

# Handbook of Research on Driving STEM Learning With Educational Technologies

María-Soledad Ramírez-Montoya  
*Tecnologico de Monterrey, Mexico*

A volume in the Advances in Educational  
Technologies and Instructional Design (AETID)  
Book Series



[www.igi-global.com](http://www.igi-global.com)

Published in the United States of America by  
IGI Global  
Information Science Reference (an imprint of IGI Global)  
701 E. Chocolate Avenue  
Hershey PA, USA 17033  
Tel: 717-533-8845  
Fax: 717-533-8661  
E-mail: [cust@igi-global.com](mailto:cust@igi-global.com)  
Web site: <http://www.igi-global.com>

Copyright © 2017 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

CIP Data Pending  
ISBN: 978-1-5225-2026-9  
eISBN: 978-1-5225-2027-6

This book is published in the IGI Global book series *Advances in Educational Technologies and Instructional Design (AE-TID)* (ISSN: 2326-8905; eISSN: 2326-8913)

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: [eresources@igi-global.com](mailto:eresources@igi-global.com).

## Chapter 20

# The Effectiveness of Computer– Aided Assessment for the Purposes of a Mathematical Sciences Lecturer

**Stephen James Broughton**  
*Oxford Brookes University, UK*

**Paul Hernandez-Martinez**  
*Loughborough University, UK*

**Carol L. Robinson**  
*Loughborough University, UK*

### **ABSTRACT**

*Computer-Aided Assessment (CAA) is becoming an increasingly popular method for assessing students in their mathematics courses in higher education. This article examines six lecturers' practices of using CAA on their mathematics courses. The interviews with these lecturers revealed that the CAA system did provide many of the benefits that were promised; however, there were some important aims not satisfied by the system, which limited the scope of its effectiveness. Using a model for effective assessment, which draws upon ideas from the assessment literature and cultural-historical activity theory, the lecturer interviews give an insight into what stops this assessment tool from remaining effective. This study shows that the CAA system was reasonably effective to an extent, and lecturers had achieved a relatively stable practice that they were satisfied to maintain; however, there were shortcomings with the existing system that limited the scope of its effectiveness, which led to diverse practices and a desire to change system.*

DOI: 10.4018/978-1-5225-2026-9.ch020

## INTRODUCTION

Computer-aided assessment (CAA) is a means of offering assessment electronically (Bull & McKenna, 2003), offering immediate and automated feedback. There are several CAA systems in use across UK higher education institutions (HEIs) that assess students in mathematics, such as HELM, QM Perception and STACK (Green, Harrison, Palipana, Pidcock & Ward, 2004, Greenhow and Gill, 2005, Sangwin, 2007). There have been calls in the literature for a systematic and rigorous review of its use (McKenna & Bull, 2000, Sangwin, 2003).

Lecturers assess students for many reasons and Samuelowicz and Bain (2002) encountered many in their study, including: to encourage students to study, to differentiate between students of different abilities, and to guide students towards a professional standard. There are also institutional demands, constraints and rules that shape lecturers' aims and choices for assessment (Hunt, 2006).

In the United Kingdom, there are changing expectations in a developing higher education landscape. Students demand more time from their lecturers and are less willing to work independently (Rolfe, 2002). The number of students attending universities has increased substantially since the 1960s (Blanden & Machin, 2004) and cohort sizes are further increasing.

Amid these changes and competing demands for time, lecturers face increasing pressure over assessment. Traditional forms of assessment used in mathematics learning that are marked by hand and offer personalized feedback take substantial time (Biscomb, Devonport & Lane, 2008). CAA offers the facility to provide feedback to students quickly and are relatively straightforward to administer. It permits lecturers to spend more time on their other commitments, such as their research interests.

However, while lecturers at HEIs have adopted CAA for use with engineering and mathematics students, it does not necessarily follow that the CAA systems they use satisfy all their aims for assessment.

This research study focuses on lecturers' use of a CAA system with first year students at one particular HEI: with one bank of questions for mathematics students and one for engineering students. The questions were in multiple choice and numerical input formats; the feedback comprised either the full solution to the question or the solution to a generalized form of the question. The lecturers were able to offer practice tests to students prior to a summative test.

Among the lecturers of these first year modules, not all had adopted CAA and practices were diverse (Broughton, Robinson & Hernandez-Martinez, 2013). The system had been in place for ten years and had not been evaluated formally in that time. The dearth of studies evaluating the effect CAA systems have had on learning and teaching in the literature suggested this situation was common. Therefore, there remained a need to establish whether these CAA systems were proving to be useful for lecturers and students. The authors have reported on how students have used this CAA system and where it has been effective (Broughton, Hernandez-Martinez & Robinson, 2012); it is the purpose of this chapter to report on the experiences of six lecturers that have been using the CAA system with their mathematics and engineering students.

The research questions for this study are:

- Why did these lecturers use the CAA system?
- How did they arrive at their current practice?
- For what aims was the CAA system effective?

This chapter highlights the importance of the lecturers' aims when determining to what extent an assessment tool has been effective for them. It also explains what is meant by effectiveness in assessment by introducing a model derived from assessment literature and a cultural-historical activity theory perspective. It does so by referring to a study in which lecturers used a CAA system with mathematics and engineering students in mathematics: the lecturers identified their motives for using the assessment tool; the contradictions and successes they encountered, and how they responded in both cases; and how they arrived at a stable practice, which came after resolving these contradictions.

## LITERATURE REVIEW

The purpose of this literature review is to offer a brief overview of the rationale for the model of effective assessment proposed in the following section. The model uses ideas from cultural-historical activity theory and the wider assessment literature. This section also briefly discusses the CAA system used at the studied institution.

