12th International Congress on Mathematical Education
Topic Study Group 18
8 July – 15 July, 2012, COEX, Seoul, Korea

USE OF INTERNET IN A MATHEMATICS ASSESSMENT SYSTEM WITH SEMI-AUTOMATIC MARKING AND CUSTOMISABLE FEEDBACK

Khoon Yoong WONG, Kwangshin OH, Qiu Ting Yvonne NG, Siew Kuan CHEONG

National Institute of Education, Nanyang Technological University, Singapore

khoonyoong.wong@nie.edu.sg

In recent years, Assessment for Learning (AfL) has become a major thrust in school assessment practices and a major topic of research. A key component of AfL is to provide meaningful feedback to students about their current performance and what steps they need to take to improve their learning. The teacher should also use data about student performance to plan subsequent teaching activities. To date, there has been little research investigating the use of the Internet to implement features of AfL for secondary mathematics education. In this paper, we describe a mathematics assessment system that is built based on the principles of AfL, delivered through the Internet. This system consists of six major components, namely: (1) delivery of extended assessment tasks; (2) capture of student answers to the tasks; (3) automatic marking of closed questions by the system and semi-automatic marking of open questions by the teachers; (4) a customisable feedback component to be used by the teachers; (5) capture of students' responses to teacher's feedback, and (6) a report system that summarises results. Most components of the system were trialled in 2011 for two extended tasks in four secondary schools in Singapore with about 400 students. This paper describes this system and provides responses from the teachers and students in the trial. The system aims to demonstrate how AfL can be achieved via the Internet for lower secondary school students.
Assessment for learning, Internet, marking system, feedback, Singapore

INTRODUCTION

In recent years, Assessment for Learning (*AfL*) has become a major thrust in school assessment practices and a major topic of research (e.g., Black & Wiliam, 1998; Hattie & Timperley, 2007; Martinez & Lipson, 1989; Sadler, 1989). One key feature of *AfL* focuses on supporting student learning by using feedback loops (Black & Wiliam, 2004; Sadler, 1989; Shavelson, 2006). Research has shown consistently that providing informative feedback to the students can enhance their performance (Black & Wiliam, 1998; Fuchs & Fuchs, 1986), and this formative assessment feature is one of the five key strategies that can help raise teaching quality (Wiliam & Thompson, 2008).

On the other hand, there has been little research investigating the use of the Internet to implement key features of *AfL* for secondary mathematics education to further promote teaching and learning. This paper considers the use of the Internet in a mathematics assessment system with semi-automatic marking and customizable feedback which might be useful for providing opportunities for students to show what they know and can do and to help teachers "close the gap between students' current learning state and the desired state by pedagogical actions" (Shavelson, 2006, p.3). This system is the core part of the Singapore Mathematics Assessment and Pedagogy Project (SMAPP)[1].

THE SMAPP ASSESSMENT SYSTEM

The logic model depicting the SMAPP assessment system is shown in Figure 1.

