Experiences of Assessment in Data and Security Courses using Personal Response Systems

ROSANNE ENGLISH, Department of Computer and Information Sciences
University of Strathclyde
JOSEPH MAGUIRE, School of Computing Science
University of Glasgow

This paper details an experience report of two interventions which explored the use of a audience response system in summative assessment in two different ways within a conversion Masters degree programme. One course explored students understanding of topics and self-assessment of ability through small multiple-choice quizzes. The other course was based around cyber security and used the audience response system to ensure engagement with the pre-class reading material. Both interventions were designed in an attempt to encourage students to engage more effectively with the material. This paper aims to identify and contrast the ways in which the audience response system was used in assessment in higher education computing science with a view to suggesting key considerations for implementing such an intervention.

1 INTRODUCTION

When considering student learning we ideally want students to become deep rather than shallow learners as discussed by Biggs [2]. We also want students to be reflective and critical learners. However, in higher education it's clear that although some students will engage in different and innovative approaches to learning, many will only choose to spend time on something if it counts towards the final grade for the course. This can mean, depending on the design of the course, that some students progress through a class feeling satisfied and assured that they understand the material and are engaging in it. This could be contrary to the belief of the lecturer teaching who may see this engagement as only engagement with the coursework, and not the course as a whole.

One way of increasing engagement with the class is to employ audience response systems (also known as classroom response systems and personal response systems amongst other names [5]) which allow the class to answer questions live during a lecture or teaching slot. This can be seen in works such as those by Martyn [19], Cutts *et. al* [8], and Lopez-Herrejon *et. al* [17]. However, what appears to be less well explored is the use of audience response systems in assessment for computing science. In an aim to address this, this paper reports two interventions in computing science courses using a personal response system for summative assessment in two different ways. The first class was an enterprise cyber security where students were asked to complete reading prior to attending class. To motivate

Authors' addresses: Rosanne English, Department of Computer and Information Sciences
University of Strathclyde, Glasgow, Scotland, G1 1XQ, rosanne.english@strath.ac.uk; Joseph Maguire, School of Computing Science
University of Glasgow, Glasgow, Scotland, G12 8QQ, joseph.maguire@glasgow.ac.uk.

Manuscript submitted to ACM

the reading, a quiz using the personal response system was issued for each of the papers. The motivation was not to determine great depth in understanding of the papers, but instead to promote engagement by reading them.

The second class was a databases class, which in the students reported satisfaction and had a normal distribution of grades for coursework. However, when presented with the unseen written exam students struggled significantly. Subsequent discussions with students identified a possible contributing factor was a perception that they understood the material. Another contributing factor identified was that aside from the coursework, students were generally disengaged from the material.

In an attempt to encourage students to engage more effectively and provide tools with which students could self-assess an intervention was designed. The intervention took the form of regular assessed multiple choice quizzes on the subject material using a personal response system.

This paper aims to compare and contrast these two uses of a personal response system in summative assessment to achieve increased engagement with the material.

2 BACKGROUND

Audience Response Systems (ARS) have many different names, but essentially comprise of a central system collecting responses from handsets. Such systems have an established history in education with oscillating popularity over decades [15][11]. The early use of PRSs in higher education required specialised hardware installed in lecture theatres collecting responses from bespoke handsets that were issued to students [4]. The approach has steadily evolved into collecting responses from dedicated applications on student smartphones [14].

Tivener and Hetzler report students demonstrated improved knowledge acquisition on an introductory athletic course that used regular quizzes delivered using an audience response system [25]. Similarly, Ma *et al.* report student performance was stronger on food science courses when using an audience response system to complete quizzes probing lecture content [18]. Song *et al.* report that student engagement and interaction was improved on language courses that utilised audience response systems for student generated questions [24].

Liu *et al.* reviewed 128 empirical studies exploring the use of audience response systems in various settings, including lectures and cooperative learning [16]. They argue that lecturers often focused on affect, achievement and participation with quizzes being the most common approach in lectures. Similarly, White, Syncox, and Alters reviewed use of ARSs and suggest quizzes are the most common application in lectures [27]. They argue four common incentives are earning grades through (1) testing and attaining the accurate answer, (2) participation where action is rewarded, (3) testing and participation with accurate answers are rewarded and (4) responding with no reward. However, Trees and Jackson argue that incentives are an important factor in the use of ARSs and non-incentives should be avoided [26].

