

Objective score versus subjective satisfaction: Impact on emotions following immediate score reporting

Lia M. Daniels¹

¹University of Alberta

Abstract:

There are positive feedback loops between students' grades and emotions. However, subjective appraisals, not grades, are theorized to trigger emotions. We extended previous research by comparing the effects of objective score and subjective appraisals of the score (i.e., satisfaction) on emotions. We used an ecologically-valid quasi-experimental design and found differences in how objective score compared to satisfaction impacted emotions. Main effects for score showed positive associations with hope, pride, relief, and negative associations with anxiety, anger, and shame. An interaction for satisfaction occurred such that students who were satisfied with their score had the same effect as objective score, but students who were unsatisfied with their score felt less hope, pride, relief, and more anger and shame. Implications for the control-value theory of emotions as well as for instructors are discussed.

Keywords: achievement emotions, computer-based testing, immediate score reports, satisfaction, cognitive appraisals

Note: This is a pre-copyedited, author-produced PDF of an article accepted for publication in the Journal of Experimental Education.

The official citation for this manuscript is: **Daniels, L. M. (2020). Objective score versus subjective satisfaction: Impact on emotions following immediate score reporting. *The Journal of Experimental Education*, 88(4), 578-594.** This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available, upon publication, via its DOI.

Because test scores deliver a message of success or failure to students, the scores can be expected to “trigger intense emotions” (Pekrun, Goetz, Perry, Kramer, Hochstadt, & Molfenter, 2004, p. 288). However, because emotions are theorized to arise from subjective cognitive appraisals of the learning environment (Pekrun, 2006), it seems that the “message of success or failure” rather than the test score itself may be a more salient predictor of emotions. Take for example two students who both receive 75% on an exam. Although objectively they have the same score, one student may subjectively consider 75% to be a satisfactory score while the other student considers it unacceptable. More as a result of the appraisals than the score itself, these two students likely experience different emotions: The satisfied student may experience heightened pride and reduced anxiety while the unsatisfied student may experience shame and anger.

The distinction between objective score and subjective score appraisal may be particularly important in computer-based testing (CBT) environments where students receive their scores immediately— even prior to exiting the testing room. Getting students their test scores as quickly as possible is generally considered a good assessment practice (Brookhart, 2008). But when “as quickly as possible” can be immediately, as is the case in CBT environments, students must make sense of (i.e., appraise) their score just as quickly. Some research suggests that immediate score reporting can be beneficial for students’ emotions (e.g., Daniels & Gierl, 2017); however, these researchers focused exclusively on objective test scores ignoring the role of students’ subjective appraisals of their scores. We sought to advance researchers’ and instructors’ understanding of objective-versus-subjective influences on students’ emotions in a CBT environment that uses immediate score reporting. Specifically, we operationalized objective influences as students’ percentage correct on an examination and subjective appraisals as student’s satisfaction with that score.

Control-value theory of emotions

The American Psychological Association defines emotion as “a complex reaction pattern, involving experiential, behavioral, and physiological elements by which an individual attempts to deal with a personally significant matter or event” (APA, online dictionary). In achievement contexts, the control-value theory of emotions specifies achievement emotions as emotions experienced in relation to achievement activities or achievement outcome (Pekrun, 2006). Students experience the full gamut of achievement emotions in achievement contexts broadly (Pekrun, Goetz, Titz, & Perry, 2002) and test-taking situations specifically (Pekrun et al., 2004). Within the control-value theory of emotions, emotions are treated as involving not only affective components but also cognitive, motivational, expressive, and physiological processes. In line with this multicomponent approach, Pekrun (2006) offers a classification system for emotions and a process model to understand the generation and effects of emotions.

For classification, Pekrun (2006) distinguishes emotions in terms of having pleasant or unpleasant valence, being activating or deactivating, and being outcome or activity focused. Outcome-focused emotions can be further divided into prospective emotions linked to an expectation of success or failure and retrospective emotions linked to a known outcome that has been appraised as success or failure. Although there are others, six outcome emotions that students may commonly experience at the conclusion of an exam are hope, relief, pride, anxiety, anger, and shame. Hope and anxiety are oppositely-valenced, prospective outcome emotions that students may experience as they conclude an examination and wait to receive their scores. Hope is linked to an expectation of success; whereas, anxiety is linked to an expectation of failure. In contrast, pride, relief, anger, and shame are retrospective outcome emotions that students may

feel after they know and appraise their score. As would be assumed intuitively, pride extends from success and anger and shame tend to be associated with failure. According to Pekrun (2006) relief is experienced when an expected failure is avoided.

For process, Pekrun (2006) proposes that students cognitively appraise components of their learning environment and then the appraisals give rise to their emotions, which in turn impact achievement outcomes. Two cognitive appraisals are central to the control-value theory: Subjective control refers to the expectation that a behavior can be enacted to lead to success, and subjective value refers to the perceived importance of the outcome. Researchers have indeed shown that subjective appraisals are proximal predictors of emotions (e.g., Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010; Tze, Klassen, & Daniels, 2014). Pekrun and colleagues (2011) showed that hope, pride, and relief were positively associated with perceptions of academic control and task value; whereas, anger, anxiety, and shame had negative associations. Although named directly in the control-value theory, the two cognitive appraisals are considered subjective to the student, to the achievement situation, and to each student's personal expectancies for success (Pekrun, 2006). Thus, there may be many possible ways to capture students' cognitive appraisals of their score, including simple measures such as score satisfaction.

Concluding this linear process, pleasant emotions including hope, pride, and relief have been regularly shown to have positive associations with grades ranging from $r = .14$ to $.34$ and unpleasant emotions including anxiety, anger, and shame tend to have negative associations ranging from $r = .14$ to $.37$ (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011).

In addition to this linear process, Pekrun (2006) describes feedback loops between antecedents, emotions, and effects. In particular, researchers have revealed positive feedback loops among test scores and emotions (e.g., Daniels & Gierl, 2017; Pekrun, Hall, Goetz, & Perry, 2014; Pekrun, Lichtenfeld, Marsh, Murayama, & Goetz, 2017; Putwain, Becker, Symes, & Pekrun, 2018). For example, using data from the Project for the Analysis of Learning and Achievement in Mathematics (PALMA), involving more than 3,000 students, Pekrun and colleagues (2017) showed that pride and enjoyment positively predicted achievement, which predicted future reports of pride and enjoyment—a pattern that continued across Grades 5 to 9. The authors found the opposite pattern for anger, anxiety, shame, boredom, and hopelessness. These feedback loops are not only present over many years but can occur within minutes. This was the case for Daniels and Gierl (2017), who found positive feedback loops between test scores and emotions in a CBT environment following receipt of a test score almost immediately following the completion of an exam.

