Comparing online and offline administration of multiple choice question assessments to psychology undergraduates: do assessment modality or computer attitudes influence performance?

CLAIRE HEWSON1, JOHN CHARLTON
University of Bolton, UK
MARK BROSnan
University of Bath, UK

Recent interest in and use of online assessments across a range of disciplines has raised a number of issues relating to student perceptions and performance when using this medium. The validity of such online tests is a crucial consideration, especially if they are to be used for summative as well as formative assessment. The current study examined higher education students’ performance on online assessments which they were required to take as part of an undergraduate psychology course. The marks obtained by students required to take the same multiple choice question (MCQ) assessment online and offline (in pen-and-paper format) were compared, and relationships between performance and computer anxiety and computer engagement measures were explored. The results indicate minimal influence of assessment modality, computer anxiety and computer engagement on MCQ test scores, with only very small and non significant effect sizes being observed overall. No evidence of any significant relationships between gender and age and computer attitudes was observed. To conclude, the results provide promising initial support for the use of online summative assessments in contexts similar to the one used here, allaying some concerns about disadvantaging certain groups of students. However, further research is needed to explore possible performance differences across different contexts, assessment types, and student cohorts.

As part of a general move towards increasing the role of IT and web based methods in higher education (HE) teaching and learning, practitioners have started to make use of online assessments in both formative (e.g., Buchanan, 2000; Henley, 2003) and summative (e.g., Ricketts and Wilks, 2002) assessment contexts, across a range of disciplines. We use the term “online” throughout this paper to refer to assessments which have been made available via a computer network (e.g., the world wide web), for access from a range of locations at the respondent’s convenience. Other authors, however, have used the term to refer to more restricted computer assisted assessment contexts (e.g., Ricketts and Wilks, 2002); we flag up these deviations from our current usage where they occur. A number of tools designed to assist in implementing online assessments have become available, both commercially (e.g., WebCT2, Questionmark’s Perception3, Blackboard4) and as freely available and often ‘open source’5 programmes (e.g., Hot Potatoes6).

A number of advantages of using online assessment methods are apparent, including possibilities for automating the marking process and storing students’ responses in a data file, thereby reducing tutor workload, enhancing efficiency, and eliminating the potential for human error. This may benefit students as well as tutors, who may for example appreciate the enhanced speed of turnaround of coursework which online assessment methods can facilitate, as well as the convenience of being able to submit an assignment remotely without the need to come into college. Students may also find online assessments novel and engaging. Online methods can also help facilitate the provision of distance learning courses, and thus play a potential role in widening participation in education. In formative contexts, the ability to provide immediate feedback may have important pedagogical benefits, which may create an advantage over pen-and-paper methods.

However, a number of potential disadvantages with online assessment methods are also apparent. Firstly, implementing an online assessment requires a certain

---

1 Correspondence concerning this article should be addressed to the first author at: Psychology and Life Sciences, University of Bolton, Deane Rd., Bolton BL3 5AB, UK. Email ch5@bolton.ac.uk.
2 http://www.webct.com
3 http://questionmark.com/us/perception
4 http://blackboard.com. At the time of writing Blackboard and WebCT are in the process of merging.
5 ‘Open source’ refers to code which has been made freely available to use and modify. For a discussion of the reasons for supporting the development of shared, open source resources see Malloy, Jensen and Regan (2002).
6 http://www.gla.ac.uk/~jwalk001/dwresourcesweb/assessments.htm
level of technological expertise, or at least technical support (Hemard and Cushion, 2003), as well as the ready availability of appropriate equipment and/or software. This may at best mean substantial time investment by tutors and/or technical support staff, and at worst might preclude the use of online assessments due to a lack of available resources. Though implementation packages such as WebCT exist, these are costly, and not always straightforward to learn how to use (Henley, 2003).

