An Evaluation of Computer-Based Interactive Simulations in the Assessment of Statistical Concepts

By David L. Neumann, Michelle Hood & Michelle M. Neumann

School of Psychology, Griffith University, Australia
D.Neumann@griffith.edu.au.

Received: 25 August 2010    Revised: 22 October 2010

In a previous report, Neumann (2010) described the use of interactive computer-based simulations in the assessment of statistical concepts. This assessment approach combined declarative knowledge of statistics with experiences in interacting with computer-based simulations. The aim of the present study was to conduct a systematic evaluation of the approach. A stratified random sample of students (n = 38) was selected and participated in an interview to provide qualitative data that was coded into themes. The students reported that the assessment approach improved their understanding of statistics and its practical application, gave them a way to practice statistics, motivated them to complete it, was interesting and engaging, and gave a visual aid to learning. Students also commented on the immediate feedback, the unique nature of the assessment, and the technological requirements involved. The results suggest that the use of computer-based interactive simulations in assessment can be a positive addition to assessment practices in a university statistics course.

1 INTRODUCTION

The increased use of technology has led to significant improvements in mathematics education. In the field of statistics, in particular, there has been a proliferation of computer-based activities to teach theoretical concepts and their application. Students can use software to conduct virtual studies and analyse the resulting data (Malloy and Jensen, 2001) and use various interactive media (e.g., JAVA applets) to learn about statistical concepts (e.g., Lane, 1999; Morris, Joiner and Scanlon, 2002; Meletiou-Mavrotheris, 2003; Dunn, 2004) either in isolation or embedded within an on-line textbook (e.g., the Computer-Assisted Statistics Textbook or CAST, see http://cast.massey.ac.nz). Researchers have also begun to explore how to use interactive multimedia in assessing student learning. Crisp (2002) gave guidelines on how JAVA applets could be used to develop on-line assessment across a range of disciplines (see also, Baxter, 1995; Baker and Mayer, 1999; Boyle, 2007). Recently, Neumann (2010) described the use of computer-based interactive simulations in the assessment of statistical concepts in a university introductory statistics course.

The assessment approach developed by Neumann (2010) was based on the principle of ensuring a constructive alignment between teaching and assessment (Biggs and Tang, 2007) and the use of an assessment approach that encouraged the application of knowledge. It was argued that if statistics instructors want to teach students how to apply their knowledge, assessment tasks must also be designed that can assess this application. The use of assessment that is linked to interactive computer-based simulations was suggested as one way to do this. In one example that was offered, a student could use an interactive simulation to describe a real data set by using a histogram. The student could select different class widths for the histogram and see how their interpretation of the data set was affected. In this way, it required students to apply their declarative knowledge of statistics (e.g., knowledge of the terms that describe skewness and the number of peaks in a distribution) to experiences gained by interacting with the simulation. Moreover, students were provided with immediate feedback on their performance. Several interactive assessment tasks, each focussing on a different statistical concept, were designed along these principles (see Neumann, 2010 for details).

Neumann (2010) gave only a limited evaluation of the use of the interactive assessment approach. It was reported that the interactive assessment was used by a large proportion of students (e.g., 91.1% passed at least one component). Neumann (2010) also reported the outcomes of an evaluation of written feedback on a generic course evaluation questionnaire by Neumann, Hood and Neumann (2008). Coding of the feedback related to the interactive assessment resulted in seven themes: Helps learning, helps confidence, practice concepts, alternative learning tool, more exercises, make compulsory and see mistakes. Based on this data, it was argued that using interactive assessment tasks that were interactive and encouraged application of knowledge had a range of beneficial effects. However, further interpretations of these results are limited by the methods used. The course evaluation questionnaire was a voluntary task and was completed by only 17.8% to 45.5% of students across three years. Not all students who completed the questionnaire made comments on the interactive assessment because there was no specific question on the questionnaire about it. The low response rate and selectivity of the responses (for example, only high achieving students may have commented on it) could have potentially biased the outcomes. The use of written feedback was also limited in that there was no opportunity to seek expansion or clarification of issues mentioned by a student.

