Individual differences in emotion elicitation in university examinations: A quasi-experimental study

Dino Krupić a, Philip J. Corr b,*

a Josip Juraj Strossmayer University of Osijek, Croatia
b City University London, United Kingdom

ARTICLE INFO

Article history:
Received 31 May 2014
Received in revised form 29 July 2014
Accepted 1 August 2014

Keywords:
Test anxiety
Reinforcement sensitivity theory
Individual differences
Examination performance
Personality
Reward
Punishment

ABSTRACT

In a quasi-experimental study, we examined the role of individual differences in the elicitation of emotional states in university examinations. Specifically, we assessed emotional states (a) before the first examination (baseline), (b) after receiving positive or negative feedback, and (c) then, again, before a series of subsequent examinations. We also measured effort in examination preparation and interest for studying. Data were collected during a university course that consisted of seven examinations in one semester; and 94 female students completed the BIS/BAS scales and SPSRQ (to measure sensitivity to punishment, SP, and reward, SR). Results revealed that higher BAS, but not SR, individuals experienced higher positive affect (PA) following positive feedback and they also showed higher levels of interest in studying. More generally, higher BIS and SP individuals experienced higher level of negative affect (NA) and they invested more effort in examination preparation; and both higher levels of SP and SR correlated positively with NA after receiving negative feedback. In addition, following negative feedback, higher BAS individuals experienced lower levels of PA, and higher SR individuals invested less effort in examination preparation. Results are discussed in terms of the reinforcement sensitivity theory (RST) of personality and directions for future research.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Academic tasks are prone to evoke a variety of emotions in students, and these emotional experiences impact academic performance. In this context, test anxiety is one of the most common emotional experiences (Pekrun, Goetz, Titz, & Perry, 2002), and it correlates negatively with: (a) cumulative grades-point average (Diener, Schwarz, & Nickerson, 2011); (b) academic performance (Pekrun, Elliot, & Maier, 2009); and (c) students’ health (Conley & Lehman, 2012). Test anxiety and emotional reactivity to test feedback are influenced by both situational and trait factors (Putwain, Woods, & Symes, 2010). In addition to the main effect of examination-related situational anxiety, individual differences in reactions involve achievement goals (Putwain & Daniels, 2010; Putwain & Symes, 2012; Putwain et al., 2010), neuroticism (Chamorro-Premuzic, Ahmetoglu, & Furnham, 2008), perfectionism (Stoeber, Feast, & Hayward, 2009), locus of control (Davis & Davis, 1972), and even a birth order (Sarason, 1969).

Test anxiety is a multidimensional construct consisting of cognitive and emotional factors (Cassidy & Johnson, 2002), with worry being the most important cognitive feature (Øktdalen & Hagtvet, 2011). Recent studies call attention to the role played by metacognitive aspects, such as beliefs about cognitive competence, uncontrollability and danger, and cognitive self-consciousness or automatic thoughts, such as fear of failure and fear of disappointing parents (Živčić-Becirević, Juretic, & Miljević, 2009). Besides metacognitive factors in test anxiety, the most studied and important personality factors are neuroticism (Chamorro-Premuzic et al., 2008; Dobson, 2000; Halamandaris & Power, 1999) and trait anxiety (Beidel & Turner, 1988; Elliot & McGregor, 1999). Moreover, it seems that personality traits play the most important role in test anxiety. Specifically, neuroticism is a better predictor of test anxiety than core self-evaluation (Chamorro-Premuzic et al., 2008), such as self-efficiency, self-esteem and locus of control (Judge & Bono, 2002). This is not surprising given that general anxiety is one of facets of neuroticism in the five-factor model (Costa & McCrae, 1992). Thus, as test anxiety shares many conceptual characteristics with traits reflecting anxiety or negative emotionality, correlations between them may be a result of conceptual similarities rather than an effect of some explanatory mechanism.

There are a number of unresolved issues in the test anxiety literature. First, there has been little research on other emotional states experienced during situations that evoke test anxiety,
specifically broader positive and negative affectivity. Secondly, little attention has been paid to the role of individual differences that could explain variation in these states. Thirdly, we do not know how past test feedback influences emotion elicitation in subsequent examinations – this is likely to be important in terms of emotional and motivational knock-on effects. To address these issues, we explore: (a) the role of trait individual differences in emotional states prior to an examination (test anxiety); (b) after the examination feedback is provided; and (c) once again just before the student sits subsequent examinations.