Secondly, reliability, integrity and validity are all currently of particular concern in the use of online assessment procedures, due to the relative novelty of the approach. Reliability issues emerge due to the potential for technological failures such as server crashes, which may result in assessments going offline, not functioning correctly, or worse, data being lost. Integrity and validity may be compromised for a number of reasons, including not being able to verify who has completed and submitted an online assessment, and an increased potential for respondents to obtain information (e.g., by hacking into the system to obtain correct answers) which may enhance performance levels without the relevant learning outcomes having been met. Also, it is possible that respondents with higher levels of computer literacy may perform better simply because they are at greater ease with computers, thus questioning the ability of these tests to measure what is intended.

Because of such issues, some authors have suggested that online assessments should be used in a limited way only, their importance in contributing to overall course grades being restricted (e.g., Roy and Armarego, 2003). Having said this, however, some of these threats do not appear to be obviously greater than for comparable offline assessments, where it may also be difficult to verify the author of a submitted piece of work. Even in time-limited examination settings it may not always be possible to be certain that the person sitting the exam is whom they say they are, and extreme levels of exam anxiety may interfere with exam performance. Nevertheless, the involvement of technology in online assessment procedures does add a further dimension to these issues, the implications of which require exploration.

**Validity of Online Assessments**

Issues of integrity, reliability and validity in the use of online assessments have, to date, received little empirical attention. Studies have tended to focus on the efficacy of online assessment procedures, particularly in formative contexts. For example, several authors have reported that students who access online formative assessments more frequently perform better in end of term summative assessments (e.g., Henley, 2003; Van Hooff and Porteuous, 2004), thus supporting the benefits of these approaches, though sometimes the reported effect sizes have been small (e.g., Buchanan, 2000). Also, it is not entirely clear the extent to which the process of taking an online formative assessment itself enhances learning, or whether student motivation levels may influence both access of online formative tests and performance on summative assessments (keener students making greater use of resources, and performing better, overall). Nevertheless, these results are encouraging.

When considering the use of online summative assessments, however, it is more crucial to determine that these tests are able to provide a valid and fair measure of a student’s level of ability, without interference due to the modality of the online testing environment. Few studies have addressed this issue, though some evidence has been provided by Ricketts and Wilks (2002), who report that being required to scroll through questions in an “online” in-class (multiple choice question (MCQ)) exam led to worse performance than either taking the exam in pen-and-paper format, or being presented with questions via a computer screen one at a time.

This result, if valid, is important. However, it is difficult to evaluate due to a range of methodological problems, including no reporting of sample sizes, effect sizes, or statistical significance levels, and the groups differing – as well as in modality of presentation – in the exam questions received, and amount of prior practice on similar computer based tests (though we do recognise that it can be difficult to generate well controlled experimental designs in this area, since the demands of the teaching, learning and assessment process impose constraints on what can be implemented in empirical studies of this nature).

Also, these authors’ usage of the term online deviates somewhat from ours; whilst strictly speaking the exam may have been available via a computer network (though this is not made clear), the fact that students completed it in a time-limited classroom setting restricts the focus of this study compared with the type of online assessments in which we are interested – that is, those which students may access from a range of locations at their leisure.

Thus, extending Ricketts and Wilks’ (2002) work, a key aim of the present study was to further clarify the possible role of modality effects in a truly online assessment context, by comparing the performance of students enrolled on the same module who were (pseudo-) randomly assigned to take the same MCQ assessment either online or offline. Since our online assessment was placed on the web, students could access, complete and submit the assignment from any internet ready computer with web access, within a 3 week time frame. Obviously, findings relating to “computer assisted” assessment procedures (e.g., in-class exams) may not necessarily generalise to truly online assessment procedures, and vice versa.
The Role of Computer Attitudes

As well as modality, other factors, such as attitudes towards computers, may potentially impact upon performance in online assessments, thus presenting a further possible threat to the validity and equity of this approach. While there has been little research on this issue which relates directly to online assessment procedures, there is a large body of literature which addresses the role of computer anxiety in performance on computer related tasks.