### Model for Effective Assessment

Assessment exemplifies the key constructs of activity proposed by the architects of the three generations of activity theory. Vygotsky introduced the notion of tool-mediated social activity: assessment is a tool in a learning environment that is used to determine the extent of learning that has taken place. Leont'ev described such use of a tool by a subject to pursue an object as a unit for analysis. Engeström made the important link between activity and learning — in expansive learning — and added the concept of contradictions to describe the problems that occur within an activity system that cause change. Engeström also described the relationships between different 'nodes' in an activity system, often drawn in a triangle: the subject, the object of the activity, the tools used, the community engaged in the activity, the rules of the activity, and the division of labour.

At the core of activity theory is the notion that “needs direct activity” (Leont'ev, 1977, p. 89). That is, one is driven to act to fulfil needs. From a learning perspective, students are most likely to engage in a learning activity if it is appropriate for their needs.

Individuals interpret their own needs when forming goals and are influenced by a number of factors, including other people involved in the activity system. Many students may first experience having learning goals set on their behalf; however, as Tikhomirov (1988) argues, they may not adopt these goals for themselves. Instead, students may formulate their own goals, which is expected of students that have developed skills in self-regulation and autonomy of learning (Nicol & Macfarlane-Dick, 2006).

Both Leont'ev's (1977) concept of individual activity and Tikhomirov's (1988) understanding of individual selection of goals imply the need to evaluate assessment with respect to individual students and their individual goals. If the student is able to articulate goals and verify whether those goals have been achieved or not, it offers a measureable outcome of the assessment.

Indeed, it suggests that an assessment would work best if it is appropriate for the student's goals and, conversely, assessments whose goals are incompatible with the student's learning goals might be less effective. This supports an assertion made by Black and Wiliam (2009, p. 10) that “actions that would appear to be the most likely to produce learning might not do so in a particular situation”. Engeström

and Sannino (2010) would describe this incompatibility as a *secondary contradiction*. The emergence of contradictions in assessment is important: it is inevitable that such contradictions change the activity (Engeström, 2001).

The student's learning goals appear important, but they are liable to change over time. Engeström describes activity as being "recurrent and cyclic" (1999, p. 33); but actions are "linear and anticipates a finite termination" (Engeström, 2001, p. 33).

Tikhomirov (1988) described actions as being linear, but essentially form a cycle: "an action is, as it were, doubly productive: the reaction of a new material product is preceded by the creation of a new ideal product" (Tikhomirov, 1988, p. 120). By this, Tikhomirov suggested that individuals do not achieve a sense of accomplishment and closure upon completing an action; rather, an opportunity for improvement immediately presents itself.

Ideally, the student achieves learning goals and moves on to more challenging learning goals. The assessment must provide information to the student to indicate performance in relation to the learning goals. It is a clear requirement that the assessment provides feedback to the student for this purpose. The nature of this feedback must be such that the student is able to interpret performance in relation to the learning goals.

It is inevitable with any assessment that not every student will achieve the learning goals in the first instance, as Black and Wiliam (2009) noted. Feedback should be an opportunity to support students in making the most appropriate decisions if the assessment confirms that the learning goals have not been met.

In order to make the best decision about the next steps, students should be able to evaluate their positions with respect to the learning goals. Assuming that the assessment is appropriate for the learning goals, then either the student would be able to achieve the learning goals, with the benefit of feedback, or the student is better placed to formulate more appropriate learning goals, given the benefit of feedback.

Therefore, feedback should advise students on the next steps in the following ways:

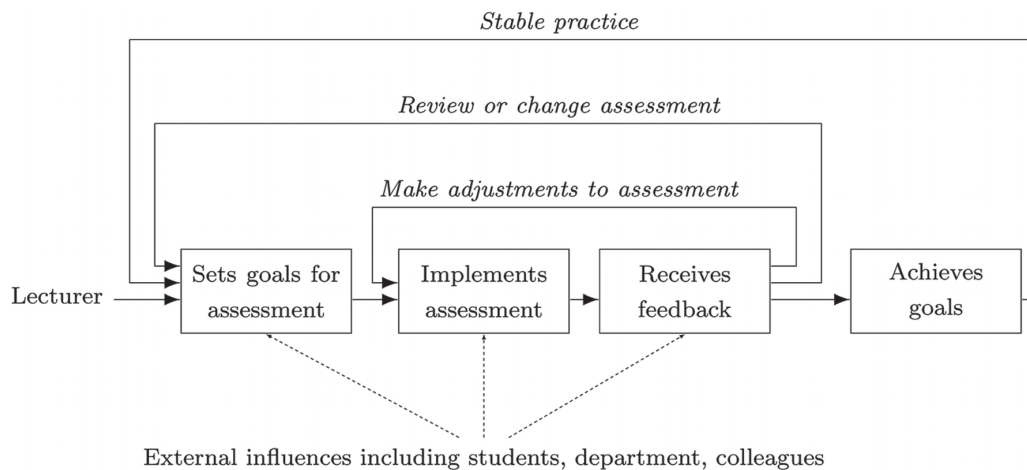
1. Feedback gives students sufficient information for them to be able to determine whether they have achieved their learning goals.
2. Feedback guides students towards their learning goals, given the work they have produced in the assessment.
3. Feedback offers information to students that allows them to formulate more appropriate learning goals.

Earlier work on this study introduced a model for effective assessment with respect to the student (Broughton et al., 2013), showing the paths that students should take in learning cycles, using assessment and feedback to evaluate and inform their learning. It also highlights the importance of individual students' aims and goals for learning, as well as the influencers on the formation of these goals.

This model was constructed from ideas within cultural-historical activity theory — including the notions of individual goal formation (Tikhomirov, 1988), expansive cycle of learning actions (Engeström, 2000), and the transfer of guild knowledge through internalising and externalising knowledge (Engeström, 1999) — and from the wider assessment literature — such as self-regulation (Nicol & Macfarlane-Dick, 2006).

This chapter proposes an analogous model for lecturers, whereby lecturers strive to achieve their aims for assessment and learn whether those aims have been achieved by delivering that assessment (Figure 1).

