Helen M. Roe[1] | David P. Robinson[2]

1: School of Geography, Archaeology and Palaeoecology, Queen's University of Belfast

2: Media Services Unit, Queen's University of Belfast

The use of Personal Response Systems (PRS) in multiple-choice assessment: benefits and pitfalls over traditional, paper-based approaches

Abstract

In this paper we discuss our experiences of using wireless keypads for multiple-choice (MCQ) assessment and evaluate the benefits over traditional, paper-based MCQ assessment methods. A Personal Response System (PRS), *TurningPoint*®, was trialled for two years for a Geography undergraduate class test, designed to provide both formative and summative assessment. The benefits of PRS are considered in terms of i) student performance; ii) students' experiences, which were assessed by questionnaires; iii) quality of feedback; and iv) staff time inputs. Of the students surveyed, 74% stated that they strongly preferred PRS over paper-based MCQ assessment, whilst 6% of students disagreed. Students praised the immediacy and graphical nature of feedback associated with PRS, but some criticised the inflexibility of PRS to change answers after a fixed 'per question' time interval. Staff time inputs were substantially reduced (ca. 75% less) for PRS than for the equivalent paper-based test. PRS was found to be a useful tool in MCQ assessment, but it has some disadvantages. We highly recommend PRS for formative assessment but caution its use for summative assessment in its current form and recommend that it only be used for assessments carrying a small mark-weighting.

Introduction

Studies over the last decade have repeatedly shown that 'clicker' or Personal Response System (PRS) technologies offer many benefits in higher education. They can be used to test students' knowledge and understanding (Dufresne *et al*., 1996), help students to engage interactively (Boyle and Nicol, 2003; Draper and Brown, 2004), and help students to learn concepts more easily and retain them for longer than students who sit learning passively (Wood, 2004). The anonymity of responses facilitated by PRS technology also allows staff to initiate class debate on sensitive topics that might otherwise be difficult to explore (Zhu, 2007). Indeed, the applications of PRS are so broad that in many undergraduate teaching institutions in North America the technology has led to a transformation in science teaching, particularly for small classes. There has been a move from a format dominated by lectures to more interactive, seminar-style courses that require students' active participation (Wood, 2004).

Many PRS systems are now available, and whilst each differs slightly in its capabilities, most comprise a hand-held keypad (a transmitter) which communicates with a receiver on a 'base' computer via radio or infra-red transmission. Presentation software (e.g. *Microsoft PowerPoint*®) is typically used for audience communication. Questions are projected on-screen and members of the audience 'click' their responses, whilst a keypad light or LCD screen-indicator signals that a response has been registered. The responses are then processed, typically within seconds, providing the audience with results charts.

Whilst the benefits of PRS for improving student engagement have been widely reported, there has been less emphasis in the recent literature on the utility of PRS as an assessment tool, particularly for summative assessment. This may reflect a number of factors, including: i) technical difficulties that have hindered the configuration of PRS systems for formal assessment; ii) concerns over the reliability of PRS technology for recording students' answers; and iii) a

reticence on the part of practitioners to explore the medium for summative assessment, possibly reflecting a combination of the above. Notwithstanding this, there are studies dating back to the 1970s that describe the use of classroom clickers for assessment, particularly in physics and the mathematical sciences (for a review see Judson and Sawada, 2002). However, it has been acknowledged recently that the potential use of PRS as a tool for formal assessment requires further research (Elliott, 2003). Given the relatively high costs associated with purchasing PRS, it is important that the full range of applications and limitations of each system are explored by users prior to investment. Like any technology, PRS can be used creatively or skilfully, bringing many teaching and learning benefits, or it can be used clumsily or destructively (Wood, 2004). Communication of experiences, both good and bad, are vital if practitioners are to refine and enhance this form of technology-supported learning.

