British Journal of Educational Technology (2011)
doi:10.1111/j.1467-8535.2011.01221.x

Applause as an achievement-based reward during a computerised self-assessment test

Christos N. Moridis and Anastasios A. Economides

Christos N. Moridis received his PhD in Information Systems from the Information Systems Department at the University of Macedonia, Thessaloniki, Greece. His present research explores different types of affective feedback for tutoring systems, such as animation, embodied conversational agents and biofeedback. Anastasios A. Economides is an associate professor and chairman of the Information Systems Department at the University of Macedonia, Thessaloniki, Greece. His research interests include e-learning, e-services and networking techno-economics. He has published over 150 peer-reviewed papers. He has been the plenary speaker in two international conferences. He has served on the editorial board of several international journals, on the program committee of many international conferences and as a reviewer for many international journals and conferences. He is an IEEE Senior member. Finally, he has been the principal investigator of several funded projects and participated in many funded projects. Address for correspondence: Dr Christos Moridis, University of Macedonia, 156 Egnatia Avenue, Thessaloniki 54006, Greece. Email: papaphilips@gmail.com

Abstract
Affective feedback during a self-assessment test could help induce the learner to an optimal emotional state regarding the learning material. However, there is a lack of experimental evidence concerning the influence of affective feedback during a self-assessment test. This paper is a step towards this direction. The effect of achievement-based reward feedback on students' state and trait anxiety was examined. Ninety-two students participated in an experiment. Half of these students received an applause sound after a correct answer to a question. Results highlight gender differences concerning this emotional type of feedback.

Practitioner Notes
What is already known about this topic

- Excessive anxiety has been shown to seriously impair students' performance and learning procedures.
- The affective learning community has indicated achievement-based reward feedback as an efficient way to express affective feedback.
- Gender differences have been indicated to be essentially important concerning the perception of feedback during learning or testing.

What this paper adds

- It provides experimental evidence concerning the effect of achievement-based reward (the sound of applause after a correct answer to a question) on students' state anxiety during a computerised self-assessment test.
- It experimentally examines how the interaction of this particular feedback with students' state and trait anxiety would relate to students' performance.
- It clearly reveals that this particular kind of reward has a different effect on males and females.

Implications for practice and/or policy

- Computerised assessment procedures would greatly benefit from the development and integration of affective feedback strategies.
- An appropriate computer response (affective feedback) to students' affective state also requires evolving and integrating new pedagogical models into computerised learning environments.
- The gender-specific effects of the achievement-based reward feedback, revealed in this study, may also apply to other kinds of affective feedback. Therefore, the 'gender dimension' should probably be seriously taken into account when designing and implementing affective feedback strategies.

Introduction

Self-assessment is a key process in training students to self-regulate their cognition, motivation and behaviour towards their learning goals (Nicol & Macfarlane-Dick, 2006). Cassady and Gridley (2005) highlighted the importance of the preparation period, prior to exams, where anxiety may lead students to procrastinate and fail to use adequate cognitive skills to master the learning material. They proposed that training through preparation tests could induce higher-test performance in students suffering from test anxiety and test failure. This sense also relates to stress-inoculation theory, whereby persons exposed to small doses of a frightening experience eventually experience less fear and anxiety (Snooks, 2004). Moreover, preparation tests do not impose the stressful experience of a typical exam, and so the detrimental effects of anxiety on students' cognitive skills are not triggered (Cassady, 2004). Thus, preparation tests could help students acquire an effective learning strategy and establish self-confidence before final exams are administered.

In addition, it has been shown that some students would prefer assessment delivered through a computer (Croft, Danson, Dawson & Ward, 2001; Ricketts & Wilks, 2002a; Sambell, Sambell & Sexton, 1999). Furthermore, Sim and Holifield (2004) provided evidence that students' prior experience of using computers in assessment does not seem to significantly influence their preference for having their tests computerised. In addition, Ricketts and Wilks (2002b) provided evidence that computerised assessments could be more user friendly for certain groups of disabled students. Moreover, they have indicated that computerised assessments may cause less anxiety in dyslexic students than paper-based assessments. Besides, there is increasingly more research evidence indicating that the majority of students consider testing through computers less stressful than paper-based assessment (Bocij & Greasley, 1999; Ricketts & Wilks, 2001). Reasons for students to prefer computerised assessments include being able to take a test at the time and place they choose (particularly when the test can be delivered online), being judged without bias (the computer does not know who you are) and being able to have their test results delivered immediately (Bocij & Greasley, 1999; Cassady & Gridley, 2005). Consequently, a computerised self-assessment test system could help students improve their knowledge and acquire a positive attitude towards the learning material.

A way of implementing computerised self-assessment is through a multiple choice question (MCQ) test that could be used to support the learner by providing adequate feedback (Nicol, 2007). Even though MCQs have been criticised for being less effective in capturing higher-order cognitive skills (Gronlund, 1988; Welsh, 1978), it has been established that multiple choice tests can be used to assess higher-order knowledge if the questions are designed appropriately (Brown, Bull & Pendlebury, 1997; Bull & McKenna, 2001). Furthermore, MCQs offer numerous advantages, including the chance to test a greater part of the syllabus (the writing time is minimised in

the multiple choice format), the chance to provide more effective feedback and the ability to mark the assessments online (Ramesh, Sidhu & Watugala, 2005).

However, Ricketts and Wilks (2002b) found that the presentation of questions of a MCQ test may affect students' performance. Median marks of students answering questions presented on-screen in a scrolling format were 14% lower when compared with questions presented on paper and 20% lower when compared with questions presented on-screen one at a time. Using a scroll bar to navigate through questions may have interfered with the students' concentration. Other studies have also indicated that displaying questions one at a time helps students to focus on a single item (Clariana & Wallace, 2002).

More importantly, as mentioned earlier, a computerised MCQ test could incorporate effective feedback practices. Research suggests that feedback integrated in the learning process could significantly advance performance and learning gains (Black & Wiliam, 1998a; Taras, 2003). Nicol and Macfarlane-Dick (2006) suggested seven principles of good feedback practice that support self-regulation: (1) clarify what good performance is (goals, criteria and expected standards); (2) facilitate the development of self-assessment (reflection) in learning; (3) deliver high-quality information to students about their learning; (4) encourage teacher and peer dialogue around learning; (5) encourage positive motivational beliefs and self-esteem; (6) provide opportunities to bridge the gap between current and desired performance and (7) provide teachers with information that can be used to help shape the teaching.