Whether examined over many years or just a few minutes, these models require grades to be conceptualized not just as an achievement outcome positioned at the end of the control-value theory of emotions (Pekrun, 2006) but also as part of the learning environment positioned at the beginning of the model. When grades (or any feedback) are conceptualized as part of the learning environment then it is important to recognize that they are subject to cognitive appraisal just as is the case with any other component of the environment. To date, however, none of the studies on emotions and feedback have considered students' subjective appraisal of their test scores. Thus, we argue that students' subjective appraisals of test scores are at least as important an antecedent as the scores themselves (Pekrun, 2006). The current research advances the field by testing both the role of objective test scores (i.e., percentage) and subjective appraisals of test scores (i.e., satisfaction) on students' emotions.

Computer-based testing taking as part of the achievement environment

Typically, an undergraduate bachelor's degree requires about 40 completed university courses. If each course requires students to write one midterm and one final examination, then students write a minimum of 80 examinations during their undergraduate program. This also means that students have at least 80 instances during which they experience emotions as they prepare for an examination, sit and actually write the exam, reflect on the exam after it is completed, receive a score indicating their performance, and come to terms with that performance (Schutz & Davis, 2000). The emotions related to sitting the exam and coming to terms with performance may be dramatically "sped up" through the use of computer-based testing, which is touted as the hallmark of contemporary 21st-century assessment (Sireci & Zenisky, 2016). The popularity of CBT is understandable because it offers instructors many advantages not available with traditional paper-based testing, including immediate score reporting.

Grades: objective and subjective indicators of success. Immediate score reporting in CBT environments is most common for multiple-choice exams, which remain commonplace in postsecondary settings (e.g., DiBattista & Kurzawa, 2011; Mavis, Cole, & Hoppe, 2001). A raw score indicating how many questions the student answered correctly or a percentage correct typically denotes performance on the exam. In both pencil-and-paper and CBT environments, there is theoretical reasoning (Pekrun, 2006) and empirical evidence documenting positive feedback loops between discrete emotions and test scores (Daniels & Gierl, 2017; Pekrun, Hall, Goetz, & Perry, 2014; Pekrun et al., 2017). However, from an emotion-elicitation perspective, subjective cognitive appraisals are arguably more relevant than objective score and thus these reciprocal models may be considered somewhat misspecified. In other words, an important shortcoming of the research is that it does not consider students' subjective appraisal of the examination score on emotions.

What subjective appraisals might students make in response to receiving an immediate test score? Control and value appraisals as typically described by Pekrun (2006) may neither be appropriate nor practical to measure in a CBT situation with immediate score reporting. In the moment, rather than focusing on the extent to which they exerted effort or their value of the test score precisely, heuristics such as score satisfaction may be a more ecologically valid indicator of students' subjective appraisal of their score and by extension an appropriate antecedent to their emotions. Indeed, on the occasions that perceived success has been considered with emotions, it tends to correlate more strongly with emotions than students' objective grades (e.g., Daniels, Haynes, Stupnisky, Perry, Newall, & Pekrun; Daniels & Gierl, 2017), a finding that aligns with the proximity of cognitive appraisals to emotions in the control-value theory (Pekrun, 2006). Therefore, in order to understand the impact of immediate score reporting on students' emotions, researchers and instructors must take into account not only the objective score but also students' subjective appraisal of the score, which we have chosen to operationalize as students' satisfaction with their score. In other words, although many stakeholders care greatly about whether students receive an A or a B+, students' satisfaction with their score may be more relevant for their own emotions.

Objectives of the current study: replication and extension

The current study is a conceptual replication (Schmidt, 2009) of Daniels and Gierl (2017). In their Study 2, Daniels and Gierl sought to answer the following research question: Is the effect of immediate score reporting on students' relief, hope, pride, anxiety, anger, and shame dependent on the objective examination score? They chose these six emotions to capture both positive and negative valence emotions that students may experience in response to positive or

negative performance feedback. In a computer-based testing environment, students completed a real examination for a required undergraduate course. At the completion of the examination students could choose to complete items measuring their emotions, then see their score, and then complete the emotion items again. In this way, Daniels and Gierl treated objective examination score as a continuous naturally manipulated independent variable theorized to bring about rapid shifts in emotions. They found that scores presented immediately had (a) a positive influence on relief, hope, and pride, (b) a negative influence on anxiety and shame, and (c) no influence on anger even after accounting for the effect of emotions prior to seeing the score. The authors concluded that “the emotions that students experienced after seeing the score presented immediately were directly impacted by the valence of the score itself, with higher scores as ubiquitously better for emotions than lower” (p. 32).

Although Daniels and Gierl (2017) made an important contribution by looking at feedback loops between scores and emotions over one of the briefest timeframes recorded, there were two main shortcomings of their research that gave rise to the objectives of the current research. First, Daniels and Gierl (2017) suffered from major attrition (62%) from Time 1 to Time 2, resulting in an overrepresentation of high achievers in their final sample. Consequently, their results may not accurately reflect the impact of the full range of objective test score on emotions. To improve on this limitation, we used an incentive strategy to retain a more complete sample of student performance and asked the following replication research question: Do original main effects of immediate score reporting on students’ relief, hope, pride, anxiety, anger, and shame documented by Daniels and Gierl replicate when a wider, more complete range of objective scores is tested? We hypothesized that the results would replicate with direct positive effects between objective score and pleasant emotions and direct negative effects between objective score and unpleasant emotions. Second, the Daniels and Gierl model may have been misspecified by not including some type of subjective appraisal of the exam score. Addressing this, our extension research question was, Is the effect of immediate score reporting on students’ relief, hope, pride, anxiety, anger, and shame moderated by score satisfaction? We offer the following three hypotheses (H):

H1. When students cognitively appraise their objective score as satisfactory, they will experience an increase in hope, pride, and relief and a decrease in anxiety, anger, and shame.

H2. When students report low satisfaction with their objective score, they will experience a decrease in hope, pride, and relief and an increase in anxiety, anger, and shame.

H3. When students report average satisfaction with their objective score, their emotions will not be impacted beyond the feelings experienced by simply having completed the exam.

Method

This study is a conceptual replication (Schmidt, 2009) of Daniels and Gierl (2017). As such, we utilized the same naturally occurring quasi-experimental design within a CBT environment and assessed the same emotions as Daniels and Gierl.

Procedure

Participants were enrolled in two sections of a required undergraduate course, Adolescent Development and Learning, at a Canadian research-intensive university. The course was highly coordinated such that students in the two sections had the same textbook, course requirements, lecture notes, and examinations even though they had different instructors. As approved by the institution’s ethical review board, the information letter was posted on the classes’ online

management system and explained that at the conclusion of their first computer-based examination students could voluntarily participate in a very short study on emotions and immediate score reporting. As a small incentive, the researchers offered each student a \$5 coffee card if they were able to complete both questionnaires at the end of the exam. Informed consent was implied when a student chose to complete questionnaires.

