Designing Formative Assessment in Mathematics

Malcolm Swan

Centre for Research in Mathematics Education, School of Education, University of Nottingham

Formative assessment is the process by which students and teachers gather evidence of learning and then use it to adapt the way they learn and teach in the classroom. In this paper I describe a design research project in which we are attempting to develop and integrate "formative assessment lessons" into classrooms across the US. In this paper, I focus on some of the issues that arose as we attempted to design lessons that would develop students' capacity to tackle non-routine problems. Particular formative aspects of lesson design are highlighted; the important roles of pre-assessment, formative feedback questions and sample work for students to critique are described.

INTRODUCTION

The potential power of formative assessment for enhancing learning in mathematics classrooms was brought to widespread attention by the research review of Paul Black and Dylan Wiliam (Black, et al. 2003; 1998; Black, et al. 1999). They launched programs of work that aimed to turn these insights into impact on practice, but found that regular meetings over a period of years were needed to enable a substantial proportion of teachers to acquire and deploy the "adaptive expertise" (Hatano & Inagaki 1986; Swan 2006a) needed for self-directed formative assessment. This is clearly an approach that is difficult to implement on a large scale. Since their research was published, the term "formative assessment" has entered common parlance where it has often been mutated to mean more frequent testing, scoring and record keeping. This, however, corrupts Black and Wiliam's original use of the term where it is taken to include:

> "… all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged. Such assessment becomes 'formative assessment' when the evidence is actually used to adapt the teaching work to meet the needs." (Black & Wiliam, 1998, p91)

Here lies the challenge: for assessment to be truly formative the teacher must develop expertise in becoming aware of and adapting to the specific learning needs of students, both in planning lessons and moment-by-moment in the classroom.

In 2009, the Bill & Melinda Gates Foundation approached us to develop a suite of formative assessment lessons to form a key element in the Foundation's program for "College and Career Ready Mathematics" based on the Common Core State Standards for Mathematics (NGA & CCSSO 2010). In response, the Mathematics Assessment Project (MAP) was designed to explore how far well-designed teaching materials can enable teachers to make high-quality formative assessment an integral part of the implemented curriculum in their classrooms, even where linked professional development support is limited or non-existent. The research-based design of these lessons, now called *Classroom Challenges*, forms the core of this paper. To date, we have designed and developed over one hundred formative assessment lessons to support US Middle and High Schools in implementing the new Common Core State Standards for Mathematics. Each lesson consists of student resources and an extensive teacher guide. About one-third of these lessons involve the tackling of non-routine, problem-solving tasks. They are available on the website: http://map.mathshell.org.uk/materials/index.php.

METHODOLOGY

Our methodology is based on design research principles, involving theory-driven iterative cycles of design, enactment, analysis and redesign (Barab & Squire 2004; Bereiter 2002; Cobb, et al. 2003; DBRC 2003, p. 5; Kelly 2003; van den Akker, et al. 2006). Each lesson was developed, through three iterative design cycles, with each lesson being trialled in three or four US classrooms between each revision. Revisions were based on structured, detailed feedback from experienced observers of the materials in use in classrooms. We thus have over 700 observer reports of lessons using these materials.

The objective of these trials was to give the design team a detailed picture of what happened in the use of the materials by teachers. The aim is to learn more on questions including:

- Do the teacher and students understand the materials?
- How closely does the teacher follow the lesson plan?
- Are any of the variations damaging to the purpose of the lesson?
- What features of the lesson proved awkward for the teacher or the students?
- What unanticipated opportunities arose that might be included on revision?

This process enabled us to obtain rich, detailed feedback, while also allowing us to distinguish general implementation issues from idiosyncratic variations by individual teachers.

THEORETICAL BACKGROUND

The theories that have underpinned our designs go back to our “Diagnostic Teaching” program of design research in the 1980s. This was an example of formative assessment of the kind identified as effective by Black and Wiliam (See e.g. Bell 1993; Swan 2006a). This approach to teaching mathematical concepts was more effective, over the longer term, than either expository or guided discovery approaches. This result was replicated over many different topics: decimal place value, rates, geometric reflections, functions and graphs, and fractions (Bassford 1988; Birks 1987; Brekke 1987; Onslow 1986; Swan 1983). From these studies it was deduced that the value of diagnostic teaching appeared to lie in the extent to which it valued the intuitive methods and ideas that students brought to each lesson, offered experiences that created inter- and intra-personal ‘conflicts’ of ideas, and created opportunities for students to reflect on and examine inconsistencies in their interpretations. A phase of ‘preparing the ground’ was found necessary, where pre-existing conceptual structures were identified and examined by students for viability. The ‘resolution’ phase, involved students in intensive, reflective discussions. Indications were that the greater the intensity of the discussion, the greater was the impact on learning.

