Using e-Assessment to Address Mathematical Misconceptions in Engineering Students

Indunil Sikurajapathi, Karen Henderson, and Rhys Gwynllyw

***Abstract*—Students, when answering a mathematical question, may make a mistake in their answer for a variety of reasons. For example, not reading the question properly, making a mistake due to carelessness or due to a mathematical misconception. It is this latter category, which is of particular interest to us in this paper. When such mistakes occur in handwritten work then, in general, the teacher is able to identify the mistake(s) during the marking process and give written detailed feedback on the student's script. The disadvantage of this approach is the time and effort it takes to mark and to get feedback back to the student. As a result, e-assessment is becoming a standard means of providing formative and summative assessment of mathematical techniques. The research problem that we have identified is how to detect mathematical misconceptions when students answer e-assessment questions incorrectly, and how to improve the feedback provided to the student in such cases. By analyzing students' rough paper-based workings for an e-examination, we have captured mathematical misconceptions made by first year engineering students. This has enabled us to catalogue common student errors made by students. By amending the e-assessment feedback code, students who make these errors will subsequently benefit from enhanced, tailored feedback, highlighting the mathematical misconception/error made. In addition, detailed guidance on how to improve their knowledge related to the topic will be given. The aim of our work is to improve the e-assessment experience for students as well as addressing and tackling misconceptions in a timely fashion.**

***Index Terms*—Common student errors, Dewis, e-assessment, engineering mathematics.**

I. Introduction

An understandable but incorrect implementation of a process resulting from a student's misconception is called a mal-rule [1]. Mal-rules can be classified as manipulative, parsing, execution/clerical and random [2]. In this paper, we focus on mal-rules or, in other words, common student errors (CSEs) in Engineering Mathematics - a subject in which students tend to make CSEs due to misconceptions in mathematics. For example, a typical CSE students make is to answer a^2+b^2 when asked to expand $(a+b)^2$. Booth [3] states that "Students hold many misconceptions as they transition from arithmetic to algebraic thinking, and these misconceptions can hinder their performance and learning in the subject." This is particularly the case in Engineering, which is a subject that requires a strong mathematics foundation.

Mathematics Education research; see for example [4]-[6], has explored possible causes and effects of certain mathematical misconceptions and the impact that they have on students' future learning. As an example, there has been recent research into theorizing student errors supported by empirical studies in the topics of natural number bias [7], visual saliency [8] and over-generalization [9]. More recently, Rushton [10] conducted a study of common errors in Mathematics made in certain General Certificate of Secondary Education mathematics papers taken by candidates in England, including an internationally available version, as referenced by examiner reports, and errors were catalogued into themes and sub-themes. Khiat [11] looked specifically at the mathematics learning of engineering students at undergraduate level and the focus of the work was on conceptions of understanding using grounded theory methodology.

E-assessment has become a standard method to provide formative and summative assessments in many universities all around the world [12]. A few of the advantages of e-assessment are that it can provide instant tailored feedback to help students to improve their knowledge and performance, they can access it in different geographical locations at different times, and undertake online tests many times to assess and refine their knowledge. Moreover, it allows educators to identify areas in which more help is needed and then to take necessary action to address difficult areas in the subject. Research has found that students learn from e-assessment feedback and enhance their technical knowledge by using it [13]. Therefore, e-assessments that provide effective feedback and select questions based on pedagogic principles should be promoted as a learning resource [14].

Research [15] shows that feedback has to be quick to be effective, while students still remember clearly the work they were engaged in and using e-assessment is one way of achieving this. A computer cannot act flexibly like a human marker when faced with ill-posed or unanticipated student responses [14]. However, if an e-assessment system could detect and report CSEs, it would behave more like a human marker and provide very effective and tailored feedback instantly for the students by pointing out their mal-rule [16]. Providing such tailored feedback will help students to learn from their misconceptions. The CalculEng system [17] has been developed to address this need for Calculus based problems that engineering students encounter, but the development still requires expert teachers with mathematical knowledge to anticipate the errors that students might make.