Viewing feedback merely as a cognitive process limited to transmission of information ignores the way feedback interacts with motivation and beliefs (Economides, 2005, 2009; Nicol & Macfarlane-Dick, 2006). The idea of a computer system that is capable of taking into account users' affective states has introduced a relatively new research field—that of affective computing. According to Picard (1997), affective computing is 'computing that relates to, arises from or deliberately influences emotions'. Affective computing could apply to a variety of computer applications to enhance users' experience. Concerning e-learning, researchers of AI in education considered developing systems, which are capable of modelling students' affective states and providing an adequately tailored response (affective feedback) based on pedagogical models (Moridis & Economides, 2008). Previous studies proposed the integration of various types of affective feedback in computerised tutoring systems (Economides, 2005, 2006; Heylen, Vissers, Akker & Nijholt, 2004; Lester, Towns & FitzGerald, 1999; Picard *et al.*, 2004). However, few studies experimentally evaluated the effectiveness of each feedback type (Hone, 2005; Kang, Gratch, Wang & Watt, 2008; Klein, Moon & Picard, 2002; Lankes *et al.*, 2008; Whang *et al.*, 2007).

Economides (2005, 2006) proposed an emotional feedback framework, taking the Computer Adaptive Testing systems as a field of application in order to manage emotions. When the effect of negative emotions is too strong, the student's performance can be significantly hampered. Errors could produce the expectation of more errors, arousing negative emotions, causing even more wrong answers until the student's performance eventually collapses. Nevertheless, positive emotions may also necessitate instruction. For example, answering a hard question correctly could not only produce positive emotions such as joy and enthusiasm, but also cause loss of concentration if too much attention is given to the elicited emotions. Without pedagogical feedback, positive emotions can make students focus on the excitement and undervalue the effort required to obtain a successful result (Efklides & Volet, 2005).

The emotional feedback can occur before and after the test, during the test, and before and after a student's answer to a question (Economides, 2006). In all these cases, emotional feedback can be provided either automatically according to the student's emotional state, either upon the student's or the teacher's request. The core channels/methods for recognizing emotions are the following: (1) questionnaire, (2) personal preference information (emotions could be predictable

if users' goals and perception of relevant events were known), (3) emotional speech recognition, (4) physiological data and (5) facial expressions. From all of these channels, researchers of AI in education are attempting to infer the student's emotional state (Moridis & Economides, 2008). The student's recognised emotional state should be properly managed from an affective test system based on pedagogical models that integrate our knowledge about emotions and learning. The system would assess whether the process is developing at a healthy rate. If there is a positive development, the system should help the student maintain this emotional state. If not, the system should induce the student to an emotional state beneficial to the process.

An affective test system could implement emotional feedback by using constructively positive emotions while preventing, controlling and managing negative emotions. Moreover, the emotional feedback can also suitably use negative emotions so as to raise the student's devotion and engagement. Humour and jokes, amusing games, expressions of sympathy, reward, pleasant surprises, encouragement, acceptance, and not only praise but also criticism (Economides, 2005) are some of the possible actions that could be practised by an affective testing system.

As stated earlier, however, it has been shown that some students would prefer computerised assessment, there is research evidence that computerised tests may cause fear and anxiety in some people (Abdelhamid, 2002; Igbaria & Chakrabarti, 1990; Rosen, Sears & Weil, 1993), which makes the need for affective feedback provided by testing systems even more intense. Personalising the testing experience requires building systems able to adapt to a wide variety of characters (Moridis & Economides, 2009a).

More importantly, Weedon (2000) showed that students' perception of individualised feedback may vary between individuals. Some students may find feedback (as opposed to mere +/− marking) more useful with regard to successes, while others may find it more useful with regard to failures. For instance, students may not perceive negative comments as criticism but as helpful insight for improving their work (Weedon, 2000).

Moreover, gender differences have been indicated to be essentially important concerning the perception of feedback. For instance, research by Burleson and Picard (2007) has revealed gender-specific impacts of feedback and affective support strategies on children. Consequently, the 'gender dimension' should be further investigated and adequately integrated into affective tutoring systems (Burleson, 2006).

Students often fail to successfully complete a learning task because of emotional pressure taking such forms as emotional pressure exerted by parents, teachers, schoolmates and even themselves. Moreover, academic success appears to hold the key to students' success in later life. All these factors may seriously impair students' learning performance, as well as their psychological and physical health (Pekrun, Goetz, Titz & Perry, 2002). Students who encounter difficulties are led to believe that they lack ability, and this belief leads them to attribute their difficulties to a defect in themselves about which they cannot do much (Black & Wiliam, 1998b). Individuals who doubt their capabilities and experience high levels of fear of failure are less likely to set and work towards goals, thus giving them no opportunities to increase levels of self-efficacy (Caraway & Tucker, 2003).

Furthermore, numerous studies have demonstrated that anxiety can seriously impair academic performance (Seipp, 1991). It is a well-known fact that learning or remembering something in a state of anxiety, anger or depression can be difficult for any individual (Goleman, 1995). Anxiety can be categorised under the two different headings of state anxiety and trait anxiety. State anxiety could be specified as a negatively alarming emotional experience in the face of menacing demands or threats (Spielberger, 1972, 1983). According to Lazarus (1991), a cognitive evaluation of threat is a precondition for the occurrence of this emotional state. However, trait anxiety reflects a predetermined level of anxiety experienced by an individual who has a predisposition to

be more anxious and respond less suitably to anxiety-arousing stimuli (Eysenck & van Berkurn, 1992). In other words, trait anxiety indicates the presence of established individual differences in the tendency to react with state anxiety in anticipation of endangering situations. On the other hand, the proper amount of anxiety or fear can help individuals to focus the mind, thus reducing distractions. It is when the negative effect is too strong that learning tasks are inhibited (Bower, 1992).

A self-assessment test system that is capable of handling students' affective and cognitive needs would help students improve their knowledge and acquire a positive attitude towards learning. According to Moridis and Economides (2009b), a system of this kind would be able to (1) recognise the current affective state of the student, (2) distinguish when and how to intervene in order to influence the student's affective state based on a holistic educational pedagogy integrating affective models in learning and (3) produce the optimal affective state for learning. Additionally, the tutoring system's affective handling has to be successfully combined with the student's cognitive handling. Consequently, further research related to the aforementioned needs is required for the effective development of such a system (Moridis & Economides, 2008).