Computer anxiety constitutes “a real phenomenon” (Moldafsky and Kwon, 1994, p. 301), and is defined as “an irrational anticipation of fear evoked by the thought of using (or actually using) computers, the effects of which result in avoiding, or minimizing, computer usage” (Brosnan, 1998a, p. 17). Research examining the extent of aversive reactions to technology highlights a sliding scale from ‘uncomfortable user’ through to ‘phobic’ (Meier, 1985; Rosen, Sears and Weil, 1993) with around 5% falling into this latter category and reporting symptoms such as sweaty palms and heart palpitations (Rosen et al., 1993). Thorpe and Brosnan (in press) demonstrate that the anxiety experienced in this latter group is similar in many respects to the anxiety experienced by those with a formal (DSM-IV) diagnosis of a more traditional phobia (e.g., spider phobia). Rosen and Maguire (1990) highlight that although an inverse relationship tends to exist between computer experience and computer anxiety (Choi, Ligon and Ward, 2002; Maurer, 1994; Shashaani, 1997), this is confounded by ‘phobic’ individuals actively avoiding computer interaction.

The more common ‘uncomfortable user’ description applies to around a quarter to one third of most occupational groups, including students within HE (Brosnan, 1998a). The proliferation of technology has shifted the strategies of the ‘uncomfortable users’ from avoidance (as this became increasingly untenable) to a strategy of minimising interaction with technology and suffering an aversive state during this interaction (Marcoulides, 1988). There is evidence that this anxiety is more prevalent in females than in males (e.g., Abdelhamid, 2002; Brosnan, 1998b, 1998c, 1999; Durnell and Haag, 2002), though Rosen and Maguire (1990) found no evidence for this.

Computer anxiety has been found to relate directly to success on computer based tasks (Brosnan, 1998c), and perceptions of computer related technology have been found to relate to performance on online versions of a task, but not to a paper-and-pencil version of the same task (Brosnan, 2005). This is consistent with some research which has examined computer based course assessment (Szanja, 1994), although the results in this area are equivocal (Mahar, Henderson and Deane, 1997). Mahar et al. (1997) found that computer anxiety was linked to poorer performance on a database editing task (regardless of levels of prior computer experience), but that few effects were found when using course related computer based assessments, as motivation and academic ability became more salient variables.

In relation to the above discussion, a second aim of the current study was to examine whether there was any relationship between computer anxiety levels and performance (in terms of total scores) on the online MCQ assessments administered. Based on previous findings (as summarised above), it could be hypothesised that if the computer based task is demanding enough, in terms of computer related performance factors, to introduce any computer anxiety related effects, then these would appear as higher levels of computer anxiety interfering with the online task, but not the offline task, thus leading to poorer performance on the former but not the latter.

However, if online assessments are to be valid, it is important to ensure that the computing technology used to administer them is as transparent to the student as possible, and the online MCQ assessment procedures under consideration in the present research were designed with this in mind. Thus, it was considered that the absence of any computer anxiety effects associated with the present online MCQ tests would provide evidence in favour of the validity of the online procedures used to administer them.

Since computer anxiety has also been shown to be related to mathematics anxiety (Weil, Rosen and Wugalter, 1990), the performance of students taking a mathematics related module, and those taking a nonmathematics related module, was compared, in order to explore whether a different relationship between computer attitudes and performance on online assessments emerged in each context. Also, students’ performance on a previous research methods/mathematics related module was considered, in order to take into account variations due to differing levels of ability and motivation.

In addition to considering the possible implications of computer anxiety for the implementation of online MCQ assessments, the present study also considered the implications of computer engagement. Whereas computer anxiety involves the extent to which people are anxious about using computers, computer engagement concerns the extent to which people show, at one extreme, apathy towards using computers, or, at the other extreme, a high (but nonpathological) involvement with computers. The concept of computer engagement was first developed by Charlton and Birkett (1995), and subsequent work showed that students on programming orientated HE courses are more highly engaged than those on business orientated computing courses (Charlton and Birkett, 1998), and that for programming orientated students there is a positive relationship between engagement and course performance (Charlton...
and Birkett, 1999).