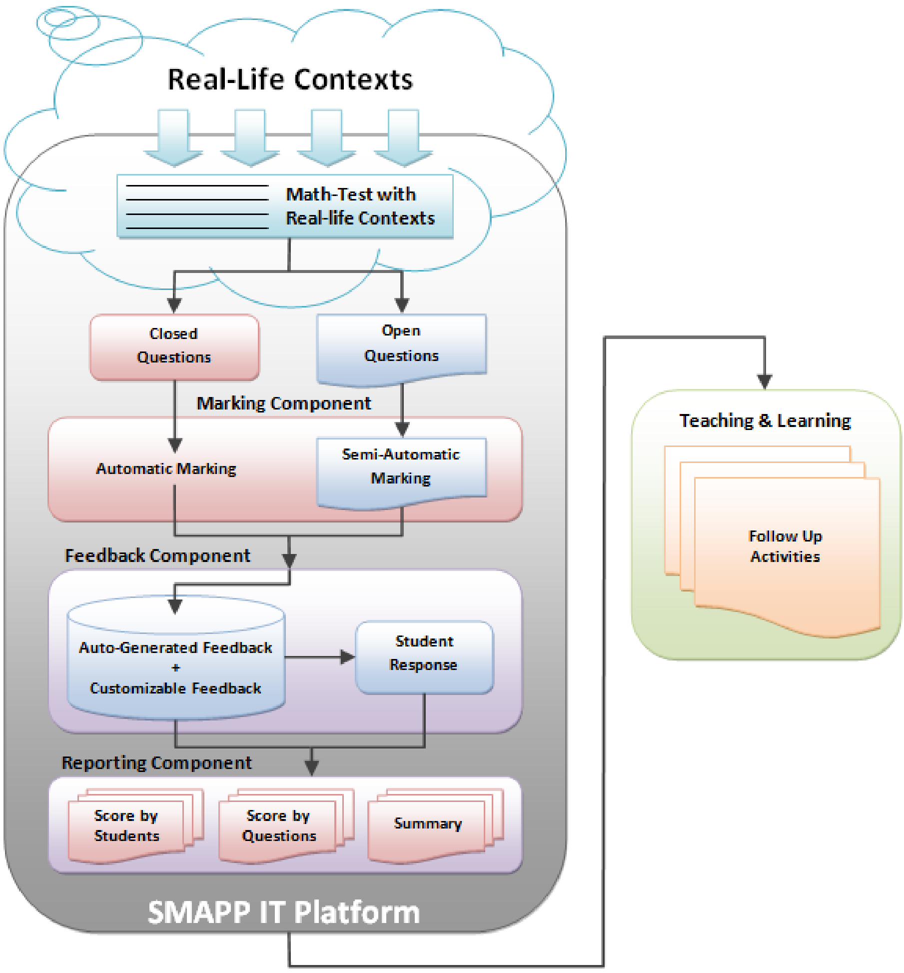


Figure 1. The major components of the SMAPP assessment system

The system follows an *AfL* cycle comprising the following six components:

1) delivery of extended mathematics assessment tasks that cover real-life contexts,
2) capture of student answers to the tasks,
3) automatic marking of closed questions by the system and semi-automatic marking of open questions by the teachers,
4) a customisable feedback component that allows teachers to select pre-designed feedback and to add their own feedback to individual student's answers,
5) capture of students' responses to the feedback given by their teachers, and

[1] http://smapp.nie.edu.sg/teamsite

6) a report system that summarises results about student performance.

Delivery of Extended Mathematics Tasks and Capture of Student Answers

The extended tasks are designed to include real-life contexts and to assess students' reasoning and communication skills. The principles used to design these tasks are given in Fan, Zhao, Cheang, Teo, and Ling (2010) and Zhao, Cheang, Teo, and Lee (2011). Paper versions are converted to web pages which are delivered through Mozilla Firefox or Internet Explorer. Students are given their own identification number and password to log into the SMAPP website to tackle the tasks. Each task takes about one hour to complete.

Closed Questions

A closed question requires students to choose from multiple choice options or to enter a final answer into blank boxes. These are illustrated in Figures 2 and 3.

