Conceptual Statistical Assessment Using JSXGraph

https://doi.org/10.3991/ijet.v18i01.36529

Curtis Hooper(✉), Ian Jones
Loughborough University, Loughborough, UK
C.Hooper@lboro.ac.uk

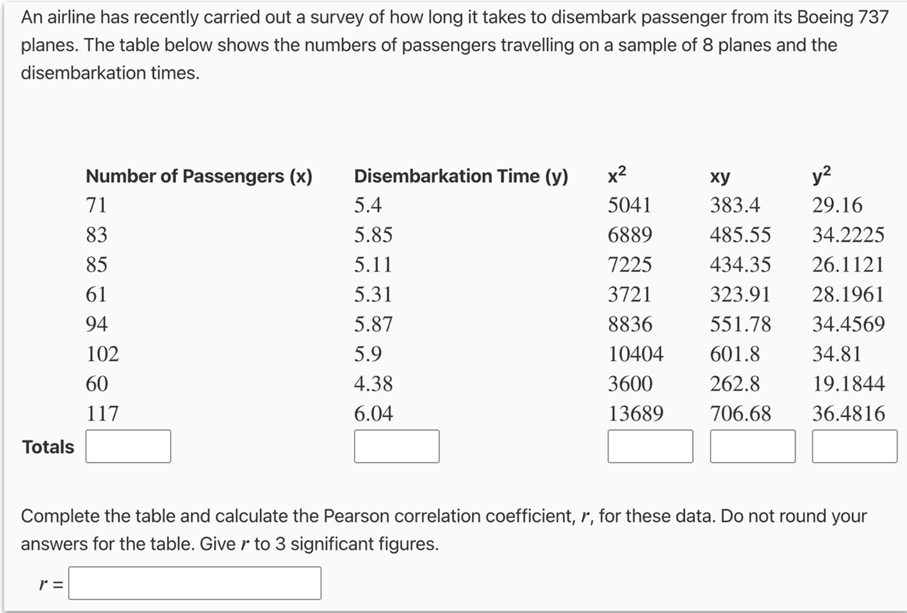
Abstract—Traditionally online assessments tend to focus on topics that require students to input algebraic and numeric responses. As such there is a paucity of questions that test students' knowledge of statistics, and what questions there are in our experience focus on computing specific values (mean, standard deviation, and so on). Through making use of a technology called JSXGraph that is supported within the STACK environment for online assessment of mathematical knowledge, we have developed statistics questions that aim to test conceptual knowledge. For example, by requiring students to adjust the bars in a graph in order to produce a dataset that has a required mean, median, mode and range. With careful design this approach enables open-ended questions that have more than one correct answer. In this paper we describe the questions we have designed, and report responses from a sample of students.

Keywords—statistics, conceptual understanding, online assessment

1 Procedural and conceptual understanding

A long-standing distinction in the mathematics education research literature is between procedural and conceptual understanding [1]. The former is often defined in terms of facts and procedures, and tends to be straightforward to assess [2]. For example, if we want to know if a student can calculate the arithmetic average (mean) given a set of numerical data we can ask them to do so and check their answer. Conceptual understanding is often defined in terms of knowledge of fundamental mathematical ideas and how they connect together. This can be difficult to assess, especially in online contexts, as it does not lend itself to traditional test items [3].

At Loughborough University we have had success developing online tests for assessing procedural knowledge of statistics. For example, the authors have worked with colleagues from Loughborough's School of Business and Economics to develop test items such as that shown in Figure 1. Such items have been used to assess the procedural knowledge of hundreds of students across various statistical topics such as hypothesis testing, sampling, and regression. Online assessments can also be used to support student learning of statistical procedures [4], which is important given the reported student difficulties in this area [5]. For example, we recently embedded test items in online interactive worksheets as part of the Helping Engineers Learn Mathematics (HELM) project developed by the universities of Loughborough and Edinburgh [6]. These online resources allow students to practice statistical procedures, and receive instant feedback.



An airline has recently carried out a survey of how long it takes to disembark passenger from its Boeing 737 planes. The table below shows the numbers of passengers travelling on a sample of 8 planes and the disembarkation times.