With respect to online assessment methods, it is possible that, all other things being equal, more highly computer engaged students will have an advantage over less computer engaged students and that if any such effects are large this will compromise the validity of such assessment methods. Just as was done for computer anxiety, the present project examined this possibility in order to assess the validity of the present online MCQ assessment procedures.

**The Role of Demographic Factors**

There is some evidence that traits such as computer anxiety may have demographic correlates; for example, as already noted, several studies have shown that females display higher levels of computer anxiety than do males (see Brosnan, 1998a, and Cooper and Weaver, 2003, for reviews). Other studies have found that both computer anxiety levels and performance on computer based tasks are related to levels of prior experience using computers (see, for example, Brosnan, 1998a; Charlton and Birkett, 1999), in which case it may be reasonable to speculate that mature students, or international students from countries which may not have ready availability of computers in schools and universities, might be disproportionately disadvantaged by the introduction of online assessment due to having less computer experience.

Although level of computer experience was not directly measured in the present study (all users being relatively experienced due to prior course requirements), it was of interest, following the above findings, whether relationships between gender, age, and computer attitudes and performance measures would be evident, and this constituted the third research question in the present study.

**METHOD**

**Design**

Independent variables were mode of administration (online or pen-and-paper) of the MCQ assignment (assessment modality), cohort (see below), age and gender. Dependent variables were the MCQ assignment scores, computer attitude scale scores (which provided two subscale scores: anxiety and engagement), and some additional attitudinal questions probing evaluations and perceptions of online assessment methods.7

Three cohorts of students took part, differentiated by being enrolled on different psychology module occurrences. Cohorts 1 and 2 were both enrolled on the same accredited module, and thus received the same lectures delivered by the same lecturers, but took this module in different academic years. Students within both of these cohorts were pseudorandomly assigned to take the same compulsory course-based MCQ assignment either online or offline (for Cohort 1 surnames A–K were assigned to the online condition and the remainder to the offline condition; for Cohort 2 this was reversed). Cohort 3 was enrolled on a different accredited module, at a different university, and took a noncompulsory course-related MCQ assignment online. All three cohorts were invited to complete the same pen-and-paper computer attitudes questionnaire (see below), for which they received a monetary reward.

**Participants**

Students at the University of Bolton enrolled on a Further (Research) Methods in Psychology module in 2003–2004 (Cohort 1, \(n = 45\) [27 females, 3 males, 1 nonresponse to this question8; age range 18 to 55 years]), and in 2004–2005 (Cohort 2, \(n = 49\) [37 females, 8 males9; age range 19 to 55 years]) participated. A third cohort consisted of students from the University of Bath taking a Controversies in Cognition module during 2003–2004 (\(n = 21\) [19 females, 2 males; age range 21 to 23 years]).

**Equipment and Materials**

The online MCQ assessments were constructed using a dedicated tool developed at the University of Bolton for this purpose, which consists of an HTML form with embedded javascript commands, and a CGI script which processes the data submitted10. Participants were able to access the assignment from any internet connected computer. Cohorts 1 and 2 received an identical assignment related to the Further Methods in Psychology module, and Cohort 3 received a different assignment which related to the Controversies in Cognition module. These assessments differed in their

---

7 Responses to these additional questions are not considered in the current paper, since focus is upon performance measures and the role of computer anxiety and computer engagement.
8 Gender and age were known for the 31 out of 45 participants who completed the computer attitudes scale. For the remaining 13 participants in this cohort information about gender and age was not obtained.
9 Gender and age information was not obtained for the 4 participants who did not complete the computer attitudes scale in this cohort.
10 This tool is available at: http://www.clairehewson.co.uk/LTSNproject/MCQtool, and is described in detail in Hewson (in press). At present, the tool does not embody a graphical user interface. A set of files are provided, along with an instruction manual which describes how to customise and install these files on a web server. Three implementations are described, each with varying levels of functionality and technological requirements (the final implementation requires access to a server that can execute CGI scripts). All the HTML and programming code provided is freely available and open source.
degree of mathematical/statistical content, the former module containing some such content, the latter none.