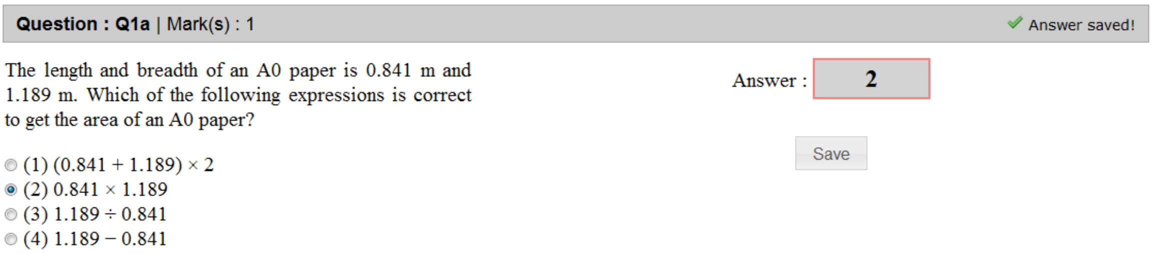
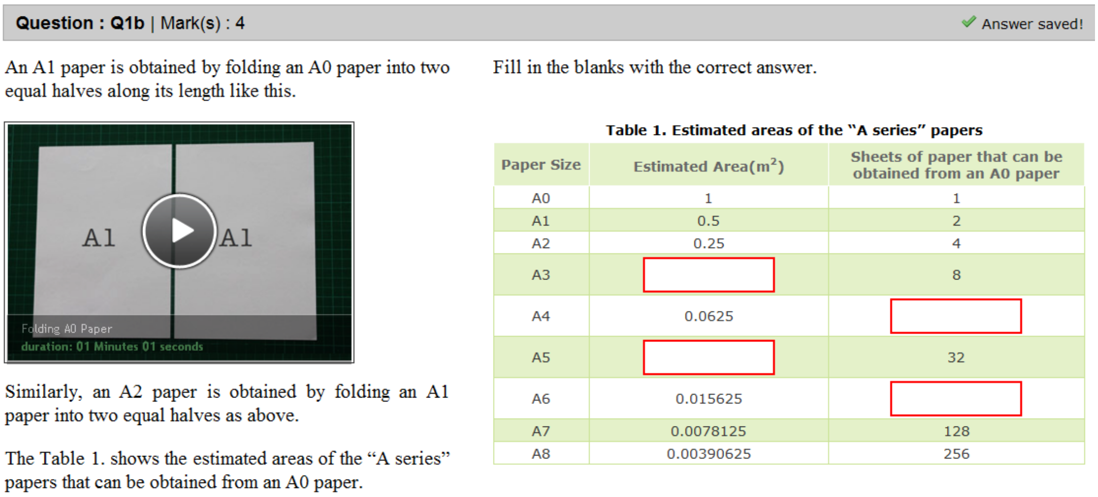


Figure 2. Closed question: Multiple choice



Paper Size	Estimated Area(m^2)	Sheets of paper that can be obtained from an A0 paper
A0	1	1
A1	0.5	2
A2	0.25	4
A3		8
A4	0.0625	
A5		32
A6	0.015625	
A7	0.0078125	128
A8	0.00390625	256

Figure 3. Closed question: Filling in blanks

Open Questions

These open questions require students to type in their mathematical workings and answers into text boxes. The text box editor, as shown on the right hand side of Figure 4, consists of a toolbar to enable students to enter mathematical symbols by clicking on the required ones. An online calculator is also provided. Besides mathematical problems, some open questions ask students to type in texts, for example, to describe two benefits of paper cycling. Such questions do not have strictly right or wrong answers, but they aim to assess students' ability to communicate their ideas clearly.