This paper focuses on the affective handling capabilities of such a system and particularly on emotional feedback. This study examined how an achievement-based reward would influence the students' state and trait anxiety during a self-assessment test, overall and in relation to gender. Moreover, it examined how the interaction of this particular feedback with students' state and trait anxiety would relate to students' performance, overall and in relation to gender. The sound of applause was heard after a correct answer to a question as an achievement-based reward. The applause was chosen as a reward sound because it is socially perceived as providing an intense sense of support (Lupyan & Rifkin, 2003; Neda, Ravasz, Brechet, Vicsek & Barabasi, 2000). Moreover, applause is a social behaviour that has existed in various cultures for centuries. Furthermore, applause can be expressed by individuals from different social backgrounds who are attending the same social occasion (eg, a politician delivering a speech). In addition, applause can be observed in many different occasions, ranging from fans clapping to support their team to opera's audience clapping as an act of kindness to the singers (Lupyan & Rifkin, 2003). Therefore, we assumed that this kind of feedback would probably have an effect on students.

Although students may need more support when they provide an incorrect answer to a question, achievement-based rewards could also help a student recover from a detrimental emotional state and a low performance. When a student provides a correct answer, this is also a chance to present affective feedback (as a reward) so as to diminish the negative effect of previous wrong answers by increasing the student's intrinsic motivation and self-efficacy (McQuiggan, Mott & Lester, 2008; Schunk, 1984, 1989).

Research evidence suggests that achievement-based rewards during learning or testing increase individuals' intrinsic motivation (Cameron, Pierce, Banko & Gear, 2005). According to Ormrod (1995), motivation can support learning in a range of ways: (1) encourage behaviour towards certain goals, (2) facilitate increased effort and energy, (3) enhance initiation of, and persistence in, activities, (4) improve cognitive processing and (5) lead to improved performance. Moreover, Rebolledo-Mendez, du Boulay and Luckin (2006) indicated a positive effect of the motivational scaffolding, particularly on unmotivated students who showed higher-learning gains. The results of the current study could be useful for the design of future systems.

Aim of the study

This study aimed to examine the influence of applause sound as an achievement-based reward after a correct answer to a question during a self-assessment test. In this context, two issues were addressed:

1 The influence of this particular feedback on students' state anxiety levels, overall and in relation to gender.
2 How the interaction of this particular feedback with students' state and trait anxiety would relate to students' performance, overall and in relation to gender.

Experiment procedure

System

An online MCQ test system, built for a previous experiment (Moridis & Economides, 2009b), was adjusted to serve the needs of the current study. Throughout the test, a student selected his/her answer among four possible answers and confirmed his/her choice by clicking the 'submit' button. After each question, the system informed the student whether his/her answer was right or wrong and presented his/her score. In case of students assigned to the applause group, the system would additionally provide an applause sound when the answer was right. Then, the student could proceed to the next question by clicking the 'next' button.

Participants

First-year undergraduate students of a Greek university in an Introductory Informatics course were recruited. During the test, one group heard applause after each correct answer to a question and the other did not. Participants were 92 students randomly assigned to each group; 46 (22 males and 24 females) participants received applause and 46 (13 males and 33 females) participants did not. Subjects were randomly distributed into groups in a way that the applause and the nonapplause groups would be equal in size. However, this resulted in a disparate distribution of male and female participants. Subjects' age varied from 18 to 21, mean age 19.10 (standard deviation = 1.02).

Measurement scales

Students' anxiety levels were measured using the State-Trait Anxiety Inventory (STAI) (Spielberger, 2005). The self-report inventory includes the State (Y1) and the Trait (Y2) scales, 20 items to assess state anxiety and another 20 items to assess trait anxiety. These two parts differ in the item wording, the response format and the instructions on how to respond. Respondents are instructed to complete the items of the Y1 scale according to their current feeling (eg, right now). When responding to the Y1 scale, participants blacken the number on the test form to the right of each item statement (eg, I feel at ease) that best describes the intensity of their feelings: (1) not at all; (2) somewhat; (3) moderately so and (4) very much so. On the other hand, respondents are instructed to complete the items of the Y2 scale according to how they feel generally. When responding to the Y2 scale, participants blacken the number on the test form to the right of each item statement (eg, I lack self-confidence) that best describes the frequency of their feelings: (1) almost never; (2) sometimes; (3) often and (4) almost always. Scores for both the state and trait anxiety scales can vary from a minimum of 20 to a maximum of 80. The STAI is the most frequently used scale in research to measure state and trait anxiety and has received many foreign language adaptations and citations.

The reliability and validity of the Greek translation have been shown (Fountoulakis *et al.*, 2006). The state and trait scores for healthy participants were 34.30 ± 10.79 and 36.07 ± 10.47 respectively. Scores for the depressed individuals were 56.22 ± 8.86 for state anxiety and 53.83 ± 10.87 for trait anxiety. Both state and trait scores were normally distributed in control subjects. Cronbach's alpha was 0.93 for the state and 0.92 for the trait subscale. The Pearson's correlation coefficient between state and trait subscales was 0.79. Test–retest reliability was outstanding, with the Pearson's coefficient being between 0.75 and 0.98 for individual items and equal to 0.96 for Y1 and 0.98 for Y2.

Material

The MCQs were focused on basic computer knowledge and skills based on material taught in lectures. The content of questions was pre-specified by the course instructors prior to the study. The test was consisted of 45 questions. The order of questions presented was randomly altered among students.

Procedure and data collection methodology

Students were told that they could optionally participate in a self-assessment test to help them assess their knowledge before the final exam. About one third of the class took up this option. The final exam was partly MCQs. Moreover, the final exam covered the same topics covered by the self-assessment test. Those who wished to participate in the self-assessment test filled in an application form in order to arrange for an appointment. As the purpose of the self-assessment test was to help the students assess their knowledge before exams, it was left to students to decide when they would feel that such a test would be helpful to them. The duration of the self-assessment test was approximately 45 minutes.

Y1 questionnaire was distributed both before and after the test. Y2 questionnaire was then distributed. The Y2 scale has been shown to remain uninfluenced even after stressful conditions (Auerbach, 1973; Spielberger, 2005). Thus, distributing Y2 both before and after the test would be pointless because trait anxiety would probably have remained constant. However, it would be a mistake to administer Y2 before Y1. The State Anxiety scale is designed to be responsive to the particular conditions under which the questionnaire is distributed, and thus the emotionality that may be created if Y2 is distributed first could have an influence on Y1 scores (Spielberger, 2005).