Figure 1. The cycle for effective assessment for lecturers



This model provides the means to evaluate the effectiveness of an assessment with respect to a lecturer's aims, by helping to identify points of departure from the cycle in the model. The points of departure are where the assessment tool no longer assists the lecturer in satisfying the aims of assessment. The following section describes the methods for the collection of data with respect to some UK higher education lecturers in mathematics that were using this one particular CAA system.

### **Computer-Aided Assessment Used in the Study**

The CAA system used by the institution at the focus of this study comprises two similar subsystems: one with a question bank for engineering students studying mathematics; and one with a question bank for mathematics students. The former bank of questions was largely developed in the HELM (Helping Engineers Learn Mathematics) project (Green et al., 2004), while the latter was developed after this project with questions designed for mathematics modules.

Questions delivered through this CAA system could accept numerical and multiple choice responses. As a consequence, algebraic responses were offered as multiple choice options, or students may be asked to give a numerical response in a particular format. Students usually received static feedback, giving either a general solution or a full-worked solution.

The CAA system also allowed students to take formative practice tests before undertaking a summative test. The practice test questions were usually very similar to those in the summative test; however, the questions were never identical. Once students have been granted access to the tests, they may do so via the Internet at any location and at any time; typically, the summative test is made available for between a few hours and a day.

### **METHODOLOGY**

Fourteen lecturers taught mathematics modules to first year students in the mathematical sciences and engineering departments at the university across eleven modules. From these, four declared that they

did not use CAA with their students and another lecturer was involved with this study and thus did not participate.

Following the distribution of a questionnaire to the remaining nine lecturers, six lecturers indicated a willingness to participate in interviews (hereafter denoted by codes L1–L6), which followed a semi-structured format of questioning. The questions covered their reasons for using CAA, their current CAA practice, their views of the system, and what they believed to be the successes and failures of the system. The interviews lasted between 27 and 54 minutes. Each interview was audio-recorded and transcribed by a professional transcriber.

Their responses were coded in three ways: first, with respect to the six nodes of Engeström’s (1987) triangular model of activity, in order to build a profile of each lecturer’s activity; second, with respect to the model for an effective assessment (Broughton et al., 2013); and third, with respect to the research questions. For example, L1 was asked in an interview, “What are your goals for assessment?” An excerpt from L1’s response is given below.

“I want to know at some level how well my students are doing. Now, that would encompass a number of things. So, if I look at what I teach, I teach a large unit on functions, I teach a unit on matrices, I teach a unit on differentiation. So, if I look at each of those: What do I want my students to know? What do I want them to be able to do? What do I want them to understand? So, if we take matrices, for example, well certainly, I want them to be able to add and multiply matrices, and I want them to be able to find inverses, which requires knowing how to find determinants. All of these can be achieved. It’s how to do them. They can be achieved fairly instrumentally. And assessment can contribute to that. So, I think the CAA is pretty good in achieving that.” (L1)

With respect to the activity system, L1 places herself as the “subject” of the activity of assessing students and the questions that she asks form the “object” for the activity. CAA is identified as a “tool” for this aim. The students form a part of the “community” in which the assessment takes place, and L1 is clear that they are an important consideration in deciding the course of her activity. She is clear that she has the responsibility to choose the assessment tool and what is assessed; subsequently the students are responsible for completing the assessment. This defines the “division of labour” of the activity. The “rules” are not identified here.

This excerpt provides some background for the model to analyse the effectiveness of the assessment: L1 is the “subject” and has “aims” for assessment and ultimately concludes that CAA is “pretty good in achieving that”. This indicates that L1 continued to find CAA effective for these aims.

With respect to the research questions, L1 indicated that she used CAA to establish what her students had grasped in relation to the study material. The excerpt did not address the second research question, but it suggested that the CAA was effective for establishing how much the students had learned.

The following analysis addresses the research questions using quotes from the transcripts.

## **ANALYSIS**

### **Why Did These Lecturers Use the CAA System?**

All six lecturers had continued to use CAA after inheriting modules from lecturers that had implemented the system previously. This proved to be an overwhelming reason for using CAA in the first place, since there was a significant effort required to implement the system. L1 noted, “If there had been no CAA



when I came [to teach this module], I might have thought about having it but I would have been put off immediately by the fact that there is no way that I can set it up myself”.

Although the existing implementation of CAA was sufficient reason to continue using the system, the system allowed them to offer frequent and timely assessments without the burden of marking. All of the lecturers believed that frequent testing allowed students to practice using the mathematical knowledge they were acquiring in lectures; some lecturers believed that the students might not practice otherwise (L2 and L4).

Table 1 gives the aims that the lecturers identified for assessment during their interviews. It shows that while many of their aims were addressed by using the CAA system, there were some aims that were not achieved this way.

Although only L2 appeared to have his aims satisfied by the CAA system, all of the lecturers had reached a stable practice that incorporated the CAA system. The following subsection describes how the lecturers arrived at their current practice.

### **How Did They Arrive at their Current Practice?**

An earlier finding from this study was that the lecturers' CAA practices were diverse (Robinson, Hernandez-Martinez & Broughton, 2012). There was no consensus between interviews of best practice, despite informal conversations between colleagues.

There were several practical choices that the lecturers made when adopting CAA practice: whether to use the summative testing facility; whether to invigilate the summative test if they did so; to what extent students should have access to the practice testing facility; and so on. In the interviews, the lecturers gave their reasons for the choices they made when arriving at their current practice.

In some cases, lecturers continued the practice adopted by their predecessors. L1 said, “That’s the way it has been done. So... I followed what other people had done.” L5 suggested a similar process: “When I first started to teach here... I was following general advice given on how it was done the previous year, so I followed the system [practice] that was in place.”

However, in most cases, the lecturers made adjustments to existing practice in order to overcome contradictions between their aims for assessment and what the system offered. For example, L2 replaced the CAA summative test with a paper test because he believed they lacked challenge; he noted, “I looked at some of these [CAA summative] tests myself, and I can see that the questions are mostly repeated and are simple... I want to differentiate between someone who could do a normal task and some who could do something more difficult.”