	Number of Passengers (x)	Disembarkation Time (y)	x^2	xy	y^2
	71	5.4	5041	383.4	29.16
	83	5.85	6889	485.55	34.2225
	85	5.11	7225	434.35	26.1121
	61	5.31	3721	323.91	28.1961
	94	5.87	8836	551.78	34.4569
	102	5.9	10404	601.8	34.81
	60	4.38	3600	262.8	19.1844
	117	6.04	13689	706.68	36.4816
Totals					

Complete the table and calculate the Pearson correlation coefficient, *r*, for these data. Do not round your answers for the table. Give *r* to 3 significant figures.

r =

Fig. 1. A test item for assessing students' procedural knowledge of correlation coefficients in a business context. (Designed by the first author)

Assessments of students' conceptual understanding of statistics often focus on interpreting (rather than computing) results, and their automation is challenging. For this project we turned to the ingenious paper-based learning activities developed by the late, great mathematical educator Malcolm Swan. In particular, we were inspired by Swan's [7] collaborative card-based activities, see Figure 2, which are used in lectures by the second author. Swan's intention is for students to discuss in small groups how the representation of data as bar charts and as summary statistics (mean and so on) relate to one another. Successfully matching the cards requires thinking strategically, qualitatively, and quantitatively about different representations of mathematical objects and this promotes conceptual understanding [8].

During the global pandemic that started in 2020 the second author developed online drag-and-drop versions of these activities for remote teaching, see Figure 3. However, the drag-and-drop versions were sub-optimal because they did not work well through mobile browsers as, once the statistics are dragged onto the bar charts the latter could no longer be seen, the cards could not be written on by students to fill the blanks, and different versions (e.g., with varied datasets, or similar questions addressing related topics) could not readily be produced. We therefore sought a technology that could be harnessed to address these limitations.

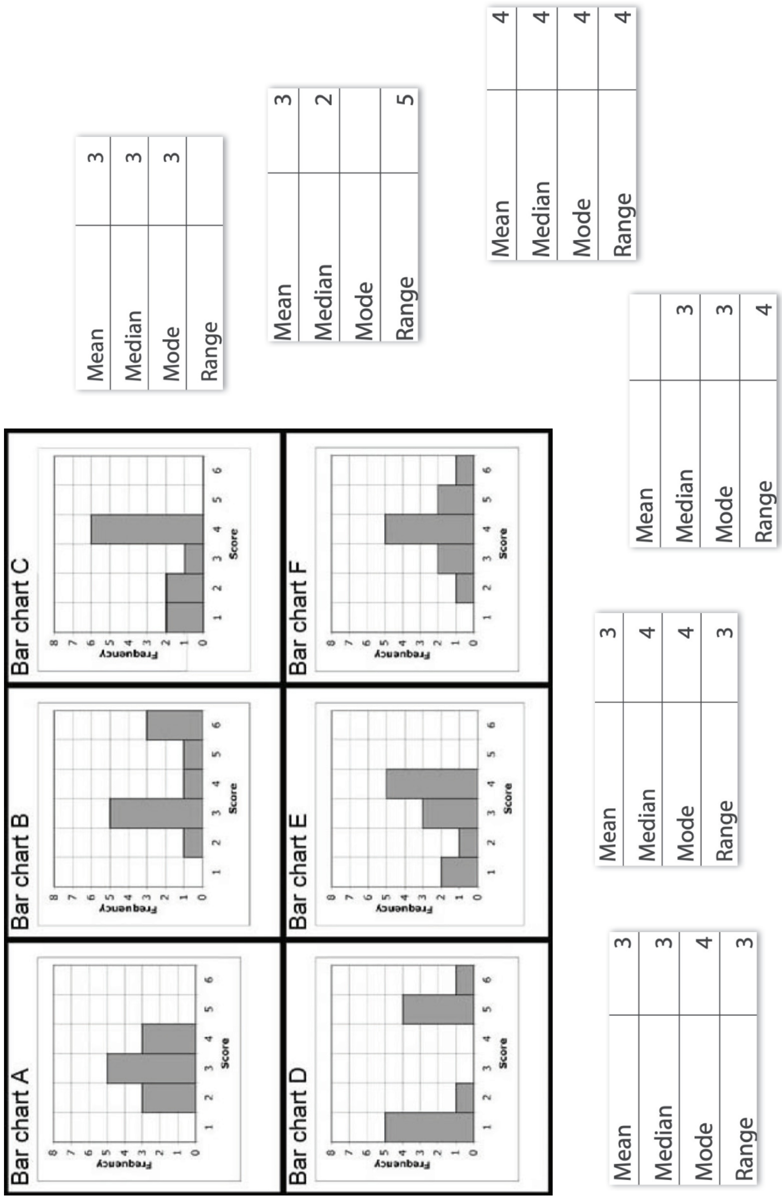


Fig. 2. Paper-based learning activities designed to promote conceptual understanding of statistics, in this case averages and range. (Designed by Malcolm Swan [7])