Regarding students assigned to the applause group, they were informed immediately before Y1 questionnaire was distributed that after a correct answer to a question an applause sound would be played. Then, Y1 questionnaire was distributed, and the students took the test. The sound was played through headphones. Students not assigned to the applause group proceeded directly to Y1 questionnaire.

Results

Tables 1 and 2 present the basic results of this study, concerning males and females receiving and not receiving applause: average, median, and maximum and minimum values of state anxiety scores (measured before and after the test), trait anxiety scores and test performance scores are presented for each group. Then, a more sophisticated analysis is presented, including the two-way analysis of variance (ANOVA) and calculations of Spearman's correlation coefficient.

The first goal of this study was to examine whether affective feedback in the form of applause could influence students' state anxiety, overall and in relation to their gender. A two-way ANOVA was conducted. The two independent variables that crossed each other are gender and applause

Table 1: Males receiving and not receiving applause: average (AVG), median (MED), and maximum (MAX) and minimum (MIN) values for state anxiety measured before (state anxiety-1) and after (state anxiety-2) the test, trait anxiety and test performance

	Males applause (n = *22*)				*Males no applause* (n = *13*)			
	AVG	*MED*	*Max*	*Min*	*AVG*	*MED*	*Max*	*Min*
State anxiety-1	36.73	34	63	23	38.46	39	46	27
State anxiety-2	22.7	21.5	33	20	30.54	30	58	20
Trait anxiety	31.36	29	46	20	38.31	36	57	27
Test performance out of 45	36.36	36	44	30	36.23	38	40	25

Table 2: Females receiving and not receiving applause: average (AVG), median (MED), and maximum (MAX) and minimum (MIN) values for state anxiety measured before (state anxiety-1) and after (state anxiety-2) the test, trait anxiety and test performance

	Females applause (n = *24*)				*Females no applause* (n = *33*)			
	AVG	*MED*	*Max*	*Min*	*AVG*	*MED*	*Max*	*Min*
State anxiety-1	39.9	37.5	65	22	40.91	39	64	23
State anxiety-2	28.25	27	51	20	29.7	28	51	20
Trait anxiety	36.71	33.5	53	25	37.48	37	59	22
Test performance out of 45	36.37	37.5	42	25	36.4	37	43	25

Table 3: The study design

	Applause		
Gender	*Present*	*Absent*	*Total*
Female	$n = 24$	$n = 33$	57
Male	$n = 22$	$n = 13$	35
Total	46	46	92

(Table 3). This study has a 2×2 design with the two genders (male–female) and two types of applause (present–absent). The dependent variable was students' state anxiety measured before and after the test.

Data in almost all groups were positively skewed; thus, a logarithmic transformation was used to satisfy the normality assumption. Accordingly, the Anderson–Darling statistical test was used to detect departures from normality. Prior to the ANOVA test, Levene's test for equality of variances was performed to confirm that homogeneity of variance assumption was not violated. An unweighted mean analysis was selected to treat the different sample sizes in groups (unbalanced two-way ANOVA).

Concerning students' state anxiety measured before the test, the two-way ANOVA showed no significant results. The row, column and interaction effect failed to reach statistical significance. This means that state anxiety before the test for males and females was not significantly influenced by the expectation or nonexpectation of the reward sound of applause (column effect) or gender (row effect) and that the expectation or nonexpectation of applause did not significantly influence state anxiety before the test in relation to gender (interaction effect).

Concerning students' state anxiety measured after the test, the column effect was significant ($F[3.95] = 11.17$, $p < 0.01$) while the row effect was not ($F[3.95] = 0.39$, $p > 0.05$), indicating that state anxiety after the test was not significantly influenced by gender but by the presence or absence of applause. Interestingly, the interaction effect was also significant ($F[3.95] = 15.39$, $p < 0.01$), representing a significant interaction between the two factors (gender and applause). This means that the presence or absence of applause did have a significant influence on state anxiety after the test in relation to gender.

Post hoc comparisons followed, using the Bonferroni correction, to identify which means differences were significant (Table 4). Males not receiving applause were significantly more anxious than females not receiving applause. Furthermore, males receiving applause were significantly less anxious than males not receiving applause.

Table 4: Post hoc comparison results of Y1 mean scores measured after the test in each group

Comparison	*Significant? (p < 0.05?)*	T
1: Females–males receiving applause	No	2.467
2: Females–males not receiving applause	Yes	3.059
3: Females receiving applause—males not	No	2.536
4: Males receiving applause—females not	No	2.179
5: Males receiving applause—males not	Yes	4.578
6: Females receiving applause—females not	No	0.478

Given that there is an interaction along with a main effect of applause, the main effect must be re-examined for its consistency. According to the main effect, students that received applause had less state anxiety after the test. However, when looking at the cell means that we see that this is only significant for males receiving applause feedback. Hence, in this case, the main effect is not consistent. Moreover, the main effect of gender is not significant; however, the interaction between males and females not receiving applause is. Males not receiving applause had significantly higher-state anxiety after the test than females not receiving applause. According to Tables 1 and 2, for males not receiving applause, the median for state anxiety after the test is 30, whereas for females not receiving applause, the median for state anxiety after the test is 28. Furthermore, the maximum state anxieties after the test in those two groups are 58 and 51, respectively. Consequently, in this case, the statistically significant small difference in average values (30.54 vs. 29.7) of state anxiety measured after the test between males not receiving applause and females not receiving applause is important, and it deserves further discussion. Additionally, males receiving applause were significantly less anxious after the test than males that did not receive applause.

The second aim of this study was to search for correlations between STAI scales and students' performance, overall and by gender (Table 5). The Spearman's correlation coefficient was used as the non-parametric equivalent of the Pearson's correlation as data concerning score were far from being normally distributed, even after transformation efforts.

A significant correlation was found between males' state anxiety before the test and their test performance, $r = 0.454$ (33), p (two tailed) < 0.01. This shows that there is some association between males' state anxiety before the test and their test performance, indicating that, to some extent, the higher the males' anxiety before the test was, the better males performed. A low, although significant, negative correlation was found between females' trait anxiety and their test performance, $r = -0.296$ (55), p (two tailed) < 0.05. This significant negative correlation between females' trait anxiety and their test performance reflects a small degree of negative association between the two variables, proportionately to which the higher the trait anxiety was, the worse females performed.