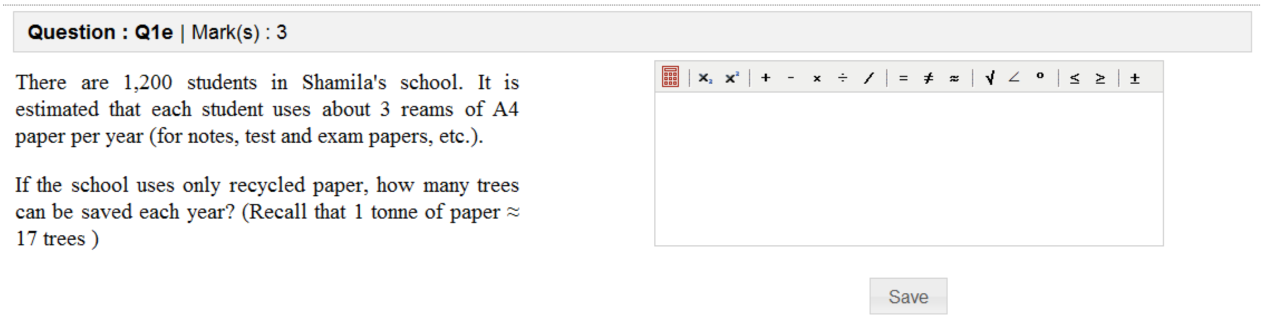


Figure 4. Open question: Show working; text editor on the right

Automatic and Semi-automatic Marking

The automatic marking of mathematics answers is still a challenge for many computer-mediated systems (e.g., Bennett, Morley, & Qurardt, 2000; McGuire & Youngson, 2002; Soh & Subramanian, 2008). Closed questions are automatically marked by the system. Short open questions have several expected answers, and the system suggests the most likely mark to be assigned to the student's answer, but the teacher can review this. Other open questions that have multi-step workings or require text input will not be marked by the system; instead, the teacher has to mark them guided by the suggested marking scheme. An example is given in Figure 5. The teachers can choose to mark by questions or by students. This semi-automatic marking of open questions serves the important pedagogy purpose of having the teachers pay closer attention to the work of their students.

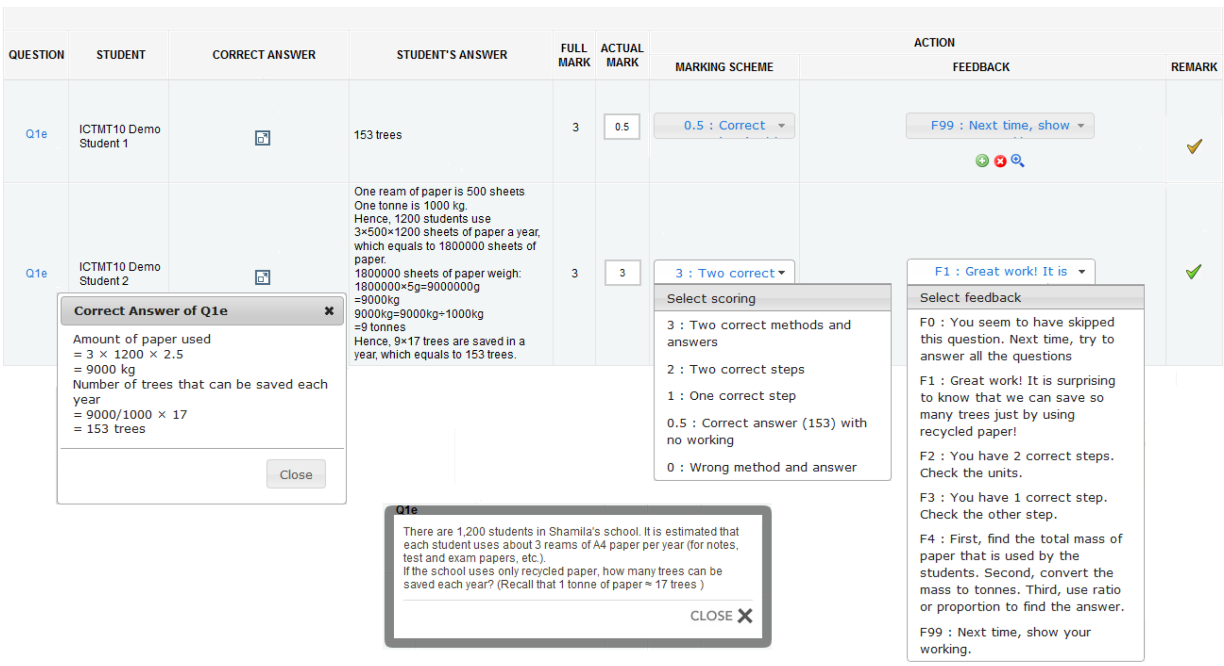


Figure 5. Marking and feedback components: Teacher version

Customisable Feedback

Every question is tagged to several feedback comments. If the student answer is correct, the feedback comment gives a general praise and reiterates the correct procedure to enable the student compare his/her working with the given one or to provide an enrichment activity. If the student answer is wrong, several comments are given, some more detailed than others, to give students hints on how to proceed further, but not the complete solution. See examples in Wong, Oh, Ng, and Cheong (2012). Although these feedback comments are specific and fairly detailed in terms of the mathematics workings, they are not too complex for the students (Shute, 2008).

