RESEARCH ARTICLE

Correct for the wrong reason: why we should know more about Mathematical Common Student Errors in e-Assessment questions

Indunil Sikurajapathi, School of Computing and Creative Technologies, University of the West of England, Bristol, UK. Email: indunil.sikurajapathi@uwe.ac.uk
Karen L. Henderson, School of Computing and Creative Technologies, University of the West of England, Bristol, UK. Email: karen.henderson@uwe.ac.uk
D. Rhys Gwynllyw, School of Computing and Creative Technologies, University of the West of England, Bristol, UK. Email: rhys.gwynllyw@uwe.ac.uk

Abstract

Students may arrive at an incorrect answer when answering a mathematical question due to several reasons, such as random errors, calculation errors or misreading the question. Such errors are sometimes referred to as Common Student Errors (CSEs). This article explains why it is important to know more about Mathematical CSEs in e-Assessment questions, using several examples encountered while conducting the CSE Project at the University of the West of England (UWE Bristol). The CSE Project at UWE Bristol began with an aim of developing a technique to detect CSEs and provide tailored feedback in e-Assessment questions delivered via Dewis, UWE Bristol's in-house e-Assessment system. In this research article, we present one important finding of this project that is related to the parameter selection(s) of e-Assessment questions which have at least one CSE. We highlight why, in this digital era, it is more vital than ever to know more about mathematical CSEs.

Keywords: Mathematical Common Student Errors, Dewis e-Assessment system, e-Assessment Parameters

1. Introduction and Background

Students may make a mistake when answering a mathematical question for a variety of reasons. For example making a mistake in their calculation, misconceptions or misreading the question. When the same error is made by several students, those errors are sometimes referred to as common errors (Rushton, 2014).

Several different terms are used in the literature to refer to either mathematical errors or misconceptions. VanLehn (1982) use the term '*bug*' to refer to a systematic error resulting from wrong steps in the calculation procedure. The term, '*mal-rule*' is used by Rees and Barr (1984) to refer to an understandable but incorrect implementation of a process resulting from a student's misconception. For example, a classic *mal-rule* students make is to answer $a^2 + b^2$ when asked to expand $(a + b)^2$. In this article we use the term Common Student Error (CSE) to refer to an error made by several students.

This article is concerned with CSEs in e-Assessments. Assessment is a key element of teaching and learning and is used widely in higher education. It enables educators to assess the extent of students' skill and knowledge and to ascertain whether students have achieved the desired learning outcomes (Stödberg, 2012). Assessments also give students the opportunity to receive feedback on their work. Race (2014) suggests that, in order for feedback to be effective, it should be available while students still remember clearly the work they were engaged in. Using e-Assessments is one way of achieving this. A comprehensive review of the advantages of e-Assessment to the student, teacher, institution and education aims can be found in Alruwais et al (2018).

The use of E-Assessment for the formative and summative assessment of procedural mathematical techniques has become standard practice in many UK higher education institutions (Sangwin, 2013). Several e-Assessment systems allow the creation of equivalent but different assessments through the use of random variables. One disadvantage is that, typically, in answering an e-Assessment question the student does not enter their intermediate workings, as would be the case for a paper-based assessments. This, together with the fact that each student takes an equivalent but different assessment, makes detecting CSEs in e-Assessment questions harder than for traditional paper-based submissions.

A technique for detecting CSEs and providing tailored feedback in e-Assessment questions has been developed for several Dewis e-Assessment questions used in a first year Engineering Mathematics module (Sikurajapathi, Henderson and Gwynllyw, 2020; Sikurajapathi, Henderson and Gwynllyw, 2021; Sikurajapathi, Henderson and Gwynllyw, 2022a; Sikurajapathi, Henderson and Gwynllyw, 2022b). This research forms part of The CSE Project at UWE Bristol (2019) and further details of the methodology used can be found in the next section.

2. Methodology

2.1 CSE data collection

For the work presented in this article we use Dewis as the e-Assessment system and a first year Engineering Mathematics (EM) Module for the data collection. Dewis (2012) is well-established, was developed at UWE Bristol by a team of mathematicians, statistics and software engineers and uses an algorithmic approach to question generation, marking and feedback. Dewis is lossless, this means that the data for every assessment attempt is recorded and stored on the Dewis server (Gwynllyw and Henderson, 2009). The EM module has used Dewis to deliver e-Assessments since 2009 and as such a huge amount of e-Assessment data is available. This, together with the fact that between 2017 and 2020 the assessment of the EM module included a controlled conditions e-examination were two of the reasons it was selected for the collection of CSEs.