Further analysis between females' trait anxiety and their test performance and males' state anxiety before the test and their test performance revealed that the above correlations were even higher but were significant only for those females and males that did receive applause, $r = -0.47$ (22), p (two tailed) < 0.05 and $r = 0.507$ (20), p (two tailed) < 0.05 respectively (Table 6).

Nevertheless, no significant correlations were found in any of the groups between state anxiety measured after the test and test performance.

*Table 5: Spearman's correlations between test performance score and Y1 (measured before the test)–Y2 scales, overall and by gender. Values with an * declare a significant correlation*

	Score
Males and females-Y1	0.119
$n = 92$	
Significant (two-tailed)	0.26
Females-Y1	–0.053
$n = 57$	
Significant (two-tailed)	0.697
Males-Y1	0.454*
$n = 35$	
Significant (two-tailed)	0.006
Males and females-Y2	–0.195
$n = 92$	
Significant (two-tailed)	0.063
Females-Y2	–0.296*
$n = 57$	
Significant (two-tailed)	0.025
Males-Y2	–0.018
$n = 35$	
Significant (two-tailed)	0.92

*Table 6: Spearman's correlations between females' Y2 receiving–not receiving applause and test performance scores, and males' Y1 (before the test) receiving–not receiving applause and test performance scores. Values with an * declare a significant correlation*

	Score
Females not receiving applause-Y2	–0.19
$n = 33$	
Significant (two-tailed)	0.289
Females receiving applause-Y2	–0.47*
$n = 24$	
Significant (two-tailed)	0.02
Males not receiving applause-Y1	0.263
$n = 13$	
Significant (two-tailed)	0.386
Males receiving applause-Y1	0.507*
$n = 22$	
Significant (two-tailed)	0.016

Discussion

Anxiety may seriously harm test performance, resulting in scores that do not correctly reveal a candidate's true levels of the relevant attributes (Arvey, Strickland, Drauden & Martin, 1990). This may eventually result in the selection of less talented candidates (Arvey *et al.*, 1990; Spielberger & Vagg, 1995). A self-assessment test system could help individuals to be psychologically prepared to face anxiety and other emotional states that can impair performance during the final test.

A step towards this direction is to provide these systems with adequate emotional feedback techniques. There are many types of feedback which need to be investigated before they can be purposely integrated into an affective self-assessment test system. Our experiment was designed to

explore the influence of a reward sound (applause), after a correct answer to a question, on students' state anxiety levels, overall and in relation to gender. Moreover, this experiment examined how the interaction of this particular feedback with students' state and trait anxiety would relate to students' performance, overall and in relation to gender.

There is research evidence that reward stimulus promotes memory development via dopamine release in the hippocampus prior to learning (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006). Both males and females were informed just before the test started whether they were going to be in the group receiving the sound of applause or not.

Results showed that males not receiving applause had significantly higher-state anxiety after the test than females not receiving applause. Additionally, males receiving applause were significantly less anxious after the test than males that did not receive applause. In relation to this, research evidence has suggested that females are more likely to have coping resources available to deal with anxiety before tests and are more likely to use adaptive coping behaviours during a test (McCarthy & Goffin, 2005).

However, there was no significant difference between males' and females' state anxiety before the test. The expectation or nonexpectation of applause did not significantly influence state anxiety before the test in relation to gender. It was after the test when the reward sound of applause had been experienced by males and females assigned to the applause group, which the interaction between presence–absence of applause and gender showed significant differences in state anxiety.

Nevertheless, when males were informed before the test that a reward sound of applause was going to follow each correct answer to a question, their state anxiety before the test was significantly positively correlated with scores. We propose that state anxiety before the test might have been a 'motivational device' for males faced with the challenge of answering correctly to the questions of the test in order to hear the rewarding sound of applause as many times as possible. We assume that the higher the males' state anxiety was at that point (before the test started), the more they were challenged by the possibility of being rewarded for providing correct answers. Similarly, the more they tried to provide correct answers during the test (so as to be rewarded), the better they performed.

Several studies on gender differences in self-evaluations of the intelligence quotient (Furnham & Rawles, 1995) and performance (Beyer, 1990) have indicated that males tend to exaggerate concerning their capacities, whereas females tend to underestimate themselves. Moreover, it has been shown (Beyer & Bowden, 1997) that when an educational task is considered 'masculine' (eg, mathematics and computer science), females' self-evaluations of performance are unfairly low. Accordingly, it has been indicated that male college students have more confidence in using computers than females and that even female Computer Science majors could be less confident about computer use than would be male non-majors (Beyer, Rynes, Perrault, Hay & Haller, 2003).

More importantly, self-perceptions can have an influence on behaviour or mental health even if they are not accurate (Beyer & Bowden, 1997; Taylor & Brown, 1988). Thus, inaccurate self-evaluations can diminish intrinsic motivation to compete on a challenging task (Beyer, 1998). This may provide some justification for experimental evidence suggesting that women are not as motivated as men in competitive environments (Gneezy, Niederle & Rustichini, 2003).

Accordingly, there is evidence that males show greater sensitivity to whether they receive reward or not than females in a social as well as in a personal context (Li, Huang, Lin & Sun, 2007). Concerning this, researchers have shown that during a video gameplay, the part of the brain that generates rewarding feelings is more active in men than in women (Hoeft, Watson, Kesler, Bettinger & Reiss, 2008). We suggest that this could also be observed during a self-assessment

test. Nevertheless, this is only our hypothesis, and future research will show if such an assumption stands true. Besides, as already discussed earlier, males' communication has a tendency to express components of social status, seeking to establish and maintain reputation and dominance in their group (Maltz & Borker, 1983). Thus, the reward sound of applause could be interpreted as a competitive characteristic of the system that satisfied the male need for distinction, producing for males a 'motivational device' prior to the test.

Yet, no significant correlations were found in any of the groups between state anxiety measured after the test and scores. We assume that as correlation quantifies how well these two variables are associated, a reason for not finding any significant correlation may be that the second variable (test performance) was not 'active' when the first variable (state anxiety) was measured (after the test). It is likely that some factors (eg, test preparation, personal characteristics, time and place of taking the test, etc.) influence both variables. However, while any influence on test performance had ceased to exist (the test was over), several factors may have continued to influence state anxiety at the time it was measured (after the test).