In this paper, we demonstrate how we have built up a

Manuscript received July 14, 2019; revised March 10, 2020.

The authors are with the Department of Engineering Design and Mathematics, University of the West of England, Bristol, United Kingdom (e-mail: indunil.sikurajapathi@uwe.ac.uk, karen.henderson@uwe.ac.uk, rhys@gwynllyw.uwe.ac.uk).

doi: 10.18178/ijiet.2020.10.5.1389

collection of CSEs made by Level 1 engineering students in their mathematics module and give an overview of our findings to date. We have achieved this by scrutinizing students' answers to e-assessment questions and by looking at their rough workings to a mid-module e-examination.

II. Methodology

A. Module Overview

Engineering Mathematics (EM) is a 30-credit module making up a quarter of the credit for Level 1 and is delivered to a large and diverse student cohort at the University of the West of England, Bristol (UWE). Students learn mathematical techniques that will support their engineering studies, including learning to program in Matlab. As well as the Matlab weekly PC sessions in Semester 1, all students receive two hours of lectures, supported by a one-hour tutorial each week. In addition, all students have a scheduled weekly two-hour Peer Assisted Learning (PAL) session [18] run by Level 2 PAL tutors and which offer whole course support, not just help with EM. The module is assessed through coursework (25%) and examination (75%). The coursework is designed to encourage engagement in the module. The Matlab assignment comprises 50% of the coursework mark, whilst e-assessments delivered throughout the year comprise the remaining 50% coursework mark. Further details of the e-assessment system used is given in Section II B and the e-assessment implementation is expanded on in Section II C.

B. The Dewis e-Assessment System

The Dewis system was used to deliver all the e-assessments on this module [19], [20]. Dewis is a fully algorithmic open-source web-based e-assessment system that was designed and developed at UWE. It was primarily designed for the assessment of mathematics and statistics and supports a range of inputs, such as numeric entry, algebraic entry, matrix entry, computer programs, multiple choice and multiple selection. Using an algorithmic approach enables the separate solution, marking and feedback algorithms to respond dynamically to a student's input. The question parameters are randomized and generated at the point of delivery; therefore, no two students receive exactly the same question. Students can practice the same question several times with different parameters in order to gain mastery. All Dewis questions have full feedback bespoke to that question and its specific randomly generated parameters. The feedback not only supplies the correct answer but a fully worked solution showing how that the correct answer was obtained. An example of an e-assessment question used for EM is illustrated in Fig. 1 together with the full feedback received.

All data relating to every assessment attempt is recorded on the Dewis server. This enables the academic to track efficiently how a student or cohort of students has performed on a particular e-assessment [16]. The highly developed reporting system enables tracking at module cohort level, tutorial group level and individual student level. Fig. 2 shows a reporting session for a particular e-assessment, in this case viewing the mark awarded for each individual question in the test. Each mark is a web link, which contains the realization of a particular question as delivered to that student, the student's answer and the resulting feedback given to them.