In this paper we discuss the results of a two year pilot of PRS for a second year Geography undergraduate class test (multiple-choice) and consider the advantages and disadvantages over a more traditional, paper-based test format used in previous years. The class test forms 10% of the mark for a core module, '*Field and Research Techniques in Geographical Practice*' (GGY2024), taken by 100-150 students each year. The primary aims of the module are to develop students' field data collection and analysis skills, which are achieved via a week-long Mediterranean field course run during the Easter vacation. The class test is held a few days before the field course to assess students' knowledge of a series of background topics relating to the geographies of the Mediterranean. These are covered in lectures and help students better contextualise issues that are developed in the field. The test significantly improves student attendance at these preparatory lectures. Feedback is provided in the lecture after the test, taking the form of a question-by-question run-through of the results. The test assessment is thus both formative and summative.

Our motivation to trial PRS for the GGY2024 test was partly influenced by a desire to reduce staff marking time; in paper format, the multiple-choice (MCQ) test typically takes nearly two days to mark, and as this also needs to be completed before the field course, this puts pressure on staff. We also felt that the graphical nature of feedback associated with PRS might improve students' understanding and retention of concepts prior to field discussion.

For our study we used a PRS technology purchased by Queen's University's Media Services Department in 2005 (*TurningPoint*®, 2005-6 version). This software embeds into *Microsoft PowerPoint*® for audience communication. The system was used to run the test in two years, 2007 and 2008. In our evaluation of PRS we consider a range of factors, including its impact on student performance, students' experiences and perceptions of the PRS interface, and impacts on staff time. Each stage of the set-up and marking process was timed and compared with the equivalent time inputs for the paper test, whilst students' opinions were assessed via anonymous questionnaires. These included questions on i) the usefulness of a PRS practice session before the tests; ii) time allocation per question; iii) ease of using the keypads; iv) quality, speed and usefulness of feedback; and v) students' perceived impact of PRS on performance. Students were also asked to compare their experiences of PRS tests with paper-based MCQ tests. In 2007 44% of the class (n=62) responded; in 2008, 85% (n=85).

System set up and preparation

TurningPoint® software (version 1.1.2) was installed on a Macintosh computer (OSX 10.4.2) and MCQs were prepared as *PowerPoint*® slides. A number of routine steps associated with session set up were followed. These are summarised below to provide an overview of the process. Readers are referred to the *TurningPoint*® or other PRS system operational instructions for specific guidance with session preparation and data processing (*TurningPoint 2010*).

Steps:

1. *Preparation of a 'Participants' (class) list which includes: a) students' identification numbers; and b) a designated keypad number for each student.*
2. *Preparation of MCQ questions (one question per slide) and identification of the correct answer for each question.*
3. *Allocation of a time limit for each question.*
4. *Performance of a trial run using a small number of keypads (2-3) to check that the software is functioning correctly.*
5. *Inspection of all keypads to ensure they are functioning correctly.*

The keypads used for the test were the second generation of keypads issued with the *TurningPoint*® system (Figure 1). These have fewer functions than some newer handsets that are currently available. A major limitation of the early generation keypads is that it is not possible to customize or 'tag' them to users prior to a *TurningPoint*® session. It was thus

necessary to supplement Step 1b by attaching a paper label to each keypad with students' identification details. This rather primitive method of tagging allowed students to locate their keypads at the start of the test, which were distributed in alphabetical order on desks around the exam room. For more advanced PRS systems this step is not necessary; users can key in or 'authenticate' their identification details at the start of a session via the handsets. This process of 'paper label' tagging was time-consuming, accounting for *ca.* 1.5 hours of the total preparation time in 2007 and ca. 1 hour in 2008 when class size was smaller (Table 1). It nevertheless had one advantage in that by placing keypads in alphabetical order this effectively randomised student distribution in the exam room. This prevented friends from sitting near to each other and copying answers. Previous studies have shown copying or cheating can be an issue with PRS, particularly if keypads do not have a shield to cover keys (Zhu, 2007). The spacing issue is therefore important.

Figure 1: Picture of a *TurningPoint*® keypad and dongle receiver.

Question preparation

In preparing MCQ questions (Step 2 above), care was taken to limit the number of answer choices to five or less per question in accordance with 'best practice' guidelines (cf. Zhu, 2007 p. 5). Questions were devised so that there was only one correct answer per question. This answer format was partly constrained by software limitations. It was not possible to use some more complex answer formats with the software available, for example, asking students to rank or order their answers or enter an absolute numerical value in response to a question, although some of these formats are available with the newer version of *TurningPoint*® and with some other systems. These alternative answer-formats had not been used previously for the class test so their lack of availability was not a disadvantage. In total 20 questions were used for the one hour test.

