12th International Congress on Mathematical Education

Topic Study Group 33

8 July – 15 July, 2012, COEX, Seoul, Korea

TEACHERS' VIEWS OF USING AN ON-LINE, FORMATIVE ASSESSMENT SYSTEM FOR MATHEMATICS

Vicki Steinle and Kaye Stacey

University of Melbourne, Australia

In this paper we discuss some issues arising from teachers' use of an on-line system for formative assessment of their students' understanding of mathematics. The 'smart tests' cover many topics in secondary school mathematics, and are programmed to provide an automated diagnosis of students' stages of learning in specific topics, to inform teaching. In particular, we highlight teachers' views about: the desirability of formative assessment; the use that they make of this information; the provision of feedback; and the effect of using the system on their pedagogical content knowledge and subsequent changes to their practice .Overall, teachers report that use of the tests improves their knowledge of the achievement of individual students and students in general, and that they can use this information in several ways to improve their teaching.

Keywords: computer-based assessment; formative assessment; diagnosis; pedagogical content knowledge

THE SMART TEST SYSTEM

A 'smart test' is a '**s**pecific **m**athematics **a**ssessment that **r**eveals **t**hinking'. The tests are embedded in an on-line system, all of which have been created by the authors and colleagues (Stacey, Price, Steinle, Chick, Gvozdenko, 2009) and are accessed through an intelligent on-line environment (www.smartvic.com/smart/index.htm). These innovative tests provide teachers with an informative diagnosis of their student's conceptual understanding of most of the topics in junior secondary school mathematics. The research-backed diagnosis, described in terms of stages of understanding, identified misconceptions and gaps in knowledge, is provided back to teachers immediately. As well as this diagnosis, the system provides teachers with explanations of the diagnoses in the particular mathematics topic, and teaching suggestions for dealing with misconceptions and for taking students to the next level of understanding. We intend that this information will be concise enough to be readily useable by teachers, deep enough to make a real difference to lesson content, and linked to targeted teaching resources. Figure 1 provides details of how teachers interact with the smart test system and the two expected outcomes; higher achievement for students and improved mathematics pedagogical content knowledge for teachers. We expect that pedagogical content knowledge will improve as teachers come to work with the stages of learning, misconceptions and gaps, as they apply to their own students. Smart tests are currently being used by about 150 teachers, and we are processing approximately 1700 student tests per month. They can be used by teachers and students, or researchers anywhere.

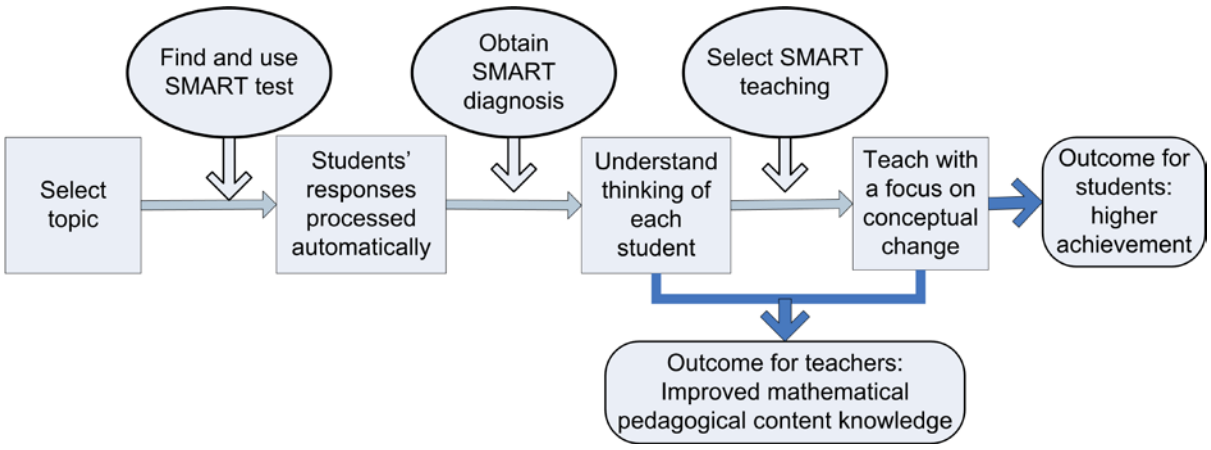

Figure 1. How teachers use the smart test system.