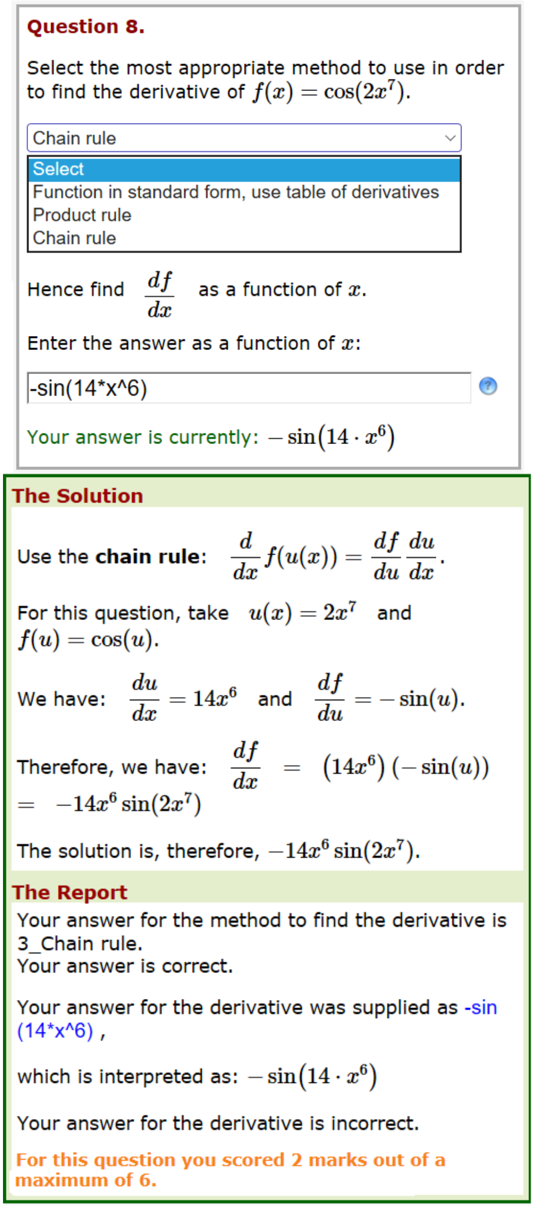
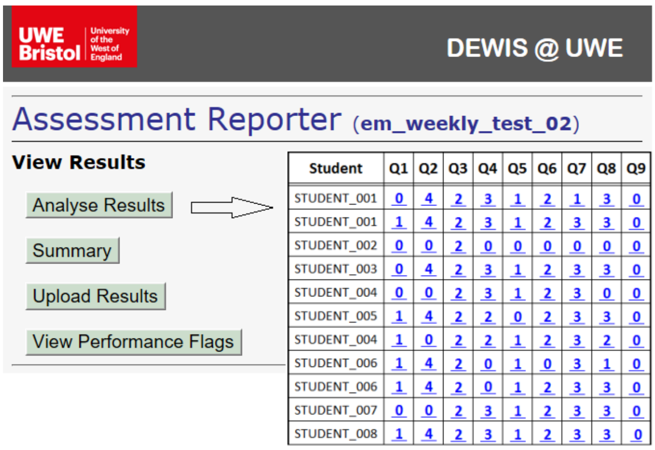


Fig. 1. An example Dewis question, together with feedback and marking bespoke to the random parameters used in this question. This question illustrates partial marking; the method of solution is correct, but the implementation was not.



Student	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
STUDENT_001	0	4	2	3	1	2	1	3	0
STUDENT_001	1	4	2	3	1	2	3	3	0
STUDENT_002	0	0	2	0	0	0	0	0	0
STUDENT_003	0	4	2	3	1	2	3	3	0
STUDENT_004	0	0	2	3	1	2	3	0	0
STUDENT_005	1	4	2	2	0	2	3	3	0
STUDENT_004	1	0	2	2	1	2	3	2	0
STUDENT_006	1	4	2	0	1	0	3	1	0
STUDENT_006	1	4	2	0	1	2	3	3	0
STUDENT_007	0	0	2	3	1	2	3	3	0
STUDENT_008	1	4	2	3	1	2	3	3	0

Fig. 2. The assessment reporter. (Student details have been anonymized.)

C. E-Assessment Delivery Implementation

We have used e-assessment at UWE since 2000 and migrated to the Dewis e-assessment system in EM in 2009. Over that time, we have built a substantial library of Dewis questions to support the teaching of engineering mathematics. The question library resource has enabled us to try out different delivery patterns of e-assessment in order to improve year-long student engagement with the module and hence improve attainment levels.

Since the 2015/2016 academic year, the module has used 22 weekly e-assessments and students are given access to these e-assessments throughout the year and are allowed unlimited attempts. The e-assessment coursework mark is calculated from the top 20 marks from the 22 weekly tests (twelve tests in Semester 1 and ten in Semester 2). All weekly tests are open from the start of the module. Each test can contribute two marks to the coursework mark, comprising one engagement mark and one attainment mark.

