students the impression that this placement is randomly chosen and not a deliberate content-based choice.

In conclusion, the third lesson could not verify the group's hypothesis that increased contrast between the different amounts would help students to discern the concepts of double and half, since simultaneity was not used as was agreed upon. The theory was thus not used as planned and the students' learning outcomes did not improve. However, the teachers became even more aware of how small changes might have a great impact on students' performance and in what way theoretical assumptions can be handled in the classroom.

5.2 Theme B: The undivided learning space

One of the assumptions of variation theory is that the learner and what is going to be learnt cannot be separated from each other. This means that the way the learner experiences the object of learning is unique and this has to be taken into consideration in a teaching situation. Finding powerful ways to develop student learning requires both deep knowledge about the content being taught and deep insight in the students' pre-knowledge. Otherwise it becomes difficult to design powerful learning situations. It is clear from the analysis of the planning meetings that the teachers' insights into the students' understanding of double and half increased during the study, both as a result of the repeated analyses of the pretests and posttests, and as a result of the discussions based on the video recordings of the lessons. One example concerns the teachers' opinions about the children's abilities to understand the concept of "half". The teacher who performed the first research lesson (T1) declared, during the first planning meeting, that the problem children had with understanding doubling and halving was not the concept of half, but the use of the expression "half as many".

> Excerpt 6, first planning meeting (T1 = Teacher 1):
>
> T1: The difficulty with "half" is not really the concept of "half". As long as you stick to saying "half", then you can do it. Because these things, they can *divide* them. What makes it complicated, as we have understood it, is when you say half as many. For that, of course, makes it difficult because there should be *fewer* and yet you say *many*.

Analysis of the pretest and posttest results associated with the first lesson showed that before the lesson, more students gave a correct answer to the questions asking for "twice as much" than to the questions asking for "half as many". After the lesson, more students were better able to answer questions about halving, but still not as well as they could answer questions about doubling. Five items in the pretest and posttest concerned doubling, and five concerned halving. The students' mean scores on the pretest were 0.45 for the items on doubling and 0.33 for those on halving, while the corresponding scores on the posttest were 0.51 and 0.44. However, despite this data, at the second planning meeting, the teacher quoted above still believed that the concept of "half" was easier for students.

Excerpt 7 (T1 = Teacher 1, T2 = Teacher 2, T4=Teacher 4, R1 = Researcher 1):

T1: However, it feels like *half* is so much easier.

R1: Yes, but we thought last time that half was harder … half as many.

T2: However, I think you learn to take "half" before you learn how to "double", that is … to share with a sibling or with a friend … they have known this since before …

R1: However, they're not better at "half" than "double".

T2: Aren't they?

R1: They do worse at half in all the items.

T4: It's the b-alternative, which …

R1: Yes, exactly, and the b-alternative is the most difficult straight through, except for the question with the squares.

T2: Halving is more difficult than doubling.

T1: The question is whether it would have been easier if they had had this stuff [in front of them], then I think certainly they can take half of it, but only if they can actually see it …

R1: Yes, but they cannot either. They have the stuff there, but still …

T1: However, don't they get it more right when they have to split it in half?

This example shows that eliciting the students' understanding through fine-grained analysis of the scores on diagnostic test questions gradually changed the teachers' view of what the students found hard to understand. Before the first test was conducted (during the first research lesson), the teachers assumed that halving was easier for the students to understand than doubling, but when the tests revealed that this was not the case, their opinion was slowly altered. The excerpts indicate, however, that their initial opinion was quite persistent and difficult to change; they were not easily convinced of the contrary even if the test results showed this. The design of instruction thus risks focusing on aspects that are not problematic for the students and neglecting the aspects that are critical.

The test was designed on the basis of variation theory, and by contrasting halving with doubling, it was possible to compare the students' knowledge of these concepts. The results of the tests informed the teacher's formative assessment, as long as the teachers accepted what the tests really said about the

students' knowledge and let go of their own beliefs. As exemplified in the exchange quoted above, it was however difficult for the teachers to take the students' perspectives when these perspectives contradicted the teachers' own beliefs. Previous research results have indicated that formative assessment might not always have an impact on students' learning outcomes, and if teachers' instruction depends on their own beliefs instead of students' knowledge, it is obvious that the impact of formative assessment is weaker than if they take the students' knowledge into consideration.

5.3 Theme C: Theoretically guided formative feedback

In this theme, the teachers' formative feedback in the classroom is in focus. The analysis of the teachers' feedback about their respective lessons focuses on how they used variation theory to understand and respond to the students' understanding of the content.

In spite of the teacher's obvious engagement with teaching in the first lesson and in the evaluation of this lesson, the subsequent analysis of the video recording showed that the lesson rarely included direct feedback to the students. The teacher listened to the students' answers, but seldom commented on whether the answers were correct.