The use of ARSs to deliver quizzes is not surprising. Dochy argues that quizzes can provide formative feedback to both the learner and educator in efficiently and effectively [10]. The use of quizzes is often valued by students as they can motivate students to engage with course content [22].

McNulty et al. reports that providing several formative quizzes were able to predict performance on summative assessments [20]. McNulty et al. warn though that only small performance benefits were perceived from providing multiple quizzes versus a single quiz. Adkins and Linville reports that while increased frequency of testing may not increase performance, students report less anxiety as they are receiving more feedback [1]. Similarly, Ngai state student students reported less anxiety and frustration through the effective use of student self-assessment and quizzes to support students in attaining the expected outcomes of an introductory programming course [21].

Nevertheless, McNulty et al. argue though that if unstructured, students will perform quizzes only a few days before summative assessments [20]. Delaram et al. report nursing students utilising structured weekly quizzes demonstrated stronger performance on summative assessments [9]. Similarly, Brusilovsky and Sosnovsky report that performance on summative assessments was improved through regular self-assessment quizzes on an introductory programming course [3].

However, Giest and Soehren report that while quizzes are valuable in enhancing short-term performance and improved course evaluation, there are not necessarily long-term benefits to knowledge acquisition [23]. However, this does not mean that quizzes can not be valuable in supporting deeper engagement within a course. Clark et al. suggests lecture attendance can be improved with the use of quizzes, but the impact comes from students appreciating the importance of attendance, rather than the quiz itself [6]. Consequently, quizzes can be used effectively to get students to participate and engage in important content within courses.

This is a fraction of the available literature on quizzes and ARS. It is clear from the research that audience response systems have the potential to effectively and efficiently support quizzes in lectures. Quizzes can be utilised differently depending on the course. Two prominent approaches are (1) continuous feedback on course content and (2) motivating students to engage with supplementary course content. The remainder of this paper reports experiences with those two approaches, and concludes with suggested considerations for educators planning such interventions.

Both courses were delivered to the same cohort, a conversion Masters course designed for students with a good undergraduate degree in a non-computer science subject such as History, or Physics for example. The class size for the Enterprise Cyber Security course was 175, and for the Databases course the class size was 107. The majority of students in these courses have little or no computing science background, though there are a small number with some experience in elements of computing science professional or otherwise.

3 CLASSROOM RESPONSE FOR DEPTH OF UNDERSTANDING

This section explores how the personal response system was used in summative quizzes for a databases course. The course employed interactive lectures which use the approach defined as interactive windows by Huxham [13]. In this approach, students are presented with short bursts of a traditional lecture followed by hands-on practical exercises. In addition to the interactive lectures, tutorials (where problem sheets were worked through by students) and labs (with at computer exercises to further embed the material) were delivered. The material for the course was split into ten distinct topics such as basic SQL commands, normalisation, and database design. Each of these topics was taught in a single week, and the following week presented a quiz on that topic before the next was started. Each quiz had four unseen multiple choice questions. This resulted in a total of 40 questions which tested student comprehension. The quizzes were worth a total of 20% of the students mark for the course.

Before each quiz took place, students were asked to rate how well they felt they understood the topic. They were asked to complete this in line with the marking system of the university. A meant excellent, B very good, C good, D satisfactory, and E was poor. After completing the questions for that topic, students were asked once more how well they felt they understood the topic. The intention behind these questions was to get students to reflect on their understanding. After completing the quiz in class, a discussion of the correct answers and the reasons the distractors were incorrect took place.