At the end of the first semester, students are required to take a two-hour e-examination, sat under controlled conditions and questions on this e-examination are based on the questions students have already encountered in their weekly e-assessments [21]. Due to the lack of available computers, this January e-examination was delivered in two sessions. Approximately half the students were timetabled for the morning session and the other half for the afternoon. For each separate run of the e-examination, we fixed the parameters of the questions in order to ensure fairness. This approach also meant that, at the start of the exam, students were given a hardcopy of the specific questions that they were attempting. Students valued this, as some found it easier to work from a paper copy than from the screen. In this paper, we have focused on the January 2018 e-examination. Each version (morning and afternoon) contained 19 questions. Both exam versions contained a mixture of input types: numerical, algebraic and dropdown. The question structure and subject content were the same for both papers but different numeric parameters were used in each case to make the two tests different but of comparable difficulty. A total of 298 students sat the e-examination, 148 in the morning and 150 in the afternoon. The official submission was electronic but students were given exam booklets in order to write their rough workings to questions and these booklets were collected at the end of the e-examination.

D. Detection of Common Student Errors

In terms of detecting CSEs, it was natural to start by analyzing the submissions from the January e-examination. This was because all morning/afternoon students sat the same paper and provided written solutions in booklets, as well as submitting their final answer electronically. We examined these written answer scripts along with the corresponding Dewis answers for all instances in which the students had inputted an incorrect answer in the e-examination. Firstly, the Dewis Reporter output was used to select the most common incorrect answers to each question. Secondly, the written answer scripts of the students who inputted the same mistake were carefully examined. The aim of this process was to understand what kind of mistake had led the students to arrive at that common wrong answer. Having access to the students' workings was invaluable for this process.

III. Results

We analyzed all 19 questions from both the morning and afternoon versions of the e-examination and 17 questions were found to exhibit CSEs. We found several of the questions to have more than one CSE associated with them and we catalogued 40 CSEs in total.

For each question, we designated the principal CSE to be the one that was triggered by the largest proportion of students. This quantity was measured as a percentage of the number of students who made the CSE compared to the total number of students who answered that question incorrectly. The results of these principal CSEs are illustrated in Fig. 3. In this chart, the height of the rectangle represents the number of students, aggregated over the two sittings, who answered the question incorrectly whilst the height of the shaded rectangle represents the number of students who triggered the principal CSE for that question. Please note that there is no shaded box for Questions 1 and 4 because no CSE was found for either question. We can see that Question 14 was the least well-answered question (197 incorrect responses) and the principal CSE for this question was triggered by 34% of students. Question 5 was the one for which the least number of incorrect responses were submitted (13 in total) and the principal CSE for that question was triggered by 38% of students. The principal CSE that was triggered by the largest proportion of students occurred for Question 12, namely 70%.

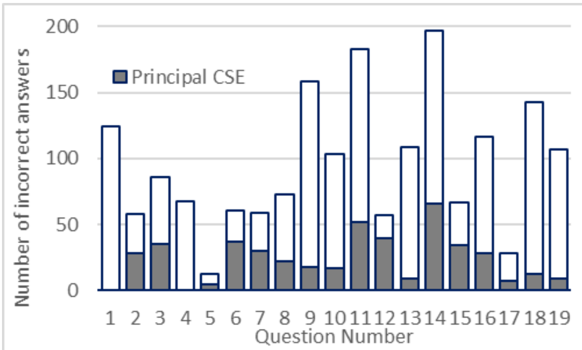


Fig. 3. Total number of incorrect responses to the e-examination questions together with the number of these that are attributable to the principal CSE found (shaded box).

In this paper, we illustrate in detail the principal CSEs found in three questions, (namely questions 3, 11 and 7 of the e-examination) and further details of these are shown in Sections III A-C.

A. CSE Example 1

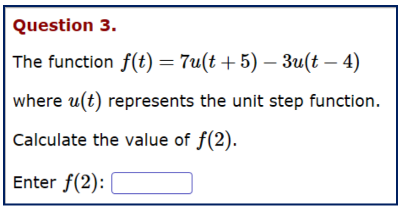


Fig. 4. Question 3 from the morning e-examination.