Since it is not practical to provide numerous feedback comments, the system has included a customizable feedback function to enable the teachers to enter their own feedback comments that they deem to be more appropriate for their students than the given ones. These teacher feedback comments can be saved and used by them and their colleagues in the same school. Through this flexible system, the teacher can save time and effort when they wish to enter the same feedback comment for the same answer. During the trial, some teachers entered several feedback comments of their own, while others just used the given ones. For the paper recycling task, the teachers created 45 new feedback comments; see the example in Figure 6. In this example, besides a general comment (be careful), which is often not specific enough to be helpful, the teacher had included a hint for the main step of the solution. Later the student responded that she now understood the correct step to follow (right most column in Figure 6).

Student's Answer	FULL MARK	Score	Teacher's Feedback for Students	Student's Feedback
Mass of A4 papers each student uses per year = 2.5kg× 3 = 7.5kg Total mass of A4 papers used by the 1200 students per year = 7.5kg× 1200 = 9000 1000kg = 1 tonne 9000kg = 9 tonne Number of trees saved per year = 9 × 17 = 158 158 trees can be saved each year, if the school uses only recycled paper.	3	2	Careful. 9x17=?	Now, I understand.

Figure 6. A new feedback comment added to the Customisable Feedback component

To use the marking and feedback components, the teacher first selects the relevant mark for the student's answer and then selects a feedback comment through a pull-down menu showing default feedback comments. They can also enter their own comments, as explained above. These feedback comments are postulated to "close the gap" in students' mastery of the work, hopefully increasing their mathematics achievement.

Students' Responses to Teacher Feedback

When the teacher has completed marking all the students' answers, the students are asked to log into the system to view their own solutions and the teacher feedback. For the wrong answers, the students can read the correct solutions in relation to the teacher feedback comments, but they are not shown the marks given to their answers. This approach is based on the counter-intuitive findings of Butler (1988) that giving students only comments tends to enhance their interest and performance compared to giving marks or comments and marks. At this stage, as shown in Figure 7, the student selects one of three options to indicate how well they have understood the solutions and comments: (1) Now, I understand. (2) I still do not understand, so I will discuss with my teacher. (3) I still do not understand, so I will discuss with my friends. These are brief responses as we do not wish to burden the students with more elaborate choices.

TASK REPORT			
SCHOOL :	| NAME :	| CLASS :	| TASK 5: PAPER RECYCLING
Question	**Correct Answer**	**My Answer**	**Teachers' Feedback**
Q1e	Amount of paper used = 3 × 1200 × 2.5 = 9000 kg Number of trees that can be saved each year = 9000/1000 × 17 = 153 trees	2.5×1200=9000	First, find the total mass of paper that is used by the students. Second, convert the mass to tonnes. Third, use ratio or proportion to find the answer. (1) Now, I understand.

Figure 7. Feedback and student responses: Student version

Finally, the teacher can look at the students' responses and decides what follow-up activities can be implemented to reinforce the students' learning. This is a critical function of using assessment *for* learning.

Report System

The report system can be used to generate various types of reports about individual students and the whole class: summary report, student report, individual student scores, mean scores, and frequency of responses by class. Similar to the suggestions by Schofield and Ashton (2005), all the data and reports can be downloaded into EXCEL format for subsequent analyses, if necessary. The teacher can examine the output by students or by questions. Examining the reports by individual students enables the teacher to know more about the performance of a particular student and can provide more focused remediation for this student. On the other hand, output by class informs the teacher about the overall performance of the class, and the teacher can decide whether to provide remediation for the whole class or individual students. Finally, by studying the responses by questions, the teacher can take note of alternative answers given by their students and identify common errors in a systematic way. These options provide teachers with the evidence to plan subsequent lessons, hence, putting formative assessment into practice.