For L1, a regime of assessment that consisted solely of CAA summative testing would leave mathematics learning “impoverished”, because they encouraged “instrumental learning”; the author suggests that this is an indication that L1 believed that the CAA summative tests were linear in nature and encouraged students to perform just well enough to pass the test without significant and long term learning gain. To counter this problem, L1 introduced group coursework tasks, which encouraged communication and exploration of key concepts to a deeper level.

Some lecturers believed the CAA summative tests to be unreliable as a measure of ability or performance. As well as the issue raised by L1, in that there were aspects of mathematics that some of the lecturers felt were not tested, there were some concerns that the students could collaborate during the invigilated test.

For large cohort sizes, particularly for the mathematics cohort that L3, L4 and L6 taught, it was unrealistic to invigilate the whole group in a single computer laboratory. L1, L3, L4 and L5 all reduced either the number of the summative tests used in their modules or reduced the weighting that the tests had against the overall module mark. L4 did not believe that invigilation would be worthwhile: “For 2.5% [of the module mark], I just don’t think it’s worth putting up a major police operation to find out what students actually do.”

All six of the lecturers had reached stable practices, albeit they were diverse. Nonetheless, the stable state of the lecturers’ assessment regimes, and CAA in particular, permits an evaluation of the effectiveness of the CAA system at achieving the lecturers’ aims for assessment.

### **For What Aims is the CAA System Effective?**

The nature of the model suggests assessment tools will be effective to a point. For example, this CAA system would not be effective for teaching second year mathematics students since there are no questions available for those topics. Therefore, it is the point at which the assessment tool is no longer effective that is of interest.

Each lecturer presented their own stories, describing the issues and limitations they encountered when using the CAA system and when the system is no longer appropriate for assessing students. Figures 2–7 show the points of departure from the cyclic model for effective assessment.

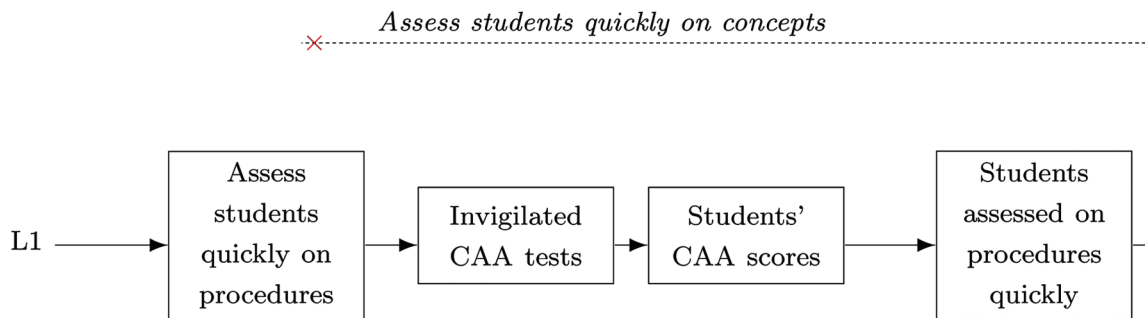
L1 (Figure 2) wished that she could have devoted more time to developing questions for the system: she believed that they encouraged “a more instrumental way of seeing mathematics” — that is, the view that mathematics involves following procedures to arrive at a solution. She had a desire to encourage and assess students’ learning of underlying concepts; however, the questions available to her were not able to assess her students to the level of understanding she desired.

This proved to be the point of departure from the cyclic model of effective assessment; however, she noted that CAA had an important role within her assessment regime: “I think that they fulfil a very

*Table 1. Lecturers’ aims satisfied and not satisfied by the CAA system*

	<b>Aims Satisfied</b>	<b>Aims Not Satisfied</b>
L1	<ul style="list-style-type: none"> <li>● Enable students to practice</li> <li>● Test students quickly and save time</li> </ul>	Test conceptual understanding
L2	<ul style="list-style-type: none"> <li>● Encourage student to work between lectures</li> <li>● Enable students to practice</li> </ul>	<ul style="list-style-type: none"> <li>● Differentiate between students</li> <li>● Prepare students for a variety of contexts</li> </ul>
L3	<ul style="list-style-type: none"> <li>● Test a large number of students quickly and efficiently</li> <li>● Test students on their ability to carry out procedures</li> <li>● Enable students to practice</li> </ul>	<ul style="list-style-type: none"> <li>● Establish the extent to which the students have grasped the content</li> <li>● Test whether students are capable beyond the lecture content</li> </ul>
L4	<ul style="list-style-type: none"> <li>● Encourage students to work between lectures</li> <li>● Consolidate learning that happens in lectures</li> </ul>	Expand and develop students’ view of what “doing mathematics” involves
L5	<ul style="list-style-type: none"> <li>● Provide a fair means of assessment</li> <li>● Enable students to practice</li> </ul>	Encourage students that are struggling with the content
L6	<ul style="list-style-type: none"> <li>● Allow students to monitor their progress</li> <li>● Enable students to practice</li> <li>● Provide a convenient form of assessment for students</li> </ul>	Test conceptual understanding

Figure 2. Point of departure for L1



useful purpose in enabling students to practice and become familiar and confident with ways of doing things... More confident students are, in my view, more likely to be open to more conceptual thinking.”

Therefore, the CAA system was effective for L1 until she needed to test the students’ conceptual understanding. However, her practice remained stable from this point since there were no contradictions between her practice and her aims: she could address her aim to assess conceptual understanding by employing group coursework.

L2 (Figure 3) aired similar concerns that the CAA system encouraged “instrumental learning”: he referred to context-dependency and concerns that the students would not be able to apply their knowledge in new situations. By using a paper test in place of the summative test offered by the CAA system, L2 could set his own questions so that he could “impose the factor of making the students think”.