Fig. 4 shows the morning version of question 3 from the 2018 e-examination paper. This question required students to input a single integer answer. The afternoon paper contained a similar question but with different parameters. We only detected one CSE for this question and it involved students' misunderstanding of the unit step function. Instead of treating $u(t)$ as a function, the detected CSE involved students setting u to take the value of one and misinterpreting the purpose of the brackets. Hence, the student incorrectly evaluated $f(2)$ as $7(2+5)-3(2-4)=55$ whilst the correct answer is $f(2)=7u(7)-3u(-2)=7$. In the morning version, 12 students, out of the 44 who answered this question incorrectly (27%), triggered this CSE whilst in the afternoon 23 from 42 (55%) did. This resulted in an aggregate of 35 students from 86 (41%) making this mistake as confirmed in Fig. 3.

B. CSE Example 2

Fig. 5 shows the morning version of question 11 from the 2018 e-examination paper. This question required students to input a single floating-point answer. The afternoon paper contained a similar question but with different parameters. We detected three CSEs for this question. The principal CSE involved students performing the integration step correctly but incorrectly using the calculator in degree mode when evaluating the antiderivative of the integrand at the two limits. In the morning version, 22 students, out of the 86 who answered this question incorrectly (26%), triggered this CSE whilst in the afternoon 30 from 97 (31%) did. This resulted in an aggregate of 52 students from 183 (41%) making this mistake as confirmed in Fig. 3. The second CSE, which was triggered by 32 students (17%), involved them directly substituting the midpoint of the range of t into the integrand, $4\cos(3t)$, in the morning version of the paper, and using the calculator in degree mode to evaluate their answer. The third CSE, which eight students (4%) triggered, involved them taking the average of the integrand evaluated at the integration limits and using the calculator in degree mode to evaluate their answer.

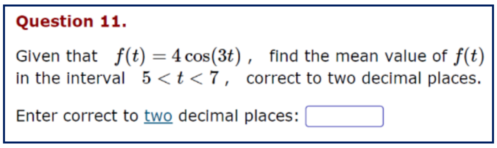


Fig. 5. Question 11 from the morning e-examination.

C. CSE Example 3

Fig. 6 shows the morning version of question 7 from the 2018 e-examination paper. This question required students to input a single answer in algebraic form. The afternoon paper contained a similar question but with different parameters. Only one CSE was detected for this question and involved students' incorrectly differentiating $\ln(ax)$ as $(ax)^{-1}$. So students making this mistake incorrectly inputted $3/(5x)$ instead of $3/x$ as their answer. In the morning version, 15 students, out of the 27 who answered this question incorrectly (56%), triggered this CSE whilst in the afternoon 15 from 32 (47%) did. This resulted in an aggregate of 30 students from 59 (51%) making this mistake as confirmed in Fig. 3.

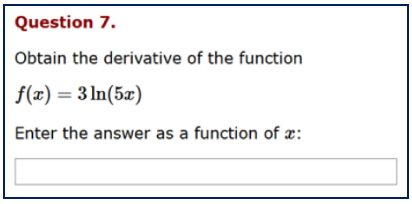


Fig. 6. Question 7 from the morning e-examination.

D. Weekly e-Assessments

All the questions from the 2018 January e-examination had been included in one of the 12 weekly first semester e-assessments taken by the same cohort in the 2017/2018 academic year. For each of the 17 questions on the e-examination paper, for which CSEs were found, we have subsequently altered the e-assessment question to consider each particular CSE. Dewis uses Performance Indicators (PIs) in the Reporter that enable the academic to view the performance of a student on each question attempt [16]. This is particularly useful in order to differentiate between a student scoring zero by not answering the question or by answering the question incorrectly. Additional PIs have been introduced into the altered question code to capture CSEs when they are triggered.

Using the re-mark feature in Dewis [16], academics are able to re-mark e-assessments using the altered question source code. By re-marking the weekly e-assessments with the new question source code, the additional PIs can identify if students made any CSEs in a particular e-assessment, prior to them taking the e-examination. For the three CSEs illustrated in this paper, we found the results as shown in Table I.