The aim of the present study was to extend the preliminary evaluation of the use of computer-based interactive simulations in the assessment of statistical concepts reported by Neumann (2010) and Neumann et al. (2008). The main focus of this study was on the student experiences that resulted from using an assessment task that was interactive and required deep processing of information.
We expected that students would be motivated to complete the assessment and would find the approach engaging. We also expected that the new experiences students had in interacting with the simulations would have beneficial learning outcomes by improving their understanding and application of statistics. To ensure that the sample was more representative of the entire student cohort, a stratified random sample of students was selected so that the distribution of grades in the sample was the same as that obtained across all students in the course. An interview approach was also employed to allow for follow up questioning and a richer data set. The importance of collecting qualitative data concerning the learning process and student experiences has been emphasised by several researchers (Jones, Scanlon, Tosunoglu, Ross, Butcher, Murphy and Greenberg, 1996; Morris and Scanlon, 2000; Morris et al., 2002). The student feedback during the interview was categorised into themes. It was expected that some of the themes identified by Neumann et al. (2008) would also emerge. However, additional themes were also expected because of the more representative sample and a more in-depth data collection approach that was used.

2 METHOD

2.1 The Statistics Course

The course in which the interactive assessment was employed was a university introductory-level statistics course with an enrolment of approximately 200 students who were predominantly studying psychology and other social sciences. The course teaches the topics of basic research, descriptive statistics, correlation, z-scores, probability, sampling distributions, hypothesis testing for one sample and two samples, and nonparametric hypothesis tests (for example, the Mann-Whitney U test). It is taught over a 13 week semester. The course is taught through an on-campus mode of teaching with a class structure made up of a 2 hour weekly lecture with smaller 1 hour tutorial classes. The tutorial classes include small group exercises and computer-based exercises in which the SPSS statistical package is taught. To supplement the in-class teaching, there were also resources that students could access on-line via the BlackBoard system.

2.2 The Use of the Computer-based Interactive Simulations in Assessment

The course contained four assessment items: a midterm exam, assignment, end-of-semester exam, and the interactive assessment. The interactive assessment contributed 10% towards the final grade. It was not compulsory that students complete or even attempt the interactive assessment, although a high proportion of students did (Neumann, 2010). Details about the design and implementation of the interactive assessment are given in Neumann (2010). Briefly, students accessed the interactive assessment on-line via the course website. Seven interactive assessment tasks were developed. Each assessment task focused on a specific topic such that the following topics were covered: graphing qualitative and quantitative data, sampling data and sampling error, sampling distribution of the mean, confidence intervals of the mean when the population standard deviation is and is not known, errors in hypothesis testing, and correlation and regression.

The interactive assessment contained three components: a simulation, instructions, and multiple choice questions (Neumann, 2010). The computer-based simulations were JAVA-based or FLASH-based applications that allowed basic data analysis and interpretation (for example, calculate histogram and stem plots from a data set), showed a statistical concept in action (for example, the sampling distribution of the mean; confidence intervals; sampling error), or showed what factors influence statistics and statistical decisions (for example, hypothesis testing). In many cases, a simulation had more than one function. The important aspect of the simulations was that they were interactive and dynamic. The student could use the interface to change values or simulate events and the display changed dynamically. Interactive simulations of this type have been developed by other authors. The Rice Virtual Lab in Statistics (Lane, 1999) is a compilation of such simulations (see: http://onlinestatbook.com/rvlabs.html). Similar computer-based simulations have also been described in articles (e.g., Morris et al., 2002; Meletiou-Mavrotheris, 2003; Dunn, 2004). To the right of the simulation were instructions that explained its main features and guided the student through its functions. The assessment component consisted of 10 multiple choice questions. Importantly, the questions could not be answered merely through knowledge of statistics, but required students to combine this knowledge with the experiences in interacting with the simulation. For example, one question may have asked the student how a statistic changed (for example, the probability of a Type II error) when one or more factors were changed (for example, when the mean value in a distribution was changed).