FINDINGS ABOUT ONE TASK ON PAPER RECYCLING

The Mathematics Task

This extended task is about paper recycling. The introduction mentions that a girl named Shamila is to help her school saves paper through recycling, and the students are asked to help her achieve this goal. The topics are “Numbers and the four operations”, “Prime numbers”, and “Mensuration”. There are altogether 11 closed questions and 7 open questions. Students are required to look for patterns, perform the four operations on numbers, convert units, and search the web for the benefits of paper recycling.

Sample and Method

A total of 364 students in 11 Grade 7 classes from four different schools worked on this task in the first half of the year 2011. The students worked on the task individually in their school computer laboratory during normal school hours. Although they were asked to complete the task individually, many of them had brief discussions with their seat neighbours during the administration. Ninety nine students were selected to complete an online survey after they had completed the task and 55 students were interviewed immediately after the session was over. Ten teachers were also interviewed about their perceptions of the task and the IT platform.

Results of Online Survey (*n* = 99)

The online survey consists of six Likert items on the 5-point scale: 1 = Strongly Disagree; 5 = Strongly Agree. The results are reported in Table 1. The overall ratings of the items were between the neutral and positive ends. Although the students found this paper recycling task quite realistic, they did not find it particularly interesting or useful in helping them gain new mathematical knowledge. They tended to prefer to work on the task on paper rather than on the IT platform. Some had difficulty navigating between the various screens of texts on the computer and using the IT tools to enter mathematics online.

Table 1: Descriptive statistics of student online survey (*n* = 99)

Questions	Mean	SD
I find the scenario provided in the task realistic.	3.70	1.00
I prefer to do this task on paper rather than doing it using the IT platform.	3.49	1.18
I have no difficulty following the instructions given.	3.43	1.22
I find the scenario provided in the task interesting.	3.34	1.14
I gained some new mathematical knowledge/skills by working on the task.	3.12	1.04
The IT platform is helpful for me to do this task.	3.02	1.55

About 43% of them mentioned having difficulty understanding the questions, and a similar percentage expressed difficulty in understanding the given information. These two findings are consistent with the general impression that many students find word problems difficult, and the teachers can pay more attention to helping their students to develop proficiency in this area.

Results from Student Interviews (*n* = 55)

The above findings were corroborated by what the interviewees said about the task and the IT delivery platform. They felt that the task was realistic but the questions were challenging. They mentioned that the questions were linked so that if they got the first question wrong, they very likely got the other questions wrong as well. Some of them felt that the task was good and could help them in their mathematics grade.

There was also a strong preference for working on paper rather than online. They were not used to typing equations and operations on the screen using the keyboard and the mathematics editor. They preferred scribbling rough workings on paper and they found it difficult to do that on the computer. The online calculator was found to be harder to use than a physical one.

Responses from Teachers (*n* = 10)

The teachers also answered a 5-point Likert scale about the implementation of this task. Their responses are summarised in Table 2. The teachers generally appreciated the links between the task and real-life situation. They found the feedback component quite easy to use and the default feedback comments seemed to cover most of the errors made by the students. During interviews, some of them mentioned that they may have difficulty finding the time to integrate such extended task into regular teaching.

Table 2: Descriptive statistics of teachers' responses (*n* = 10)

Questions	Mean	SD
Appreciate the connections of mathematics with real life situations.	4.20	0.92
It is easy to add my own feedback to the system.	3.75	1.04
By focussing on open-ended questions, I know more about my students' thinking, e.g., different methods used.	3.44	1.13
I prefer this system to manual marking.	3.33	1.12

Reliability of Marking and Feedback

There was substantial agreement between the marks awarded by the teachers and the research assistants from the National Institute of Education (NIE) for the open questions in this task. The correlation was .99 (for 364 student marks), and the mean marks by the two groups were 20.95 (teachers) and 20.76 (NIE). This shows that the marking scheme could be used reliably and the online semi-automatic marking system is a reliable tool.