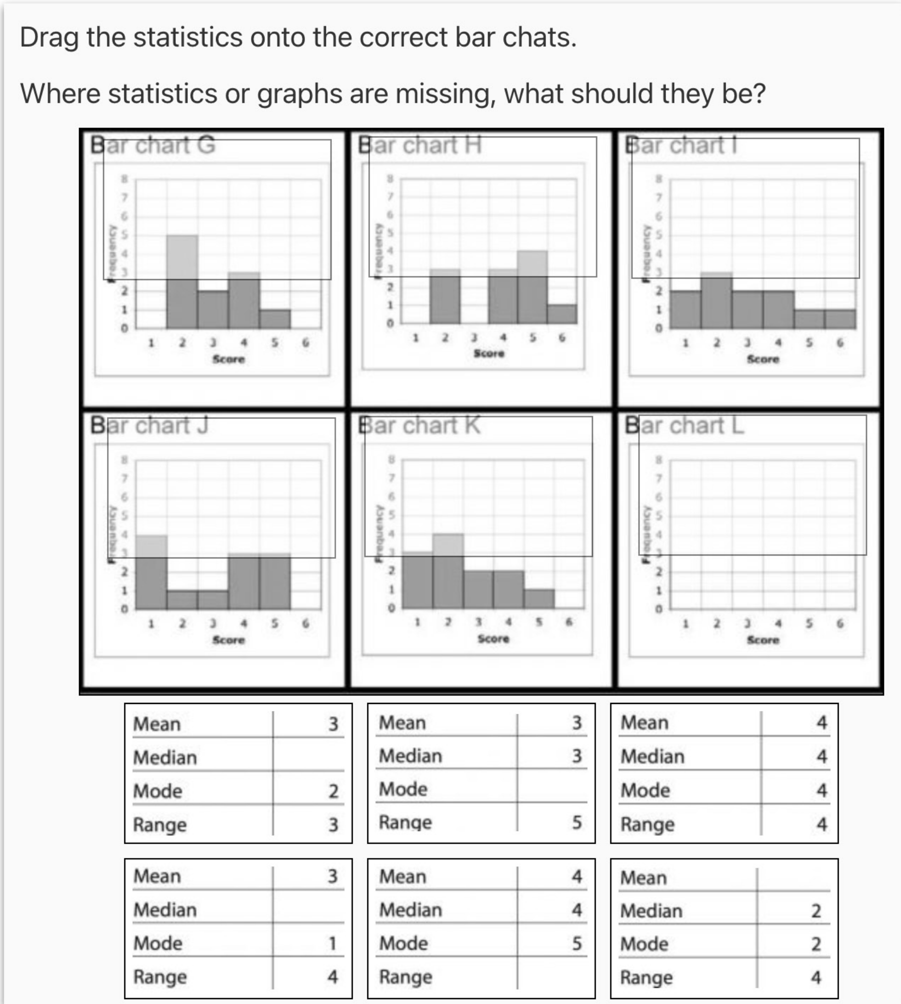


Fig. 3. Drag-and-drop version of Swan's [7] statistical learning activity

2 STACK and JSXGraph

The program we have used to create such resources is the System for Teaching and Assessment using a Computer Algebra Kernel, better known as STACK [9]. It is available for the Moodle Virtual Learning Environment and can produce online mathematics questions that can be used as standalone learning resources or part of a formal assessment.

A key feature of STACK is to generate randomised question variants within the constraints set by the question author. This creates a more rigorous test environment when such questions are used as part of a formal assessment as the correct answers from one of test will not necessarily be the correct answers to another test, thus making it harder for cheating to occur.

Another useful feature in STACK is that questions with both numerical and algebraic answers can be marked. Furthermore, feedback can be given instantly on completion of a question. This can range from a simple "correct" or "incorrect" statement, to a more tailored response which is dependent on the answer given. Fully worked solutions to the question can also be displayed. Students can reflect upon the feedback received and, if permitted by the author, re-attempt the question. As the feedback is immediate, the marking of formal assessments set through STACK is far more efficient than marking the test would otherwise be with a human, which would require significantly more time and money.