The students answered the questions on-line and submitted the answers to a database to gain immediate feedback on their mark. A pass mark was deemed to be 8 or more out of 10. If a student did not pass a topic, they were allowed to repeat the topic again for a maximum of three attempts. The assessment task was linked to the grade book database for the course, in which the marks obtained by each student were saved against their name. Each interactive assessment topic contributed 2% towards the final grade, such that students could obtain a maximum of 10% for the interactive assessment component.

2.3 Participant Selection

The participant selection and interview procedure was granted approval by the institutional ethical review board. A stratified random sample of students were selected to ensure that the distribution of grades (Fail, Pass, Credit, Distinction or High Distinction) were the same in the final sample as they were in the total course enrolment. The first step in obtaining the sample was to randomly select students from each grade and make contact with them via telephone. This was done after the semester had finished and the grades known. Furthermore, the selection and contact was made by an individual that was not involved in the course and no names were released to the teachers of the course. This
ensured anonymity and reduced any undue pressure to participate felt by the students. Upon contact, the aim of the study and the requirements for participation were described. Participants were told that they would be interviewed by a different researcher approximately one week later. Participants were also informed that a cafè voucher (valued at AUD$7.00) would be offered for participation. Fifty students in total had to be selected to obtain a final stratified sample of 40. The attrition of 10 students was due to 5 students that were unable to be contacted (e.g., incorrect or old telephone number) and 5 students that declined to participate (of these 2 received a pass grade and 1 each received a credit, distinction, and high distinction grade). Those students that agreed to participate were mailed a copy of the information sheet and consent form and arrangements were made to conduct the interview at a later date.

At the later date agreed to, an interviewer contacted the student via telephone. This interviewer was a different individual to that who made the initial selection and contact with the student and was also not involved in the course. This ensured that the interviewer was not aware of the grade that the student had obtained or any other aspect of their participation in the course. Of the 40 students that originally agreed to participate, 1 could not be contacted again and 1 declined to participate. The final sample consisted of 38 students (27 females and 9 males) with a mean age of 23.97 years (SD = 7.33). Thirty students had no prior post-secondary school education, four students had completed a diploma, three students had completed a certificate, and one student had completed a bachelor’s degree. The students reported attending 91% (SD = 13.14) of the classes in the course and all had attempted the interactive assessment topics.

2.4 Interview Procedure

The interviews were conducted over the telephone and lasted for approximately 20 minutes. The interviews were digitally recorded for later transcription. Some of the interview was devoted to collecting demographic information and to collect feedback regarding other initiatives used in the course (see Neumann, Hood and Neumann, 2009; Neumann, Neumann and Hood, 2010). The interview used a semi-structured format. The student was asked a series of closed and open-ended questions with follow up questions used to gain further details as required. The students were asked Did you complete any of the interactive simulations for assessment? The student was further asked the questions: What general comments do you have about the interactive simulations? Would you recommend that it be used in the future and why? Should the simulations be compulsory and part of assessment or optional? Follow up questions were also used and consisted of: Did it help you engage with the material and how? Did it help motivate you to learn about statistics and in what way? What were some positive aspects to it? What were some negative aspects to it? After the interview, the student responses were transcribed and given a numerical code unique to that student.

3 DATA CODING AND ANALYSIS

The qualitative data were coded into conceptual themes following the methods described by Neuman (2006). First, open coding was conducted in which the transcribed interviews were initially examined by an individual that had not been involved in the interview or the teaching of the course. The purpose of this initial phase was to condense the 111 unique statements into preliminary analytic themes. This resulted in 12 themes. To confirm the reliability and validity of the coding that was done, a second coder next reviewed the statements in each theme to evaluate whether each statement belonged to that theme, should be moved to a different existing theme, or required a new theme to be established. There was agreement on 77.5% of the coding decisions that were made and this increased to 100% following discussion. The majority of the coding disagreements (10.8% in total) related to the themes later coded as Helped understanding and Visual aid to learning. One coder was more likely than the other to code any reference to visual-based learning in the latter category (for example, “Just that it gave everyone a physical representation of what we were doing”). After discussion, it was decided to code such items in the Visual aid to learning category. The next stage of analysis was axial coding. This process involved reviewing and linking the themes, assigning a descriptive label to each theme, and constructing a definition. In this process, it was noted that a group of statements from one student was coded in its own category (examples were: “Good for group interaction”, “It was good to see people in groups helping each other”). Due to these statements being made from only one student and that it did not seem relevant to the individual nature of the task, it was discarded. In the final pass through the data, selective coding was used to look for statements that illustrated each theme and to highlight the contrasts between related themes.