It was possible to provide students with immediate exam scoring because they wrote their exams in a computer laboratory using a computer-based testing system created at the university. Students wrote their exams during a flexible window of time when they felt ready and this was possible because the testing environment was securely monitored and offered flexible proctoring and the instructional team had created multiple equivalent forms of the exams (Gierl, Daniels, & Zhang, 2017). Students were given 90 minutes to write the exam, which contained 48 multiple-choice questions based on the first three chapters of the required textbook and the content presented during lectures. The exam contributed 15% toward the student's final grade.

After students submitted their exam responses, the computer automatically redirected to a new screen that explained the student was finished with the examination and free to leave unless the student was interested in participating in the research study. First, students who chose to participate answered 18 items measuring six emotions (Time 1: relief, hope, pride, anxiety, anger, and shame). Second, students selected between two options: to see their percentage score immediately or wait to access their score later through the online management system. Third, students who chose to see their score were then asked how satisfied they were with their score and then all students were asked to complete the 18 emotion items again (Time 2).

Participants

A total of 148 students wrote the exam in the CBT environment and 100% completed the first questionnaire directly following their exam and 100% chose to view their score immediately. Then 138 students completed the second questionnaire. Thus, with a small incentive we were able to retain 93% of participants, which is a marked improvement over Daniels and Gierl (2017), who only retained 34% of their original sample using the same methodology without the incentive. The ecological nature of this study meant that it was not possible to collect demographic information from participants. However, based on enrollment information there are typically more women (approximately 70% to 75%) in the class than men and students tended to be on average 23 to 25 years old. No ethnicity data were collected by the researchers nor tracked by the institution.

Measures

Independent variables: grades and satisfaction. We used two indicators of success as naturally occurring independent variables. Students' actual exam scores were considered an objective indicator of success (i.e., a component of the learning environment) and ranged from 52% to 93%. The class average was 75.78% (SD 14.93). For the subjective cognitive appraisal, students indicated how satisfied they were with their score on a 5-point scale where 1 = not at all satisfied, 3 = neutral, and 5 = very satisfied. Because there was only one indicator for this item, we chose to collapse it into a three-level categorical variable and thus grouped participants based on their response as low satisfaction, $n = 45$ (indicated 1 or 2); average satisfaction, $n = 42$ (indicated 3); and high satisfaction, $n = 45$ (indicated 4 or 5).

Dependent variables: achievement emotions. Participants reported their emotions twice—once prior to viewing their score (Time 1) and once after viewing their score and indicating their satisfaction (Time 2). We used a reduced number of items from the Achievement Emotions Questionnaire (AEQ; Pekrun et al., 2011) to assess hope, relief, pride, anxiety, anger,

and shame (three items each). All items were answered on a 11/4strongly disagree to 51/4strongly agree rating scale. The descriptive statistics for the scales are presented in Table 1 and demonstrated adequate reliability.

Plan for analysis

As preliminary analyses we examined all descriptive information for the variables and zero-order correlations. We conducted the same analyses as Daniels and Gierl (2017) to test the effect of objective test score on emotions. Specifically, we used six separate hierarchical regression analyses with forced entry of blocks of variables (Field, 2009) to test relationships between immediate exam scores and students' emotions. The Time 1 emotion matching the dependent variable and objective exam score were entered in Step 1. To test for interactions, in Step 2 we entered a multiplicative interaction term consisting of centered Time 1 Emotion X centered exam score. All average VIFs were close to 1.0 (range 1/4 1.01–1.51), all tolerance statistics greater than .20, and no diagnostic condition indices were less than 30, suggesting that multicollinearity was not a problem (Kleinbaum, Kupper, Nizam & Muller, 2008).

To answer the new research question regarding subjective score satisfaction as a cognitive appraisal that predicts emotions, we began by examining the equivalence of the satisfaction groups on Time 1 emotions and objective exam scores at a Bonferonni-corrected significance level of $p \leq .05/6 \leq .008$. Then, because satisfaction was a categorical variable, we used repeated measures analysis of covariance (RM-ANCOVA) to test for changes in emotions moderated by score satisfaction. We entered Time as the two-level within-subject factor (T1 1/4 emotions prior to viewing score; T21/4emotions after viewing score) and score satisfaction as the three-level between-subject factor. We included objective exam score as a covariate and it met the assumption of homogeneity of regression for all dependent variables except anger, as determined by testing the interaction between score and group (Field, 2009). Because of this we conducted the main analysis for anger without the covariate (Tabachnick & Fidell, 2001). In terms of other assumptions, for the within-subject factor sphericity was assumed because there were only two levels. For the between-subject factor Levene's test of equality of error variances was nonsignificant for variables at each time point except Time 2 anxiety, shame, and hope and Box's test of equality of covariance matrices was nonsignificant for all variables except shame (Field, 2009). The data did not show any major signs of nonnormality and we expected some extreme scores at Time 2 given that these scores represent emotional responses to real exam performance. There were no outliers on any Time 1 assessment of emotions and only 10 scores were considered extreme across all Time 2 emotions. We probed all significant interactions with paired samples t tests conducted within each satisfaction group separately with a stringent significance level $p \leq .002$ to account for multiple tests.

Results

Preliminary analyses

Descriptive statistics including coefficient alpha (range a 1/4 .69 to .94) for all emotions at Time 1 and Time 2 are presented in Table 1. Directly following their exam and prior to seeing their score, students endorsed hope, relief, pride, anxiety, anger, and then shame in descending intensity. At Time 1 all pleasant emotions were endorsed above the midpoint of the scale and all unpleasant emotions were endorsed below the midpoint, suggesting that overall students felt more pleasant than unpleasant emotions. At Time 2, after seeing their score, students still reported above-midpoint levels of hope, relief, and pride, but anger was also endorsed above the midpoint, suggesting that anger may increase in response to seeing an exam score in a CBT environment.

The patterns of correlations (Table 2) provide strong evidence of validity for the emotion scales and the single-item satisfaction score. The Time 1 and Time 2 assessments of each emotion were moderately positively correlated. Moreover, the matched correlation (e.g., Time 1 hope and Time 2 hope) was stronger than any nonmatched correlation (e.g., Time 1 hope and Time 2 pride). Within a measurement period, similarly valenced emotions were positively correlated (e.g., Time 1 hope and Time 1 pride) and negatively correlated with oppositely valenced emotions (e.g., Time 2 relief and Time 2 anxiety). Finally, the correlations among variables at Time 2 were noticeably stronger than at Time 1 (M difference in correlations $1/4 .246$, range of differences $.09$ to $.47$), suggesting that perhaps after seeing their exam score students experienced similarly valenced emotions as more similar and experienced oppositely valenced emotions as more separate than they did prior to seeing their score.