He noted that he continually refines his assessments each year and resolved to introduce more practical, real-world contexts to his questions. His practice had reached a stable state, with no outstanding contradictions. The CAA system’s practice facility was effective for training his students, though it was not effective for providing a satisfactory summative test for L2’s needs.

L3 (Figure 4) encountered two particular contradictions in the course of her experience with the CAA system. The feedback that the system provided was too static for her taste, noting: “I don’t think it’s good quality feedback in the sense of being individual, or being able to give hints, or anything like

Figure 3. Point of departure for L2

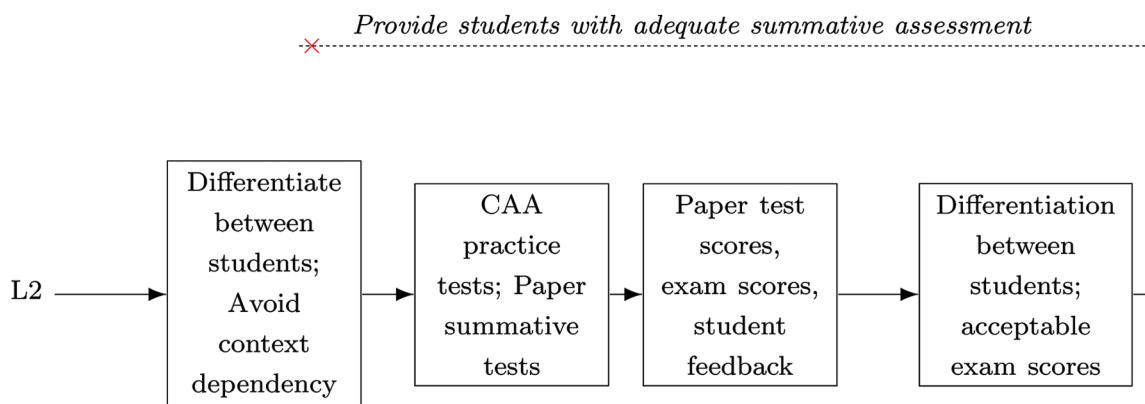
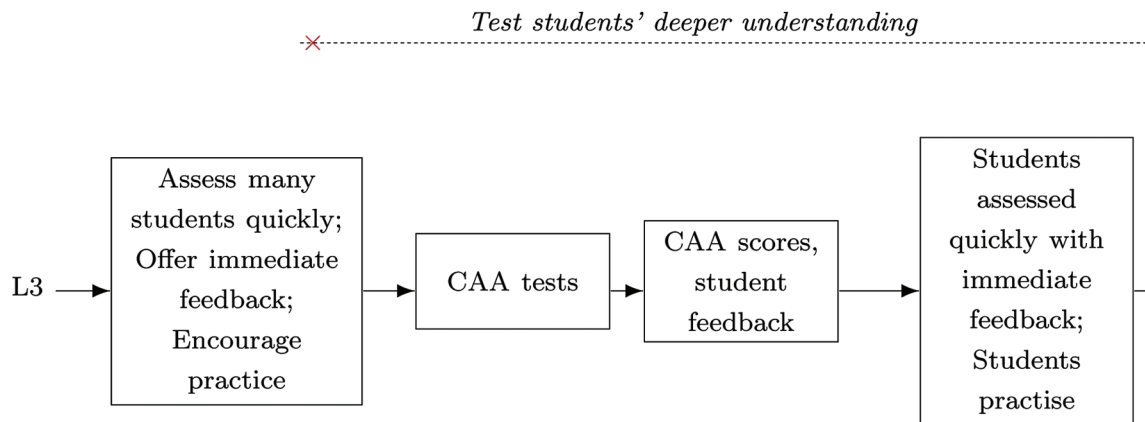


Figure 4. Point of departure for L3



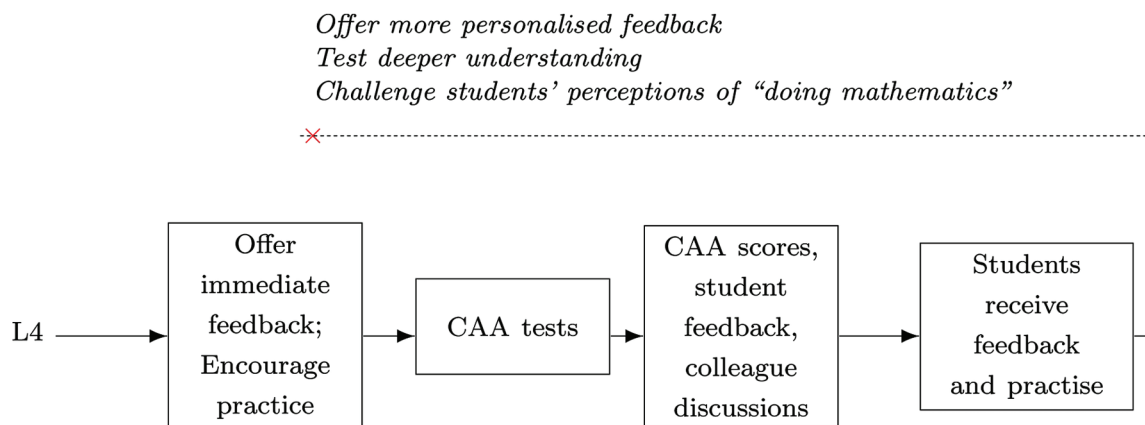
you would do if you were in the room with a person”. Although the feedback was offered immediately, L3 felt that the lack of a personalized response was a compromise. She also felt, like L1, that the tests lacked the ability to assess students on their conceptual knowledge.

In response to these contradictions, L3 reduced the number of CAA summative tests in her module to a single test, and replaced the other test with an alternative assessment tool. This tool permitted her to assess her students on a conceptual level and to receive more personalized feedback.

Consequently, she had reached a stable practice and the CAA system was effective for her aims to encourage students to practice and to allow them to receive immediate feedback. However, the system was not effective at testing her students’ deeper knowledge or for providing personalized feedback.