There was no significant difference in the levels of trait anxiety between males and females in this experiment. This is consistent with earlier research stating that there are no significant gender differences concerning trait anxiety in science (Zoller & Ben-Chaim, 1989) or in any of the disciplines (Westerback & Long, 1990).

However, when females heard applause, there was a significant negative correlation between females' trait anxiety and their scores, indicating that the higher the trait anxiety was, the worse females performed. We suggest that the reason for this observation in the female group that received applause may result from the way females interpreted the sound of applause. Thus, we propose that it could be implied that females related the rewarding sound of applause to the already anxiety provoking experiences they had concerning a test. Pursuing reward through the sound of applause was a competitive characteristic of the self-assessment test that worked for males but not for females.

Several studies related to this (Gorriz & Medina, 2000; Hartmann & Klimmt, 2006; Lucas & Sherry, 2004) have pointed out that there is a gender gap concerning video games. One main finding of this research was that female players were unlikely to enjoy game-play situations involving games that incorporated competition features. Usually, winning is less important for females than it is for males. Women and girls tend to have a preference for games involving exploration and narrative (Gorriz & Medina, 2000).

Female preference for less competitive tasks may be explained by socialisation factors that prescribe gender-specific expectations concerning anxiety provoking stimuli (McLean & Anderson, 2009). Thus, males may have learned that the male role involves courage and determined coping behaviour when encountering anxiety provoking situations (Bem, 1981), while females may have learned that the female role typically tends to avoid competition. Nevertheless, gender differences in terms of competitive behaviour could also be attributed to biological factors (McLean & Anderson, 2009). Gender differences must be further examined and better interpreted so that the development of successful affective learning systems can be facilitated (Burleson, 2006).

Although the results of this study are in line with other research on gender differences in education, they should be confirmed by studies targeting the settings of a self-assessment test platform. The environment of a self-assessment test platform is a separate case where affective feedback has not been examined sufficiently. Moreover, the content of the self-assessment test, in this study, was focused on basic computer knowledge and skills based on material taught in lectures. Possibly, results might have been different if the test involved a different course (not involving computers and technology). In addition, students took the self-assessment test in a university laboratory. They might have behaved differently if the test had taken place at the

convenience of their home. This is an important factor as the aim of self-assessment test platforms is online use, helping students prepare for exams at the time and place they choose. Also, the effect of applause should be confirmed by exposing students to this kind of experience during more than one self-assessment test session.

Furthermore, this study focused on the specific sound of applause as a reward after a correct answer to a question. Maybe a different kind of reward (eg, a musical rhythm) would have provided different results. Thus, we should keep in mind that these results relating to reward concern only this particular kind of feedback.

Conclusion

Despite the limitations of this study, it is clear that this particular kind of reward (the sound of applause after a correct answer to a question) has a different effect on males and females. This may also be the case for other kinds of affective feedback. Therefore, research into affective feedback should also take into account gender differences. Moreover, other parts of an affective learning system (eg, the interface) may require a gender-specific approach. Thus, research in the direction of gender differences concerning different parts of affective learning systems will greatly contribute to the personalisation of the learning experience while using such systems.

As discussed earlier, males not receiving applause were significantly more anxious after the test than females not receiving applause and than males that received applause. Nevertheless, no such effect was observed on state anxiety prior to the test. Moreover, a significant positive correlation was observed between state anxiety before the test and test performance by males assigned to the applause group. However, no such effect took place after the test. Finally, a significant negative correlation was found between trait anxiety (measured after the test) and test performance by females assigned to the applause group. Some interpretation of these results is given at the discussion section of this paper. We envisage these results as a step towards providing self-assessment test platforms with affective feedback capabilities.

In this paper, we particularly focused on the effect of achievement-based reward feedback on students' state and trait anxiety in the context of a self-assessment test platform. We consider this kind of feedback to be one of the many tools that should be added to the platform's 'menu of options for affective feedback'. Because it is not clear so far which kind of affective feedback and under what conditions would be most effective, empirical findings from future research could help develop various kinds of affective feedback that could be effective under different conditions (eg, trying to help students with excessive fear of failure).

Affective feedback is only one of many capabilities that a self-assessment test system could offer to become more efficient. A flexible system would take into consideration the student's current knowledge and learning preferences to generate individualised tests. In addition, the system would try to introduce students to an emotional state conducive to learning by providing adequate feedback (Economides, 2005, 2006). Furthermore, affective handling has to be successfully combined with cognitive handling. Moreover, the use of this platform should be effectively embedded within each educational context.

Thus, we are dealing with several different elements that need to be combined effectively to produce a new generation of self-assessment test systems. Therefore, a way to do this is to formulate and establish each part in separation and then try to determine how they can all be combined to produce optimal results. Clearly, this has to be a joint effort, bringing together researchers from various fields.

Acknowledgement

We thank the anonymous reviewers for their constructive comments.

References

Abdelhamid, I. S. (2002). Attitudes toward computer: a study of gender differences and other variables. *Journal of the Social Sciences*, *30*, 285–316.

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B. & Gabrieli, J. D. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, *50*, 507–517.

Arvey, R. D., Strickland, W., Drauden, G. & Martin, C. (1990). Motivational components of test-taking. *Personnel Psychology*, *43*, 695–717.

Auerbach, S. M. (1973). Trait-state anxiety and adjustment to surgery. *Journal of Consulting and Clinical Psychology*, *40*, 264–271.

Bem, S. L. (1981). Gender schema theory: a cognitive account of sex typing. *Psychological Review*, *88*, 4, 354–364.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluation of performance. *Journal of Personality and Social Psychology*, *59*, 960–970.

Beyer, S. (1998). Gender differences in self-perception and negative recall bias. *Sex Roles*, *38*, 103–133.

Beyer, S. & Bowden, E. M. (1997). Gender differences in self-perceptions: convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin*, *23*, 157–172.

Beyer, S., Rynes, K., Perrault, J., Hay, K. & Haller, S. (2003). Gender differences in computer science students. Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education, Reno, NV.

Black, P. & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, 5, 1, 7–74.

Black, P. & Wiliam, D. (1998b). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, *80*, 2, 139–148.

Bocij, P. & Greasley, A. (1999). Can computer-based testing achieve quality and efficiency in assessment? *International Journal of Educational Technology*, *1*, 1. Retrieved July 24, 2010, from http://www.ed.uiuc.edu/ijet/v1n1/bocij/index.html

Bower, G. (1992). How might emotions affect learning? In C. Svenake & E. Lawrence (Eds), *Handbook of emotion and memory: research and theory* (pp. 3–31). Hilsdale, NJ: Erlbaum.