In terms of indicators of success, exam score and score satisfaction were moderately positively correlated ($r 1/4 .47$, $p < .001$), suggesting that although high scores were generally viewed as more satisfactory the subjective appraisal is by no means a direct translation of exam score. Moreover, both indicators of performance were positively correlated with pleasant emotions and negatively correlated with negative emotions at Time 2.

Objective test score

The results of the six separate regression analyses are presented in Table 3. As was the case in Daniels and Gierl (2017), for the positive emotions of hope, relief, and pride, positive significant direct relationships were found between scores presented immediately following the exam and the positive emotions (bs $1/4 .36$, $.25$, and $.35$, all $ps < .01$, respectively) even after accounting for the effect of emotions prior to seeing the score. Also replicating the original study, we found that shame was further reduced when students saw a high score immediately following the exam (bs $1/4 \text{-----}.20$, $p .01$) even after accounting for the effect of shame prior to seeing the score.

The original findings did not replicate exactly for anger and anxiety. Originally, there were no significant interaction terms for any of the emotions, suggesting that immediate test scores exerted a direct effect on subsequent emotions. In the current study, the direct negative effect of exam score on anxiety and anger was qualified by a significant interaction effect between test score and the Time 1 respective emotion (b $1/4 \text{-----}.15$, and b $1/4 .17$, $ps < .05$, respectively). We examined simple slopes at ± 1 standard deviation around the mean and found that anxiety was reduced for those who scored ± 1 SD above the mean but not for other groups and that anger increased sharply for students who score -1 SD below the mean but not others. It is possible that these significant interaction effects were uncovered because we retained a larger sample with a wider, more complete range of objective exam scores. Although the interactions were statistically significant, they yielded only a 2% increase in the explained variance in each outcome.

Subjective satisfaction as cognitive appraisal

Baseline equivalence. The three satisfaction groups differed in terms of their feelings of hope, $F(2, 135) 1/4 3.40$, $p 1/4 .004$, and pride, $F(2, 135) 1/4 6.72$, $p 1/4 .002$, at the end of the test and prior to students' seeing the examination score (Time 1). Specifically, students in the low-satisfaction group felt less hope and less pride than students in the average-satisfaction or high-satisfaction group. This may imply that the experience of taking the test, even prior to receiving their exam score, was less positive for students who ended up in the low-satisfaction group than for their peers. The groups also differed on their objective exam score. The low-satisfaction group scored 10.58% lower than the high-satisfaction group and 9.51% lower than

the average-satisfaction group. Despite this mean difference, the range of objective exam scores was quite similar for each group: the low-satisfaction group ranged in scores from 52% to 89%; the average-satisfaction group ranged in scores from 52% to 85%; and the high-satisfaction group ranged in scores from 59% to 93%. The similarity in range but difference in mean scores gives further evidence of validity to the single-item satisfaction rating and reinforces that students can cognitively appraise any objective score as more or less satisfactory.

Change in emotions. The results of the six separate RM-ANCOVA analyses are presented in Table 4. For anger, we chose to exclude score as a covariate due to violation of the assumption of homogeneity of regression. Anger was the only emotion to reveal a significant main effect for Time. None of the other analyses had significant main effects for Time or interaction with exam score as a covariate. However, there were significant Time X Score Satisfaction interactions for each emotion. Follow-up paired samples t tests within each satisfaction group separately (Table 5 and Figure 1) showed that when students cognitively appraised their exam score as low satisfaction, their anger and shame increased and their hope, pride, and relief decreased. When students reported average satisfaction their anxiety decreased. When students appraised the score as highly satisfying, their anxiety decreased and their hope, pride, and relief increased.

Discussion

We conducted a conceptual replication and extension of Daniels and Gierl (2017) to examine the extent to which students' relief, hope, pride, anxiety, anger, and shame at the end of an examination change in response to their objective examination score (i.e., percentage) as compared to their subjective score satisfy-9action as a type of cognitive appraisal. For objective exam score, the results largely replicated Daniels and Gierl (2017) such that higher examination scores were directly positively associated with hope, pride, and relief and negatively associated with shame. A similar pattern emerged when students cognitively appraised their score as being high in satisfaction: hope, pride, and relief increased and anxiety decreased. However, a much more negative emotional experience emerged when students perceived their score as low satisfaction. As a result of a low-satisfaction cognitive appraisal, students' hope, pride, and relief decreased and their anger and shame increased.

The results of this research suggest that score satisfaction as one type of cognitive appraisal is key to understanding the impact of immediate score reporting on students' emotions. We focus on three particularly important findings that contribute to the literature. First, we describe the extent to which we were able to replicate the findings by Daniels and Gierl (2017), pointing to similarities and differences in methodology and results. In this regard, we highlight methodological implications dealing with incentives. Second, we compare the emotional responses of students who felt satisfied with their immediate exam scores with those who did not and discuss implications of these results for both the control-value theory of emotions and for instructors' practice. Within this, we highlight relief as an understudied emotion requiring greater attention. Finally, we discuss limitations of the current research and make suggestions for future research.

Impact of immediate test score reporting on emotions: replication

We conducted a conceptual replication (Schmidt, 2009) of Daniels and Gierl (2017) that involved the same measurement tools and analyses with purposefully altered procedures to improve the retention of participants and answer an additional research question. Several points of convergence emerged that increase confidence in the original results as they pertain to students' experiences of emotions in a computer-based testing environment and the effects of immediate score reporting. First, the Time 1 mean emotion scores reported by Daniels and Gierl

(2017) are largely within the same confidence interval of the scores recorded in the current study. This continues to suggest that at the end of a computer-based test, students feel more pleasant emotions on the whole than unpleasant. Given the ongoing student preference for traditional over contemporary learning environments (Clayton, Blumberg, & Anthony, 2018), this is a promising position for CBT and one that administrators can highlight for students or instructors who may be opposed to modern testing environments.

Following the immediate score reporting, however, the mean Time 2 emotion scores reported in Daniels and Gierl (2017) were often outside of the confidence intervals found in the current study. Specifically, the original Time 2 pleasant emotions were higher than the current confidence intervals and the unpleasant emotions were lower than the current confidence intervals. One explanation for this difference is that in the original Daniels and Gierl study, only 34% of students who completed the Time 1 assessment of emotions went on to complete the Time 2 assessment after immediately receiving their test score. Moreover, Daniels and Gierl explained that students who completed both time points scored on average 5% higher than those who chose to leave the testing room after seeing their score but without indicating their subsequent emotions. This attrition likely made their final sample somewhat more representative of high achievers reporting more pleasant and fewer unpleasant emotions. The average Time 2 emotion scores recorded in the current study are likely a better representation of the emotional responses of a full class of students and show a wider range of emotional responses to immediate scores.