> Excerpt 8, first lesson (T1 = Teacher 1, S1 = Student):
>
> S1: We had six, then we thought that half of them was all of them.
>
> T1: You thought half was all and so you took them all away. So, we have some different answers. Let us see what the others thought. What did you come up with, [student name]? (Turns to another group.)

After listening to the first student, the teacher turned directly to the next group without revealing the correct answer, and this answer was not revealed until summing up the instruction at the end of the lesson. This lack of feedback was not highlighted or discussed when the group evaluated the first lesson during the second planning meeting, so we do not know whether it was common for this teacher or not; it is possible that the teacher conducted the lesson differently from how she usually would due to her awareness of being filmed. Nevertheless, it is evident that the students were left in doubt several times.

According to the videotapes, the teacher who conducted the second lesson gave the students more detailed feedback regarding critical aspects than did the first teacher. It is probable that the discussion during the evaluation of the first lesson had an impact on this teacher's awareness of which aspects of the concepts of doubling and halving the children might not have discerned, which may have led her to challenge the students' answers with a more developed formative assessment than the first teacher was able to. In the example below, the teacher worked formatively with a student who was asked to show half of eight. The teacher clearly contrasted the doubled and halved amounts, comparing them

with each other to help the students discern the difference between double and half, taking advantage of the experiences from the previous lesson.

Excerpt 9, second lesson (T2 = Teacher 2, S = student):

T2: Was he asked to put down half or double the amount?

S: Half!

T2: Half. Is there the half of the amount on that side here? (Points to the tray and one of the piles with four chestnuts.)

(The student nods.)

T2: How many are here? (Points to the tray and one of the piles with four chestnuts.)

S: Four!

T2: And how many are here? (Points to the tray and the other pile with four chestnuts.)

S: Four!

T2: However, that's just as much! Where is half the amount?

Here, the teacher made a serious effort to understand which aspects the student had experienced and which the student had not yet discerned. She evidently used knowledge gained from the analysis of the first lesson about how the aspects might be discerned and how to use patterns of variation to make aspects discernible. She noticed that the student had taken four chestnuts from the original amount, resulting in two piles with the same amount of nuts, i.e. four chestnuts in each pile. As she wants the student to discern both eight and half of eight (four), she puts a question to make it possible for the student to discern the difference. Because the student had taken four items from the original pile of eight items when halving, the distinction between four and eight was not visible anymore; the child ended up with two piles with equal amounts instead. Based on such mistakes, to draw attention to the difference between the original amount and half this amount, the teacher repeatedly used simultaneous contrast in her instruction. There were, however, also examples during the lesson in which the teacher did not give a clear indication of whether a student's answer was correct.

During the third research lesson, the students were given some direct feedback during their performance of the tasks, but not after, and often this feedback neither confirmed nor rejected the students' answers. The lack of correspondence between the intended design of the third lesson and the teacher's actual performance of it could be one reason for this lack of

constructive feedback. The communication between teacher and students was mainly based on giving instruction at a procedural level:

> Excerpt 10, third lesson (T3 = Teacher 3):
>
> T3: You also think that 4+4 is 8. Have a look here … I have 8 pieces from the beginning. That is my quantity, my amount that I am not allowed to touch. If I am going to put down the double amount, I have to put down as many as I have here (points at the 8 pieces in the middle area) + the same amount. This means I have to put down 8+8, and this means?

The feedback in the third lesson thus ends with the teachers telling the students the correct way to solve the problem, a process, instead of taking into consideration how the students might have understood the problem and what aspects had been discernible.

6. Discussion and Conclusions

This study explored teachers' development of formative assessment strategies guided by variation theory during a classroom-based professional development project. The result of the analysis found three themes for how the teachers used theoretically informed strategies for assessment in the classroom in their planning, conducting, and evaluating of lessons. These three themes demonstrate how formative assessment contributed to the development of the teaching: theme A concerns the critical aspects for learning the specific content; theme B focuses on how teachers understand their students' knowledge about the content; and theme C brings A and B together by exploring the teachers' formative feedback in the classroom. The results indicate that throughout the study, the teachers became increasingly aware of how even small changes in teaching could affect the students' learning and how to use this knowledge in further instruction. For example, after finding out that the students tried to find a definite figure to use when doubling or halving, such as in all cases choosing to add one, two, or another number to the initial amount instead of seeing the relationship between the original amount and what to add, the teachers used different ways of explaining the concepts to the students. Two of the teachers realized the importance of varying the initial amount, based on the theoretical assumptions, and varying this aspect made it easier for the students to understand that the number added changes when the original amount changes. However, the third teacher was not aware of the importance of this because she had not accepted the common theoretical base and thus used this knowledge neither in her instructions nor in her feedback to the students, resulting in no significant increase in learning. The learning outcomes of the students' tests strengthened the other two teachers' analyses, and they became aware that even small changes in instruction might have an impact on student learning, such as what examples to use in the teaching (type and number of different objects), how these are to be used in the teaching and exercises (simultaneously contrasted or not), or what expressions to use when explaining the tasks ("same amount twice" or not). Eliciting and clarifying students' ideas and understandings at a fine-grained level based on theoretical assumptions thus

seems to have helped the teachers to understand the difficulties students may have in comprehending the object of learning, and to give relevant feedback to the students.