Brown, G., Bull, J. & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.

Bull, J. & McKenna, C. (2001). *Blueprint for computer assisted assessment*. London: RoutledgeFalmer.

Burleson, W. (2006). Affective learning companions: strategies for empathetic agents with real-time multimodal affective sensing to foster meta-cognitive and meta-affective approaches to learning, motivation, and perseverance (Ph.D. thesis, Massachusetts Institute of Technology).

Burleson, W. & Picard, R. W. (2007). Gender-specific approaches to developing emotionally intelligent learning companions. *Intelligent Systems*, *22*, 4, 62–69.

Cameron, J., Pierce, W. D., Banko, K. M. & Gear, A. (2005). Achievement-based rewards and intrinsic motivation: a test of cognitive mediators. *Journal of Educational Psychology*, *97*, 4, 641–655.

Caraway, K. & Tucker, C. M. (2003). Self-efficacy, goal orientation, and fear of failure as predictors of school engagement in high school students. *Psychology in the Schools*, *40*, 4. Retrieved February 12, 2008, from http://onlinelibrary.wiley.com/doi/10.1002/pits.10092/pdf

Cassady, J. C. (2004). The impact of cognitive test anxiety on text comprehension and recall in the absence of salient evaluative pressure. *Applied Cognitive Psychology*, *18*, 3, 311–325.

Cassady, J. C. & Gridley, B. E. (2005). The effects of online formative and summative assessment on test anxiety and performance. *Journal of Technology, Learning and Assessment*, *4*, 1, 1–31.

Clariana, R. & Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with test mode effect. *British Journal of Educational Technology*, *33*, 5, 593–602.

Croft, A. C., Danson, M., Dawson, B. R. & Ward, J. P. (2001). Experience of using computer assisted assessment in engineering mathematics. *Computers and Education*, *37*, 1, 53–66.

Economides, A. A. (2005). Personalized feedback in CAT (Computer Adaptive Testing). *WSEAS Transactions on Advances in Engineering Education*, *2*, 3, 174–181.

Economides, A. A. (2006). Emotional feedback in CAT (Computer Adaptive Testing). *International Journal of Instructional Technology & Distance Learning*, *3*, 2. Retrieved February 8, 2008, from http://itdl.org/Journal/Feb_06/article02.htm

Economides, A. A. (2009). Conative feedback in computer-based assessment. *Computers in the Schools*, *26*, 3, 207–223.

Efklides, A. & Volet, S. (2005). Feelings and emotions in the learning process. *Learning and Instruction*, *15*, 5, 1–10.

Eysenck, M. W. & van Berkurn, J. (1992). Trait anxiety, defensiveness, and the structure of worry. *Personality and Individual Differences*, *13*, 1285–1290.

Fountoulakis, K., Papadopoulou, M., Kleanthous, S., Papadopoulou, A., Bizeli, V., Nimatoudis, I. *et al.* (2006). Reliability and psychometric properties of the Greek translation of the State-Trait Anxiety

Inventory form Y: preliminary data. *Annals of General Psychiatry*, *5*, 2. Retrieved May 16, 2009, from http://www.annals-general-psychiatry.com/content/5/1/2
Furnham, A. & Rawles, R. (1995). Sex differences in the estimation of intelligence. *Journal of Social Behaviour and Personality*, *10*, 741–745.
Gneezy, U., Niederle, M. & Rustichini, A. (2003). Performance in competitive environments: gender differences. *Quarterly Journal of Economics*, *118*, 1049–1073.
Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.
Gorriz, C. & Medina, C. (2000). Engaging females with computers through software games. *Communications of the ACM*, *43*, 1, 42–49.
Gronlund, N. E. (1988). *How to construct achievement tests*. Englewood Cliffs, NJ: Prentice-Hall.
Hartmann, T. & Klimmt, C. (2006). Gender and computer games: exploring females' dislikes. *Journal of Computer-Mediated Communication*, *11*, 4, 910–931.
Heylen, D., Vissers, M., Akker, R. & Nijholt, A. (2004). Affective feedback in a tutoring system for procedural tasks. In E. André, L. Dybkjaer, W. Minker & P. Heisterkamp (Eds), *Affective dialogue systems* (pp. 244–253). Lecture Notes in Computer Science. Heidelberg: Springer-Verlag.
Hoeft, F., Watson, C. L., Kesler, S. R., Bettinger, K. E. & Reiss, A. L. (2008). Gender differences in the mesocorticolimbic system during computer game-play. *Journal of Psychiatric Research*, *42*, 4, 253–258.
Hone, K. (2005). Empathic agents to reduce user frustration: the effects of varying agent characteristics. *Interacting with Computers*, *18*, 2, 227–245.
Igbaria, M. & Chakrabarti, A. (1990). Computer anxiety and attitudes towards microcomputer use. *Behaviour and Information Technology*, *9*, 229–241.
Kang, S., Gratch, J., Wang, N. & Watt, J. H. (2008). Does the contingency of agent's nonverbal feedback affect users' social anxiety? Proceedings of 7th International Conference on Autonomous Agents and Multiagent Systems, 120–127.
Klein, J., Moon, Y. & Picard, R. W. (2002). This computer responds to user frustration: theory, design and results. *Interacting with Computers*, *14*, 2, 119–140.
Lankes, M., Riegler, S., Weiss, A., Mirlacher, T., Pirker, M. & Tscheligi, M. (2008). Facial expressions as game input with different emotional feedback conditions. *Proceedings of International Conference on Advances in Computer Entertainment Technology, ACM International Conference Proceeding Series*, *352*, 253–256.
Lazarus, R. S. (1991). *Emotion and adaptation*. London: Oxford University Press.
Lester, J. C., Towns, S. G. & FitzGerald, P. J. (1999). Achieving affective impact: visual emotive communication in lifelike pedagogical agents. *The International Journal of Artificial Intelligence in Education*, *10*, 3–4, 278–291.
Li, C. R., Huang, C.-Y., Lin, W. & Sun, C. W. V. (2007). Gender differences in punishment and reward sensitivity in a sample of Taiwanese college students. *Personality and Individual Differences*, *43*, 457–483.
Lucas, K. & Sherry, J. L. (2004). Sex differences in video game play: a communication-based explanation. *Communication Research*, *31*, 5, 499–523.
Lupyan, G. & Rifkin, I. (2003). Dynamics of applause: modeling group phenomena through agent interaction. Poster presented at the 25th Annual Conference of the Cognitive Science Society.
Maltz, D. N. & Borker, R. A. (1983). *A cultural approach to male–female miscommunication*. New York: Cambridge University Press.
McCarthy, J. M. & Goffin, R. D. (2005). Selection test anxiety: exploring tension and fear of failure across the sexes in simulated selection scenarios. *International Journal of Selection and Assessment*, *13*, 4, 282–295.
McLean, C. P. & Anderson, E. R. (2009). Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review*, *29*, 6, 496–505.
McQuiggan, S., Mott, B. W. & Lester, J. (2008). Modeling self-efficacy in intelligent tutoring systems: an inductive approach. *User Modeling and User-Adapted Interaction*, *18*, 81–123.
Moridis, C. N. & Economides, A. A. (2008). Towards computer-aided affective learning systems: a literature review. *Journal of Educational Computing Research*, *39*, 4, 313–337.
Moridis, C. N. & Economides, A. A. (2009a). Mood recognition during online self-assessment tests. *IEEE Transactions on Learning Technologies*, *2*, 1, 50–61.
Moridis, C. N. & Economides, A. A. (2009b). Prediction of student's mood during an online test using formula-based and neural network-based methods. *Computers & Education*, *53*, 3, 644–652.
Neda, Z., Ravasz, E., Brechet, Y., Vicsek, T. & Barabasi, A.-L. (2000). Self-organizing processes: the sound of many hands clapping. *Nature*, *403*, 849–850.
Nicol, D. J. (2007). E-assessment by design: using multiple-choice tests to good effect. *Journal of Further and Higher Education*, *31*, 1, 53–64.
Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*, 2, 199–218.