Despite the fact that their sample overrepresented high achievers, Daniels and Gierl (2017) cautiously concluded that immediate score reporting enhanced pleasant emotions and mitigated unpleasant emotions above and beyond the emotions reported just by having completed the exam. We largely replicated these results, at least for pleasant emotions, which increased based on objective score. However, the main effects for anxiety and anger were qualified by significant interactions between Time 1 emotion and objective exam score, which were not noted in the original Daniels and Gierl study. These interactions suggest that emotional responses can change depending on how students cognitively appraise the score they receive. We believe this interaction points toward the possible role of cognitive appraisals such as satisfaction that may be masked by focusing on only objective score.

Methodological implications. Ecological education studies often suffer from limitations that researchers fail to foresee or do not believe they can overcome. In the case of Daniels and Gierl (2017), the researchers could not have predicted such a large attrition over such a short time span (< 5 minutes). With one simple and relatively cost-effective modification to the original methodology, we were able to prevent almost all attrition and retain almost the full class of students as participants. We encourage researchers to actively pursue ecological research settings, but to maximize the rigor of their design within the setting. In this case, a small remuneration for every participant who completed both Time 1 and Time 2 questionnaires overcame a major limitation of previous work. Researchers should be willing to ask for random assignment, to use small individual incentives, or make anonymous charitable incentives to increase the retention of participants and thereby the rigor of their design.

Satisfaction as a cognitive appraisal of exam score: Extension

The crux of Pekrun's (2006) control-value theory of emotions is that cognitive appraisals give rise to emotions. By examining the impact of objective exam scores on students' emotions, Daniels and Gierl (2017) overlooked this central premise of the theory by assuming that the score and the appraisal function equivalently. The same has been true of many researchers

pursuing reciprocal relationships between grades and emotions (e.g., Pekrun et al., 2017; Putwain et al., 2018). We empirically tested the assumption that objective score functioned similarly to subjective appraisals operationalized as score satisfaction by examining the impact of both on emotions. The results suggest that students' score satisfaction functioned as an important moderator.

For students who appraised their objective score as highly satisfying, they experienced an increase to their hope, relief, and pride and a decrease in anxiety above and beyond their feelings from simply having completed their exam. There was no change in their anger or shame, which were already low at the completion of the exam. In other words, high score satisfaction and objectively high scores led to similar emotions: pleasant emotions increased and unpleasant emotions decreased or remained low. Although students in the high-satisfaction group did obtain exam scores higher than those in the low-satisfaction group, these students still experienced a wide range of objective scores reinforcing the notion that any objective score can be appraised as satisfactory. Thus, overall it seems that immediate score reporting benefits students' emotions when students feel satisfied. However, only a third of students perceived their scores as highly satisfying and the emotional consequences for the remaining two-thirds of students were not as positive.

For students who perceived a low level of satisfaction with their objective score, they experienced a decrease in their hope, relief, and pride and increases in their anger and shame relative to the emotions they felt after completing the exam but prior to seeing the score immediately. By looking at score satisfaction, our understanding of the impact of receiving test scores immediately on students' emotional experience changes. Instead of being generally beneficial for emotions, students who appraised their score as unsatisfactory suffered a loss of pleasant emotions and an increase in unpleasant emotions. Low satisfaction seemed to particularly sacrifice relief and exacerbate feelings of anger. Each of these emotions crossed the midpoint of the scale in response to an unsatisfactory immediate score. Thus, low-satisfaction students traded relief, a pleasant-deactivating emotion that may provide them with space to recuperate their resources prior to another achievement-striving event, for anger, a negative-activating emotion that could cause the students to ruminate on their current performance (Pekrun, 2006). Future research could examine the impact of these changes of emotions on students' future motivation and the exertion of effort on subsequent exam preparation and performance.

For students in the average-satisfaction category, there was little difference between the emotions experienced simply by completing the exam and those experienced after seeing and appraising their score. They experienced a decrease in anxiety from Time 1 to Time 2. Arguably, a decrease in anxiety is beneficial for students even if no other emotions are impacted by the appraisal of the immediate score. Therefore, for students who reported an average level of satisfaction, the experience of receiving their score immediately may have a slight emotional advantage by reducing anxiety, and this is a benefit of CBT.

Implications for theory. In terms of the control-value theory of emotions (Pekrun, 2006), our results remind researchers that subjective cognitive appraisal are important to students' experience and can function differently than the objective components of the learning environment. Research that uses objective learning environments, including test scores, as predictors of emotions without considering students' appraisal of the environment are missing an important distinction and assuming that all students appraise an environment in similar ways. Existing research continuously shows this is not the case: Students perceive learning

environments and teachers' instructional methods very differently even within the same class (e.g., Babad, 1990; Wentzel, 1997). Ames (1992) stated that "to predict and examine motivated cognitions, affect, and behavior of a student it is necessary to attend to how that student perceives and gives meaning to classroom experiences" (pp. 267–268). Our results reinforce this assertion and highlights exam scores as one category of classroom experience that can be interpreted differently by students.

Relief emerged as a particularly interesting emotion with potential theoretical implications. According to Pekrun (2006) relief is experienced when an expected failure is avoided. In our study and the original Daniels and Gierl (2017) studies, relief was the most strongly endorsed emotion at the conclusion of the exam before students even saw their immediate score. This suggests that, contrary to theory, relief may not be exclusively tied to an outcome but instead a pleasant space that exists at the completion of an achievement activity. Inasmuch as the current education system is being sped up by technological advances, it is possible that relief may become an endangered emotion that students rarely experience because the next achievement activity and potential threat of failure (or success) appears quite rapidly. An additional theoretical consideration is that relief increased for students in the high-satisfaction group and decreased for students in the low-satisfaction group, meaning its link to cognitive appraisals may be particularly important. Extant research has shown instances when relief functioned more like an unpleasant emotion than a pleasant one. Specifically, whereas enjoyment, hope, and pride were positively influenced by anticipating self-referent feedback, relief was positively influenced by anticipating normative feedback goals, as was the case for anxiety, hopelessness, and shame (Pekrun, Cusack, Murayama, Elliot, & Thomas, 2014). In contrast, Fong and colleagues found relief to load onto the same factor as other pleasant emotions that students expected experienced in response to hypothetical negative, positive, or constructive feedback (Fong, Warner, Williams, Schallert, Chen, Williamson, & Lin, 2016). Researchers may want focus more on students' experiences of relief and its role in the control-value theory of emotions.

