Implementing computer-based assessment – A web-based mock examination changes attitudes

Tobias Deutsch*, Kristin Herrmann, Thomas Frese, Hagen Sandholzer

Department of Primary Care, Leipzig Medical School, Philipp-Rosenthal-Str. 55, 04103 Leipzig, Germany

A R T I C L E   I N F O

Article history:
Received 26 April 2011
Received in revised form 6 October 2011
Accepted 15 November 2011

Keywords:
Computer-based assessment
Technology acceptance
Attitudes
Gender
Higher education

A B S T R A C T

Interest in the educational use of information and communication technologies (ICT) in higher education is increasing. For successful implementation, it is important to know students’ attitudes and reservations and how they can be positively influenced. The objective of the present study was to examine possible attitudinal changes towards computer-based assessment (CBA) in students, after undergoing one such assessment. A web-based mock examination was provided to all fourth year medical students at Leipzig Medical School in 2008 and 2009. Before and after the web test, students were asked to document their agreement with statements concerning CBA. A large number of students made use of the offered web-based assessment. 383 participants could be analysed for the pre-post comparisons. The majority of the students rated their computer self-efficacy as high. In summary, students’ attitudes towards CBA in higher education tended to be positive. Gender differences seemed to be substantially influenced by differences in computer self-efficacy and were reduced considerably after only one practical experience. The actual experience had a positive influence on attitudes towards CBA. Nevertheless there were strong reservations about technical problems influencing the test performance when used for summative assessment. These concerns should not be ignored when trying to implement CBA. Optional formative CBA, perhaps early in higher education, seems to be a promising possibility of attracting students to computer- or web-based examination and learning methods, and may be a useful component of a successful implementation strategy.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The educational use of information and communication technologies (ICT) for different purposes in higher education is increasing. Computers and the internet open up new possibilities for delivering learning content (e-learning) and examinations. The investigation and further development of computer-based assessment (CBA) is seen as a major challenge in higher education today (Whitelock, 2009). The advantages of CBA or web-based assessment (WBA) include flexibility in time and place, usability of a greater variety of media and test types, immediate scoring and feedback and, once set up, low staff requirements. CBA can serve as summative assessment, allowing the teacher to evaluate a student’s abilities at a given point, as the basis for a decision about pass or failure. Used as formative assessment – defined as judgement combined with immediate feedback – it enables the learner to identify and close possible ‘gaps’ between the actual level of work and the required standard (Cantillon et al., 2004; Gikandi, Morrow, & Davis, 2011; Taras, 2005). Several recent studies suggest that formative self-assessments, implemented in the curriculum, are perceived as useful by the students (Cassady & Gridley, 2005) and improve their performance in the final exam (Cassady & Gridley, 2005; Ibabe & Jauregizar, 2010; Wilson, Boyd, Chen, & Jamal, 2011).

Despite the numerous advantages, the adoption and integration of ICT in higher education is progressing relatively slowly in many cases, due to several concerns und barriers on institutional, pedagogical and individual levels (Birch & Burnett, 2009). On an individual level, a successful implementation strategy for educational (technological) innovations needs to consider the students’ acceptance. Results of
2. Material and methods

2.1. Web-based mock examination

In 2008 and 2009 an optional, web-based primary care mock examination in preparation for the regular written test was provided to all fourth year medical students (8th semester) at the Leipzig Medical School after attending primary care lectures. The online test was announced in good time at one of the primary care lectures and on the website of the medical faculty. Information about technical and organisational aspects was given in the process. The structure of the mock examination was identical with the final in class test. It consisted of a total of 30 questions, whereby 29 questions had a multiple choice response format and one question had to be answered in free text. In order to avoid inequalities concerning the preparation for the final exam and to ensure comparable conditions with regard to possible attitudinal changes all students edited the same questions in the same order (no test variations or item pool).

Students could take the mock exam in a defined time slot a few days before the final written test. While in 2008 the time slot was two days, it was reduced to 2 h in 2009 after positive results of a simulated server-applied load-test. During that time slot all components (pre-evaluation, mock examination, post-evaluation) were unlocked for editing. In the context of a greeting on the first screen seen by students after login, the idea of the test and of the evaluation was communicated again and students were asked to edit all components in the correct order (pre-evaluation, web test, post-evaluation). For the editing of the evaluation components there was no time limit.