4 RESULTS

The labels, definitions and percentages of the sample that contributed statements for each theme are shown in Table 1. As can be seen, 11 themes were identified among the students. The most common theme was Helped understanding with 45% of students endorsing this theme. It indicated that students reported pedagogical benefits to the use of the interactive assessment. Related to this theme was Visual aid to learning (13%). This was coded as a separate theme because it seemed to reflect the specific advantage of the computer-based simulations in helping students to visualise the statistical concept being taught. Other themes that may also be expected to contribute to student learning were Practice (24%), Immediate feedback (21%), and Practical application (13%), although these were coded into a unique theme in their own right.
<table>
<thead>
<tr>
<th>Label</th>
<th>Definition</th>
<th>Percent</th>
<th>Example statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helped understanding</td>
<td>The interactive assessment helped students to comprehend and learn about research methods and statistics</td>
<td>45%</td>
<td>“It helped me show how stats worked”, “It definitely enhanced what we learnt in class”, “It helped me understand what I knew and what I didn’t know”</td>
</tr>
<tr>
<td>Motivation</td>
<td>The use of interactive assessment and making them a compulsory part of assessment motivated students to learn about statistics</td>
<td>26%</td>
<td>“To get marks you were motivated to do it. It makes you also do and keep up with the lectures and stuff”, “Knowing what level you are at and what you did know and didn’t know will help to motivate you to learn and be better at it and improve”</td>
</tr>
<tr>
<td>Practice</td>
<td>The interactive assessment gave a way for students to practice their application of statistics</td>
<td>24%</td>
<td>“After what I had learned, it gives me a chance to practice”, “I thought it reinforced what we learnt”</td>
</tr>
<tr>
<td>Immediate feedback</td>
<td>The use of the interactive assessment gave students immediate feedback on their learning and application of statistics</td>
<td>21%</td>
<td>“Being able to see want you were doing instantly was good”, “You could tell whether you were getting a right or wrong answer and could see whether you were on the right track or not”.</td>
</tr>
<tr>
<td>Created interest</td>
<td>The use of interactive assessment increased student interest in statistical concepts</td>
<td>18%</td>
<td>“I found them interesting”, “It was a bit more interesting”, “They were unusual”</td>
</tr>
<tr>
<td>Technology</td>
<td>The use and effectiveness of the interactive assessment was influenced by the technological requirements to access and use them</td>
<td>16%</td>
<td>“It’s on the computer; some people are good with that and some people could be hindered by it”, “It was easy to log into the system, I didn’t have any problems technology wise”, “I definitely would use them if I had easier access”</td>
</tr>
<tr>
<td>Visual aid to learning</td>
<td>The interactive assessment helped students to better visualise the statistical concept being learnt</td>
<td>13%</td>
<td>“By doing the graphs it makes it more visual which helps with learning”, “It was good that you were able to see the graphs being made and everything and how it came up quite graphically”</td>
</tr>
<tr>
<td>Practical application</td>
<td>The use of interactive assessment increased the connection between statistics and real concepts and showed examples of practical applications</td>
<td>13%</td>
<td>“It helped put the theory into practice”, “It just made it more materialized and more hands on”</td>
</tr>
<tr>
<td>Engagement</td>
<td>The interactive nature of the learning approach engaged students</td>
<td>10.5%</td>
<td>“It was much more interactive learning”, “You are actually getting to muck around with it and play with it and it was much more interactive learning”</td>
</tr>
<tr>
<td>Different to exam or assignment</td>
<td>The interactive assessment was unique and had not been encountered before as an assessment task in this or other courses</td>
<td>8%</td>
<td>“Because it was interactive it was something different to exams and assignments”, “It was probably a bit less stressful than having to do an exam or assignment”, “It was good that it was another way to learn and test your skills”</td>
</tr>
<tr>
<td>Easy to do</td>
<td>The interactive assessment was a relatively easy assessment task in the context of the course</td>
<td>8%</td>
<td>“They are easy to do”, “It didn’t seem so overwhelming”, “It’s helpful and easy marks”</td>
</tr>
</tbody>
</table>