More recently, these results have been replicated on a wider scale. UK government funded the development of a multimedia professional development resource to support diagnostic teaching of algebra (Swan & Green 2002). This was distributed to all FE colleges, leading to research on the effects of implementing collaborative approaches to learning in 40 GCSE retake classes. This again showed the greater effectiveness of approaches that elicit and address conceptual difficulties through student-student and whole class discussion (Swan 2006a, 2006b; Swan 2006c). The government, recognizing the potential of such resources, commissioned the design of a more substantial multimedia PD resource, 'Improving Learning in Mathematics' (DfES 2005). This material was trialled in 90 colleges, before being distributed to all English FE colleges and secondary schools.

In our design of lessons for problem solving we have also drawn inspiration from the Lesson Study research in Japan and the US (Fernandez & Yoshida 2004; Shimizu 1999). In Japanese classrooms, lessons are often structured with four key components: *hatsumon* (the teacher gives the class a problem to initiate discussion); *kikan-shido* (the students tackle the problem in groups or individually); *neriage* (a whole class discussion in which alternative strategies are compared and contrasted and in which consensus is sought) and finally the *matome*, or summary. Among these, the *neriage* stage is considered to be the most crucial. This term, in Japanese refers to kneading or polishing in pottery, where different colours of clay are blended together. This serves as a metaphor for the considering and blending of students' own approaches to solving a mathematics problem. It involves great skill on the part of the teacher, as she must select student work carefully during the *kikan-shido* phase and sequence the work in a way that will elicit the most profitable discussions. In the *matome* stage of the lesson, the Japanese teachers will tend to make a careful final comment on the mathematical sophistication of the approaches used. The process is described by Shimizu:

> "Based on the teacher's observations during Kikan-shido, he or she carefully calls on students to present their solution methods on the chalkboard, selecting the students in a particular order. The order is quite important both for encouraging those students who found naive methods and for showing students' ideas in relation to the mathematical connections among them. In some cases, even an incorrect method or error may be presented if the teacher thinks this would be beneficial to the class. Once students' ideas are presented on the chalkboard, they are compared and contrasted orally. The teacher's role is not to point out the best solution but to guide the discussion toward an integrated idea." (Shimizu 1999, p110)

In part, perhaps, influenced by the Japanese approaches, other researchers have also adopted similar models for structuring classroom activity. They too emphasise the importance of: anticipating student responses to cognitively demanding tasks; careful monitoring of student work; discerning the mathematical value of alternative approaches in order to scaffold learning; purposefully selecting solution-methods for whole class discussion; orchestrating this discussion to build on the collective sense-making of students by intentionally ordering the work to be shared; helping students make connections between and among different approaches and looking for generalizations; and recognizing and valuing and students' constructed solutions by comparing this with existing valued knowledge, so that they may be transformed into reusable knowledge (Brousseau 1997; Chazan & Ball 1999; Lampert 2001; Stein, et al. 2008).

Each of these aspects presents a substantial challenge for teachers in a problem-solving context. Normally in the course of teaching mathematical skills, student reasoning is predictable and short. When problem solving, students construct chains of reasoning that may not be well-expressed nor easily predicted. In the busy classroom, teachers have little time to spend listening over the shoulders of students as they discuss alternative problem solving strategies. Often students' sharing of their methods in whole class discussions are reduced to mere 'show and tell' occasions and do not reveal the thinking behind the approaches in any depth. Frequently, students' presentations are poorly expressed and remain incomprehensible to their peers and teachers appear more concerned with giving everyone a chance to share than in analysing the quality of the reasoning. Merely accepting answers, without attempting to critique and synthesise individual contributions can constrain the development of mathematical thinking (Mercer 1995).

THE DESIGN OF THE CLASSROOM CHALLENGES

We now illustrate how this research has informed the products of our design research using one of the Classroom Challenges, focused on problem-solving: "Counting Trees" (Figure 6). Further lessons may be downloaded from http://map.mathshell.org.

Counting Trees

The diagram shows some trees in a tree farm.

The circles ● show old trees and the triangles ♦ show young trees.

Tom wants to know how many trees there are of each type, but says it would take too long counting them all, one by one.

1. What method could Tom use to estimate the number of trees of each type?
2. Use your method to estimate the number of:
   (a) Old trees (b) New trees.