According to Wiliam (2007), "For teacher questioning to be effective, teachers need to know what kinds of conceptualizations students are likely to have, and need tools to identify them …Items that reveal unintended conceptions – in other words that provide a "window into thinking" – are not easy to generate, but they are crucially important to improve the quality of students' mathematical learning", (p. 1069). Smart tests are built on exactly the sorts of questions that Wiliam refers to; sets of questions that together provide a window into student thinking, wherever possible derived from research literature.

A question that provides a window into student thinking can often be very valuable to stimulate mathematical investigations or classroom discussion where students and the teacher can explain and elaborate answers (Wiliam, 2007). However, if they are intended to be used as part of an automated diagnostic assessment, modifications often need to be made because a computer is still very limited in its processing of free response items. Hence smart tests often include multiple choice questions with alternatives based on research evidence, rather than free response. Importantly the student diagnosis is made on responses to a series of questions rather than a single question. It is time-consuming for a person to look for predicted patterns of responses to a series of questions, but computers can be programmed to look for many of these predicted patterns and do this very quickly. At this point it is important to emphasise that the power of this system is the potential to diagnose student thinking from *response patterns* not just *accuracy patterns*. In other words, *wrong in which way* is more revealing than just *wrong*. For example, Steinle, Gvozdenko, Price, Stacey & Pierce (2009) indicates how we have used response patterns to diagnose student misconceptions in algebra. This test uses two types of items. One begins with the scenario that some students had to find the values of x that make true the equation $x + x + x = 12$. Test takers then have to grade as correct or incorrect supposed answers from imaginary students such as '$x = 2$ and $x = 5$ and $x = 5$'. The other type of item is concerned with the equation $x + y = 16$ and possible values of x and y. By looking for patterns in the responses, test takers are allocated to one of four stages for interpreting

the letters in equations. For example, at the second stage, students understand letters stand for numbers but overgeneralise to conclude that different letters must stand for different numbers. Because the aim is to diagnose student thinking in a particular topic to assist the teacher plan more effective teaching in that topic, smart tests are focused on that one topic and typically take students less than 10 minutes to complete.

The smart test system embeds research in mathematics education into artefacts that are easy for practitioners to use, creating what Pea (1993) calls ‘distributed intelligence’ in tools for teaching. When planning the teaching of a new topic, the diagnostics from the smart test provides teachers with better knowledge of the mathematical thinking of their current students. Incidentally, yet importantly, teachers will also learn about the mathematical knowledge of students more generally.

Smart tests are not limited to the multiple choice format. Drag-drop items allow students to participate in an *activity* in much the same as a sorting card task in an interview. There are considerable benefits if a teacher is able to conduct interviews with students on their mathematical understanding, but clearly this is a resource-intensive option; smart tests can provide teachers with a comparable amount of information in a more efficient manner. Sliders provide another very flexible interactive format for smart test items.

For the above reasons, smart tests are different to the tests that teachers normally set for summative assessment. It is because of this difference that teachers who use the smart test system are likely to initially experience some disequilibrium. In this paper, we explore two of the issues that teachers have raised during the development of the smart test system: teachers’ assumptions about the purpose of the assessment (summative or formative); and the requirement for the system to provide feedback directly to students. The additional two issues to be reported are teachers’ views about the effect of using the smart test system on their knowledge for teaching and how it has changed their practice. It is achieving the latter two goals that will really make the system worthwhile.

DATA

The data reported in this paper comes from three sources: notes from 10 focus groups held with teachers at schools involved in the development of the smart test system in its first two years; on-line surveys completed by volunteer teachers after they have used a smart test; and spontaneous emails that teachers have sent to us on an ad-hoc basis. The two questions in the on-line surveys that are reported are listed below and respondents were invited to elaborate on their answers.