The teachers' formative assessment, which increased over the course of the learning study, was informed by the results of the students' tests and the video recordings of the lessons, both of which proved to be powerful tools for the teachers' assessment of the students' learning and understanding. Thus, there is evidence within the study that the learning study design encourages teachers to work formatively and that the formative assessment is strengthened by the theoretical framework used. This is in line with the argument put forward by Black and Wiliam (2009, 2012) and Wiliam (2009) that in order to be able to give relevant feedback to a student, the teacher needs a theory of how students learn and the ability to apply this theoretical understanding in a specific context. This is also in line with the work of James and Pedder (2006) and Pedder and James (2012), who suggest that the concept of 'research lesson' could be a useful strategy to promote assessment for learning in classrooms.

The evaluation of the last lesson showed that the simultaneity of double, half, and the original amount had not been presented in the third lesson as intended. During the planning meeting before this lesson, the teachers developed a new design, which was then not implemented during instruction. It was also clear through subsequent analysis that direct feedback to the students differed between the teachers and the groups; the teacher conducting the second lesson gave more frequent and more specific feedback than the other teachers by explicitly using the concept of simultaneous contrast. It is therefore possible that the improvement in scores after the second lesson was due to the combination of increased feedback with a lesson design that followed the assumptions of variation theory with regard to determining which aspects of the content should be in focus and which should be kept in the background. On the basis of this, we can conclude that the outcome of the learning study as a whole might have been even better had the joint evaluation of the lessons focused more strongly on the teachers' feedback (or lack thereof) to the students, both individually and to the class as a whole, along with discussing the effectiveness of the lesson design, as revealed by the students' test results and comments and behavior during the lessons.

7. Limitations of the Study

The small scale of the study—four meetings with four teachers over a limited period of time and with a limited number of students—means that it is not possible to make any generalizations from the results. However, these limitations made it possible to perform an in-depth study, giving some insights into how a theory of learning can guide teachers in their planning, implementation and evaluation of a content-specific topic. Since the study is class-room based we consider it however to have rather high ecological validity (Brewer, 2000), especially regarding the use of variation theory in designing, implementing and evaluating teaching, and the impact this is likely to have on

the opportunities for children in similar situations and settings to develop knowledge about doubling and halving.

8. Acknowledgments

We would like to thank the participating teachers, who generously shared their time and teaching, the Learning Design research team (LeaD) at Kristianstad University, for their encouragement and support; and the Department of Pedagogical, Curricular and Professional Studies, University of Gothenburg, Sweden, for partly financing this study. We are also grateful for valuable comments provided by the reviewers and for the thorough language review carried out by Dr. Catherine MacHale, which increased the clarity of the paper.

9. References

Adamson, B., & Walker, E. (2011). Messy collaboration. Learning from a learning study. *Teaching and Teacher Education, 27,* 29-36.

Bennet, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*: 1, 5 – 25.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education: Principles, Policy and Practice, 5*(1), 7–73.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31.

Black, P., & Wiliam, D. (2012). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 206–229). (2nd ed.) London: Sage Publications.

Boyatzis, R.E. (1998). *Transforming qualitative information: Thematic analysis and code development*. London: SAGE.

Brewer, M. (2000). Research Design and Issues of Validity. In H.T. Reis & C.M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 3-16). Cambridge, U.K.: University Press.

Elliott, J., & Yu, C. (2008) "Learning Studies as an Educational Change Strategy in Hong Kong: An independent evaluation of the 'Variation for the Improvement of Teaching and Learning'(VITAL) Project", funded by the Dept for Education and Manpower, Hong Kong. Published by the Centre for Learning Studies, HK Inst Educ. http://www.ied.edu.hk/cls/resources.htm

Elliott, J., & Yu, C . (2013) Learning Studies in Hong Kong Schools: A Summary Evaluation Report on 'The Variation for the Improvement of Teaching and Learning' (VITAL) Project, *Education and Didactique, 7*(2), 147-163.

Gustavsson, L. (2008). *Att bli bättre lärare. Hur undervisningsinnehållets behandling blir till samtalsämne lärare emellan.* [To become better teachers. How the handling of the lesson content becomes a topic of conversations between teachers.] Umeå: University of Umeå.