The e-Assessment profile for the mathematical techniques learnt in EM, for the period of interest for the CSE Project, is as follows:

- 22 weekly e-Assessments available throughout the year, with students being allowed unlimited attempts. The e-Assessment coursework mark was calculated from the top 20 marks from these 22 weekly tests;
- A two-hour mid-module e-examination, sat under controlled conditions in January. All of the questions in this e-examination were based on questions students had already encountered in their weekly e-Assessments;
- Formative revision e-Assessments, made available to students a few weeks before the e-examination. Students were allowed unlimited attempts.

Due to a lack of computer rooms, each January e-examination was delivered to a morning and afternoon cohort of students. For each cohort, the parameters of the e-examination questions were fixed, so each cohort sat the same test. Although the official submission was via Dewis, each student was given an examination booklet for their rough workings and these were collected at the end of each e-examination.

A total of 298 and 321 students sat the January e-examination in 2018 and 2019 respectively. Output from the Dewis Reporter was scrutinised in order to select the most common incorrect answers (MCIAs) to each question on the 2018 and 2019 January e-examinations. Once the MCIAs were identified, the rough workings booklets of those students who submitted each of the MCIAs were carefully examined. Having access to the students' workings allowed us to work out what mistake(s) had been made by students resulting in each MCIA.

For each MCIA, the CSE percentage is calculated as follows:

$$CSE\ percentage = \frac{Number\ of\ CSE\ answers}{Number\ of\ incorrect\ answers}\ \%$$

If the CSE percentage is 4% or more, then that MCIA is considered as a CSE in this study.

Through this process, a bank of CSEs has been found and further details of the data collection process and results can be found in Sikurajapathi et al. (2020). Furthermore, this collection of CSEs has been taxonomically classified by Sikurajapathi et al. (2022a) using the taxonomy coding described in Ford et al. (2018) as a guideline.

2.2. CSE capture

In Dewis, the marking of each e-Assessment question, populates performance indicators (PIs). These PIs contain information on how a student has answered a question and are used to allocate marks, report outcomes and provide feedback. For example, for a question that requires one integer input, the three possible PI values would be 1 (correct), 0 (incorrect) and -1 (not answered). In order to capture the identified CSEs within Dewis, each e-Assessment question was amended and an additional PI was introduced, typically taking the value of 1 if the CSE was triggered and 0 if not. This not only allowed Dewis to provide enhanced feedback to the student to address the potential CSE (Sikurajapathi et al., 2021) but also allows the academic, through the Dewis Reporter, to identify all of the students in a cohort that made that CSE.

Since the data for every assessment attempt is recorded and stored on the Dewis server, it is possible to re-mark an assessment, for example, using an amended marking or feedback algorithm for one or more questions in that assessment. The amended CSE capture code for each question was validated by re-marking the e-examinations for the 2017-2018 cohort. This was done by checking that the additional PIs were populated for those students who had already been identified as making CSEs on the e-examination. Once this process had been completed satisfactorily, the weekly e-Assessments and revision tests were also re-marked, using the amended question code. In this research article, we present one important finding from this process, which is related to the parameter selections of e-Assessment questions which have at least one CSE. Details of the prevalence of CSEs made by EM students in e-examinations is available from Sikurajapathi et al. (2022a).

3. Results

During the re-marking of the weekly assessments for the 2017-2018 cohort, some restrictions related to the random parameter selections of the questions which have CSEs were found. Specifically, for some questions, there were particular parameters for which the correct answer and the CSE answer were the same. In these cases, in the marking of the e-Assessment, some students may have been awarded full marks and hence thought that they had answered the question correctly when in fact they had made a CSE.

In this section, several cases in which this happened are presented. For each case, we present a generic form of the question, an example of the parameter selections that lead to the correct and CSE answer being the same, the correct method of solution and the CSE. We use tilde (~) on the CSE answer to differentiate it from the correct answer.

3.1. Case 1

An instance of the first question considered is shown in Figure 3, which requires the student to find the value of the difference between two Unit Step functions at a given point (It should be noted that, the Unit Step function, $u(t)$ is equal to 1 for $t \geq 0$ and 0 for $t < 0$). The generic form of this question involves the function, $f(t) = a\,u(t+b) - c\,u(t+d)$ and the value of $f(p)$ is asked for, where parameters a, b, c, d, p are all integers, and created randomly for each instance of the question.