Students could participate with the quizzes either using the University provided personal response system, which allowed them to use their own internet connected devices, or using paper if they had no device. As the time was short for the quizzes and the questions had to be shown on a lecture slide, it seemed appropriate to reduce cognitive load on

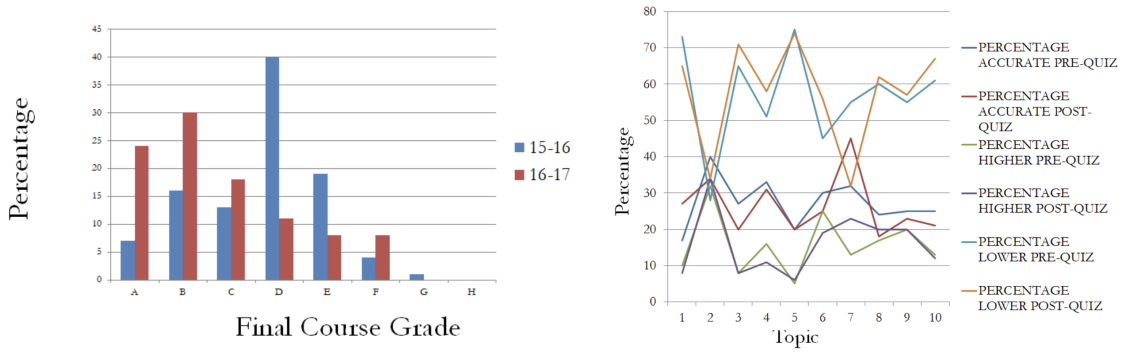

Fig. 1. (A) Left: Final Grade Differences, (B) Right: Self-assessment Over Time

students, so a range of scenarios were constructed to centre the questions around. When constructing questions, care was taken to select distractors which reflected typical mistakes, and concepts which were known to be a challenge for students. For example, the distinction between a fan trap and a chasm trap is often a difficult concept for students. This made it more challenging for the student, and more likely to assist them in self-evaluation and correction of misconceptions.

3.1 Results

Does students' self-assessment of their ability match their performance? In most instances this was due to students who were rating themselves as lower than their ability initially as illustrated in Figure 1 (B), reducing their grade further after completing the quiz. This indicates that perhaps those students had issues with confidence in their ability. In contrast those who were indicating a higher grade than their ability tended to stick with their self-assessment with the exception of two topics. These topics were data definition in SQL and query optimisation. It is likely that the distractors for these topics were more obviously incorrect after completing the quiz.

Did they get better at judging their abilities? As one can probably infer from the previous section, students generally did not improve in their self-assessment of their ability. This can be seen in Figure 1 (A) where we would expect to see the percentage accurate post-quiz to be higher than the percentage accurate pre-quiz if this was assisting students to become better at self-assessing their abilities. Despite this aspect of the quizzes note being as effective as hoped, it still had benefit which was evidenced by both an increase in marks when compared to the previous year, and also student feedback. These aspects are discussed in the following two sub-sections.

3.2 Student Feedback

After running the course, students were asked to give feedback on the quizzes. 15 out of 107 students completed the feedback, so the sample size is too small to draw any conclusions. They were asked to rate the following statements on a 5 point Likert scale from 'Strongly Disagree' to 'Strongly Agree'.

Of those who responded only one student said they were neutral about whether the continuous assessment through the quizzes helped, the other 14 identified as agreeing (4) or strongly agreeing (10). This response generally reflected the overall informal feedback from students who noted on several occasions how much they liked them.

The results on whether the process helped students reflect on their understanding were less clear cut. Whilst 11 out of 14 either strongly agreed or agreed, 3 were neutral and 1 disagreed. Some of the feedback around this question noted that some students always thought they were not good (as demonstrated in the results of self-assessment earlier), and others felt they could gauge their own understanding without use of the quizzes.

The results for whether students felt they improved at self-assessment as they progressed 12 out of 15 agreed or strongly agreed, 2 were neutral and 1 disagreed. Those who were neutral or disagreed felt this was down to the specific questions as to whether they could accurately self-assess their understanding.

In addition to the Likert scale questions, participants were also asked if they would prepare for these quizzes even if it did not contribute to their final mark. Students were given the choices "yes", "no", and "maybe". 20% of respondents said they would, 46.67% said maybe, and 33.33% said they would not. This perhaps hints at there being less engagement if the quizzes are not assessed.

Overall students responded well to the quizzes, the main point of contention being the weighting which was set at 20%. The intention had been to reflect the effort required to keep up with the material, but students generally felt that 20% was too much for multiple choice questions and 10% as with the enterprise cyber security class was deemed more appropriate. As a lecturer it seemed to benefit the students learning as they were asking more insightful questions much earlier in the course, and not just before the exam as had been the case in previous years.