Implications for practice. For instructors, our results suggest that although CBT environments appear safe for emotions, immediate score reporting is not without emotional risks for some students. Thus, instructors will need to weigh the pros and cons of immediate score reporting in light of the exam characteristics and their students' expectations. Although we are not able to determine why students appraised their objective scores as satisfactory or unsatisfactory, certain characteristics of the examination such as low reliability or high difficulty may come into play along with student characteristics. In this vein, instructors may want to familiarize themselves with the principles of fair assessment (e.g., Rogers, 1996) as a standard for good test making. On the student side, in addition to supporting good study strategies (Gurung, Weider, & Jeske, 2010) instructors may want to consider how to help students feel more satisfied with their score regardless of the absolute number.

Limitations and directions for future research

The results of this study need to be interpreted with the following three limitations in mind. First, although we retained almost all participants from Time 1 to Time 2, there is no way to know that participants provided genuine responses on the Time 2 questionnaire. The higher correlations between similarly valenced items at Time 2 may mean that students experienced these emotions as more similar or it may reflect a bias in their reporting. This is likely a product of the ecological setting and one that is difficult to overcome. Second, although score satisfaction represents students' subjective appraisal of their objective score, it does not specifically measure students' sense of control or value specifically as defined by the control-value theory of emotions

(Pekrun, 2006). Thus, our inferences-based satisfaction could be improved by examining how characteristics of the student and exam mix to give rise to the student's sense of control and value appraisals once presented with the objective score in a CBT environment. Also, the course is required and we have no measures of students' interest or expectations for their achievement, both of which could be influencing their satisfaction of any specific objective score. Third, our results cannot be interpreted in comparison to traditional test-taking environments. Although the comparison may be interesting, the prominence of CBT environments and their association with immediate score reporting make this an important context to explore independent of emotions in traditional testing environments.

Rather than focusing on differences between testing environments, we suggest that future research focus on components that might be similar between environments to better understand the impact of tests and scores on emotions. For example, future research may want to examine how test characteristics such as item difficulty, test length, time provision, and reliability influence students' cognitive appraisals, perhaps more directly related to control and value, and thus their emotions. Another exciting area of advancement involves examining the effects of other types of tests and other types of score reporting. For example, in multiple-choice tests it may be possible to provide students with descriptive score reporting forms that provide information alongside their objective scores. This information could be tailored to the student's performance and provide direction for future improvement. Alternatively, programmers are beginning to offer automatic scoring of essay questions (Shermis, Burstein, Higgins, & Zechner, 2010) for instructors who are willing to build the scoring guide and keys. Finally, future research will need to examine not just what happens to students' emotions when they are presented with their test scores but if and how emotions are implicated as the students engage with the feedback (Hattie & Timperley, 2007; Winstone, Nash, Parker, & Rowntree, 2017).

The opportunities within a CBT environment are almost limitless. However, that does not mean that all choices are beneficial for students, even though nearly all students appear to want their scores immediately. The results of the current study remind researchers and instructors that students are real people who can have both positive and negative emotional responses depending on how satisfied they are with their immediate test score. Therefore, in making decisions about immediate score reporting, the choice needs to focus on both the timing and mode of delivery of the score and the meaning of the score to students.

References

- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3), 261–271. doi:10.1037/0022-0663.84.3.261
- APA, online dictionary. Retrieved from <http://dictionary.apa.org/emotion>
- Babad, E. (1990). Calling on students: How a teachers' behavior can acquire disparate meanings in students' minds. *The Journal of Classroom Interaction*, 25, 1–4.
- Brookhart, S. M. (2008). Feedback that fits. *Educational Leadership*, 65, 54–59. Retrieved from: <http://www.ascd.org/publications/educational-leadership/dec07/vol65/num04/Feedback-That-Fits.aspx>
- Clayton, K. E., Blumberg, F. C., & Anthony, J. A. (2018). Linkages between course status, perceived course value, and students' preference for traditional versus non-traditional learning environments. *Computers & Education*, 125, 175–181. doi:10.1016/j.compedu.2018.06.002
- Daniels, L. M., & Gierl, M. J. (2017). The impact of immediate test score reporting on university students' achievement emotions in the context of computer-based multiple-choice exams. *Learning and Instruction*, 52, 27–35. doi:10.1016/j.learninstruc.2017.04.001
- Daniels, L. M., Stupnisky, R. H., Pekrun, R., Haynes, T. L., Perry, R. P., & Newall, N. E. (2009). A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes. *Journal of Educational Psychology*, 101(4), 948. doi:10.1037/a0016096
- DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, 2(2), 1–23. doi:10.5206/cjsotl-rcacea.2011.2.4
- Field, A. (2009). *Discovering statistics using SPSS*. Sage publications: Los Angeles.
- Fong, C. J., Warner, J. R., Williams, K. M., Schallert, D. L., Chen, L. H., Williamson, Z. H., & Lin, S. (2016). Deconstructing constructive criticism: The nature of academic emotions associated with constructive, positive, and negative feedback. *Learning and Individual Differences*, 49, 393–399. doi:10.1016/j.lindif.2016.05.019
- Gierl, M. J., Daniels, L. M., & Zhang, X. (2017). Creating parallel forms to support on-demand testing for undergraduate students in psychology. *Journal of Measurement and Evaluation in Education and Psychology*, 8, 298–303.
- Gurung, R. A. R., Weider, J., & Jeske, A. (2010). Focusing on how students study. *Journal of the Scholarship of Teaching and Learning*, 10, 28–35.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Muller, K.E. (2008). *Applied Regression Analysis and Multivariable Methods* 4th., Thomson: Belmont, CA.
- Mavis, B. E., Cole, B. L., & Hoppe, R. B. (2001). A survey of student assessment in U.S. medical schools: The balance of breadth versus fidelity. *Teaching and Learning in Medicine*, 13(2), 74–79. doi:10.1207/S15328015TLM1302_1
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341. doi:10.1007/s10648-006-9029-9
- Pekrun, R., Cusack, A., Murayama, K., Elliot, A. J., & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction*, 29, 115–124. doi: 10.1016/j.learninstruc.2013.09.002

- Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology, 102*(3), 531. doi:10.1037/a0019243
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), 36–48. doi:10.1016/j.cedpsych.2010.10.002
- Pekrun, R., Goetz, T., Perry, R. P., Kramer, K., Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: Development and validation of the test emotions questionnaire (TEQ). *Anxiety, Stress, & Coping, 17*, 287–316. doi:10.1080/10615800412331303847
- 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. doi:10.1207/S15326985EP3702_4
- Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*(3), 696–710. doi:10.1037/a0036006
- Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development, 88*(5), 1653–1670. doi:10.1111/cdev.12704
- Putwain, D. W., Becker, S., Symes, W., & Pekrun, R. (2018). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learning and Instruction, 54*, 73–81. doi:10.1016/j.learninstruc.2017.08.004
- Rogers, T. (1996). Principles for fair student assessment practices for education in Canada. *Assessment in Education: Principles, Policy & Practice, 3*, 397–400. doi:10.1080/0969594960030308
- Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology, 13*(2), 90–100. doi:10.1037/a0015108
- Schutz, P. A., & Davis, H. A. (2000). Emotions and self-regulation during test taking. *Educational Psychologist, 35*(4), 243–256. doi:10.1207/S15326985EP3504_03
- Shermis, M. D., Burstein, J., Higgins, D., & Zechner, K. (2010). Automated essay scoring: Writing assessment and instruction. In E. Baker, B. McGaw, & N.S. Petersen (Eds.), *International Encyclopedia of Education* (3rd ed., pp. 75–80). Oxford, England: Elsevier.
- Sireci, S., & Zenisky, A. (2016). Computerised innovative item formats. In S. Lane, M. Raymond, & T. Haladyna (Eds.), *Handbook of Test Development* (2nd edition, pp. 313–334). New York: Routledge.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* 4th Edn., Needham Heights, MA. Allyn and Bacon, :
- Tze, V. M., Klassen, R. M., & Daniels, L. M. (2014). Patterns of boredom and its relationship with perceived autonomy support and engagement. *Contemporary Educational Psychology, 39*(3), 175–187. doi:10.1016/j.cedpsych.2014.05.001
- Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology, 89*(3), 411–419. doi:10.1037/0022-0663.89.3.411
- Winstone, N., Nash, R., Parker, M., & Rowntree, J. (2017). Supporting learners' engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17–37. doi: 10.1080/00461520.2016.1207538

Table 1. Descriptive statistics for emotion variables by whole sample and satisfaction group.

Group		Hope		Relief		Pride		Anxiety		Anger		Shame	
		Time 1	Time 2										
Total Sample (n = 138)	Mean	3.32	3.27	3.28	3.18	2.69	2.59	2.41	2.21	2.35	2.60	1.84	2.10
	95% CI Mean	[3.18, 3.44]	[3.12, 3.41]	[3.15, 3.40]	[3.03, 3.33]	[2.58, 2.81]	[2.45, 2.74]	[2.23, 2.58]	[2.05, 2.37]	[2.20, 2.50]	[2.44, 2.75]	[1.72, 1.95]	[2.27, 2.03]
	SD	0.79	0.85	0.72	0.91	0.66	0.88	1.01	0.96	0.89	0.93	0.69	1.01
	Skewness	-0.24	-0.29	-0.31	-0.05	-0.14	0.08	0.35	0.73	0.35	0.3	0.56	0.86
	Kurtosis	-0.25	-0.29	-0.48	-0.6	0.1	-0.39	-0.67	0.04	-0.25	-0.33	-0.43	-0.02
Low Satisfaction (n = 51)	α	0.83	0.87	0.69	0.83	0.8	0.87	0.86	0.88	0.84	0.78	0.81	0.94
	Mean	3.03	2.54	3.17	2.49	2.44	1.79	2.58	2.81	2.60	3.22	2.01	2.80
Average Satisfaction (n = 42)	SD	0.81	0.71	0.76	0.68	0.65	0.54	1.10	1.05	0.98	0.80	0.74	1.13
	Mean	3.46	3.40	3.33	3.24	2.78	2.60	2.32	1.96	2.27	2.46	1.82	1.88
High Satisfaction (n = 45)	SD	0.72	0.53	0.69	0.74	0.64	0.50	1.00	0.69	0.76	0.70	0.64	0.68
	Mean	3.50	3.96	3.35	3.91	2.90	3.50	2.29	1.77	2.15	2.02	1.66	1.51
	SD	0.74	0.53	0.71	0.66	0.61	0.53	0.92	0.72	0.85	0.85	0.64	0.55

Table 2. Zero-order correlations for all variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Hope Time 1													
2. Relief Time 1	.21*												
3. Pride Time 1	.53***	.23**											
4. Anxiety Time 1	-.48***	-.16	-.24**										
5. Anger Time 1	-.35***	-.24**	-.23**	.40***									
6. Shame Time 1	-.60**	-.09	-.34***	.56***	.44***								
7. Examination Score	.42***	.06	.21*	-.09	-.17	-.31***							
8. Score Satisfaction ^a	.26**	.09	.29**	-.10	-.21*	-.21*	.47***						
9. Hope Time 2	.48***	.15	.32***	-.27**	-.24**	-.37***	.50***	.73***					
10. Relief Time 2	.10	.31***	.14	.09	-.18*	-.08	.27**	.66***	.67***				
11. Pride Time 2	.29**	.15	.45***	-.11	-.12	-.22*	.43***	.83***	.79***	.70***			
12. Anxiety Time 2	-.41***	-.08	-.22*	.63***	.35***	.55***	-.25**	-.44***	-.61***	-.37**	-.46***		
13. Anger Time 2	-.25**	-.21*	-.21*	.31***	.75***	.31***	-.27**	-.56***	-.54***	-.53***	-.49**	.52***	
14. Shame Time 2	-.41***	-.09	-.25**	.40***	.26**	.56***	-.35***	-.52***	-.69***	-.45***	-.62***	.73***	.62***

^aSpearman's rho coefficients for categorical score satisfaction variable; all other coefficients are Pearson correlation coefficients.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 3. Regression analyses with objective score only ($n = 138$).

Criterion Variable (T2)	Step 1					Step 2		
	Emotion T1		Test Score		Adjusted R^2	Emotion (T1) x Score ^a		Adjusted R^2
	<i>t</i>	95% CI	<i>t</i>	95% CI		<i>t</i>	95% CI	
1. Hope	4.32***	[.19, .52]	4.61***	[.02, .05]	.33***	-.78	[-.03, .01]	.33
2. Relief	3.69***	[.17, .56]	3.15**	[.01, .04]	.14***	-.76	[-.03, .01]	.14
3. Pride	5.23***	[.32, .70]	4.88***	[.02, .05]	.31***	-.05	[-.02, .02]	.31
4. Anxiety	9.40***	[.46, .70]	-2.93**	[-.03, -.01]	.42***	-2.33*	[-.03, -.002]	.44*
5. Anger	13.03***	[.65, .88]	-2.77**	[-.03, -.004]	.59***	3.12**	[.01, .03]	.61**
6. Shame	6.80***	[.51, .93]	-2.74**	[-.04, -.01]	.34***	-.62	[-.03, .02]	.34

Note. Boldface items represent nonreplication of Daniels & Gierl (2017).