The expected length of time for editing was announced previously. Analogously to the final exam, the time for the editing of the mock examination itself was limited. Once activated, students had 45 min to complete it. After that time the test would have been transmitted in an incomplete form for assessment. Since the free text answer could not be assessed by the computer, the database with results was locked two days after the test, but in time to be helpful for preparation of the final written examination. Results included total score as well as a feedback on item level. Login could be carried out from any computer under specified conditions for a consecutive number for each person before sending data from the computer centre to statistical analysis. Unique use of all components was guaranteed by locking each of them for the respective matriculation number after it was edited and transmitted. To give students the possibility to check the technical assumptions and the editing modalities prior to the online test, we provided a user interface try-out in 2009. Students could test their Login and check the editing of the two response formats during one week prior to the web test.

The described interrelations suggest that positive CBA experiences might have positive effects on the respective attitudes and the intention to use CBA.

The objective of the present study was to provide students with a web-based mock examination in preparation for the regular written exam, in order to consider questions of attendance and acceptance of the CBA offered. Specifically of interest were attitudinal changes towards CBA. To detect possible changes we used a pre-post design. To our knowledge there were no published pre-post comparisons addressing this topic until now. In consideration of the described differences between male and female students with regard to computer self-efficacy and computer use, a possible influence of gender and computer self-efficacy was of a particular interest.

2. Material and methods

2.1. Web-based mock examination

In 2008 and 2009 an optional, web-based primary care mock examination in preparation for the regular written test was provided to all fourth year medical students (8th semester) at the Leipzig Medical School after attending primary care lectures. The online test was announced in good time at one of the primary care lectures and on the website of the medical faculty. Information about technical and organisational aspects was given in the process. The structure of the mock examination was identical with the final in class test. It consisted of a total of 30 questions, whereby 29 questions had a multiple choice response format and one question had to be answered in free text. In order to avoid inequalities concerning the preparation for the final exam and to ensure comparable conditions with regard to possible attitudinal changes all students edited the same questions in the same order (no test variations or item pool).

Students could take the mock exam in a defined time slot a few days before the final written test. While in 2008 the time slot was two days, it was reduced to 2 h in 2009 after positive results of a simulated server-applied load-test. During that time slot all components (pre-evaluation, mock examination, post-evaluation) were unlocked for editing. In the context of a greeting on the first screen seen by students after login, the idea of the test and of the evaluation was communicated again and students were asked to edit all components in the correct order (pre-evaluation, web test, post-evaluation). For the editing of the evaluation components there was no time limit.

The expected length of time for editing was announced previously. Analogously to the final exam, the time for the editing of the mock examination itself was limited. Once activated, students had 45 min to complete it. After that time the test would have been transmitted in an incomplete form for assessment. Since the free text answer could not be assessed by the computer, the database with results was locked two days after the test, but in time to be helpful for preparation of the final written examination. Results included total score as well as a feedback on item level. Login could be carried out from any computer under specified conditions for a consecutive number for each person before sending data from the computer centre to statistical analysis. Unique use of all components was guaranteed by locking each of them for the respective matriculation number after it was edited and transmitted. To give students the possibility to check the technical assumptions and the editing modalities prior to the online test, we provided a user interface try-out in 2009. Students could test their Login and check the editing of the two response formats during one week prior to the web test.

The described interrelations suggest that positive CBA experiences might have positive effects on the respective attitudes and the intention to use CBA.

The objective of the present study was to provide students with a web-based mock examination in preparation for the regular written exam, in order to consider questions of attendance and acceptance of the CBA offered. Specifically of interest were attitudinal changes towards CBA. To detect possible changes we used a pre-post design. To our knowledge there were no published pre-post comparisons addressing this topic until now. In consideration of the described differences between male and female students with regard to computer self-efficacy and computer use, a possible influence of gender and computer self-efficacy was of a particular interest.
were highlighted by flashing orange in this structure. Information about the current processing status and time left was given throughout the entire examination. After editing of all questions the test could be transmitted for assessment by a click on the appropriate button.