We can see that for all three questions the principal CSE percentage in the e-assessment was less than occurred in the e-examination. A possible explanation for it being lower is that students typically attempt the weekly tests with fresh knowledge, that is, soon after or while they are learning the new concept. It could also be due to students being under more pressure in the e-examination due to it being a high-stakes assessment and sat under controlled conditions.

TABLE I: RESULTS OF THE PRINCIPAL CSES FROM RE-MARKING THE WEEKLY E-ASSESSMENTS WHICH INCLUDED QUESTIONS 3, 7, 11 FROM THE E-EXAMINATION

	Principal CSE Occurences	Incorrect Answers	Principal CSE Percentage
Question 3	85	335	25%
Question 7	25	123	20%
Question 11	73	324	23%

IV. DISCUSSION

Re-marking the weekly e-assessments with the CSE software capture included raised some interesting points. Firstly, we found that for some questions, there were particular random parameters for which the correct answer and the CSE answer were the same. This occurs for example, for Question type 3 (as shown in Fig. 4) for the function $f(t)=2u(t+7)-5u(t+1)$ when the value of $f(4)$ is asked for. In this case, both the correct answer and the CSE answer are equal to -3. Therefore, if a student is presented with this realization of the question and entered -3, Dewis would mark them as having answered the question correctly but it could be that the student had erroneously arrived at that answer by performing a CSE instead. This finding shows the importance of awareness of CSEs related to a problem when coding an e-assessment question. In order to mitigate against such scenarios, the random parameters for the question should be selected such that the correct answer differs from the CSE answer(s).

Secondly, we found that it is possible for more than one CSE to be triggered for some questions. This occurred for the question type detailed in Section III B. In the question presented in the morning version of the e-examination (as illustrated in Fig. 5) the second and third CSEs described in that section result in the same incorrect value, namely 3.80. During the CSE collection process, it was straightforward to determine which CSE students had made by examining their written scripts. However, for instances when the same phenomenon occurs in the weekly e-assessments (when no intermediate workings are available) it is not clear how to decide which CSE led the student to obtain that incorrect answer. Again, this finding shows the importance of awareness of CSEs when selecting parameters for a question. Further, when coding a question, if it is difficult to avoid parameters which trigger several CSEs, careful decisions need to be made when providing enhanced feedback.

V. Conclusion

Having catalogued the 40 CSEs found on the 2018 e-examination and introduced Performance Indicators to capture them in the 17 e-assessment questions the next steps are to create detailed feedback based on students' answers. Thus, for future uses of the questions, if one of the pre-coded CSEs is triggered, the student will be provided with information about what could have gone wrong in their calculation together with extra supportive resources for them to work through. In order to assess the impact of the improved feedback, the enhanced e-assessment questions will be integrated into the weekly e-assessments for the Engineering Mathematics module from the 2019/20 academic year. Using the Dewis Reporter, we will be able to see which students triggered CSEs and hence received the enhanced feedback. These students will be asked to fill in a short questionnaire giving information on the tailored feedback that they received and this will allow us to improve the feedback given. The first semester weekly e-Assessments contain over 100 questions in total. As we build a taxonomy of CSEs, our goal is to enhance a significant proportion of these questions. Through the generation of this additional personalized feedback, our aim is to improve the e-assessment experience for students as well as addressing and tackling misconceptions in a timely fashion.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Indunil Sikurajapathi conducted the research, analysed the data and wrote the first draft of the paper. Karen Henderson took responsibility for the final version of the paper. Karen Henderson and Rhys Gwynllyw organised and supervised the research; all authors discussed progress and results of the research and approved the final version.

References

[1] R. Rees and G. Barr, *Diagnosis and Prescription in the Classroom: Some Common Maths Problems*, Harper & Row, London, UK, 1984.

[2] D. Sleeman, "Mis-generalization: An explanation of observed mal-rules," presented at the Annual Conference of the Cognitive Science Society, 1984.

[3] J. L. Booth, C. Barbieri, F. Eyer, and E. J. Pare-Blagoev, "Persistent and pernicious errors in algebraic problem solving," *Journal of Problem Solving*, vol. 7, pp. 10-23, 2014.