^aAll centered scores are based on means of the total sample.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 4. Repeated measures analysis of covariance for each emotion.

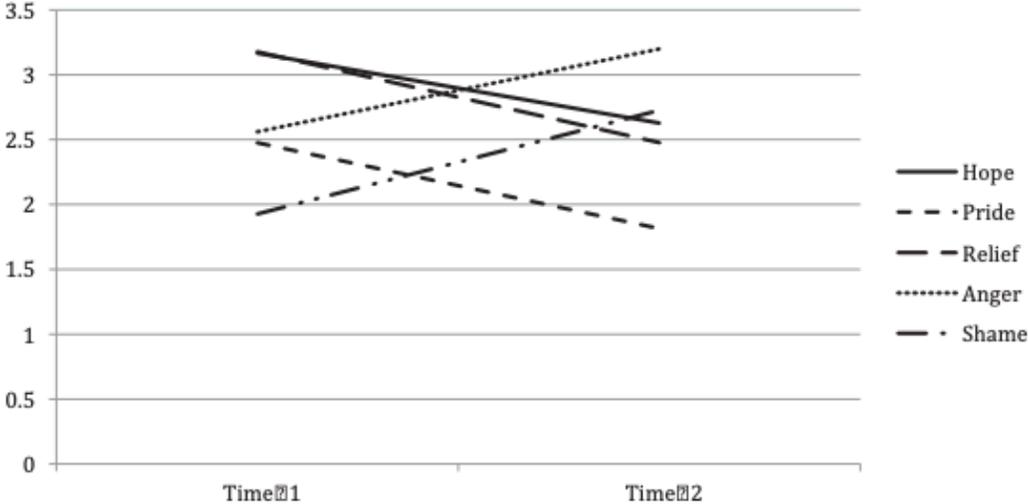
Effect	<i>MS</i>	<i>df</i>	<i>F</i>	<i>p</i>	Partial eta squared
Hope					
Time	.85	1	3.16	.08	.02
Time x Test Score	.91	1	3.38	.07	.03
Time x Score Satisfaction	5.46	2	20.27	<.001	.23
Relief					
Time	.12	1	.36	.55	.003
Time x Test Score	.17	1	.49	.49	.004
Time x Score Satisfaction	7.78	2	22.43	<.001	.25
Pride					
Time	.01	1	.02	.88	<.001
Time x Test Score	.02	1	.09	.77	.001
Time x Score Satisfaction	7.40	2	35.32	<.001	.35
Anxiety					
Time	.001	1	.001	.97	<.001
Time x Test Score	.03	1	.09	.76	.001
Time x Score Satisfaction	3.11	2	9.88	<.001	.13
Anger					
Time	3.58	1	22.63	<.001	.14
Time x Score Satisfaction	3.38	2	21.37	<.001	.24
Shame					
Time	.001	1	.005	.94	<.001
Time x Test Score	.03	1	.10	.75	.001
Time x Score Satisfaction	5.22	2	18.77	<.001	.22

Table 5. Paired samples *t* tests within each satisfaction group.

Comparison	Low Satisfaction (<i>n</i> = 51)				Average Satisfaction (<i>n</i> = 42)				High Satisfaction (<i>n</i> = 45)			
	<i>M</i> diff	<i>SE</i>	<i>t</i>	95% CI	<i>M</i> diff.	<i>SE</i>	<i>t</i>	95% CI	<i>M</i> diff.	<i>SE</i>	<i>t</i>	95% CI
Hope1-Hope2	.48	.11	4.47	[.26, .70]	.06	.11	.56	[-.16, .29]	-.45	.10	-4.32	[-.67, -.24]
Relief1-Relief2	.67	.12	5.56	[.43, .92]	.09	.13	.67	[-.17, .35]	-.55	.11	-4.80	[-.79, -.32]
Pride1-Pride2	.65	.08	7.96	[.49, .82]	.17	.10	1.66	[-.03, .39]	-.60	.10	-5.92	[-.80, -.39]
Anxiety1-Anxiety2	-.23	.12	-1.94	[-.47, .01]	.36	.11	3.17	[.13, .59]	.52	.11	4.61	[.29, .75]
Anger1-Anger2	-.62	.08	-7.49	[-.79, -.45]	-.19	.10	-1.96	[-.39, .01]	.13	.07	1.87	[-.01, .26]
Shame1-Shame2	-.79	.13	-6.25	[-1.04, -.53]	-.05	.11	-.49	[-.28, .17]	.15	.08	1.91	[-.01, .30]

Note. All *t* values $\geq \pm 4.32$, *ps* < .001; *t* = 3.17, *p* = .003.

Panel A: Low Satisfaction



Panel B: High Satisfaction

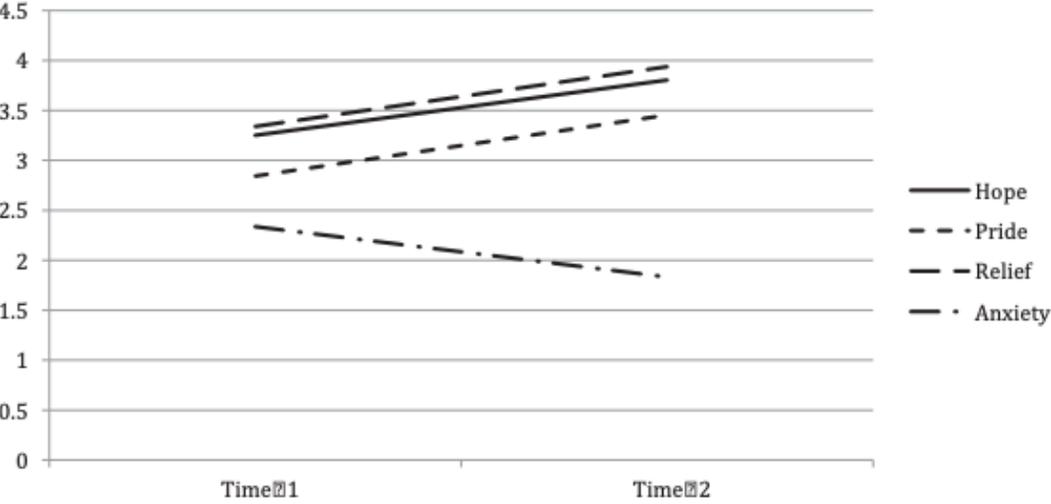


Figure 1. Within-subject change in emotions between Time 1 and Time 2 for low- and high-satisfaction groups.