The computer attitudes questionnaire consisted of a computer anxiety (CA) subscale (17 items, a higher score indicating higher anxiety, with minimum and maximum possible scores of 17 and 85 respectively) and a computer engagement (CE) subscale (12 items, a higher score indicating higher engagement, with minimum and maximum possible scores of 12 and 60 respectively), which were derived from factors in the factor analysis of Charlton (2002), which in turn consisted predominantly of items from the Computer Apathy and Anxiety Scale (Charlton and Birkett, 1995). On these two subscales, students gave responses to statements concerning their computing related thoughts and behaviours using a 5-point Likert type response format ranging from strongly disagree to strongly agree. The present data yielded Cronbach’s alphas of .91 for the Anxiety subscale and .78 for the Engagement subscale (n = 97).

Procedure
Cohort 1 and 2 were administered the computer attitudes scale in pen-and-paper format (to avoid the possible influence on scores of computer based administration), and completed and returned this at the start of the first lecture for the module. As part of their module assessment, students within each of these two cohorts were required to complete a non-time-limited 20-item MCQ assignment (either online or offline, depending on which condition they had been assigned to) which tested knowledge of the material delivered in Weeks 1 and 2, and was completed anytime between Weeks 4 and 7, although for Cohort 2 the assignment was given out in Week 2 with a due date of Week 5. There was no reason why this difference should have influenced performance, both cohorts having completed all relevant lectures before the assignment was received, and having the same amount of time (i.e., 3 weeks) for completion. The mark obtained for this assignment did not contribute towards the overall module mark.

Cohort 3 were invited to take a noncompulsory online MCQ assessment which they were advised would be a useful activity in preparation for the formal (written) assessment. These participants took the assessment individually, and completed the pen-and-paper computer attitudes scale during a 1-week period outside of lecture time.

RESULTS
Assessment Modality and MCQ Scores
To examine modality effects, the performance of the students who took the offline and online versions of the same MCQ test (i.e., Cohorts 1 and 2 combined) was compared. Since the relative performance of students taking the module in different years was also of interest, a two-way ANOVA was carried out with assessment modality and cohort as independent variables, and MCQ test score as the dependent variable. Table 1 shows the means and standard deviations for all four groups, and the marginal means for the online/offline and Cohort 1/ Cohort 2 groups.

As shown in Table 1, scores for the online and offline groups were similar, there being no significant main effect of assessment modality ($F(1, 85) = 0.93, ns$)$^{11}$.

However, there was a significant main effect across cohorts, with Cohort 2 scoring significantly lower than Cohort 1 ($F(1, 85) = 6.88, p = .01, partial \( \eta^2 = .08 \)$).

To investigate whether this could be attributed to Cohort 2 being weaker overall, marks for a previously taken strongly related research methods module, Methods in Psychology 2, and overall marks for the Further Methods in Psychology module$^{12}$, were examined. No significant differences were observed between Cohorts 1 and 2 on either of these measures.

There was no significant interaction between cohort and assessment mode on MCQ test scores, both cohorts scoring marginally worse in the offline condition. For all the above comparisons, post hoc analyses using the

| Table 1 |
| Mean multiple choice question (MCQ) test scores (out of 20), and standard deviations, for the online and offline groups, Cohorts 1 and 2 |

<table>
<thead>
<tr>
<th>MCQ Score</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$n$</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online</td>
<td>14.4</td>
<td>3.3</td>
<td>27</td>
</tr>
<tr>
<td>Offline</td>
<td>14.1</td>
<td>2.0</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>14.3</td>
<td>2.8</td>
<td>44</td>
</tr>
</tbody>
</table>

$^{11}$Where effect sizes and exact $p$ values are not reported, these can be obtained from the first author upon request.