An important consideration in the preparation of question slides is the amount of time to display, and allow students to answer, each question (Step 3). This issue gains greater significance for summative rather than purely formative assessment and so requires careful consideration. PRS systems invariably include a 'countdown' timer-facility that enables a fixed time to be allocated to each question or slide. After the time elapses, answering or 'voting' closes. Voting can also be closed 'manually' by staff when all students have provided an answer, which can be tallied on-screen as students 'click' their responses. A major bonus of PRS is that answers can be changed as many times as keypad users wish within the 'voting' period. We felt that it would disadvantage students to close voting when all students had 'clicked' a response, as it would be impossible to assess whether these represented students' final answers. We also recognised that if too short or long a time was allocated then this might pressure students into making a wrong answer or, conversely, promote an air of distraction in the exam room as students waited for the next question. As a compromise, 2 minutes was allocated to each question during the 2007 test. This was reduced to 1 minute and 30 seconds in 2008 following overwhelming feedback from the student surveys that 2 minutes was too long.

As the class occasionally includes students with visual disabilities, questions were read out loud as well as being displayed on screen. Over 90% of the students surveyed found this helpful. This took on average ca. 20 seconds per question and the 'voting' or answering period was activated after this, giving students a few additional seconds to consider their responses.

Software-related difficulties

A number of technical difficulties arose when implementing Step 3 in both years of running the test. These were associated with running the *TurningPoint*® software on a Macintosh platform. The test was thus transferred to a PC and no further problems were encountered.

Implementation

As few students registered for the module had any previous experience of PRS, a 15 minute practice session was run in a lecture prior to the tests to enable students to familiarise themselves with the interface (Figure 2). The majority of students surveyed stated that they found the session increased their confidence in using PRS, although a small number of students (3% in 2007, 9% in 2008) stated that the system was so easy to use that they did not need a run-through. This session also enabled keypads to be re-checked, augmenting Step 5. Reliability and issues of resistance to tampering have been cited as important factors in selecting a response system (Burnstein and Lederman, 2001). A small number of keypads were found to be non-functional during the practice sessions, resulting from battery problems, and were replaced.

Early in the 2007 test it was noticed that many students were answering questions very quickly (within 30 seconds) of voting opening. Whilst this implied that students were confident in their answers, it also indicated that some students may have had a poor perception of the time left. The number of students answering within 30 seconds was noted during both tests. In 2007 this was 62%; in 2008, 54%.

Data processing was carried out immediately after each test. In both years it took less than 30 minutes to produce a spreadsheet with marks for distribution to students. This represented a significant reduction in marking time over the previous paper-based test, which took over 11 hours to mark for a class of 150 students (Table 1).

Figure 2: Photograph of a student practice session with PRS.

Table 1: Data relating to the paper-based and PRS class tests, 2006-2008.

Year	2006	2007	2008
Number of students who took test	151	142	100
Format of test	Paper: MCQ	Personal response: MCQ	Personal response: MCQ
Number of questions left blank per student	0.03	0.06	0.07
% students who answered questions within 30 seconds after 'voting' opened	(unknown)	62%	54%
Average Mark	60%	51%	60%
Staff time to prepare test	2 hours, 30 mins	2 hours, 30 mins	2 hours, 15 mins
Staff time to mark test, check and collate results	11 hours, 30 mins	30 mins	30 mins
Staff time to display results in graph format	n/a	1 hour	45 mins
Total staff time involved	**14 hours**	**4 hours**	**3 hours, 30 mins**

The correct answers to the MCQs were relayed to students in dedicated feedback sessions using a correct answer display facility in *TurningPoint*®. Like most PRS systems, this system permits charts to be generated showing the percentage of students who have selected the correct and incorrect answers for each question (Figure 3). These provide an excellent platform for discussion, particularly when combined with standard graphics or text slides to remind students of the context of the questions. For example, for a question relating to karst geomorphology, images of the features in question could be displayed next to the chart.