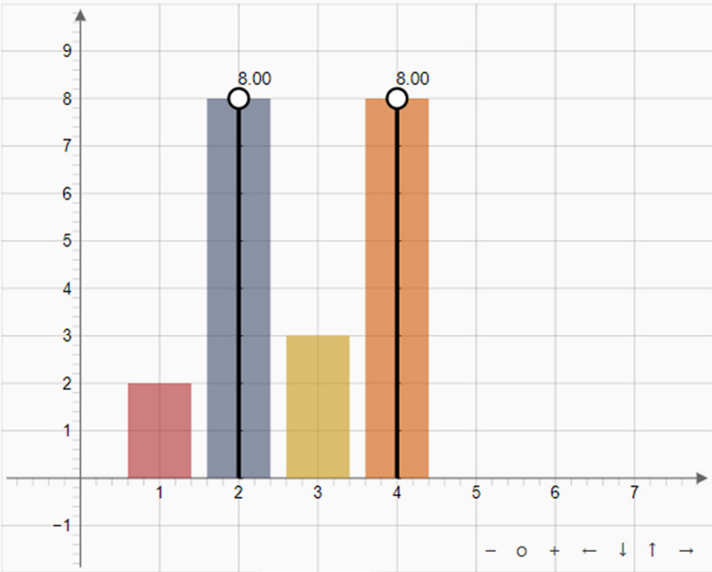
Recently, it has been possible to implement JSXGraph into STACK thanks to its compatibility with Moodle. JSXGraph is a web-based JavaScript program which has the ability to plot interactive graphs and mathematical functions [10]. This allows graphs to be plotted and edited by the student in real time through the use of, for example, sliders which control certain parameters of the function plotted. When implemented into a STACK question, the value of the slider can be submitted as the answer to a question and subsequently marked by STACK in the usual manner.

3 Conceptual statistics test questions

We now present the questions that we constructed by implementing the JSXGraph feature into STACK. The aim was to create interactive learning resources that students could use in a learning environment, such as during a lecture or in their own study time, to improve their understanding of statistics.

An example of such activity is shown in Figure 4. Here the students are given the mean, median, mode and range of some unknown data and their task is to move the two moveable bars in the bar chart such that the chart conveys the given statistics. The bars themselves are controlled by a slider which, in Figure 4, are set by default to a numerical value of 8. The student can simply click and hold the white circle to drag the bar to the desired position. The value of the slider will be displayed in the appropriate answer box as the student moves them. It will be these values that will be scored by STACK when the student submits their answer.

Change the bars corresponding to a scores of 2 and 4 such that the data in the bar chart has the statistics given in the table below.



Frequency of score $2 =$ [ ]

Frequency of score $4 =$ [ ]

Mean	3
Median	3
Mode	4
Range	3

Fig. 4. An interactive bar chart using JSXGraph embedded in STACK

When presented with this task, the student must carefully consider what each statistic means and how it relates to the bar chart. A particular case that can arise when there are two moveable bars is the possibility of multiple correct answers. This is the case in the question presented in Figure 4. Indeed, to satisfy the condition of the mean equaling 3 we arrive at the equation

$$\frac{11+2x+4y}{5+x+y}=3, \quad (1)$$

Where x and y are the user answers corresponding to "Frequency score of 2" and "Frequency score of 4" respectively as given in the question. This leads to the condition for a correct mean of

$$y = x + 4. \quad (2)$$

It can be seen that, if this condition is used correctly, the mode will always be 4 as required and the range will be 3 accordingly. If students did not take this algebraic approach to solve the question, the feedback given on a correct answer asked them to consider if there were any other correct answers. It was hoped then that students would try to find other solutions thus getting a better understanding of the question.

An alternative question that we made was one which involved the students moving two nodes connected by a straight line. The task was to draw the correct regression line for data shown in the form of a scatter plot. An example of this question type is shown in Figure 5.