2.2. Sampling and design

The web-based mock examination was provided to all fourth year medical students (8th semester) of the years 2008 and 2009 at the Leipzig Medical School. Before and after the online test students were asked to document their agreement with some statements concerning computer- or web-based assessment and training. Main outcome criterion was the attitude towards CBA overall. Further items addressed the acceptance of computer- or web-based educational methods, the perceived ease of use of CBA, the perceived objectivity of CBA, the acceptance of the results of a CBA and the extent of reservations about technical problems. Items were self-created. The response format was always a ten-point Likert scale (“do not agree at all” to “agree absolutely”). In the course of the pre-evaluation students were furthermore asked to document their age and gender. Computer self-efficacy was assessed with a single self-created item. The students rated on a ten-point Likert scale how sure they feel in dealing with computers (“feel very unsure” to “feel very sure”). Pre- and post-evaluation each took 3 min. As there was no technical possibility to force students to edit all offered components in the right order, they were previously asked emphatically to do so. Only data from subjects who complied with this instruction were included in the statistical analysis. Furthermore there have been no completeness checks. The test as well as the evaluation components could have been transmitted incompletely. Log file analyses were conducted to check the utilisation of the user interface try-out and the recall of the web test results. Comparisons were made between pre- and post-evaluation, to some extent in consideration of gender, age and computer self-efficacy. Furthermore the two years (2008 and 2009) were compared concerning the pre-evaluation, regarding possible influence of the provided user interface try-out in 2009. We thereby followed the “intention-to-treat” principle, meaning that all students of the year 2009, regardless of whether they made use of the user interface try-out or not, were compared with all students of the year 2008.

2.3. Statistical analysis

Data was analysed with IBM SPSS Statistics 18 for Windows (SPSS Inc., 2009). In addition to descriptive analysis, a one-sample Chi-square test and a one-sample t-test to examine the representativeness of the sample with regard to gender and age, we conducted Chi-square tests to analyse frequency differences between groups, one-sample Kolmogorov–Smirnov tests to test for normal distribution, Mann–Whitney U tests to analyse mean differences between groups and Wilcoxon signed-rank tests as well as two-way repeated measures analyses of covariance (ANOVA) to check pre-post differences. Taking into account the multiple testing problem, local levels of significance were adjusted according to the Bonferroni–Holm method.

3. Results

The optional web-based mock test proceeded without technical problems in both runs. 449 of 687 students (65.4%) from both years (2008 and 2009) participated. All of them completed the test. In doing so 383 of 449 (85.3%) complied with the right order (pre-evaluation, mock exam, post-evaluation) and could therefore be analysed for pre-post comparisons. Fluctuating sample sizes in the different comparisons were due to several missing values. The results of the web test were recalled by 86.9% of the participants.

3.1. Sample characteristics

The sample characteristics, including gender, age and computer self-efficacy are shown in Table 1. We found no statistically significant differences between the sample and the basic population of all 687 students with regard to gender distribution (66.9% vs. 62.4% female; Chi² = 3.331, p = 0.068) and age (mean (SD): 23.9 (2.3) vs. 24.1 (2.4); p = 0.115). While overall students declared themselves to be quite confident in dealing with computers (mean 7.6, SD 1.8) we found significant differences between male (mean 8.3; SD 1.6) and female (mean 7.3, SD 1.7) students (Mann–Whitney U Test: p < 0.001). 94.0 percent of the female and 98.4 percent of the male students described a moderate or high computer self-efficacy (“high” = 7–10, “moderate” = 5–6, “low” = 1–4 points on the used Likert scale).