Ormrod, J. (1995). *Cognitive factors in motivation. Human learning* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.
Pekrun, R., Goetz, T., Titz, W. & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educational Psychologist*, *37*, 2, 91–105.
Picard, R. (1997). *Affective computing*. Cambridge, MA: MIT Press.
Picard, R., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D. *et al*. (2004). Affective learning—a manifesto. *BT Technology Journal*, *22*, 4, 253–269.
Ramesh, S., Sidhu, S. M. & Watugala, G. K. (2005). Exploring the potential of multiple choice questions in computer-based assessment of student learning. *Malaysian Online Journal of Instructional Technology*, *2*, 1, 1–15.
Rebolledo-Mendez, G., du Boulay, B. & Luckin, R. (2006). Motivating the learner: an empirical evaluation. In M. Ikeda, K. D. Ashley & T.-W. Chan (Eds), *Its 2006*. LNCS Vol. 4053 (pp. 545–554). Heidelberg: Springer.
Ricketts, C. & Wilks, S. (2001). Is computer based assessment good for students? Proceedings of the 5th International Computer Assisted Assessment Conference, 2001, Loughborough.
Ricketts, C. & Wilks, S. J. (2002a). What factors affect student opinions of computer-assisted assessment? Proceedings of the 6th International Computer Assisted Assessment Conference, Loughborough.
Ricketts, C. & Wilks, S. J. (2002b). Improving student performance through computer-based assessment: insights from recent research. *Assessment & Evaluation in Higher Education*, *27*, 5, 475–479.
Rosen, L. D., Sears, D. C. & Weil, M. M. (1993). Treating technophobia: a longitudinal evaluation of the computerphobia reduction program. *Computers in Human Behavior*, *9*, 27–50.
Sambell, K., Sambell, A. & Sexton, G. (1999). Students' perception of the learning benefits of computer-assisted assessment: a case study in electronic engineering. In S. Brown, J. Bull & P. Race (Eds), *Computer-assisted assessment in higher education* (pp. 179–191). Birmingham: SEDA.
Schunk, D. H. (1984). Enhancing self-efficacy and achievement through rewards and goals: motivational and informational effects. *Journal of Educational Research*, *78*, 29–34.
Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review*, *1*, 173–208.
Seipp, B. (1991). Anxiety and academic performance: a meta-analysis of findings. *Anxiety Research*, *4*, 27–41.
Sim, G. & Holifield, P. (2004). Computer assisted assessment: all those in favour tick here. Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications (AACE), Lugano.
Snooks, M. K. (2004). Using practice tests on a regular basis to improve student learning. *New Directions for Teaching and Learning*, *100*, 109–113.
Spielberger, C. D. (1972). *Anxiety: current trends in theory and research*. New York, NY: Academic Press.
Spielberger, C. D. (1983). *Manual for the state-trait anxiety inventory (STAI)*. Palo Alto, CA: Consulting Psychologists Press.
Spielberger, C. D. (2005). *State-trait anxiety inventory for adults*. Redwood City, CA: Mind Garden.
Spielberger, C. D., & Vagg, P. (1995). Test anxiety: A transactional Process Model. In C. D. Spielberger & P. Vagg (Eds.), *Test anxiety: Theory, assessment, and treatment* (pp. 3–14). Washington, DC: Taylor & Francis.
Taras, M. (2003). To feedback or not to feedback in student self-assessment. *Assessment and Evaluation in Higher Education*, *28*, 5, 549–565.
Taylor, S. & Brown, J. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, *103*, 2, 193–210.
Weedon, E. (2000). Do you read this the way I read this? *British Journal of Educational Technology*, *31*, 185–197.
Welsh, A. L. (1978). Multiple choice objective tests. In P. Saunders, A. L. Welsh & W. L. Hansen (Eds), *Resource manual for teacher training programs in economics* (pp. 191–228). New York: Joint Council on Economic Education.
Westerback, M. E. & Long, M. J. (1990). Science knowledge and the reduction of anxiety about teaching earth science in exemplary teachers as measured by the science teaching state-trait anxiety inventory. *Social Science and Mathematics*, *90*, 361–374.
Whang, M. C., Hyun, H. J., Lim, J. S., Park, K. R., Cho, Y. J. & Park, J. S. (2007). Effect of different modal feedback on attention recovery. Second International Conference on Usability and Internationalization Held as Part of HCI International 2007, Proceedings LNCS, 4559/2007, 631–636.
Zoller, U. & Ben-Chaim, D. (1989). Interaction between examination type, anxiety state, and academic achievement in college science: an action-oriented research. *Journal of Research in Science Teaching*, *26*, 1, 65–77.