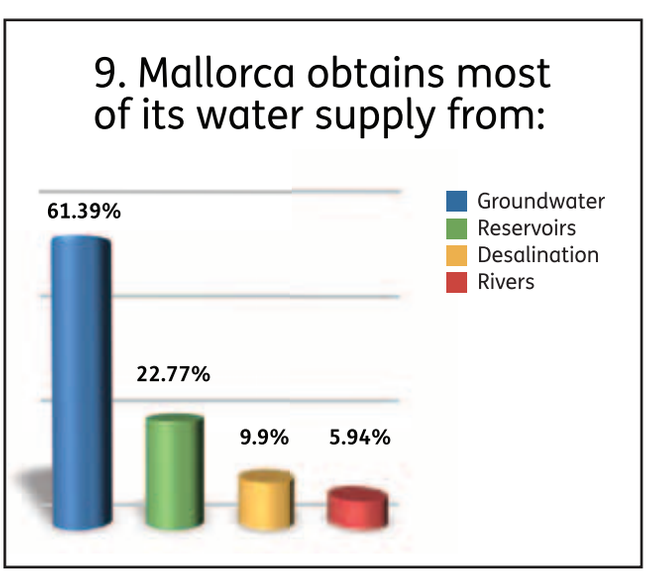

Figure 3: Example of a MCQ question and answer chart showing students' responses.

Student performance

Student performance during the two PRS test years varied, with a mean test mark of 51% in 2007 and 60% in 2008 (Table 1). For comparison, the mean marks in the two previous years when the test had run in paper format were 60% (2006) and 57% (2005). It is difficult to explain these variations as factors unrelated to test-format are likely to have influenced the results. For example, different questions were used each year which may have affected performance. In spite of this, we speculate that the dip in performance in 2007 may have been slightly influenced by students' lack of familiarity with the system and the fast speed with which many students answered the questions, perhaps without sufficient reflection on their answers. In the 2008 test, the importance of reflecting on answers before answering was stressed more fully at the start of the test and this may have contributed to the improved performance (and slower class answering time) recorded that year.

Interestingly, a slightly larger proportion of students failed to answer 100% of the questions with the PRS tests than the 2006 paper-test (Table 1). This is surprising, as students were well aware of the need to guess answers if they did not know them. We can only speculate that students may on rare occasions have thought that they had 'clicked' when they had not. They may not perhaps have checked the keypad indicator light when pressing the answer key, or they may have pressed a wrong key with no answer

allocated to it. Clearly if students fail to take sufficient care in the 'clicking' process then this severely limits the utility of PRS as an assessment tool. Whilst it is relatively easy to check that students have all answered a particular question with small classes, with a large class (>75 students) it is difficult to monitor which students have failed to 'click' during the test itself without incurring significant time delays so this was not attempted.

Students' opinions of PRS

In the surveys, a large majority of students (74% both 2007 and 2008) agreed with the statement 'I strongly prefer PRS format to a paper-based test'. When asked via an open question to state what they liked about PRS (Box 1), these students frequently cited the rapid distribution of marks and the graphical feedback. Some students noted that they liked the class response charts because they gave them reassurance that if they had got the answers wrong they were not alone. Many students commented that the technology was innovative and enjoyable to use.

"Great system, I really enjoyed using it. It was far better than doing the test on paper"

"I really appreciated getting my marks back so quickly"

"The answer graphs were very helpful. They will definitely help me remember the answers"

"The handsets were easy to use. A great technology - the way of the future. I wish we could use them more often"

Box 1: Examples of positive comments from students on PRS.

A smaller proportion of students (19% in 2007; 18% in 2008) agreed with the statement 'I prefer PRS to a paper test, but I have some reservations'. These students typically stated that they would have felt happier with PRS if they were more familiar with using it. Some also commented that they would have liked to have had reassurance that their answers had definitely been recorded by the base computer. Only a small proportion of students (6% in 2007 and 2008) agreed with the statement 'I prefer paper-based tests'. Most of these students cited the lack of ability to change their answers after the 'per-question' voting period had closed as their cause for concern.

"I think it's a good system, however, I like to look back at my answers to make sure I'm happy with them. This cannot be done with the digital system, and I was unsure of the way I left some answers, with a paper test I could have gone back"

"It didn't have the feeling of a proper test so I didn't work as hard as I would have done for a conventional test"

Box 2: Examples of negative comments from students on PRS.