4 PERSONAL RESPONSE SYSTEM FOR ENGAGEMENT WITH PRE-CLASS READING

4.0.1 Motivation. The course expected students to complete reading of research papers which expanded on content presented in the course. The expectation is this would legitimise course content as students could observe these were on-going problems, cement and deepen understanding of course content, and familiarise students with research papers.

In the initial few weeks of the course class discussions would support students in reading research papers. It became clear in these sessions that many students were not reading the material even though they were expected to prepare for class. Moreover, formative and summative assessments often lacked depth that could be observed in the research papers.

Consequently, quizzes that probed reading of researching papers were employed to motivate students.

4.0.2 Practice. A reading list of research papers was released to the cohort prior in the introductory session of the class. Students were advised that each week in the same session they would need to complete a quiz that probed reading of the research paper. The students were advised of the specific research paper the week prior to the quiz probing it.

Students had to complete 11 quizzes to earn course credit. The lowest scoring quiz was discounted, effectively determining performance over 10 quizzes. The motivation for including this discount was to effectively deal with situations where students had an 'off-day'. Each quiz compromised seven multiple choice questions with four options. Students had 90 seconds to answer each question via the ARS accessible via smartphones. If a student did not have a smartphone they were permitted to use paper.

4.0.3 Experience. The motivation for the quiz was to was timely engagement with research reading. The expectation was that performance would either oscillate between students largely performing well as they had completed the research reading or poorly as they had not reviewed the paper.

Performance on the quiz itself was not of particular interest as the focus was in enriching in-class discussion. In comparison to the previous session where quizzes were not utilised students were noticeable better prepared for in-class discussion. The primary difference was that previous in-class discussions would comprise of very few students

participating. In contrast, when quizzes were employed many more students appeared prepared and participated in the class discussions. This presented a problem in that often discussion had to be curtailed due to time constraints.

The standard course evaluation procedures revealed many positive responses to the use of research readings and the requirement to complete weekly quizzes. The primary remark was that the quizzes were an effective tool that motivated students to read research papers and participate in discussions.

5 LESSONS LEARNED

Whilst the motivation behind each of the two approaches above was different, one to encourage deeper learning and reflection and the other to encourage more engagement with pre-class reading, both had a number of similarities in terms of implementation of the intervention. Here we present these key lessons which we hope to be beneficial for those considering employing a similar approach.

5.1 Motivation

It is important to consider the motivation for the quizzes. Here we presented two possible motivations - assessment for learning, and assessment to motivate engagement. There are many other possible motivations such as assessment of learning. We believe that whatever your motivation, you should consider it carefully and clearly communicate it to the students. It was important students understood the motivation so they did not confuse something to promote course engagement with a final exam. This could be most clearly seen in the enterprise cyber security course where initially students confused the quizzes with an attempt to ensure they recalled the data of the papers such as "What does X say on page 4 " which was much more specific than the intention. It quickly became apparent to those students that the structure was more concerning the highlights of the papers, such as the main theme.

5.2 Preparation

Not to be underestimated, preparation for such an intervention is key. Preparing the questions for the database course required consideration of the common misconceptions and misunderstandings to ensure distractors which would challenge the students understanding of the material and assist in demonstrating where such knowledge was lacking. This applies even if students got the multiple choice question right. For example one student reported " I spent a couple of hours on joins last night and I thought I understood them, I managed to get them all right, but it was total fluke. Now I know I need to spend a lot more time on them!" This feedback confirmed the motivation of assessment for learning was working, as some students were identifying much earlier than the final exam that their knowledge was not as strong as they believed. The process of creating 40 such questions took approximately two weeks concentrated effort over the summer.

5.3 Practice

The ARS is key to ensuring the process works. It is important that you are able to ensure with some degree of certainty that a specific user is a given student due to the responses being used to determine a summative grade. The system used for these interventions was purpose built at the institution and allowed students to use their university account to authenticate. Thus the connection between student and responses could be accurately made. The first time such interventions were run was the first time in using the system. This meant both lecturers had to practice with the system to ensure they knew how it worked. It was also essential when students started the class that they became familiar with the system as well. This meant having practice quizzes which were not summatively assessed. This served two

purposes. Firstly, to ensure the system could run with the size of class in real time. Secondly, to ensure the students were comfortable using the system before the results contributed to their final grade for the course.