<table>
<thead>
<tr>
<th>gender [n (%)]</th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>255 (66.9)</td>
<td>126 (33.1)</td>
</tr>
</tbody>
</table>

| age [mean (SD)] | 23.9 (2.3) |

<table>
<thead>
<tr>
<th>computer self-efficacy [%]</th>
<th>female</th>
<th>male</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>74.9</td>
<td>87.0</td>
<td>78.9</td>
</tr>
<tr>
<td>moderate</td>
<td>19.1</td>
<td>11.4</td>
<td>16.6</td>
</tr>
<tr>
<td>low</td>
<td>6.0</td>
<td>1.6</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Table 1
Student characteristics (valid data: \(n = 381\)).
Overall I think CBA is a good idea.  

<table>
<thead>
<tr>
<th>N</th>
<th>Pre mean (SD)</th>
<th>Post mean (SD)</th>
<th>p value (Wilcoxon)</th>
<th>Post increased agreement [%]</th>
<th>Post decreased agreement [%]</th>
<th>Agreement remained constant from pre to post [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>358</td>
<td>6.0 (2.7)</td>
<td>6.3 (2.8)</td>
<td>&lt;0.001*</td>
<td>36.0</td>
<td>19.8</td>
<td>44.1</td>
</tr>
<tr>
<td>366</td>
<td>6.8 (2.5)</td>
<td>7.1 (2.4)</td>
<td>0.002*</td>
<td>35.5</td>
<td>21.6</td>
<td>35.8</td>
</tr>
<tr>
<td>362</td>
<td>5.0 (2.8)</td>
<td>5.6 (2.8)</td>
<td>&lt;0.001*</td>
<td>45.3</td>
<td>18.5</td>
<td>36.2</td>
</tr>
<tr>
<td>361</td>
<td>5.1 (2.7)</td>
<td>5.4 (2.5)</td>
<td>0.001*</td>
<td>32.4</td>
<td>21.6</td>
<td>46.0</td>
</tr>
<tr>
<td>362</td>
<td>6.7 (3.0)</td>
<td>6.9 (2.7)</td>
<td>0.246</td>
<td>25.7</td>
<td>25.4</td>
<td>48.9</td>
</tr>
<tr>
<td>361</td>
<td>8.1 (2.2)</td>
<td>7.9 (2.3)</td>
<td>0.021</td>
<td>17.2</td>
<td>24.4</td>
<td>58.4</td>
</tr>
</tbody>
</table>

*significant at a local level of significance after adjusting for multiple testing (Bonferroni–Holm method).

3.2. Attitudes towards CBA and pre to post changes

Students’ attitudes towards CBA and changes between the two measuring times are shown in Table 2. The attitude towards CBA overall was significantly more positive after the web test experience. The acceptance of CBA/WBA as educational method, the perceived ease of use of CBA and the perceived objectivity of CBA were significantly changed in a positive sense. While the data also showed fewer reservations about technical problems and an increased acceptance of the results of a CBA after the web test experience on a descriptive level, these changes were not found to be statistically significant at the respective local level of significance.

3.3. Pre to post changes in consideration of gender and computer self-efficacy

Table 3 shows the results of a two-way repeated measures ANCOVA comparing male and female students, adjusted for computer self-efficacy. Female students show significantly more reservations about technical problems compared to male students prior to the web test as well as after it. We found a significant interaction of gender and pre-post effect regarding the overall attitude towards e-testing. While male students attitude seem to remain constant over time on average, the previously more negative attitude of the female students seems to converge to that of their male colleagues at the second measuring time. We found a significant impact of computer self-efficacy on the variance between male and female students concerning every item with the exception of the perceived objectivity of a CBA.

3.4. Influence of a user interface try-out

There were no statistically significant differences between the two years (2008 and 2009) regarding age, gender and computer self-efficacy. The user interface try-out offered in 2009 was used by 69.6% of the participants in this year. Within the year 2009 students who made use of the user interface try-out were not significantly different from students who did not in terms of age, gender and computer self-efficacy. Comparisons of the student pre-evaluations between the two years revealed no statistically significant differences (p = 0.024 to p = 0.900, Mann–Whitney U test). On a descriptive level we found a higher agreement in 2009 concerning the perceived ease of use of CBA (mean ± SD: 5.2 ± 2.9 vs. 4.6 ± 2.5), but this difference failed to be statistically significant at the local level of significance after adjusting for multiple testing (p = 0.024).