> The function $f(t) = 2u(t+4) - u(t+2)$
>
> where $u(t)$ represents the unit step function.
>
> Calculate the value of $f(0)$.
>
> Enter $f(0)$:

Figure 3: An instance of a question on the difference of two Unit Step functions

One CSE has been identified related to this question. This CSE occurs by assuming that the unit step function, u, is equal to 1 and is not a function. Whilst re-marking the weekly tests, it was noted that for some parameter values, the correct answer and the CSE answer for this question were the same. This occurs for example, when $a = 2,\ b = 7,\ c = 5,\ d = 1$ and $p = 4$. For this particular parameter selection, the correct answer and the CSE answer can be calculated as shown in Figure 4 and both are equal to -3.

Correct Answer	**CSE Answer** ***CSE:*** *taking the unit step function, u, to be equal to 1 and not a function.*
$f(4) = 2u(4+7) - 5u(4+1)$ $= 2u(11) - 5u(5)$ $= 2 \times 1 - 5 \times 1$ $= -3$	$\tilde{f}(4) = 2u(4+7) - 5u(4+1)$ $= 2u(11) - 5u(5)$ $= 2 \times u \times 11 - 5 \times u \times 5$ $= 22u - 25u$ $= -3u$ $= -3$

Figure 4: Workings showing the correct answer and the CSE answer of Case 1

3.2. Case 2

The second question considered here is related to the Geometric Series. Students were presented with an infinite geometric series of the form $a(r) + a(r)^2 + a(r)^3 + \cdots$, where parameters a and r are generated randomly for each instance of the question. The question requires students to calculate the sum S correct to three decimal places. One CSE was identified with this question and it occurs by finding the sum of the first four terms instead of the sum of the infinite series. An example in which the CSE answer is equal to the question's answer was found during the re-marking process and occurs when the sum of the infinite series, $S = 2 + 2(0.1) + 2(0.1)^2 + 2(0.1)^3 + \cdots$ is asked for. This is illustrated in Figure 5. As shown in Figure 5, it can be seen that, to three decimal places, both the correct answer and the CSE answer are the same in this case.

Correct Answer	**CSE Answer** ***CSE:*** *finding the sum of the first four terms instead of the sum of the infinite series*
$S = \frac{a}{(1-r)} = \frac{2}{(1-0.1)}$ $= 2.22222\ldots$ $= 2.222 \quad (\text{correct to 3 dp})$	$\tilde{S} = \frac{a(1-r^n)}{(1-r)} = \frac{2(1-0.1^4)}{(1-0.1)}$ $= 2.222$

Figure 5: Workings showing the correct answer and the CSE answer of Case 2

3.3. Case 3

For this case, students were asked to find the power series expansion, $P_3(x)$, of $f(x) = e^{ax}$, up to and including the cubic term, and to use $P_3(x)$, to calculate an approximate value for $f(x)$ at $x = c$, correct to three decimal places. The parameters a and c are generated randomly for each instance of the question. One of the identified CSEs of this question is to give the exact value of e^{ax} instead of the approximate value of e^{ax} at $x = c$.

It was found that when $a = 2$ and $c = -0.1$, the correct answer and the CSE answer of this question are the same, to three decimal places, namely 0.819, as shown in Figure 6.

Correct Answer	**CSE Answer** ***CSE:*** *finding the exact value of e^{ax} instead of the approximate value of e^{ax} at $x = c$.*
$P_3(x) = 1 + 2x + \frac{(2x)^2}{2} + \frac{(2x)^3}{6}$ $P_3(-0.1) = 1 + (-0.2) + \frac{(-0.2)^2}{2} + \frac{(-0.2)^3}{6}$ $= 0.818667\ldots$ $= 0.819 \quad (\text{correct to 3 dp})$	$\widetilde{P_3(-0.1)} = e^{-0.2} = 0.818731\ldots$ $= 0.819 \quad (\text{correct to 3 dp})$

Figure 6: Workings showing the correct answer and the CSE answer of Case 3

3.4. Case 4

The question for this case, required the student to find the mean value of $f(t) = a\sin(bt)$ in the interval $p < t < q$ correct to two decimal places, where the parameters a, b, p and q are generated randomly for each instance of the question. One of the identified CSEs of this question is to evaluate the mean value of $f(t)$ using degrees instead of radians in the calculation.