CONCLUSION

The SMAPP assessment system described here is a comprehensive prototype that aims to promote formative assessment by providing components that address the “assessment → marking (evidence) → feedback” cycle. We have found that it is a workable model in terms of delivery, semi-automatic marking, and giving feedback. However, some students voiced their preference to work on the paper version, and it could be due to certain weaknesses in the Internet environment for doing mathematics and the need for a better interface. We agree with Martinez and Lipson (1989) that “assessment for learning won’t happen automatically” (p.75). Thus, more testing and development need to be conducted if this *AfL* system is to be scaled up to reach more students and teachers.

References

Bennett, R. E., Morley, M., & Qurardt, D. (2000) Three response types for broadening the conception of mathematical problem solving in computerized tests. *Applied Psychological Measurement, 24*, 294–309.

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

Black, P., & Wiliam, D. (2004). Working inside the black box. *Phi Delta Kappan, 86*(1), 8–21.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 58*(3), 1-14.

Fan, L., Zhao, D., Cheang, W. K., Teo, K. M., & Ling, P. Y. (2010). Developing disciplinary tasks to improve mathematics assessment and pedagogy: An exploratory study in Singapore schools. *Procedia Social and Behavioural Sciences, 2*(2), 2000–2005.

Fuchs, L., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*, 199–208.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Martinez, M. E., & Lipson, J. I. (1989). Assessment for learning. *Educational Leadership,* 46(7), 73 – 75.

McGuire, G. R. & Youngson, M. A. (2002) Assessing ICT assessment in mathematics. *MSOR (LTSN) Series on CAA in Mathematics*. Retrieved from http://ltsn.mathstore.ac.uk/articles/maths-caa-series/mar2002/index.shtml

Sadler, D. R. (1989). Formative assessment and the design of instructional strategies. *Instructional Science, 18*, 119–144.

Schofield, D., & Ashton, H. (2005). Effective reporting for online assessment: Shedding light on student behaviour. *MSOR (LTSN) Series on CAA in Mathematics*. Retrieved from http://www.heacademy.ac.uk/assets/documents/subjects/msor/mathscaa_feb2005.pdf

Shavelson, R. J. (2006). *On the integration of formative assessment in teaching and learning with implications for teacher education.* Retrieved from http://www.stanford.edu/dept/SUSE/SEAL

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.

Soh, P.H., & Subramanian, A. M. (2008). Is usage a missing link in explaining the perceived learning outcome of technology-mediated learning? *IEEE Transactions on Engineering Management, 55*(1), 50–66.

Wiliam, D., & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), (2008). *The future of assessment: Shaping teaching and learning* (pp. 53-82). New York: Lawrence Erlbaum Associates.

Wong, K. Y., Oh, K., Ng, Q. T. Y., & Cheong, J. S. K. (2012). Linking IT-based semi-automatic marking of student mathematics responses and meaningful feedback to pedagogical objectives. *Teaching Mathematics Applications, 31*(1), 57-63. doi: 10.1093/teamat/hrr023

Zhao, D.S., Cheang, W. K., Teo, K. M., & Lee, P. Y. (2011). Some principles and guidelines for designing mathematical disciplinary tasks for Singapore schools. In *Mathematics: Traditions and (new) practices: Proceedings of the AAMT-MERGA conference* (pp. 1107-1115). Adelaide: Australian Association of Mathematics Teachers.

Acknowledgement. This project is funded by the Centre for Research in Pedagogy and Practice, National Institute of Education, Nanyang Technological University, Singapore (Project code: OER 01/10FLH). The views expressed in this article are those of the authors' and do not necessarily represent the views of the Centre for Research in Pedagogy and Practice, the National Institute of Education, or the Ministry of Education, Singapore.