Table 3

Changes in students’ attitudes concerning computer- and web-based assessment and training in consideration of gender and computer self-efficacy (two-way repeated measures ANCOVA adjusted for computer self-efficacy).

<table>
<thead>
<tr>
<th>Mean pre (SD)</th>
<th>Mean post (SD)</th>
<th>Within-subjects</th>
<th>Between-subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>male</td>
<td>female</td>
<td>male</td>
</tr>
<tr>
<td>Overall I think CBA is a good idea.</td>
<td>6.4 (2.8)</td>
<td>5.9 (2.6)</td>
<td>6.4 (2.8)</td>
</tr>
<tr>
<td>Computer- or web-based training should play a more important role in medical school.</td>
<td>7.1 (2.5)</td>
<td>6.7 (2.5)</td>
<td>7.2 (2.5)</td>
</tr>
<tr>
<td>CBA are easier to handle than paper-and-pencil-tests.</td>
<td>5.2 (2.9)</td>
<td>4.9 (2.7)</td>
<td>5.6 (2.8)</td>
</tr>
<tr>
<td>A CBA reflects objectively a student’s performance.</td>
<td>4.9 (2.6)</td>
<td>5.2 (2.7)</td>
<td>5.1 (2.4)</td>
</tr>
<tr>
<td>I would accept the results of a CBA as well as those of a paper-and-pencil test.</td>
<td>6.8 (3.2)</td>
<td>6.7 (2.9)</td>
<td>6.8 (2.8)</td>
</tr>
<tr>
<td>I have reservations about technical problems influencing the accomplishment of a CBA.</td>
<td>7.2 (2.6)</td>
<td>8.5 (1.9)</td>
<td>7.2 (2.4)</td>
</tr>
</tbody>
</table>

*significant at a local level of significance after adjusting for multiple testing (Bonferroni–Holm method).
4. Discussion

Our optional web-based mock examination proved to be technically feasible and was used by a large number of students. The vast majority of the participants said that they felt confident in dealing with computers, with female students being significantly less convinced of their own abilities. Overall, students' attitudes towards CBA were already fairly positive prior to the web test and improved after the actual experience. Positive changes were found concerning the overall attitude towards CBA, as well as in terms of the perceived ease of use, the perceived objectivity, and the acceptance of computer- or web-based methods as teaching components in higher education. We found many reservations about technical problems that still remained after exposure to the test. Attitudinal differences between male and female students seemed to be influenced by differences in computer self-efficacy. The women's reservations about technical problems affecting the accomplishment of CBA were significantly higher. However, especially the female students' overall attitude towards CBA changed positively after undergoing the assessment. Although different prior to the test, male and female student attitudes converged after exposure. Even though not significant on a local level of significance after adjusting for multiple testing, the present data suggest that a user interface try-out prior to the web test has an effect on the perceived ease of use.

The examined student sample appeared to be representative of the basic population of all medical students in the two years with regard to gender and age. The overall high self-reported confidence in dealing with computers is in line with results of other recent studies. Students of today were found to be highly experienced in computer- and web-based technologies (e.g. Kennedy, Gray et al., 2008; Kennedy, Judd et al., 2008; Link & Marz, 2006) and mostly confident about their computer skills (e.g. Krueckeberg, Paulmann, Fischer, Haller, & Matthies, 2008). Considering the rapid development of new information and communication technologies in recent years, it is to be expected that students' computer skills, which are already good, will continue to improve. This will, in turn, have a positive influence on the acceptance of computer-based learning (Reynolds, Rice, & Udden, 2007) and examination methods. Although both female and male students in our study reported comparatively high computer self-efficacy, male students felt significantly more confident in dealing with computers. Many studies have reported comparable gender differences in computer self-efficacy (e.g. He & Freeman, 2010; Isman & Celikli, 2009; Karsten & Schmidt, 2008; Li & Kirkup, 2007; Richter, Naumann, & Horz, 2001). According to He and Freeman (2010), these differences can be explained by less computer knowledge, less computing experience, and greater computing anxiety on the part of the females. Gender differences concerning the use of computers and internet were described by Li and Kirkup (2007), for instance. Despite the gender differences described in the literature regarding computer skills and use, female students do not seem to perform worse in CBA than male (Kies et al., 2006).