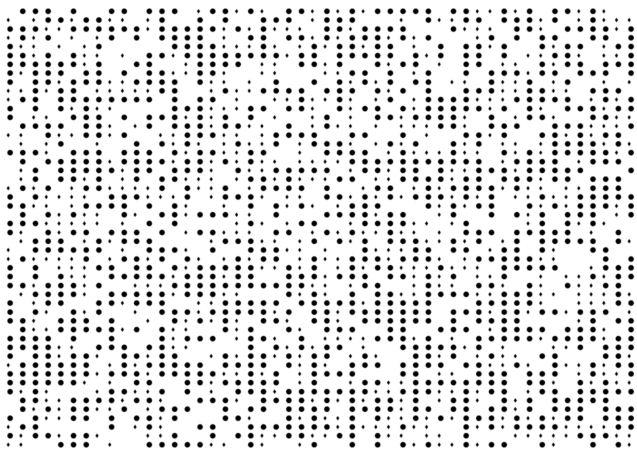


Figure 1: The "Counting trees" task

As a preliminary assessment, students are invited to tackle a problem individually. This exposes students' different approaches. Through trialling, we have developed a "common issues table" that lists for the teacher the most common difficulties that students have together with suggestions for questions that the teacher might pose to move thinking forward (Table 1). The teacher guide suggests that students' responses are collected in by the teacher and analysed, with the help of this table. The teacher may, if time permits, write some of these questions on each student's work, or alternatively prepare a few questions for the whole class to consider. This process has enabled teachers to anticipate student reasoning in the main lesson.

Table 1: A few of the common issues and suggested questions for "Counting Trees"

Common issues	Suggested questions and prompts
Student chooses a method which does not involve any sampling: E.g. student counts the trees.	• Have you done what was asked? • What assumptions have you made? Are your assumptions reasonable?
Student chooses a sampling method that is unrepresentative. E.g.: student counts trees in the first row and multiples by the number of rows.	• How could you improve/check your estimate? • Is your sample typical of the whole tree farm? How do you know?
Student makes incorrect assumptions. E.g.: student does not account for gaps.	• Is there a pattern to how the trees are distributed in the tree farm? Does your work assume there is a pattern? • What does your method assume? Is this a reasonable assumption?
Student chooses appropriate sampling method	• Can you suggest a second, different sampling method? • If you miscount your sample by 1, how does that affect your overall estimate?

The lesson itself begins with the teacher returning students' initial individual attempts along with the prepared questions. Working individually, students review their initial attempts and try to respond to the teacher's questions.

The students are now asked to work in small groups to discuss the work of each individual, then to produce a poster showing a joint solution that is better than the initial attempts. Groups are often organized so that students with contrasting ideas are paired. This activity promotes peer assessment and refinement of ideas. The teacher's role is to observe the groups and challenge students to justify their decisions as they progress and thus refine and improve their strategies.

The teacher now introduces up to four pieces of "sample student work", provided in the materials (Figure 2). This pre-prepared work has been carefully chosen to highlight alternative approaches and common mistakes. Each piece of work is annotated with questions that focus students' attention. So, for example: *Does Laura's approach make mathematical sense? Why does she halve her answer? What assumptions has Laura made? How can Laura improve her work? To help you understand Laura's work, what question(s) would you ask her?* Introducing work from outside the classroom is helpful in that (i) students are able to critique it freely without fear of other students being hurt by criticism; (ii) handwritten 'student' work carries less status than printed or teacher-produced work and it is thus easier for students to challenge, extend and adapt. A further benefit is that this work enables teachers to prepare the discussion before the lesson, avoiding the difficulty of having to select work from the class during the lesson itself.

We have found that teachers like to be flexible in the way they distribute sample student work, in response to the particular needs of their own students. For example if students have struggled with a particular strategy, the teacher may want them to analyse a similar sample student work. Conversely if students successfully solved the problem using a particular strategy, then the teacher may want to them to analyse sample student work that uses a different strategy. The teacher can thus decide if their students would benefit from working with all the sample student work or just one or two pieces.