As a result of using this quiz have you learned something useful for you as a teacher?

Did you adjust your teaching plan as a result of the diagnostic information?

ISSUE 1: THE VARIOUS PURPOSES OF ASSESSMENT

We found that the concept of formative assessment – assessment that directly feeds into lesson planning – was not understood by all teachers, even in the volunteer trial schools. During the development of the smart test system, we visited the project schools several times each year to get feedback from teachers and explain how we were improving the system based on their feedback. One of the questions regularly raised by the teachers was "How do we use this information in our reports?" Implicit in this is that teachers first wanted summative rather than formative assessment. Teachers have many demands made upon their time and some of the teachers in the project were hoping to use the smart tests data directly in the time-consuming task of writing reports. Some teachers also expected a measurement of the overall level of the state curriculum framework that students had reached in mathematics – this time both for reporting to parents and for accountability to the department of education. Some teachers took quite some time to understand the notion of this type of formative assessment and thus to broaden their understanding of useful assessment to include both summative to formative. At the other end of spectrum, there were some schools ready to embed pre-testing with smart tests into their curriculum in an attempt to personalise learning for students. Evidence that teachers are using the smart test system to inform their teaching and some examples of how this is done are provided by the teacher comments in Figure 2. Here and elsewhere, we have tried to select a range of comments which show the major themes of the comments across teachers.

I use the smart tests as a part of my diagnostic 'toolbox'. They are clear, easy to access and give a quick snapshot of where my students' prior knowledge is developed or underdeveloped. This information influences the activities I implement in class, ensuring that the students are being challenged in Mathematics.

I used the smart test "Understanding angle" with my year 7class. In my teaching I adopted an approach that best addressed the needs of the students based upon the diagnostic test. I was able to avoid certain areas that were well understood and concentrate on areas that were not.

The other end of the spectrum is that I've been more confident in moving kids, not making them go over things. I can see "alright, this child has a really good understanding of fractions". I'm not going to sit and make him (or her) repeat all of those skills so I feel more confident in moving them to something else.

When our Year 7 students did the fractions smart tests, we were surprised to find many students were at Stage 0. All these years we've always presumed that they were at a particular level but obviously that's not happening, and so that's changed our curriculum, the way we think about teaching fractions.

This quiz is a genuinely useful tool to assist in the differentiation of the curriculum. It is efficient and informative.

A very valuable tool to assess a class before teaching the topic.

Excellent formative assessment tool which allowed me as a coach to discuss the various misconceptions and student thinking within a year 8 class. It provided teachers with real data that allowed them to address the misconceptions through their teaching.

I often overlook and brush off students' misconceptions without considering the difficulty that students faced. With this assessment tool, I am able to analyse my students better, individually and correct their misconceptions on a particular topic.

Very useful as a pre-test on reading scales. I found out exactly where each student was at and that enabled me to target my teaching into the areas where it was most needed, while giving extension work to the students who had already gained a good understanding of the topic. Now I am going to retest them using another form of the test to see how effective my teaching has been.

Figure 2. Teacher comments on using smart test system as formative assessment

ISSUE 2: THE PROVISION OF FEEDBACK DIRECTLY TO STUDENTS

Another issue that arose from the trial schools was whether the computer diagnosis (i.e. feedback of the test) should be delivered to teachers only, to students or to both. We believe that there were two drivers of this request. Firstly, most of the computer games or quizzes that students use provide immediate feedback (usually right/wrong), so students expected this from smart tests. Secondly, many teachers are aware that good quality feedback, presented to students soon after the completion of a task, can lead to increases in learning. In fact, involving students in the results of assessment is often cited as a hallmark of good formative assessment.