During the re-marking process, it was found that when $f(t) = -3\sin(5t)$ and the interval is $3 < t < 7$, the value of the mean, which is $m = -0.02$, is the same as the CSE answer, $\tilde{m}$, correct to two decimal places as shown in Figure 7.

Correct Answer	**CSE Answer** ***CSE:*** *evaluating the mean value of $f(t)$ using degrees instead of radians.*
$m = \frac{1}{(7-3)}\int_3^7 -3\sin(5t)\ dt$ $= \frac{1}{4}\left[\frac{3}{5}\cos(5t)\right]_3^7$ $= \frac{3}{20}[\cos(35) - \cos(15)]$ $= \frac{3}{20}[-0.9037 + 0.7596]$ $= -0.021615 \ldots$ $= -0.02$ (correct to 2 dp)	$\tilde{m} = \frac{1}{(7-3)}\int_3^7 -3\sin(5t)\ dt$ $= \frac{1}{4}\left[\frac{3}{5}\cos(5t)\right]_3^7$ $= \frac{3}{20}[\cos(35^o) - \cos(15^o)]$ $= \frac{3}{20}[0.8192 + 0.9659]$ $= -0.022005 \ldots$ $= -0.02$ (correct to 2 dp)

Figure 7: Workings showing the correct answer and the CSE answer of Case 4

3.5. Case 5

The question in this case involves finding the volume, V, of the solid formed when the part of the curve $y = a\,x^b$ is rotated about the x −axis between $x = f$ and $x = g$ and quoting the answer to two decimal places. The parameters a, f and g are generated randomly for each instance of the question and b is selected randomly from a pre-determined list of possible values. One of the identified CSEs of this question was to calculate V without integrating the required expression, but instead substituting the upper and lower limits directly into the integrand.

During the re-marking process, it was found that for some question parameters, the correct answer and the CSE answer of this question were the same. For example, this occurs when $a = 6,\ b = 1$, $f = 0$ and $g = 3$. In this case, the correct answer and the CSE answer ($\tilde{V}$) can be calculated as shown in Figure 8.

Correct Answer	CSE Answer *CSE: finding the volume of revolution by substituting for the upper and lower limits without integrating.*
$V = \pi \int_0^3 (6x)^2\, dx$ $= \pi \int_0^3 36x^2\, dx$ $= 36\,\pi \left[\frac{x^3}{3}\right]_0^3$ $= 36\,\pi\,[3^2 - 0]$ $= 1017.88\ \text{(correct to 2 dp)}$	$\tilde{V} = \pi\,[(6x)^2]_0^3$ $= 36\,\pi\,[x^2\,]_0^3$ $= 36\,\pi[3^2 - 0]$ $= 1017.88\ \text{(correct to 2 dp)}$

Figure 8: Workings showing the correct answer and the CSE answer of Case 5

3.6. Case 6

Another identified CSE of the question presented in Case 5 was to find the volume of revolution by taking $(x^p)^q$ to be x^{p^q}. The correct answer and the aforementioned second CSE answer of this question are the same when $a = 0.6,\ b = 2,\ f = 1$ and $g = 4$. In fact, this would be the case when $b = 2$ no matter the values of a, f, g since in this case $\left(x^b\right)^2 = (x^2)^2 = x^{(2^2)} = x^4$ and from there on the workings for the CSE answer would be exactly the same as for the correct answer.

4. Resolution

Without rough workings, for the examples presented in Section 3, there is no way of ascertaining whether the student arrived at the final answer by following the correct approach or by making the identified CSE. We have resolved this issue by ensuring that, for future instances of the question, the random parameters are selected so as the correct answer and the CSE answer(s) are different. This was achieved by further amending the CSE question code. For Cases 1-5, at the parameter selection stage of the code, the correct answer and the CSE answer(s) were calculated for each set of parameters. A while loop was then used to re-select the parameters until the correct answer and the CSE answer(s) were all different to each other.

In the original question code for Case 5, b was selected randomly from the following list of values: $[0.25, 0.5, 1, 1.25, 1.5, 2]$. In order to avoid the correct answer being equal to the CSE answer identified in Case 6, the value 2 was removed from the list of possible values for b in the amended code. In addition, for Case 4, a further CSE was identified in which students neglected to divide the integral by the interval $q - p$. To avoid the correct answer being equal to this CSE answer, in the amended code q is randomly selected so that q does not equal $1 + p$.