As described in other studies (DeAngelis, 2000; Kopp et al., 2005; Lim et al., 2007; Ogilvie, Trusk, & Blue, 1999) the overall attitude towards CBA in our sample tended to be positive, albeit students did not express any enthusiasm. The web test experience led to a more positive overall attitude, even though this effect seemed to be solely due to attitudinal changes in the female students. Gender differences concerning overall attitude were found to be influenced by computer self-efficacy. On average, students in our sample endorsed increased implementation of computer- or web-based training in higher education. The endorsement increased significantly after the web test experience. Corresponding gender differences were affected by computer self-efficacy. We found the strongest descriptive alteration after undergoing the web test was with respect to the perceived ease of use of CBA, with nearly half of the students changing their attitudes positively. Also concerning the perceived ease of use, gender differences seemed to be affected by computer self-efficacy. This finding is consistent with the results of other recent studies indicating an indirect effect of computer self-efficacy on perceived ease of use (Saadé & Kira, 2009; Terzis & Economides, 2011a) and a gender difference in computer self-efficacy (Haywood et al., 2004; He & Freeman, 2010; Karsten & Schmidt, 2008; Li & Kirkup, 2007). Even though appraisals of the objectivity of a CBA were only moderate both prior to and after the web test, we found a significant positive influence of the web test experience on students' perceptions. In agreement with results of Dermo (2009), the present data do not suggest an association between perceived objectivity and gender. Acceptance of CBA results tended to be high even before the web test and did not change significantly. Once again, the corresponding gender differences seemed to be affected by differences in computer self-efficacy. A positive but still somewhat reluctant acceptance of CBA results by German students was already found by Fischer and Kopp (2006). They assumed the reluctance to be caused by the still low prevalence of computerised examinations in higher education in Germany. Reservations regarding technical problems were strong before and after the web test. Cassidy and Gridley (2005) declared testing security to be a constant concern in online testing as well. A small descriptive decline due to the web test experience was not statistically significant on a local level of significance after adjusting for multiple testing. Nevertheless, it may indicate that a positive CBA experience has a small effect on the perception of CBA security, as already assumed by Krueckeberg, Koesling, and Matthies (2008). Although male and female students expressed major technical concerns, those of the female students were considerably greater, explained partly by differences in computer self-efficacy. The gender differences we found concerning technical reservations are not compatible with the findings of Dermo (2009), who reported that students considered e-assessment in general to be secure and did not detect a corresponding gender gap. Our results on the effect of the user interface try-out support the hypothesis of a direct influence of technical support on the perceived ease of use, as described by Ngai et al. (2007). The difference between students in the two years was not significant on a local level of significance after adjusting for multiple testing. Nevertheless, it can be interpreted as at least an indication of a corresponding effect, especially in consideration of the "intention-to-treat" comparison, including more than 30 percent of the 2009 students who did not use the offered user interface try-out. Unexpectedly, the user interface try-out in our study did not reduce technical reservations prior to the CBA. We think it is expedient, however, that participants in CBA familiarise themselves with its user interface in advance.

Studies investigating attitudinal changes due to CBA experiences are rare. According to Liao and Lu (2008), experiences with e-learning increase the intention of its continued use. Fluck et al. (2009) found evidence that previous successful exposure to CBA leads to higher acceptance of the computer medium, and they concluded that creating positive CBA experiences is an important step for its adoption. He and Freeman (2010) suggest that increasing computer experiences will help women to develop more confidence in dealing with computers and possibly lead to increased utilisation of computer-based exercises. Ogilvie et al. (1999) reported that CBA raised students' levels of
comfort with computers. In summary, the results of the present study support the hypothesis of a positive CBA experience effecting positive attitudinal changes with respect to the overall attitude, the role of computer- or web-based educational methods, the perceived ease of use, and the perceived objectivity of CBA.