Hermerén, G. (2011). *Good research practice*. Stockholm: The Swedish Research Council.

Holmqvist, M. (2002). *Lärandets pedagogik*. Forskningsansökan till Vetenskapsrådet. [The Pedagogy of Learning. Research application to the Swedish Research Council.] Dnr 721-2002-3386

Holmqvist, M., Gustavsson, L., & A. (2007). Generative learning: learning beyond the learning situation. *Educational Action Research, 15*(2), 181-208.

Holmqvist, M., Gustavsson, L. & Wernberg, A. (2008) Variation Theory – An Organizing Principle to Guide Design Research in Education. In Kelly, A.E., Lesh, R., &. Baek J. (eds) *Handbook of design research methods in education, p 111-130*. New York: Routledge.

Holmqvist, M. (2011). Teachers' learning in a learning study. *Instructional Science, 39*(4), 497-511.

Holmqvist Olander, M., & Nyberg, E. (2014). Learning study guided by variation theory: Exemplified by children learning to halve and double whole numbers. *Journal of Research in Childhood Education* . 28(2), 238-260.

James, M., & Pedder, D. (2006). Professional learning as a condition for assessment for learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 27–43). London: Sage Publications.

Kullberg, A. (2010). *What is taught and what is learned: Professional insights gained and shared by teachers of mathematics*. Diss. Göteborg : Göteborgs universitet, 2010. Göteborg. http://gupea.ub.gu.se/hdl.handle/2077/22180

Lo, Mun Ling (2012). *Variation theory and the improvement of teaching and learning* Göteborg: Acta universitatis Gothoburgensis. http://gupea.ub.gu.se/handle/2077/29645¨

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54-61.

Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researche*r, 31(3), 3-14.

Lo, M.L., Pong, W.Y., & Chik, P.M. (Eds.) (2005). *For each and everyone: Catering for individual differences through learning study.* Hong Kong: Hong Kong University Press.

Lo, M. L., & Marton, F. (2012). Towards a science of the art of teaching: Using variation theory as a guiding principle of pedagogical design. *International Journal for Lesson and Learning Studies*, 1, 7–22. Doi: 10.1108/20468251211179678

Marton, F. (1981). Phenomenography - describing conceptions of the world around us. *Instructional Science*, 10, 177-200.

Marton, F. (2003). Learning Study - pedagogisk utveckling direkt i klassrummet. [Learning Study—pedagogical development directly in the classroom]. In Carlgren, I. (Ed.). *Forskning av denna världen: praxisnära forskning inom utbildningsvetenskap* [Research of this world-practice-oriented research in educational sciences] (pp. 41–46). Stockholm: Vetenskapsrådet [The Swedish Research Council].

Marton, F. (2014). Necessary conditions of learning. New York, N Y: Routledge

Marton, F., & Tsui, A. (2004). *Classroom discourse and the space of learning*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Marton, F., & Booth, S. (1997). *Learning and awareness*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Olteanu, C., & Lennerstad, H. (2011). *Lesson study och Learning study samt IKT i matematikundervisningen: en utvärdering av matematiksatsningen* [Lesson Study and Learning Study and ICT in mathematics education: An evaluation of the mathematics initiative]. Stockholm: Fritzes.

Pang, M. F., & Marton, F. (2013). Interaction between the learners' initial grasp of the object of learning and the learning resource afforded. *Instructional Science, 41*(6), pp 1065-1082.

Pedder, D., & James, M. (2012) Professional learning as a condition for assessment for learning'. In J. Gardner (Ed), *Assessment and Learning* (pp. 33-48), (2nd ed.). London: Sage Publications.

Phelan, J., Choi, K., Vendlinski, T., Baker, E., & Herman, J. (2011). Differential Improvement in Student Understanding of Mathematical Principles Following Formative Assessment Intervention. *The Journal of Educational Research,* Vol 104, pp. 330-339.

Stigler, J.W., & Hiebert, J. (1999). *The teaching gap: best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.

Stigler, J. W., & Hiebert, J. (2009). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.

Swedish National Agency for Education (2011). *Curriculum for the compulsory school system, the pre-school class and the leisure-time centre 2011*. Stockholm: Swedish National Agency for Education [Skolverket].

Wernberg, A. (2009). *Lärandets objekt: Vad elever förväntas lära sig, vad görs möjligt för dem att lära och vad de faktiskt lär sig under lektionerna* [The Object of Learning: What students are expected to learn, what is possible for them to learn and what they actually learn in class]. Umeå: University of Umeå.

Wiliam, D. (2006). Formative Assessment: Getting the Focus Right. *Educational Assessment, 11*:3-4, 283-289.

Wiliam, D. (2009). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. Andrade & G. J. Cizek (Eds.), *Handbook of Formative Assessment* (p.18-40). New York: Routledge.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation 37*, 3–14.