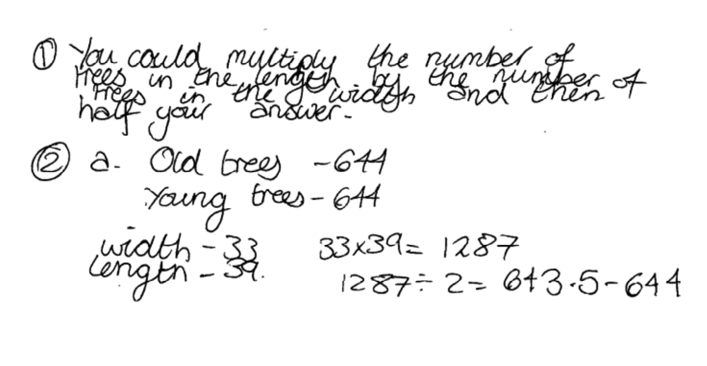
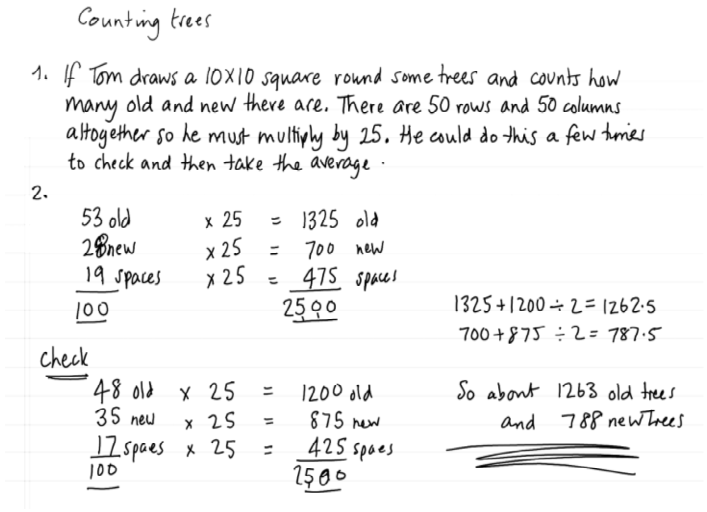
Laura attempts to estimate the number of old and new trees by multiplying the number along each side of the whole diagram and then halving. She does not account for gaps nor does she realize that there are an unequal number of trees of each kind. *Can you explain why Laura halves her answer? What assumption is she making?*	 ① You could multiply the number of trees in the length by the number of trees in the width and then half your answer. ② a. Old trees - 644 Young trees - 644 width - 33 33x39= 1287 length - 39. 1287÷2= 643.5-644
Amber chooses a representative sample and carries through her work to get a reasonable answer. She correctly uses proportional reasoning. She checks her work as she goes along by counting the gaps in the trees. Her work is clear and easy to follow, although a bit inefficient. *Can you explain why Amber multiplies by 25 in her method?*	 Counting trees 1. If Tom draws a 10X10 square round some trees and counts how many old and new there are. There are 50 rows and 50 columns altogether so he must multiply by 25. He could do this a few times to check and then take the average. 2. 53 old x 25 = 1325 old 28 new x 25 = 700 new 19 spaces x 25 = 475 spaces 100 2500 1325+1200÷2=1262.5 700+875÷2=787.5 Check 48 old x 25 = 1200 old 35 new x 25 = 875 new 17 spaces x 25 = 425 spaces 100 2500 So about 1263 old trees and 788 new Trees

Figure 2: Sample student work for discussion, with commentary from the teacher guide.

After critiquing the sample work, students are encouraged to revise their own group solutions. This process of successive refinement in which methods are tried, critiqued and adapted has been found to be extremely profitable for developing problem solving strategies.

The lesson concludes with a whole class discussion that is intended to draw out some comparisons of the approaches used; in this case the power of sampling. Students are invited to respond individually to such questions as:

- How was your group's solution better than your individual solution?
- How did you check your method?
- How was your response similar to or different from the sample student responses?
- What assumptions did you make?

CONCLUDING REMARKS

In this brief paper, I have attempted to describe how systematic design research has enabled us to tackle a significant pedagogical problem: how might we enable students to develop the skills necessary for the effective tackling of non-routine problems. This involves the development of planning, monitoring and critiquing behaviours on the part of students; aspects that are not developed in mathematics lessons that focus on routine skills. Particular features that we have found of importance are:

- **Pre-assessment**; giving students opportunity to engage with the problem individually, before group discussion takes place and giving the teacher opportunity to anticipate student reasoning in advance of the lesson;
- **Common issues tables**; that use empirical research results to inform teachers of the likely issues that students will face in the lesson and offer teachers suggested formative questions that they may ask students during the lesson;
- **Sample student work** that focuses student attention on the comparison of alternative approaches, assumptions made, representations used and offers them opportunity to develop criticality. In addition this allows the teacher to plan discussions of such strategies before the lesson.