[4] P. Nesher, "Towards an instructional theory: The role of student's misconceptions," *For the Learning of Mathematics*, vol. 7, pp. 33-40, 1987.

[5] E. Fischbein, "Tacit models and mathematical reasoning," *For the Learning of Mathematics*, vol. 9, pp. 9-14, 1989.

[6] J. Confrey, "A review of the research on student conceptions in mathematics, science, and programming," *Review of Research in Education*, vol. 16, pp. 3-56, 1990.

[7] A. Obersteiner, W. Dooren, J. Hoof, and L. Verschaffel, "The natural number bias and magnitude representation in fraction comparison by expert mathematicians," *Learning and Instruction*, vol. 28, pp. 64-72, 2013.

[8] D. Kirshner and T. Awtry, "Visual salience of algebraic transformations," *Journal for Research in Mathematics Education*, vol. 35, pp. 224-257, 2004.

[9] E. Knuth, A. Stephens, N. McNeil, and M. Alibali, "Does understanding the equal sign matter? Evidence from solving equations," *Journal for Research in Mathematics Education*, vol. 37, pp. 297-312, 2006.

[10] N. Rushton, "Common errors in mathematics," *Research Matters: A Cambridge Assessment publication*, vol. 17, pp. 8-17, 2014.

[11] H. Khiat, "A grounded theory approach: Conceptions of understanding in engineering mathematics learning," *The Qualitative Report*, vol. 15, pp. 1459-1488, 2010.

[12] C. Sangwin, *Computer Aided Assessment of Mathematics*, Oxford University Press, Oxford, UK, 2013.

[13] M. Gill and M. Greenhow, "How effective is feedback in computer-aided assessments?" *Learning, Media and Technology*, vol. 33, pp. 207-220, 2008.

[14] M. Greenhow, "Effective computer-aided assessment of mathematics; Principles, practice and results," *Teaching Mathematics and Its Applications*, vol. 34, pp. 117-137, 2015.

[15] P. Race, *Making Learning Happen*, Sage Publications, 2014.

[16] P. Walker, R. Gwynllyw, and K. Henderson, "Diagnosing student errors in e-assessment questions," *Teaching mathematics and its Applications*, vol. 34, pp. 160-170, 2015.

[17] M. Davis, G. Hunter, L. Thalaal, V. Ba, and A. Wooding-Olajorin, "Developing smart tutorial tools to assist students learn calculus, taking account of their changing preferred approaches to learning," in *Proc. International Conference on Intelligent Environments Ambient Intelligence and Smart Environments*, ed. A. Muñoz et al., 2019, pp. 227-238.

[18] N. Falchikov, *Learning Together: Peer Tutoring in Higher Education*, Routledge, London, UK, 2001.

[19] R. Gwynllyw and K. Henderson, "Intelligent marking in summative e-assessment," in *Proc. the HEA STEM Learning and Teaching Conference*, 2012.

[20] R. Gwynllyw and K. Henderson, "A computer aided assessment system for mathematics and statistics," in *Proc. the CETL-MSOR 2008 Conference*, 2009, pp. 38-44.

[21] K. Henderson, R. Gwynllyw, and A. Hooper, "Using electronic exams to provide engineering mathematics students with rapid feedback," in *Proc. the 18th SEFI Mathematics Working Group Seminar on Mathematics in Engineering Education*, 2016, pp. 105-111.

Indunil Sikurajapathi is a graduate tutor and PhD candidate in mathematics in the Department of Engineering Design and Mathematics at UWE Bristol, United Kingdom. Her research is concerned with mathematical education and computing focusing on the use of and the development of Dewis, UWE's in-house algorithmic e-assessment system.

Karen Henderson is an associate professor in technology enhanced learning in the Department of Engineering Design and Mathematics at UWE Bristol, United Kingdom. Her research is focused on supporting students' learning through the use of technology.

Rhys Gwynllyw is an associate professor in learning and teaching in the Department of Engineering Design and Mathematics at UWE Bristol, United Kingdom. He is the principal developer of Dewis. His research has concentrated on the benefits of e-assessments to the learning and teaching process.