After further consideration, we have maintained our position to provide information only to teachers who would then take action to improve student learning, so the feedback about students' mathematical thinking directly prompts teacher actions (rather than student actions) to improve student learning. The aim of the smart tests is not to provide students with tasks and then automated feedback so that learning takes place independently of the teacher. There are other challenges which lead to this

decision. The detailed, topic specific diagnoses are written for adults, and some effort, background and technical language is required to understand them. Students are not likely to be able to understand the diagnosis that is being supplied to teachers, so separate student feedback would need to be written. In addition, teachers are reluctant to give negative feedback to students, and prefer to control this personally. A vocal group of teachers warned about the consequences of students' receiving feedback indicating low performance. Evidence that some teachers are now comfortable using the smart test system without direct feedback to students is provided by the teacher comment in Figure 3. We have achieved this comfort by advising teachers that they should explain the purpose of formative assessment to students, so students do not demand it.

> *I just explain to the kids what it shows, and that it's showing me how to teach you guys better. "It's not about something that you're going to get tests back. It's just a tool that I'm using to see what you guys know so I can teach you better." They have had no issues with that. And the parents that I've spoken to on parent teacher night a couple of times think it is fantastic.*

Figure 3. Teacher comment related to managing demand for student feedback

ISSUE 3: EFFECT ON KNOWLEDGE FOR TEACHING

A major aim of the smart test system is to increase teachers' mathematical pedagogical content knowledge. We hypothesised that putting data on their own students' thinking into teachers' hands would make research results come alive for teachers, and hence build their capacity. Table 1 provides the frequency of responses to the multiple choice question: *As a result of using this quiz have you learned something useful for you as a teacher*? Of the 127 responses to this question, 117 (92%) indicated that "yes" the teacher did learn something useful, and 58 (46%) indicated that this was "very valuable learning".

Table 1: Responses to *As a result of using this quiz have you learned something useful for you as a teacher?*

Options provided	Frequency
YES, very valuable learning	58
YES, useful learning	59
NO	10
blank	16
Total	143

The high proportion of *yes* responses is likely to be due to the fact that many of the teachers are new to the smart test system. We would expect that, on subsequent use of the same smart test, teachers will be familiar with the stages of development that are reported, and after being initially alerted by the smart tests, they may come to see them in their normal interactions with students. In this way, some of the smart tests may become redundant, as teachers modify their teaching to reduce the likelihood of misconceptions, and take care to develop strong concepts and have items which reveal understanding in the specific topic at their fingertips. For example, as noted above, Steinle et al. (2009) report results of a smart test which investigates whether students think that different letters in algebra can stand for the same number, and whether one letter can stand for different numbers in the same expression. Once teachers are aware that students often make false assumptions about this, they can quite easily address it in their teaching and look for it in students' work. If the test is no longer required because of increased teacher understanding of students' thinking, then that is itself a success. Evidence that teachers perceive that using smart tests has led to improvement in their knowledge for teaching is provided by the comments in Figure 4.

It certainly has encouraged a dialogue between the student and the teacher, and looking at specific things because you as a teacher feel more confident about what you're talking about, because you've got all that information there. The smart test directs you about where to go. And also you can speak to that student about that particular misconception. It works quite well.

Well worth doing. Made me feel like an 'expert' teacher instead of just an experienced teacher.

[I] used a table structure similar to dual number line to help students with showing/organising information contained in problems and to find what 1 part represents[and also] to emphasise the use of multiplication/division.

I read the referenced research paper, which was informative and useful. The teaching suggestions were really practical, and were suitable to have a go at straight away. I used paper strips and pieces to fold and colour to estimate

percentages.

Figure 4. Teacher comments related to improvement in knowledge for teaching

ISSUE 4: EFFECT ON TEACHING PRACTICE

Formative assessment is only *formative* if it results in a change in the opportunity for a student to learn. Table 2 provides the frequency of responses to the multiple choice question: *Did you adjust your teaching plan as a result of the diagnostic information?* Of the 124 responses to this question, 87 (70%) indicated that they did adjust their teaching. Of course, adjusting is not always required. One of the teachers who did not adjust their teaching made the following comment: *I didn't adjust my teaching plan as such, because the results supported what I expected, but confirmation was valuable.*

Table 2: Responses to *Did you adjust your teaching plan as a result of the diagnostic information?*

Options provided	Frequency
YES	87
NO	37
blank	19
Total	143

The follow-up question to teachers was: *If YES: In what way did you change your teaching plan?* There were many different types of responses here. Two very frequent themes are illustrated in Figures 5a and 5b. Figure 5a contains sample comments which indicate how they changed their teaching practice by grouping students and Figure 5b provides comments indicating that teachers started their teaching at a higher level than they were intending to.