$^{12}$The overall module marks were made up of an independent research project practical report, weighted at 80%, and an oral presentation of independent research work, weighted at 20%.
Computer Anxiety, Computer Engagement and MCQ Scores
To test whether levels of computer anxiety and computer engagement were related to MCQ test scores, in both online and offline contexts, Pearson’s $r$ correlation coefficients were computed.

For this analysis, Cohorts 1 and 2 were combined, and Cohort 3 (who had received a different MCQ assignment) was considered separately. As shown in Table 2, the means and standard deviations for both CA and CE scores were similar across cohorts, and one-way ANOVAs with cohort as the independent variable and computer attitudes score as the dependent variable showed no significant difference in the mean CA $(F(2, 94) = 0.30, ns)$ and CE $(F(2, 94) = 0.72, ns)$ scores across all cohorts. Therefore combining Cohorts 1 and 2 was considered justified.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Computer Anxiety</th>
<th>Computer Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Cohort 1</td>
<td>38.7</td>
<td>11.9</td>
</tr>
<tr>
<td>Cohort 2</td>
<td>40.3</td>
<td>10.2</td>
</tr>
<tr>
<td>Cohort 3</td>
<td>40.8</td>
<td>9.1</td>
</tr>
</tbody>
</table>

As can be seen from Table 3, using Cohen’s (1988) benchmarks for small and medium effect sizes ($r = .10$ and $r = .30$ respectively), effect sizes ranged from very small to small/medium. All coefficients were nonsignificant, although the positive correlation between CE and MCQ scores for the online Cohort 1/Cohort 2 group approached significance ($p = .066$), indicating higher levels of computer engagement were associated with higher MCQ test scores. However, a test of the difference between this coefficient and the very small partial correlation was reduced to a small effect; and the very small negative correlation between CE and MCQ scores for the online group increased to a medium-sized (but still nonsignificant) negative partial correlation. Thus, overall, no significant relationships were found between CA or CE and MCQ test scores.

While the primary goal of the current research was to determine the extent to which mode of assessment and computer attitudes were key variables in determining performance on online MCQ assessments, it is of interest to note that computer attitude measures were actually more strongly related to overall module marks than to MCQ assessment marks.

Combining Cohorts 1 and 2, a significant negative relationship was found between Methods 2 marks and CA scores ($r (62) = -.34, p = .005$), with more highly anxious students performing worse on the Methods 2 module, while a near significant small/moderate positive relationship was observed between Methods 2 marks and CE scores ($r (62) = .24, p = .055$), with more highly engaged students performing better. Similarly, higher Further Methods module marks were associated with lower levels of anxiety ($r (62) = -.36, p = .004$) and higher levels of engagement ($r (62) = .28, p = .03$). Cohort 3 displayed a small/moderate, but nonsignificant, negative relationship between module marks and CA scores ($r (18) = -.24, p = .32$), but a negligible effect with CE scores ($r (18) = -.01, p = .97$).

Since these module marks were all derived from offline
Comparing Online and Offline Administration of MCQs

forms of assessment, this indicates that the computer attitudes measured in this study did not display any particular impact upon performance on online, as opposed to offline, assessment modes.

The Role of Demographic Factors
Table 4 shows Pearson's $r$ correlation coefficients between CA and CE scores, and age and gender, for all three cohorts combined.

Only a very weak nonsignificant relationship was observed between age and CA scores. A slightly stronger but still nonsignificant relationship was observed between age and CE scores, which would indicate older participants having higher engagement levels. Point biserial correlations showed gender to be only very weakly related to both CA scores and CE scores. Given that females were given a numerical code lower than males, these indicated very small effects whereby males exhibited lower anxiety and greater engagement than did females. No significant relationships were observed between MCQ test scores and gender or age.