Table 1  The labels, definitions, percentages of sample that contributed statements, and representative comments for the final themes resulting from the qualitative data coding methods
The second most common theme was labelled Motivation with 26% of students endorsing this view. The motivation seemed to be partly attributed to the way in which the interactive assessment contributed towards their course grade. However, students accessed the interactive assessment topics progressively throughout the semester and some students were motivated to complete the assessment at the same times as they were covering the topics in class. Other themes that may be related to motivation were Created Interest (18%) and Engagement (10.5%). The former of these themes indicated that the task was more interesting than the typical exam or assignment-based forms of assessment. This increased interest may have encouraged students to attempt the assessment and to complete it. The theme of engagement was largely attributed to the interactive nature of the assessment. Students commented on the ability to change values and manipulate statistics as aspects that increased engagement.

Issues concerning the use of Technology were identified by 16% of students. While the statements were related to the same theme, they contained a mixture of both positive and negative comments. Some students commented that the on-line nature and ease of logging on to the system were positive aspects to the approach. Other students commented that it required access to the Internet and that they needed to be competent with using computers in general or with using computer-based learning tasks in particular. In addition, the interactive assessment was designed for compatibility with the computers that were available for student use on the university campus. The computer-based interactive simulations did not always work on a personal computer owned by a student due to incompatibilities with the web browser or the version of JAVA that was installed on the computer.

The final two themes were that the interactive assessment was Different to exam or assignment (8%) and Easy to do (8%). The interactive nature of the assessment was said to be different to the standard exams and assignment. It was also commented that it was less stressful than sitting an exam or submitting an assignment. The ability to repeat an exercise if the pass criterion of 8 out of 10 was not achieved may have also reduced stress. The ease in completing the interactive assessment was also contrasted with the exams and assignment. The pass/fail nature of the assessment, use of multiple-choice question format, and the ability to complete the assessment at their own pace may have been contributing factors towards the perceived ease of the assessment in students.

5 DISCUSSION

The assessment approach investigated in the present study used computer-based simulations that were designed to be interactive and encourage students to engage in deep processing of information. The analysis of the qualitative data collected from the interviews showed several themes suggesting that the assessment method achieved these aims and had additional benefits for students' understanding and application of statistics. Themes that emerged indicated that the interactive assessment had benefits for student learning (Helped understanding, Practice, Immediate feedback, Visual aid to learning), student motivation and engagement (Motivation, Created interest, Engagement, Different to exam or assignment), among other issues (Technology, Easy to do). Some of these themes were similar to those identified in the preliminary investigation by Neumann et al. (2008) (that is, Helps learning, Practice concepts, Alternative learning tool, and Helps confidence). However, an overall larger number and greater diversity of themes emerged in the present study and this may be due to the interview procedure used. The one-on-one interview approach used here was more dynamic in that it could seek clarification and elaboration from students as relevant. The sampling methods also ensured a more representative group of students were interviewed and this may have also produced a wider diversity of themes.

The most common theme that was identified indicated that the interactive assessment helped students to understand and learn about statistics. Although this theme was also identified by Neumann et al. (2008), one might consider that the ability of an assessment task to also promote learning is somewhat paradoxical. However, it has been noted that computer-based simulations encourage an active learning approach and promote problem-solving skills (Bowker, 1988). Moreover, the present interactive assessment approach was developed based on the principles of combining a declarative knowledge of statistics with functional knowledge gained from interacting with the simulations. Through the interaction, students were exposed to new experiences and were forced to apply their knowledge in new ways. This generated new problem solving challenges that they had to overcome and this process of knowledge generation seemed to have increased their understanding and application of statistics. Assessment need not merely test students on knowledge that they have already memorised, but can test their ability to apply what they know in novel situations. In addition, the option to reattempt the questions if the student fell short of the pass criteria may have encouraged them to think about what questions they could have answered wrong and check their declarative knowledge before reattempting the assessment.