L4 (Figure 5) was the only lecturer in the study that attempted writing new questions for the CAA system in order to test his mathematics students on their conceptual awareness. He found that the process of doing so was onerous and was not a viable use of his time. Like many of his colleagues, L4 felt that the CAA questions were procedural in nature and too similar to what his students have experienced already: “Because I’m more or less bound to the low-level computational things, I end up asking the questions

Figure 5. Point of departure for L4



that students expect to be asked, even though I'd much rather try to move away from that expectation and change their view of what it means to do mathematics.”

Nonetheless, he declared that there were “other assignments in which I can do that”, and he used group tutorials to set and discuss mathematical problems that required more thought and application of knowledge. Although he would consider developing more questions, given the time, in the future, he said, “For now, I'm quite happy with what I have.” L4 added, “These CAA tests, for me, are essentially formative. They have some nominal credit attached to them but, essentially, they are there for students to work through them and to learn from it.”

He was satisfied that CAA was fulfilling his aims to encourage students to practice between lectures and had no plans to change his practice. He found the system effective for that purpose, but it could not provide the level of feedback and challenge he desired.

L5 (Figure 6) was the only lecturer in this study that had unresolved contradictions in her practice of using the CAA system. Her primary concern was that the weakest students in her cohort were becoming disheartened by harsh marking criteria and the lack of personalized feedback. She said, “I've come to conclude that CAA testing advantages the better students... They gain a lot of confidence if they're getting it right... If they're nervous or if they forget, ... they're ending up with zero [marks].”

In the course of her interview, L5 pondered alternatives to CAA. She concluded that she would not change her practice as a result of determining that she was “not convinced of alternatives”, although she did use other assessment tools, such as group projects. She felt that her engineering students were overburdened by coursework and projects and had no desire to add further to these time-consuming assessments. On balance, she felt that, “I'm only using two every semester, and I think that is enough.”

What she regarded as the system “failing” her weaker students left her practice prone to further changes, as the unresolved contradiction continued to challenge her. While she found the assessment effective at reassuring her students that they were able to undertake the required mathematical methods, it was not effective for all her students.

L6 (Figure 7) was keen to encourage students to practice using the knowledge they had acquired in lectures; however, like L3, she noted the difficulty in assessing mathematical conceptual understanding and felt that the CAA system was not capable of achieving this aim. She said, “I think it's very difficult to evaluate conceptual understanding without asking them to write something,” before noting that the system is not usually used to accept written answers and it would be difficult to analyse written responses for conceptual awareness.

*Figure 6. Point of departure for L5*

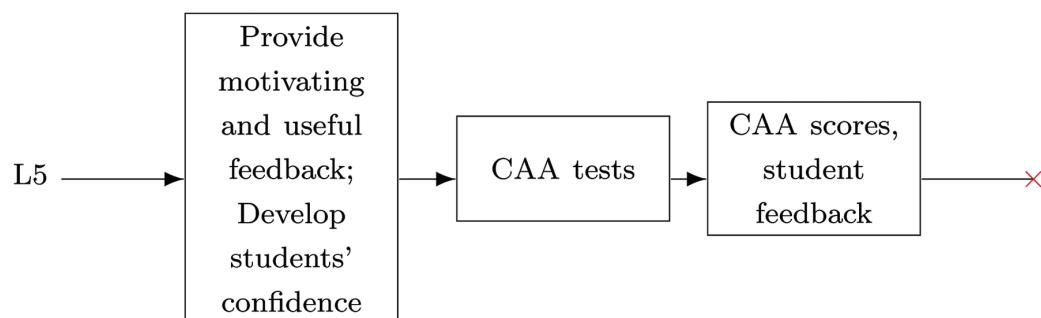
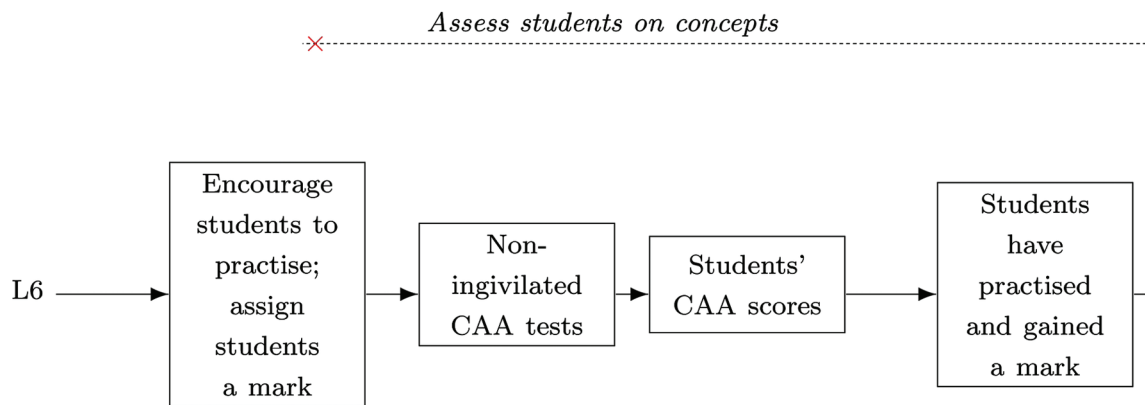


Figure 7. Point of departure for L6



The CAA system was effective, however, at ensuring that her students could perform the necessary mathematical procedures. Her practice remained stable because she could assess her students in other ways: “I just test the procedures through the computer courseworks and I test the conceptual things through written courseworks. It’s by design, partially, that I’ve done it that way.”

## DISCUSSION

The six lecturers had all acquired existing practice of the CAA system when they succeeded previous lecturers on their modules. This was a significant reason for using the CAA system on their courses, since it was viewed as a significant time outlay to implement the system.