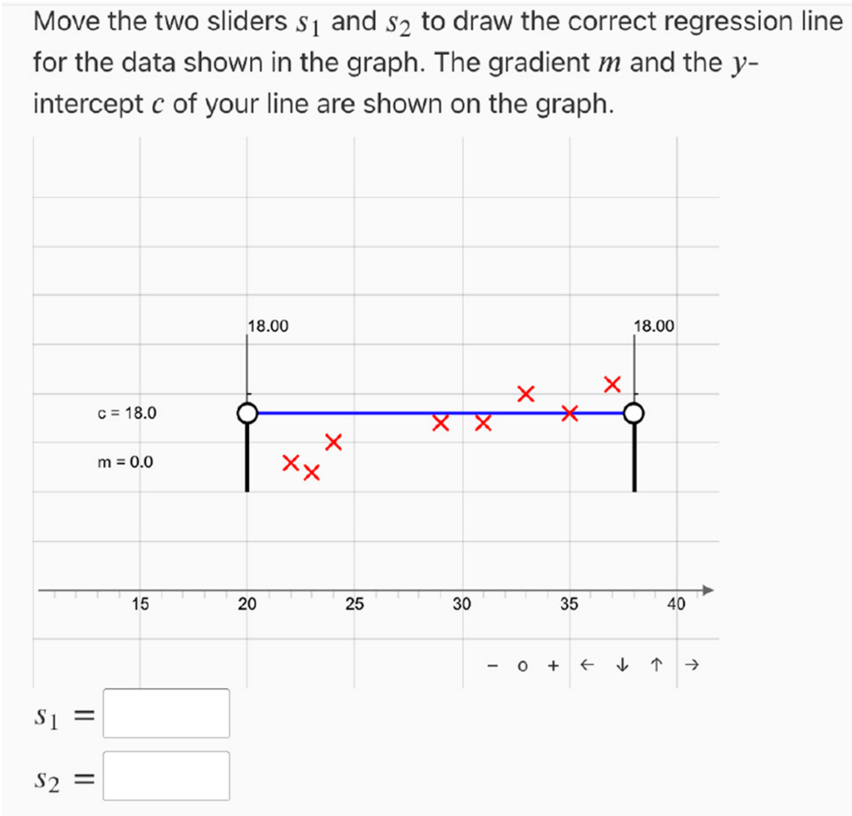


Fig. 5. An interactive straight line on a scatter graph using JSXGraph embedded in STACK

As the students move the nodes on the sliders, the y–intercept and gradient of the line (given by c and m respectively in Figure 5) are displayed in real time. As before, the values of the sliders are displayed in the boxes s_1 and s_2 respectively and are the values that will be scored on submission. Here, it is highly unlikely that the student will draw the true regression line, so it is important for a suitable tolerance to be determined and set accordingly within the coding of the question by the author. The discretisation of the sliders must also be carefully considered here as making the intervals of the slider smaller increases the number of possible answers that the student can give, hence making it more difficult for them to draw the true regression line.

4 Student responses to the test questions

We explored the performance of JSXGraph-based conceptual questions as compared to STACK-based procedural questions using opportunistic data from a practice (non-graded, optional) test administered as part of a mathematics refresher module. The test was attempted by 37 students in their first year of university. The test included five conceptual questions, all of which required students to drag one or two bars, similar to the question shown in Figure 4. The test also included five procedural questions that involved calculating the mean, median and/or mode for a presented set of data. The items contained between 1 and 3 responses each, shown as number of parts in Table 1, and each was scored out of a total maximum score of 1.

Table 1. Items, number of parts, and scores (max for each item = 1)

Item Name	Number of Parts	Mean Score
Conceptual items		
C1	2	0.82
C2	2	0.73
C3	2	0.75
C4	2	0.70
C5	3	057
Procedural items		
P1	2	0.95
P2	2	0.89
P3	2	0.91
P4	3	0.93
P5	1	0.85

We conducted a series of analyses to evaluate the validity of the conceptual items. First, we investigated the internal consistency of each set of items by calculating Cronbach's alpha, using the guideline that a value ≥ 0.70 indicates satisfactory internal consistency [11]. For the conceptual items we found $\alpha = 0.81$ and for the procedural items we found $\alpha = 0.73$. This suggests satisfactory internal consistency for both sets of items, meaning each set can be considered to measure a unidimensional construct.

Second, we calculated a total conceptual score and a total procedural score for each student, and then calculated the Spearman Rank Order correlation coefficient between them. We found a modest and significant correlation, $\rho = 0.50$, $p = 0.01$, suggesting the two sets of items assessed related but not identical constructs. This result is consistent with all the items assessing statistical knowledge, but differing as to whether the focus is on conceptual or procedural understanding [3].