Discussion
The current study has found that taking the same MCQ assessment in online or offline modes does not impact upon performance, at least within the current context. Further, only very weak evidence for any effect of computer anxiety (CA) or computer engagement (CE) on performance on online assessments was found, the strongest effect being a near-significant small/moderate positive relationship between computer engagement and MCQ assessment scores on a research methods related assessment which had a high mathematics/statistics component. A similar but much weaker nonsignificant effect was observed for a cohort taking a non-research methods related online assessment. Furthermore, these small effects were weaker than relationships observed between computer attitudes and performance on offline assessments. These results provide support for the validity of online assessments of the type used here, and should help alleviate some concerns that the use of such assessments may disadvantage certain types of students.

While these results are promising, and show that online assessment methods can lead to levels of performance equivalent to offline methods, a number of restrictions on the generalisability of these findings should be noted. Firstly, the task used here was relatively straightforward, in terms of computer related performance elements (and intentionally so), and further research is needed to determine the extent to which the current results can be replicated with different, and more complex or demanding, computer based tasks.

Also, in addition to computer related task complexity, other aspects of the online assessment context may differentially affect performance. While no mode of presentation effect was found in the current study, Ricketts and Wilks (2002) do report such an effect, such that being required to scroll down a page of questions, in an in-class computer based exam setting, impaired performance compared with both computer based presentation of questions one-at-a-time, and paper-and-pen format. It is possible (ignoring the problems of interpreting their result, noted earlier) that the assessment context used by Ricketts and Wilks was more anxiety inducing than the one used here, and that these contrasting results between their and our study, in terms of mode of presentation effects, may be attributable to this factor. The influence of such

### Table 3
Correlations between MCQ scores and computer anxiety, and MCQ scores and computer engagement, for the online group, the offline group (both from Cohorts 1 and 2 combined) and the online Cohort 3 group, separately

<table>
<thead>
<tr>
<th>Group</th>
<th>Computer Anxiety</th>
<th>Computer Engagement</th>
<th>Partial correlations*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>n</td>
<td>p</td>
</tr>
<tr>
<td>Online</td>
<td>-.18</td>
<td>42</td>
<td>.26</td>
</tr>
<tr>
<td>Offline</td>
<td>-.30</td>
<td>29</td>
<td>.12</td>
</tr>
<tr>
<td>Cohort 3 (online)</td>
<td>-.09</td>
<td>20</td>
<td>.71</td>
</tr>
</tbody>
</table>

*Note. Partial correlation coefficients controlling for marks on previous research methods module.

### Table 4
Correlations between CA and CE scores, and gender and age, for all 3 cohorts combined

<table>
<thead>
<tr>
<th>CA</th>
<th>Gender</th>
<th>r</th>
<th>n</th>
<th>p</th>
<th>CA</th>
<th>Age</th>
<th>r</th>
<th>n</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-.06</td>
<td>96</td>
<td>.55</td>
<td>.04</td>
<td>96</td>
<td>.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-.01</td>
<td>97</td>
<td>.91</td>
<td>.14</td>
<td>97</td>
<td>.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
contextual differences would be a useful line of research to pursue in this area.

A further consideration addressed in the current study was whether assessment content may play a role in moderating the relationship between computer attitudes and online assessment performance, simply because this is one of a range of contextual variations which may conceivably have an impact, and which has received limited empirical exploration to date. In general, it was the case that the small/medium relationships observed between each of CA and CE and online MCQ scores were weaker for Cohort 3, who had received a nonmathematics/statistics related assessment, than for Cohorts 1 and 2, who received an assessment with mathematical/statistical content.