The system itself offered many benefits for the lecturers and their students: once implemented, it could offer unlimited practice tests with immediate and detailed feedback; it would save lecturers’ time in the long term, since it removed the burden of marking assignments; and it was more than adequate at testing whether the students could carry out the procedures they were expected to know.

However, the system could not fulfil all the aims of the lecturers. As a result of these contradictions between these aims and the facilities this assessment tool could offer, the lecturers implemented the system and developed their practice in diverse ways to accommodate other assessment tools.

The most common complaint that resulted in curtailing the extent of the effectiveness of the CAA system was its lack of challenge. Most of the lecturers wished to assess their students on their conceptual awareness and either adopted other assessment tools to use alongside CAA or, in the case of L2, devised an alternative summative test.

Some of the lecturers felt that there was scope to improve the system by developing new, more challenging questions, which would increase the effectiveness of the system. However, all the lecturers reported that they were satisfied with their current practice, so such further development of the system was unlikely.

Lecturers endeavour to reach a relatively stable practice where there are minimal contradictions between practice and what the lecturer aims to achieve in assessment. Engeström argued that instability and contradictions are the cause of further change — contradictions generate “disturbances and conflicts, but also innovative attempts to change the activity” (Engeström, 2001, p. 134) — and that this change



affects others in the environment too (Daniels, 2004). The lecturers in this study had discussed their practices with each other, in mostly informal settings, but arrived at divergent practices. Such divergent practices may have an impact on students should they encounter different lecturers covering the same course over two semesters.

The state of reaching a stable practice does not indicate that there are no contradictions: they “can either enable learning to progress, or they can actually disable it, depending on whether or not they are acknowledged or resolved” (Murphy & Rodriguez-Manzanares, 2008, p. 445). Thus, there may be unknown contradictions, or issues that are simply not acted upon. From an analysis perspective, they are difficult to identify and discuss. However, it does make it important to note that an apparently stable practice does not necessarily mean it is a perfect practice.

L5 provided an example of such a contradiction remaining unresolved. She continued to use CAA with her students despite her reservations towards the quality and helpfulness of the feedback. L4 also felt that the feedback was lacking; he overcame this contradiction by resolving to offer further feedback to students that request it. Although both had developed a regular routine for CAA, L4 had resolved the contradiction in his practice whereas L5 remained doubtful that the system was helping her weaker students.

The extent to which the lecturers made changes to their practice to resolve their contradictions varied widely; L2 was particularly radical among his colleagues by using a paper test instead of the CAA summative test. By designing his own summative assessment tool, he was able to address these contradictions and tailor the questions and feedback according to his own needs. L1, L3 and L5 made minimal changes to their practice: they reduced their dependence on the system for gauging their students’ understanding by opting for other assessment techniques or by lowering the significance of CAA marks in the end-of-module score.

## **CONCLUSION**

The lecturers in this study felt that the CAA system was unreliable as a measure of student performance, knowledge and understanding. While the tests were useful for a number of reasons — encouraging students to practice, reducing workload, and providing immediate feedback, primarily — they could not rely on the marks that the system offered. Those lecturers that could not invigilate the tests accepted that plagiarism could occur. The feedback was not to the standard that the lecturers desired to give to their students. The questions were largely focused on the procedural aspects of the mathematics they were learning.

These contradictions could not be overcome with the given CAA system; instead, the lecturers could only mitigate against them with compromises. They resulted in three consequences.

- The lecturers were able to achieve a stable practice in which the contradictions they had faced were no longer sufficiently burdensome to warrant further change.
- The contradictions indicated the point at which the assessment tool was no longer effective.
- The contradictions identified new motives for assessment, which would extend the scope of effectiveness for the assessment tool; that is, should the CAA system be able to address these new motives, it would become more effective.

The remaining contradictions, albeit mitigated, remain motivations for change within the activity of assessment (Engeström 2001). That is, while the practice may appear to have reached a state of regularity or repetition, there remain instabilities in the activity system and the likelihood of further change. L3 and L6 were aware of other CAA systems that could address the contradictions they encountered and had made enquiries; L4 had attempted to develop the system to better suit his needs; and L1 had recently changed her practice in light of the contradictions she encountered.

## **SOLUTIONS AND RECOMMENDATIONS**

The lecturers were mindful of the aims that the CAA system could not fully address. They were the qualities that were missing from the CAA system, which would have made it more effective. In the case of the system used in this study, adding these facilities to the system would make it more effective. In order to make this system more effective, it would need to offer more personal, detailed feedback; it would have to address conceptual understanding; and summative tests would have to be undertaken under invigilation. Since this study, some module leaders have started to use a different CAA system, which is able to accept algebraic responses, more complex questioning and more detailed, responsive feedback. While these features should address some of the limitations to the effectiveness of the CAA system evaluated in this study, it remains for a similar evaluation to be performed on the new system.

Nonetheless, the lecturers in this study had arrived at their own practices in which they were relatively satisfied and stable. Since it was not practical to make adjustments to the system to address the shortcomings, the lecturers' freedom over their practices ensured that they could maximize the effectiveness of the system and achieve as many of their aims for assessment as possible. The diversity of practices caused some problems for the students encountering different practices within the same module (Broughton et al., 2013), and this study also found that lecturers rarely discussed their practices with their colleagues. Therefore, it is a recommendation of this study to encourage dialogue between lecturers to minimize the impact of diverse practice within and between modules of study.

## **FUTURE RESEARCH DIRECTIONS**

This study provided a snapshot report of the lecturers' use of the CAA system that was in use at the HEI. Since the collection and analysis of these data, the CAA system was changed for some modules. A possible extension for this study is to identify how the introduction of this new system has had an impact on teaching and the extent to which this new system overcomes the shortcomings of its predecessor.

While the model for effective assessment has been used here to identify these shortcomings, it is anticipated that the model could be used to evaluate other assessment tools. Further study could examine other assessments for effectiveness and to identify where improvements might be made to extend the scope of effectiveness. The model might also be used to design an effective assessment; in which case, a further study might examine such an assessment used as an intervention during learning.