We have found that, as might be expected, the *neriage* and *matome* stages of the lesson in which teachers select, synthesise and generalise what has been achieved in the lesson are still the most challenging and these aspects are currently being researched in a new Lesson Study Project on problem solving funded by the Nuffield Foundation.

The resulting lesson plans we have developed are extensive (for counting trees it covers seven pages), reflecting the new territory that many teachers find themselves. This has been in response to teacher requests for advice and guidance. The result has have proved very popular with teachers (to date, over two million of the lesson plans have been downloaded). To quote one of the trial teachers:

> "At my school kids have generally not been interested in mathematics. They haven't seen it as exciting, as a chance to think critically, and as a fun challenge. But I think Classroom Challenges change that. The CCs offer the right portrayal of what mathematics is about. When kids begin to experience that they see how rich and how exciting the subject really is."

References

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1-14.

Bassford, D. (1988). *Fractions: A comparison of Teaching Methods.* Unpublished M.Phil, University of Nottingham.

Bell, A. (1993). Some experiments in diagnostic teaching. *Educational Studies in Mathematics, 24*(1).

Bereiter, C. (2002). Design research for sustained innovation. *Cognitive studies, Bulletin of the Japanese Cognitive Science Society, 9*(3), 321-327.

Birks, D. (1987). *Reflections: a Diagnostic Teaching Experiment.* University of Nottingham.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.

Black, P., & Wiliam, D. (1998). *Inside the black box : raising standards through classroom assessment*. London: King's College London School of Education 1998.

Black, P., Wiliam, D., & Group, A. R. (1999). *Assessment for learning : beyond the black box*. Cambridge: University of Cambridge Institute of Education.

Brekke, G. (1987). *Graphical Interpretation: a study of pupils' understanding and some teaching comparisons.* University of Nottingham.

Brousseau, G. (1997). *Theory of Didactical Situations in Mathematics* (N. Balacheff, M. Cooper, R. Sutherland & V. Warfield, Trans. Vol. 19). Dordrecht: Kluwer.

Chazan, D., & Ball, D. L. (1999). Beyond being told not to tell. *For the Learning of Mathematics, 19*(2), 2-10.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher, 32*(1).

DBRC (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational researcher, 32*(1), 5-8.

DfES (2005). *Improving Learning in Mathematics*. London: Standards Unit, Teaching and Learning Division.

Fernandez, C., & Yoshida, M. (2004). *Lesson Study: A Japanese Approach to Improving Mathematics Teaching and Learning*. Mahwah, New Jersey: Laurence Erlbaum Associates.

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. W. Stevenson, H. Azuma & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York: W H Freeman/Times Books/ Henry Holt & Co.

Kelly, A. (2003). Theme issue: The role of design in educational research. *Educational Researcher, 32*(1), 3-4.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Mercer, N. (1995). *The guided construction of knowledge*. Clevedon, Philadelphia, Adelaide.

NGA, & CCSSO (2010). Common Core State Standards for Mathematics. Retrieved from http://www.corestandards.org/Math

Onslow, B. (1986). *Overcoming conceptual obstacles concerning rates: Design and Implementation of a diagnostic Teaching Unit.* Unpublished PhD, University of Nottingham.

Shimizu, Y. (1999). Aspects of Mathematics Teacher Education in Japan: Focusing on Teachers' Roles. *Journal of Mathematics Teacher Education 2*, 107-116.

Stein, M. K., Eagle, R. A., Smith, M. A., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*, 313-340.

Swan, M. (1983). *Teaching Decimal Place Value - a comparative study of 'conflict' and 'positive only' approaches.* Paper presented at the 7th Conference of International Group for the Psychology of Mathematics Education, Jerusalem, Israel.

Swan, M. (2006a). *Collaborative Learning in Mathematics: A Challenge to our Beliefs and Practices*. London: National Institute for Advanced and Continuing Education (NIACE) for the National Research and Development Centre for Adult Literacy and Numeracy (NRDC).

Swan, M. (2006b). Designing and using research instruments to describe the beliefs and practices of mathematics teachers. *Research in Education, 75*, 58-70.

Swan, M., & Green, M. (2002). Learning Mathematics through Discussion and Reflection. London: Learning and Skills Development Agency.

Swan, M. B. (2006c). Learning GCSE mathematics through discussion: what are the effects on students? *Journal of Further and Higher Education, 30*(3), 229-241.

van den Akker, J., Graveemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational Design Research*. London and New York: Routledge.