5.4 Logistics

There are a number of logistical aspects to consider for such an intervention. You should ensure the room is sufficiently large to seat all students ideally with room to spare. Unlike a standard lecture where many students may choose not to attend, when there is an assessment taking place most, if not all students will come. The room needs to be large enough to accommodate all students, including those who arrive late. It is also essential to ensure that the wi-fi signal is sufficiently strong in the allocated room. Many modern personal response systems such as Mentimeter and Meetoo rely on such a connection. It's best to ensure you have paper alternatives to accommodate any students who may not have access to an internet enabled device for a class. This also helps in any technical failures. Another aspect to consider is timing. How long will you give the students to complete the quiz - will you provide a timer which is consistent for each question, or be more flexible depending on the response. It is interesting to note that in both cases, the lecturers noticed many students would not submit their answer until the final seconds before the question finished. In addition, it is advisable to select the same time for the quizzes and advertise this to the students. Another situation to examine is if there are any students who may require additional time due to reasonable adjustments. It may be required to provide support for doing the quiz in a separate room.

6 DISCUSSION

Whilst both interventions provided feedback to the lecturer, and to the student, it did not necessarily promote engagement between members. The students had to sit in silence, and did not interact with each other. Another possible limitation is the use of multiple choice questions. The system was limited in that questions had to be structured as multiple choice questions, which may not be the optimal format as proposed by Couch who noted that multiple choice was limited in terms of determining the success of the mechanism [7]. However, Hubbard argues student performance on later exams depends on the initial performance on the quizzes, not the question format [12]. Despite these and other potential flaws in design of the intervention, it was felt that both students and staff responded well to the approaches. It appeared as though the students engaged more in the material for both classes much earlier on, and feedback from them was positive. Given appropriate consideration of ARS using aspects discussed in this paper it is hoped that others may learn from our lessons and employ ARS for assessment in computing science without such a large learning curve.

REFERENCES

[1] Joni K Adkins and Diana Linville. 2017. Testing frequency in an introductory computer programming course. *Information Systems Education Journal* 15, 3 (2017), 22.

[2] J. B. Biggs. 1993. From Theory to Practice: A Cognitive Systems Approach. *Higher Education Research & Development* 12, 1 (1993), 73–85. https://doi.org/10.1080/0729436930120107

[3] Peter Brusilovsky and Sergey Sosnovsky. 2005. Engaging students to work with self-assessment questions. *ACM SIGCSE Bulletin* 37, 3 (2005), 251. https://doi.org/10.1145/1151954.1067514

[4] Ray A Burnstein and Leon M Lederman. 2001. Using wireless keypads in lecture classes. *The Physics Teacher* 39, 1 (2001), 8–11.

[5] Jeff Cain and Evan Robinson. 2008. A primer on audience response systems: current applications and future considerations. *American Journal of Pharmaceutical Education* 72, 4 (2008), 77.

[6] Gordon Clark, Nick Gill, Marion Walker, and Rebecca Whittle. 2011. Attendance and performance: Correlations and motives in lecture-based modules. *Journal of Geography in Higher Education* 35, 2 (2011), 199–215.

[7] Brian A Couch, Joanna K Hubbard, and Chad E Brassil. 2018. Multiple–True–False Questions Reveal the Limits of the Multiple–Choice Format for Detecting Students with Incomplete Understandings. *BioScience* 68, 6 (2018), 455–463.

[8] Quinton Cutts, Angela Carbone, and Kelsey Leigh Van Haaster. 2004. Using an electronic voting system to promote active reflection on coursework feedback. (2004), 1251 – 1258.

[9] Masoumeh Delaram, Sahand Shams, and Hamidreza Sadeghi Gandomani. 2017. The Effect of Quizzes on Test Scores of Nursing Students for Learning Maternal and Child Health. *Journal of Medical Education* 16, 2 (2017).