## REFERENCES

- Biscomb, K., Devonport, T. J., & Lane, A. M. (2008). Evaluating the use of computer-aided assessment in higher education. *Journal of Hospitality, Leisure, Sport and Tourism Education*, 7(1), 82–88. doi:10.3794/johlste.71.181
- Black, P., & Wiliam, P. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. doi:10.1007/s11092-008-9068-5
- Blanden, J., & Machin, S. (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, 51(2), 230–249. doi:10.1111/j.0036-9292.2004.00304.x
- Broughton, S. J., Hernandez-Martinez, P., & Robinson, C. L. (2012). Using focus groups to investigate the presence of formative feedback in CAA. *Research in Mathematics Education*, 14(1), 87–88. doi:10.1080/14794802.2012.657440
- Broughton, S. J., Hernandez-Martinez, P., & Robinson, C. L. (2013). A definition for effective assessment and implications on computer-aided assessment practice. In A. M. Lindmeier, & A. Heinze (Eds.), *37th Conference of the International Group for the Psychology of Education* (Vol. 2, pp. 113–120). Norwich, UK: University of East Anglia.
- Broughton, S. J., Robinson, C. L., & Hernandez-Martinez, P. (2013). Lecturers perspectives on the use of a mathematics-based computer-aided assessment system. *Teaching Mathematics and Its Applications*, 32(2), 88–94. doi:10.1093/teamat/hrt008
- Bull, J., & McKenna, C. (2003). *Blueprint for Computer-Assisted Assessment*. London: Routledge.
- Daniels, H. (2004). Cultural historical activity theory and professional learning. *International Journal of Disability Development and Education*, 51(2), 185–200. doi:10.1080/10349120410001687391
- Engeström, Y. (1987). The emergence of learning activity as a historical form of human learning. In *Learning by Expanding*. Helsinki: Orienta-Konsultit Oy.
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamäki (Eds.), *Perspectives on Activity Theory* (pp. 19–38). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511812774.003
- Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, 43(7), 960–974. doi:10.1080/001401300409143 PMID:10929830
- Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133–156. doi:10.1080/13639080020028747
- Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*, 5(1), 1–24. doi:10.1016/j.edurev.2009.12.002
- Green, D. R., Harrison, M. C., Palipana, A. S., Pidcock, D., & Ward, J. P. (2004). *The role of CAA in helping engineering undergraduates learn mathematics*. Maths-CAA Series.

- Greenhow, M., & Gill, M. (2005). *Setting objective tests in mathematics using QM Perception*. Maths-CAA Series.
- Hunt, R. (2006). Institutional constraints on authenticity in teaching. *New Directions for Adult and Continuing Education*, 2006(111), 51–62. doi:10.1002/ace.227
- Leont'ev, A. N. (1977). *Activity and Consciousness*. Retrieved from <http://www.marxists.org/archive/leontev/works/activity-consciousness.pdf>
- McKenna, C., & Bull, J. (2000). Quality assurance of computer-assisted assessment: Practical and strategic issues. *Quality Assurance in Education*, 8(1), 24–32. doi:10.1108/09684880010312659
- Murphy, E., & Rodriguez-Manzanares, M. A. (2008). Using activity theory and its principle of contradictions to guide research in educational technology. *Australasian Journal of Educational Technology*, 24(4), 442–457. doi:10.14742/ajet.1203
- Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. doi:10.1080/03075070600572090
- Robinson, C. L., Hernandez-Martinez, P., & Broughton, S. J. (2012). Mathematics lecturers' practice and perception of computer-aided assessment. In P. Iannone & P. Simpson (Eds.), *Mapping University Mathematics Assessment Practices* (pp. 105–118). Norwich, UK: University of East Anglia.
- Rolfe, H. (2002). Students demands and expectations in an age of reduced financial support: The perspectives of lecturers in four English universities. *Journal of Higher Education Policy and Management*, 24(2), 171–182. doi:10.1080/1360080022000013491
- Samuelowicz, K., & Bain, J. D. (2002). Identifying academics orientations to assessment practice. *Higher Education*, 43(2), 173–201. doi:10.1023/A:1013796916022
- Sangwin, C. J. (2003). New opportunities for encouraging higher level mathematical learning by creative use of emerging computer aided assessment. *International Journal of Mathematical Education in Science and Technology*, 34(6), 813–829. doi:10.1080/00207390310001595474
- Sangwin, C. J. (2007). STACK: making many fine judgements rapidly. CAME 2007 — The Fifth CAME Symposium. Pécs, Hungary: University of Pécs. Retrieved from <http://www.lkl.ac.uk/research/came/events/CAME5/CAME5-Theme3-Sangwin.pdf>
- Tikhomirov, O. K. (1988). *The Psychology of Thinking*. Moscow: Progress.

## KEY TERMS AND DEFINITIONS

**Activity:** In cultural-historical activity theory (CHAT), an activity is a unit of analysis, in which an actor (the subject) undertakes actions directed towards achieving the object.

**Computer-Aided Assessment:** A test that is taken via a computer terminal, usually conducted via the Internet.

**Contradiction:** In CHAT, a contradiction occurs within or between nodes and activity systems. They are conflicts that destabilize an activity system and inspire changes to resolve the conflict.

**Formative Assessment:** Although there continue to be discussions and disagreements in the literature, formative assessment is usually taken to mean a zero-stakes assessment in which students receive feedback.

**Object:** The object of the activity in CHAT is the overarching aim of an activity, which comes from a subject's need.

**Subject:** The subject of the activity is the primary actor in an activity system. It is the subject that is striving to achieve the object of the activity.

**Summative Assessment:** Summative assessments are tests taken by a student to evaluate performance, knowledge or understanding. They are usually taken at the end of a unit of learning.

**Tool:** Tools are socially created artefacts from an activity that are used in other activities. For example, the CAA system is a tool that has been created by lecturers and developers, which is now used by lecturers to assess their students.