After finding these cases, all of the other CSE question codes were amended to avoid parameter selections for which the correct answers were equal to the CSE answers. As a further precaution, the question codes were amended so that CSE enhanced feedback is provided only when the PI value of the correct answer is zero and the PI of the CSE answer is one. Thus, the respective CSE enhanced feedback is only given to students making a CSE when their answer is incorrect.

5. Discussion and Conclusion

In this article we have shown why it is important to know more about Mathematical CSEs in e-Assessment questions, using several examples. These examples were discovered while conducting the CSE Project at UWE Bristol. We have shown how a correct answer can take the same value as a CSE answer for certain e-Assessment question parameters. In such cases, there may have been instances where some students were awarded full marks and hence thought that they had answered the question correctly when in fact they had made a CSE. We have described, how we addressed this issue by amending the original question code for all identified CSEs.

There has been a significant increase in usage of e-Assessments in higher education in this millennium. Even before the Covid 19 pandemic (World Health Organization 2020), a JISC report (2020) concluded that the archaic pen and paper assessment process is in need of a technological overhaul by 2025. We believe that, in this digital era, the work presented in our research article demonstrates why it is more important than ever to know more about mathematical CSEs.

6. References

Alruwais, N., Wills, G. and Wald, M. 2018. Advantages and Challenges of Using e-Assessment. *International Journal of Information and Education Technology* 8(1), pp. 34-37. https://doi.org/10.18178/ijiet.2018.8.1.1008

CSE Project at UWE Bristol 2019. *Diagnosing and Remediating Mathematical Common Student Errors in e-Assessment Questions: A Case Study.* Available from: https://fetstudy.uwe.ac.uk/~bin-sikurajapa/dewis/cseproject/ [Accessed 24 February 2022].

Dewis, 2012. Dewis welcome page. Available from: http://dewis.uwe.ac.uk [Accessed 02 November 2022].

Ford, S., Gillard, J. and Pugh, M. 2018. Creating a Taxonomy of Mathematical Errors For Undergraduate Mathematics. *MSOR Connections*. 18 (1), pp. 37-45.

Gwynllyw, R. and Henderson, K. 2009. Dewis: A Computer Aided Assessment System for Mathematics and Statistics. *CETL-MSOR 2008 Conference Proceedings*. pp. 38-44.

JISC. 2020. The future of assessment: five principles, five targets for 2025. Available from: https://www.jisc.ac.uk/reports/the-future-of-assessment [Accessed 02 November 2022].

Race, P. 2014. *Making Learning Happen*, Sage Publications.

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

Rushton, N. 2014. Common errors in Mathematics. *Research Matters: A Cambridge Assessment publication*, *17,* pp. 8-17.

Sangwin, 2013. *Computer Aided Assessment of Mathematics*. Oxford, UK: Oxford University Press.

Sikurajapathi, I., Henderson, K., and Gwynllyw, R., 2020. Using E-Assessment to Address Mathematical Misconceptions in Engineering Students. *International Journal of Information and Education Technology*. 105), pp. 356–361.

Sikurajapathi, I., Henderson, K., and Gwynllyw, R. 2021. Students' Perceptions of Enhanced E-assessment Feedback Addressing Common Student Errors in Mathematics. *MSOR Connections*. 19 (2), pp. 10-27.

Sikurajapathi, I., Henderson, K., and Gwynllyw, R. 2022a. *Collection of Taxonomically Classified Mathematical Common Student Errors in E Assessments – (CSE Book).* [online]. University of the West of England Bristol, United Kingdom. Available from: https://uwe-repository.worktribe.com/OutputFile/9305581 [Accessed 08 April 2022].

Sikurajapathi, I., Henderson, K., & Gwynllyw, R. 2022b. Gathering and Compiling Mathematical Common Student Errors in e-Assessment Questions with Taxonomical Classification. *MSOR Connections*, 20 (3), pp. 55-71

Stödberg, U. 2012. A research review of e-assessment. *Assessment & Evaluation in Higher Education*, 37(5), pp. 591–604.

VanLehn, K. 1982. Bugs are not enough: Empirical studies of bugs, impasses and repairs in procedural skills. *Journal of Mathematical Behavior*, 3(2), pp. 3-71.

World Health Organization 2020, Coronavirus disease (COVID-19) Pandemic. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [Accessed 02 November 2022].