However, whether this can be attributed to differences in the content of the assessments is inconclusive, due to several other differences between the cohorts: they came from different and quite distinct universities (Cohorts 1 and 2 from a new university with a high proportion of first generation university students, and Cohort 3 from a traditional university with few first generation university students), and differed in whether they were required to take the online assessment for successful module completion (this not being the case for Cohort 3 only). The role of these respective factors thus requires further investigation.

Another possible explanation for these results might be that students’ responses to the computer attitudes scale were influenced by the context in which it was completed. The two cohorts who completed this during their research methods class may have been primed by this context to focus upon their performance on and attitudes towards this particular module, with lower performing students perhaps tending to answer in a more negative way than they might have done if this particular module had not been cognitively salient. For Cohort 3 this effect may plausibly not have occurred, thus lessening any relationships between module/online assessment performance and computer attitudes. The finding that for Cohorts 1 and 2 computer attitudes were significantly related to overall module marks, but not for Cohort 3 (effect sizes being small, as well as nonsignificant), provides tentative support for this hypothesis.

It is also possible that the correlation between computer anxiety and research methods module marks represents a noncausal spurious relationship, with the correlation possibly arising because mathematics anxiety is a common causal antecedent of both computer anxiety and performance on the methods module, this speculation also leading to possible future research avenues.

Finally, it should be noted that the cohorts used in the current study had fairly high levels of computer literacy and competency, due to having had fairly extensive relevant training as part of the research methods modules they had taken, which included substantial use of statistical analysis and database software packages, as well as word processing software. It was thus quite likely that the cohorts used here had relatively low levels of computer anxiety, overall. Further, since Cohort 3 were not required to take the online assessment as part of their module assessment, they may have self-selected for those individuals with lower levels of computer anxiety and higher levels of engagement, though as indicated in Table 2, the mean computer anxiety and engagement scores of this cohort did not differ greatly from the mean CA and CE scores of the other two cohorts. The current results should therefore not be generalised to more highly anxious populations, and this highlights another area for future research.

While previous studies have reported relationships between gender and computer attitudes, no such relationships were uncovered here. Neither was age found to be related to computer attitudes in the present study.

To conclude, the current study has supported the validity of online assessments using tasks which are simple and transparent. The role of computer attitudes in more complex, less transparent tasks requires further investigation for the validity of such tasks to also be supported. Further research is also needed to validate the use of such assessments with less experienced, and more highly computer anxious, populations. However, it is not only students’ levels of performance which warrants consideration when deciding whether to use online assessment procedures; levels of student satisfaction are also important. If students, or certain groups of students (such as those with high levels of computer anxiety), find the approach unappealing or stressful then this could have a detrimental effect on the quality of the teaching/learning experience.

Students’ attitudes, preferences, and perceptions of online assessments have been little studied to date; however there is some evidence that students do appreciate features such as the flexibility of access of online assessments, and immediacy of feedback (in formative contexts) (e.g., Farrell and Leung, 2004). Features of online assessments which may feasibly diminish student satisfaction levels are a lack of mobility (students must be sitting at a computer screen to complete and submit them – though of course a student may print out an online assignment to work on offline, and then later use a computer to input and submit answers – and it cannot be taken for granted that students have ready computer access outside college), and the requirement that the test must be completed ‘in one sitting’, though this latter restriction if shown to be relevant could be addressed by technological solutions. Concerns that an
assignment might get lost may also be greater in online than traditional offline assessment contexts. While it has been beyond the scope of the current paper to explore the issue of students' perceptions and satisfaction levels in relation to online assessment procedures, this is clearly also an important issue worthy of investigation. Nevertheless, the current results do provide support for the validity of online MCQ assessments, having shown that they can provide as valid and fair a measure of student ability as traditional offline assessments.

ACKNOWLEDGMENTS

This research was supported by a grant from the Higher Education Academy Psychology Network, whose support is gratefully acknowledged.

REFERENCES


45


*Manuscript received on 26 July 2005*  
*Revision accepted for publication on 29 September 2006*