I have put the students into groups and will give them activities to focus on and correct their misconceptions. I will be looking carefully at the [suggested resources] for ideas.

Focussed teaching groups using the Stages as a starting point.

I had assumed that at year 10 my students would have a basic understanding of the idea of percentages - many of them didn't! Instead of going straight into calculating percentages of quantities and calculating whole quantities given a percentage, and then on to financial arithmetic (simple interest), I went back to basics with the students who needed it, and others who could cope with this were assigned the original tasks I had planned.

Figure 5a. Teacher comments related to changing teaching practice by grouping

I adapted the simpler task that we were going to approach in class with something that reflected the students' greater level of understanding.

I adopted an approach that best addressed the needs of the students based upon the diagnostic test. I was able to avoid certain areas that were well understood and concentrate on areas that were not.

I looked at the course outline. As many of my students were very strong in perimeter, we focused more on area and volume.

Figure 5b. Comments on changes to teaching practice by starting at a higher level

Other strong themes in the comments related to changes in their teaching practice by focussing on problem areas; by explicitly discussing misconceptions with the students; and by focussing on the use of clear language in their explanations. A few teachers also commented on classroom practices for example: *I WILL use more materials and a lot more justification from the students.*

CONCLUSION

As noted above, the intention of the smart tests is to take the results of research into students' understanding of particular mathematics topics and to embed it into an intelligent system: a tool holding distributed intelligence which amplifies what teachers can do. This paper has reported some of the views of the early users of the system. The results show that formative assessment is not yet part of the culture of all schools in our region, but that some schools are certainly ready for it, and indeed are now actively using this as a standard part of their teaching. The issue of to whom feedback should go – to teachers or students – indicated that expectations of a test format (in this case a computer test) can be managed by explaining the purpose clearly to students. Finally, data from the surveys indicates that there is considerable evidence of an increase in teachers' pedagogical knowledge resulting from use of the system,

and that teachers are incorporating the information into their subsequent lessons. The system is being further improved by using detailed feedback from teachers on individual tests and on the usefulness of the advice offered for students at different stages and also by analysing the results of each test to examine and possibly improve the statistical properties of the reporting schemes.

References

Pea, R. (1987). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47-87). Cambridge, MA: Cambridge University Press.

Stacey, K., Price, B., Steinle, V., Chick, H., Gvozdenko, E. (2009, October). *SMART Assessment for Learning*. Paper presented at Conference of the International Society for Design and Development in Education, Cairns, Australia. September 28 – October 1, 2009. http://www.isdde.org/isdde/cairns/pdf/papers/isdde09_stacey.pdf

Steinle, V., Gvozdenko, E., Price, B., Stacey, K., Pierce, R. (2009). Investigating students' numerical misconceptions in algebra. In R. Hunter, B. Bicknell, T. Burgess (Eds.) *Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 491-498) Wellington: MERGA

Wiliam, D. (2007) Keeping learning on track: classroom assessment and the regulation of learning. In F. K. Lester Jr (Ed.), *Second handbook of mathematics teaching and learning*. (pp. 1053 – 1098) Greenwich, CT: Information Age.

Acknowledgement

Smart tests were developed at the University of Melbourne by Professor Kaye Stacey, Dr Helen Chick, Dr Vicki Steinle, Ms Beth Price and Dr Eugene Gvozdenko, with initial funding from the Australian Research Council Linkage Projects Program LP0882176 in partnership with the Department of Education and Early Childhood Development, Victoria. We would like to thank all the teachers and students who have contributed to this project and continue to do so.