In terms of students' perceptions of their test performance, the majority of the students surveyed (>90% in both years) agreed that they thought that the test format made no difference to their marks. This finding, together with the lack of a clear improvement in the actual performance between the PRS and paper-test years, confirms that students were not incentivised to work harder by the PRS technology. Indeed, for a minority of students the test format may have distracted them from focusing on their answers, as we mentioned above. This is also reinforced by the fact that a small percentage (<7%) of the survey respondents stated that the lack of familiarity with PRS and/or concerns over time may have negatively impacted their performance.

Summary and recommendations

We summarise what we consider to be the major pros and cons of running a large undergraduate MCQ class test with PRS in Table 2. Of the advantages to students, the immediacy and quality of feedback associated with PRS are the most significant; the surveys show that the students found the feedback sessions helpful and there is some evidence to suggest that this improved their understanding, or at least their perceived understanding, of the topics under consideration. This finding confirms that PRS can be a useful tool for formative assessment. The answer charts are also helpful to staff, providing a clear indication of which questions students are answering incorrectly and thus aiding remedial teaching. The large reduction in marking time and to a lesser degree preparation time is also impressive; this resulted in ca. 75% fall in total staff time inputs (Table 1). This reduction would have been even greater had keypads with a user-authentication facility been available.

By far the most significant drawback of running the GGY2024 test with PRS was the problems associated with the need to allocate a fixed 'per-question' time limit, which clearly disadvantaged students who do not cope well with answering questions under pressure. The paper-based MCQ test format is better for such students. This problem cannot easily be rectified with the technology available without repeating the test and voiding the first set of answers, which would probably be unpopular with most students, and impractical.

Advantages	Disadvantages
Students	***Students***
• Innovative assessment medium praised by the majority of students • Rapid generation of results and feedback • Graphical feedback aids student engagement with course materials	• Fixed 'per-question' time interval limits the capacity for students to review answers at their own pace • Some students unfamiliar with PRS expressed lack of confidence in system • Students with physical disabilities may struggle with small buttons on keypads
Staff	***Staff***
• Significantly reduced marking time • Slightly reduced preparation time (significantly reduced if students' identification details can be authenticated via keypads) • No need for photocopying / paper • Potential for marking errors likely to be significantly less than for paper-based MCQ test • Easily-generated answer charts provide a clear visual gauge of questions which students struggled with, aiding targetted remedial teaching	• Significantly greater cost • Keypad reliability / tampering issues are a concern but considered to be of low incidence and can be detected • Software-related problems e.g., difficulties with Macintosh interface in our experience • PRS not suited to other class test answer formats, e.g. short answers, diagram plotting • Staff training in the use of the PRS interface required

Table 2: Advantages and disadvantages of running the GGY2024 test with PRS in comparison to paper-based format.

The issue of staff training in the use of PRS is also a concern to practitioners, particularly for institutions without media support staff. Although PRS technology is straightforward to use, and would probably take only a few hours to learn for anyone familiar with presentation software (e.g., *PowerPoint*®), the difficulties that we encountered relating to the lack of cross-platform software versatility (Macintosh *versus* PC) were frustrating and would have required specialised knowledge to resolve. These compatibility concerns may also be relevant for other PRS systems and should be researched before a system is purchased, particularly for institutions using both platforms. Clearly this and other issues must be weighed up alongside the more obvious cost-related factors when departments are considering PRS.

Our findings regarding student performance neither strengthen nor weaken the case for PRS, although our cautionary points regarding students being slightly distracted by the interface and concerns over time remaining should be heeded by staff contemplating running MCQ tests via PRS. Previous research on the impact of PRS on student performance remains equivocal, with some studies suggesting that the technology improves performance (e.g., Poulis *et al*., 1997; Hake, 1998) and others suggesting that it makes no difference (e.g., Brown, 1972). In considering this issue, an important distinction must clearly be drawn between improvements that stem from application of the technology *per se* (e.g., through incentivisation) and the wider pedagogic benefits gained from its skilful application.