[10] Filip Dochy, Catherine De Rijdt, and Walter Dyck. 2002. Cognitive prerequisites and learning: How far have we progressed since Bloom? Implications for educational practice and teaching. *Active learning in higher education* 3, 3 (2002), 265–284.

[11] Carmen Fies and Jill Marshall. 2006. Classroom response systems: A review of the literature. *Journal of Science Education and Technology* 15, 1 (2006), 101–109.

[12] Joanna K Hubbard and Brian A Couch. 2018. The positive effect of in-class clicker questions on later exams depends on initial student performance level but not question format. *Computers & Education* 120 (2018), 1–12.

[13] Mark Huxham. 2005. Learning in lectures: Do 'interactive windows' help? *Active Learning in Higher Education* 6, 1 (2005), 17–31. https://doi.org/10.1177/1469787405049943

[14] Isabel Hwang, Kevin Wong, Shun Leung Lam, and Paul Lam. 2015. Student Response (Clicker) Systems: Preferences of Biomedical Physiology Students in Asian Classes. *Electronic Journal of e-Learning* 13, 5 (2015), 319–330.

[15] Eugene Judson and Daiyo Sawada. 2002. Learning from past and present: Electronic response systems in college lecture halls. *Journal of Computers in Mathematics and Science Teaching* 21, 2 (2002), 167–181.

[16] Cui Liu, Sufen Chen, Chi Chi, Kuei-Pin Chien, Yuzhen Liu, and Te-Lien Chou. 2017. The effects of clickers with different teaching strategies. *Journal of Educational Computing Research* 55, 5 (2017), 603–628.

[17] Roberto Lopez-Herrejon and Morrie Schulman. 2004. Using interactive technology in a short Java course: An experience report. *ACM Sigcse Bulletin* 36 (09 2004), 203–207. https://doi.org/10.1145/1026487.1008051

[18] Sihui Ma, Daniel G Steger, Peter E Doolittle, and Amanda C Stewart. 2018. Improved Academic Performance and Student Perceptions of Learning Through Use of a Cell Phone-Based Personal Response System. *Journal of Food Science Education* 17, 1 (2018), 27–32.

[19] M Martyn. 2007. Clickers in the classroom: An active learning approach. *EDUCAUSE Quarterly* 30 (01 2007), 71–74.

[20] John A McNulty, Baltazar R Espiritu, Amy E Hoyt, David C Ensminger, and Arcot J Chandrasekhar. 2015. Associations between formative practice quizzes and summative examination outcomes in a medical anatomy course. *Anatomical sciences education* 8, 1 (2015), 37–44.

[21] Grace Ngai, Winnie W.Y. Lau, Stephen C.F. Chan, and Hong-va Leong. 2010. On the implementation of self-assessment in an introductory programming course. *ACM SIGCSE Bulletin* 41, 4 (jan 2010), 85. https://doi.org/10.1145/1709424.1709453

[22] Ana M. Pinto-Llorente, M. Cruz Sánchez-Gómez, and Francisco J. García-Peñalvo. 2017. The Use of Online Quizzes in Blended Learning Modality. In *Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM 2017*. ACM Press, New York, New York, USA, 1–7. https://doi.org/10.1145/3144826.3145430

[23] SE Soehren et al. 1997. The effect of frequent quizzes on short-and long-term academic performance. *Journal of Dental Education* 61, 4 (1997), 339–345.

[24] Donggil Song, Eun Young Oh, and Krista Glazewski. 2017. Student-generated questioning activity in second language courses using a customized personal response system: a case study. *Educational Technology Research and Development* 65, 6 (2017), 1425–1449.

[25] Kristin Ann Tivener and Tona Hetzler. 2015. The effects of an electronic audience response system on athletic training student knowledge and interactivity. *Athletic Training Education Journal* 10, 3 (2015), 212–218.

[26] April R Trees and Michele H Jackson. 2007. The learning environment in clicker classrooms: student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology* 32, 1 (2007), 21–40.

[27] Peter White, David Syncox, and Brian Alters. 2011. Clicking for grades? Really? Investigating the use of clickers for awarding grade-points in post-secondary education. *Interactive Learning Environments* 19, 5 (2011), 551–561.