Part of the increase in understanding that followed from the interactive assessment task could be due to the visual nature of the simulations. Some students endorsed the view that the interactive assessment was a visual aid to learning. It is possible that these students are those that benefit more from a visual form of learning than from an auditory form. Some students also noted that the interactive assessment allowed them to practice their application of statistics. In this respect, it is important that they were able to gain immediate feedback on what they were doing either in the form of the mark obtained on the assessment task or by viewing the interactive simulation in action and thereby seeing whether their ideas were correct or not. Immediate feedback on performance would be particularly important in online teaching in which the student may not be able to immediately ask the teacher questions or receive confirmation from them on their performance on the task.
Statistics courses can elicit negative attitudes and high anxiety, particularly in students from the social sciences (Onwuegbuzie, Slate, Paterson, Watson and Schwartz, 2000; Tremblay, Gardner and Hiepel, 2000; Neumann and Hood, 2009). The present sample of predominantly psychology students would be expected to also share these sentiments. The findings that the interactive assessment was reported to increase motivation, create interest, and were engaging are positive aspects to the approach. While the motivation associated with the assessment may be, at least in part, attributed to the link with the overall grades for the course, the themes of created interest and engagement seem to reflect other factors. The interactive nature of the computer-based simulations and their novelty seem to be important. A more interesting and engaging form of assessment is likely to encourage students to complete the required tasks and have a positive benefit on overall student grades.

It is unlikely that any assessment approach will be free of problems or disadvantages. Instructors need to weigh up the potential positive and negative aspects of any assessment approach as well as its suitability in relation to assessing learning outcomes (Biggs and Tang, 2007). Technological issues were noted by some students as important in completing the interactive assessment. Problems such as not being able to easily log on to the system, computer failure, and simulations not working were mentioned by students and are practical technological challenges that need to be met when employing this assessment approach. The relative ease of the interactive assessment may be another potential disadvantage of the assessment as implemented in this study. Ideally, the different assessment items in a course should be of comparable difficulty.

Recent years has seen an increased use of on-line technology in summative assessment at university. These approaches have included the use of on-line examinations (for example, Fluck, Pullen, and Harper, 2009), wikis (for example, Trentin, 2009), and open on-line assignment submission (for example, Geri and Naor-Elaiza, 2008). The use of computer-based interactive simulations in assessment as evaluated in this study can be added to this list. Although not implemented as an on-line system, the use of computer simulations in the assessment of science knowledge in sixth grade students has shown that students perform equally well with the computer-based assessment as they do with assessment that uses concrete materials (Baxter, 1995). Other studies have also shown that students perform as well with computer-based and paper-and-pencil examinations (Russell and Haney, 1997; Wolfson, Velosky, Robeson and Maxwell, 2001). The present application of computer simulations in assessing statistical knowledge is similarly not expected to disadvantage students any more than if it was conducted using paper-and-pencil. On the contrary, the automated marking and feedback, and the multimedia types of questions that have been identified as advantageous in computerised assessment (Cantillon, Irish and Sales, 2004) exist with the present assessment approach. As such, the use of computer-based interactive simulations in assessment may prove to be a useful addition to the assessment practices in courses on statistics and other disciplines in mathematics.

Author Notes

We acknowledge the assistance of Tiffany Middleton and Sandor Heng during participant recruitment and data coding.

REFERENCES


Biographical Notes

David Neumann is an Associate Professor in the School of Psychology, Griffith University. He has taught statistics courses in psychology and business and has research interests in improving the teaching of statistics to non-mathematicians through the use of technology and novel teaching approaches such as humour and using student data.

Michelle Hood is a senior lecturer in the School of Psychology, Griffith University. She has taught research methods and statistics courses to psychology students and has research interests in the use of innovative teaching practices in statistics, including web-based media and on-line delivery of classes.

Michelle Neumann is a recent PhD graduate in the School of Psychology at Griffith University. She is a registered school teacher that has taught K-12 mathematics and has research interests in early childhood numeracy and literacy.