Third, we explored students' performance on each set of items, as shown in the mean score column in Table 1. Overall, the students scored highly on the test, and all the procedural item scores (range 0.85 to 0.95) were higher than all conceptual item scores (range 0.57 to 0.82). A Wilcoxon signed-rank test showed this difference was statistically significant, $p = 0.01$. This analysis of student performance is consistent procedural understanding being generally more accessible by students than conceptual understanding [1].

Taken together, the above series of analyses provide supportive evidence that the JSXGraph-based items validly assessed conceptual understanding of statistics.

5 Conclusion

We have demonstrated how the use of an innovative technology, namely JSXGraph embedded in the STACK system, can be used to develop items for assessing students' conceptual understanding of mathematics, at least for the case of simple statistics. We achieved this by taking research-informed paper-based tasks in the mathematics literature [7], and adapting them into items that can be automatically marked within a Virtual Learning Environment. To investigate the validity of the items we calculated internal consistency, and compared an opportunistic sample of students' performance on the items to their performance on procedural items. Our results provided supportive evidence that the JSXGraph-based items validly and successfully assessed conceptual, rather than procedural, understanding of simple statistics.

We are currently conducting further developmental work and analyses of student responses towards generating a suite of items that automated the assessment of conceptual understanding across a range of statistical topics and contexts.

6 Acknowledgement

This work was funded by a Teaching Innovation Award from Loughborough University, UK.

7 References

[1] R. Skemp, "Relational understanding and instrumental understanding", *Mathematics Teaching*, vol. 77, 1976, Accessed: Jan. 02, 2013.

[2] J. R. Star, "Reconceptualizing procedural knowledge", *Journal for Research in Mathematics Education*, vol. 36, pp. 404–411, 2005.

[3] N. M. Crooks and M. W. Alibali, "Defining and measuring conceptual knowledge in mathematics", *Developmental Review*, vol. 34, no. 4, pp. 344–377, Dec. 2014. https://doi.org/10.1016/j.dr.2014.10.001

[4] S. Krishnasamy, L. S. Ling, and T. C. Kim, "Improving Learning Experience of Probability and Statistics using Multimedia System", *International Journal of Emerging Technologies in Learning (iJET),* vol. 15, no. 01, pp. pp. 77–87, Jan. 2020. https://doi.org/10.3991/ijet.v15i01.11349

[5] N. Chandrasena Premawardhena, "ICT in the foreign language classroom in Sri Lanka: A journey through a decade", 10th World Conference on Computers in Education (WCCE 2013), Nicolaus Copernicus University, July 2–5, 2013, Torun, Poland. pp. 223–224.

[6] K. Zerva, I. Nicholson, and A. Mateo, "Translating the HELM workbooks to STACK", *STACK Assessment*, 2021. [Online]. Available: https://stack-assessment.org/CaseStudies/2021/HELM/. [Accessed Oct. 21, 2022].
[7] M. Swan, Improving Learning in Mathematics: Challenges and Strategies. Department for Education and Skills (UK), Standards Unit, 2005.
[8] M. Swan, "Learning GCSE mathematics through discussion: what are the effects on students?", *Journal of Further and Higher Education*, vol. 30, no. 3, pp. 229–241, 2006. https://doi.org/10.1080/03098770600802263
[9] C. Sangwin, Computer Aided Assessment of Mathematics. Oxford University Press, 2013. https://doi.org/10.1093/acprof:oso/9780199660353.001.0001
[10] M. Gerhäuser, B. Valentin, and A. Wassermann, "JSXGraph-Dynamic Mathematics with JavaScript", *International Journal for Technology in Mathematics Education*, vol. 17, no. 4, 2010.
[11] L. J. Cronbach, "Coefficient alpha and the internal structure of tests", *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951. https://doi.org/10.1007/BF02310555

8 Authors

Dr Curtis Hooper is a University Teacher in the Wolfson School of Mechanical, Electrical and Manufacturing Engineering at Loughborough University. He supports module leaders across the University with the development and implementation of STACK questions in their modules. (email: C.Hooper@lboro.ac.uk).

Dr Ian Jones is a Reader in Education Assessment in the Department of Mathematics Education at Loughborough University, UK. He was a Royal Society Shuttleworth Education Research Fellow, sits on the editorial boards of *Research in Mathematics Education* and the *International Journal for Research and Method in Education*, and has won numerous awards and grants for his work on assessment and other areas of education. (email: I.Jones@lboro.ac.uk).

Article submitted 2022-08-27. Resubmitted 2022-10-29. Final acceptance 2022-11-03. Final version published as submitted by the authors.