As a result of our experiences, we would strongly recommend the use of PRS as a formative assessment tool in undergraduate teaching, particularly for large classes. PRS is also well suited to knowledge-centred assessment in the Earth and Environmental Sciences, because of the great diversity of information that can be displayed and assessed in graphical format. However, given the problems relating to the fixed per-question time issue and the anxieties expressed by a small minority of students in using PRS, we recommend that the technology in its current form only be used for formative assessments that carry a relatively small mark weighting. This may mean that the technology is more applicable to first or second year undergraduate modules than for final year assessment. We also recommend that staff contemplating using PRS as an assessment tool should: i) carefully review different systems to check that the response options available fulfill specific needs; ii) allow plenty of time for session set-up, including trouble-shooting and knowing where to turn to for technical support; iii) run practice sessions to ensure that students are comfortable and competent with the technology; iv) perform trials to consider the important issue of time allocation per question; and v) ensure that students sit at least 1 metre apart to minimise the incidence of cheating.

In the longer term, as PRS technology evolves, the problems relating to the allocation of a fixed 'per-question' time limit in MCQ testing may be overcome, making PRS a more attractive tool for summative as well as formative assessment. Indeed, alternative response systems have recently been developed (e.g., *Testingpoint*®, also manufactured by *TurningPoint Technologies*®) which allow students to work independently with a keypad to answer questions that are presented on paper under exam conditions. Responses are not relayed to the base computer until the end of the session, thus allowing students to review their answers. Whilst this new, 'self-paced' technology resolves the per-question time limit issue, it has other disadvantages. For example, it reverts back to the use of paper and it does not permit instantaneous display of results charts that can be used for discussion or for encouraging interaction with students. This technology was not available when we ran our tests. Increasing student exposure to PRS via a variety of classroom contexts may also engender a greater sense of trust in the medium, thus diminishing the technology-related concerns that some of our students expressed. This may in turn pave the way for a wider 'culture' change in the way in which PRS-supported learning and assessment are perceived, as has already been the case in many North American universities.

References

- **Boyle J. and Nicol D. 2003** Using classroom communication systems to support interaction and discussion in large class settings, *Association of Learning Technology Journal*, 11, 43-57
- **Brown J.D. 1972** An evaluation of the Spitz student response system in teaching a course in logical and mathematical concepts, *Journal of Experimental Education*, 40, 3, 12-20
- **Burnstein R.A. and Lederman L.M. 2001** Using wireless keypads in lecture classes, *The Physics Teacher*, 39, 8-11
- **Draper S. W. and Brown M. I. 2004** Increasing interactivity in lectures using an electronic voting system, *Journal of Computer Assisted Learning*, 20, 81-94
- **Dufresne R.J., Gerace W.J., Leonard W.J., Mestre J.P. and Wenk. L. 1996** Classtalk: A classroom communication system for active learning, *Journal of Computing in Higher Education*, 7, 3-47
- **Elliott C. 2003** Using a personal response system in economics teaching, *International Review of Economics Education*, 1,1, 80-86
- **Hake R.R. 1998** Interactive-engagement versus traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses, *American Journal of Physics*, 66, 1, 64–74
- **Judson E. and Sawada A. 2002** Learning from past and present: Electronic Response Systems in college lecture halls, *Journal of Computers in Mathematics and Science Teaching*, 21, 167-181
- **Poulis J., Massen C., Robens E. and Gilbert M. 1997** Physics lecturing with audience paced feedback, *American Journal of Physics*, 66, 5, 439–641
- **TurningPoint 2010** http://www.turningpointtechnologies.co.uk accessed 1st October 2010
- **Wood W. B. 2004** Clickers: a teaching gimmick that works, *Developmental Cell 7*, 6, 796–798
- **Zhu E. 2007** Teaching with Clickers, *Center for Research on Learning and Teaching (CRLT), University of Michigan, Occasional Papers*, 22, 8

Helen Roe h.roe@qub.ac.uk
David Robinson david.robinson@qub.ac.uk

GEES PHOTO COMPETITION 2009/10

Andrew Elvidge
University of East Anglia
"From out the mist"
Mount Bromo, Indonesia, 20th August 2009

HIGHLY COMMENDED