4.1. Strength and limitations

To our knowledge, the present study is the first substantial pre-post study investigating students' attitudinal changes towards CBA after undergoing an actual experience. Because computer- or web-based assessment is not yet established at Leipzig Medical School, we were able to explore a widely "uninfluenced" and inexperienced student group with regard to CBA. This seemed to be a beneficial precondition. The sample size of the study was appropriate to consider the main research questions with sufficient statistical power. One limitation of the present study concerns the participation in the mock exam by choice. Although we found the gender ratio of the sample to be representative of the students in both years, we cannot exclude selectivity concerning computer self-efficacy, interest in and acceptance of CBA. This could possibly lead to an overestimation of the students' acceptance of CBA and of their self-evaluated computer self-efficacy. Secondly, this study analysed the change of attitudes within a very short period, which requires scales sensitive to possible changes between the two times of measurement. We therefore decided to use a ten-point Likert Scale and accepted the sometimes difficult interpretation of small but significant changes in the means on such a scale. This could also be discussed as a limitation. In any case, the long-term effect of the intervention remains open. The third limitation of this study may be the simple design of the web-based assessment used. Particular advantages of CBA, for example the possible use of pictures and videos, were not exhausted. A design integrating all the advantages of the medium might have been even more persuasive. A fourth limitation could be that the construct "computer self-efficacy" was measured in quite a simple manner, with only one question. Therefore we are not able to provide sound statements with regard to validity and reliability. However, the compatibility of our results with other findings indicates validity. Finally we want to mention the use of ANCOVA even though the assumption of normal distribution was not fulfilled for a few variables. Admittedly, ANOVA and ANCOVA have been shown to be remarkably robust when failing to meet this assumption (Bortz & Schuster, 2010).

4.2. Conclusion and implications for practice

Large numbers of students are interested in the use of formative CBA for exam preparation. Student computer self-efficacy can be assumed to be high. Overall, students' attitudes towards computer- or web-based assessment methods in higher education tend to be positive. Gender differences seem to be substantially influenced by differences in computer self-efficacy and seem to reduce after only one practical experience. Exposure to an actual CBA experience seems to have a positive influence on attitudes towards CBA. Nevertheless, there are strong reservations about technical problems which could influence test performance and results in summative assessments. These concerns should not be ignored when trying to implement CBA. The positive attitudinal changes and the reduction of respective gender differences after a practical CBA experience as found in the present study suggest that early practical experiences would be a helpful component of a successful implementation strategy for CBA in higher education. Positive experiences already in school would be desirable. The implementation of optional formative CBA early in higher education could familiarise students right from the start with respective learning and assessment methods and could help to reduce reservations. The results of our study indicate that the interest of the students is high. Furthermore formative web-based assessment proved to be feasible with limited effort regarding time, staff and money. Once established, the contents could be easily modified for the use in other courses. Further research could investigate the effect of multiple positive CBA experiences on students' attitudes based on longitudinal data. It may be of interest, how attitudes and reservations change after receiving feedback about the performance and how many positive experiences would be helpful to achieve maximum effect. In summary, our results implicate that optional formative CBA, preferably early in higher education, seem to be a promising possibility of attracting students to computer- or web-based examination and learning methods, and may be a useful component of a successful implementation strategy.

Funding

None.

Contributors

TD contributed to study conception and design, data acquisition, analysis and interpretation and the draft and revision of the paper. KH contributed to the statistical interpretation and to the draft and the revision of the paper. TF contributed to the statistical interpretation and to the revision of the paper. HS contributed to study conception and design and the revision of the paper. All authors approved the final manuscript for publication.

Conflicts of interest

None.

Acknowledgements

We thank Prof. Dr. Heinz-Werner Wollersheim from the Department of Educational Sciences of the University of Leipzig for his support of our e-testing project by allowing us the use of the internet platform “elate” and by giving valuable advice. Furthermore, we are grateful to those employees of the Computer Centre of the University of Leipzig who gave us technical support.
Appendix

Fig. A1. Screenshot of the user interface test (test questions introducing different response formats).

References


CERI, & OECD. (2010). Are the new millennium learners making the grade? Technology use and educational performance in PISA. Paris: CERI.