One lens through which to view individual difference in such emotional states is afforded by the reinforcement sensitivity theory (RST) of personality (Corr, 2013). RST was originally based on the studies of reactions to punishment and reward in typical animal learning paradigms. In its current form (Coutinho & McNaughton, 2008, 2012) it postulates three general domain systems explaining reactions to reward, punishment, and their conflict. The Behavioral Approach System (BAS) mediates reactions to all appetitive stimuli (which include relief from nonpunishment). The Fight/Flight/Freezing System (FFFS) mediates reactions to all aversive stimuli (which include frustrative non-reward). The Behavioral Inhibition System (BIS) is hypothesized to be active in conflict situations that entail specific opposing approach (BAS) and avoidance (FFFS) goals, as well as goal-conflict more generally. Individuals higher on the BAS are proposed to be higher on extraversion and impulsivity; whereas individuals higher on the BIS are proposed to be higher on neuroticism and anxiety; and finally, the FFFS is a defensive mechanism that underlies fear and panic and, like the BIS, is related to neuroticism (Corr, DeYoung, & McNaughton, 2013). In general terms, variation in BAS reflects sensitivity to reward, while the variations in BIS and FFFS together reflect sensitivity to punishment (Corr, 2008; this provides a summary of RST). In this study, we do not differentiate the FFFS and BIS, but treat both as reflections of different aspects of punishment sensitivity.

The first aim of this study is to explore the role of sensitivity to reward and sensitivity to punishment in examination test settings in terms of the elicitation of positive and negative emotional states; and the second aim is to examine the emotions evoked by knowledge of previous examination performance (feedback). From an RST perspective, individuals higher on BAS and FFFS should be more reactive to cues of punishment and conflict. Two general hypotheses are tested. First, as the BIS (including the FFFS) mediates emotion and behavior in punishing situations, we expect a positive correlation with degree of negative affectivity generated following negative examination feedback. Secondly, sensitivity to reward (SR) or BAS should mediate reactions to both to reward and non-punishment stimuli and, thus, we expect that individuals higher on BAS and SR should feel more positive affect following favorable examination feedback.

When measuring emotional states in an academic context, there is the opportunity to conduct studies with real life observation, what Wallbott and Scherer (1989) describe as an ideal setting to study emotional experience. However, studies conducted in such settings entail methodological and ethical concerns. The strength of experimental studies is their internal validity, while for self-report studies (such as diary method sampling or correlations studies) the strength is external, or ecological, validity. In real life observation, internal validity can be diminished due to many uncontrollable factors that increase measurement error. On the other hand, mood induction in experimental studies is typically done by creating an artificial situation through presentation of stimuli, such as movie clips (Schaefer, Nils, Sanchez, & Philippot, 2010), pictures (Lang, Bradley, & Cuthbert, 2008), or music (Coutinho & Cangelosi, 2011) that provide a greater degree of experimental control, which gains much in terms of internal validity but loses in terms of external validity: this represents an important shortcoming in comparison to those studies conducted with diary experience sampling method.

The question is how findings from more artificial mood induction procedures can be generalized to test anxiety in more realistic academic settings. For this reason, we conducted a real life, quasi-experimental, study, since this methodology provides the greatest degree of ecological validity. In addition, there are ethical concerns with studies when the lecturer is performing a study on his or her students. In order to prevent this ethical concern, it is important to ensure anonymity of the participants’ data. This also has a beneficial impact on the validity of data collected because it goes a long way to avoiding socially desirable responding.

Our quasi-experimental study was designed with these theoretical and methodological considerations in mind.

2. Method

2.1. Participants

A total of 94 female university students, ranging in age from 19 to 24 years ($M = 21.86$ and $SD = 1.43$), participated in the study during an obligatory courses in their first and second academic years at the Department of psychology, University of J.J. Strossmayer, Croatia. All students participated in exchange to course credit.

2.2. Materials

Two questionnaires measuring personality traits, one assessing emotions, and one an evaluation list, designed specifically for this study, were administered. They were administered in the Croatian language.

2.2.1. BIS/BAS Scales

BIS/BAS Scales (Carver & White, 1994) consist of 13 items to assess reactivity of the BAS, which can be measured either on a unidimensional scale or divided into three subscales: BAS Drive (4 items; example item “When I want something, I usually go all-out to get it”), BAS Fun seeking (4 items; example item “I crave excitement and new sensations”) and BAS Reward Responsiveness (5 items; example item “It would excite me to win a contest!”) and 7 items to assess reactivity of the BIS (example item “I worry about making mistakes”) measured on a unidimensional scale. Items are answered on 4-point Likert scale ranging from 1 (Strongly disagree) to 4 (Strongly agree).

In this study, Cronbach alpha reliability coefficients for BAS Total, Drive, Fun-Seeking, Reward Responsiveness and BIS were .81, .80, .72, .68, and .80, respectively.

2.2.2. Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ)

SPSRQ (Torrubia, Avila, Moltó, & Caseras, 2001) consists of 48 items, 24 items measuring Sensitivity to Reward (SR; example item “Does the good prospect of obtaining money motivate you strongly to do some things?”) and 24 items measuring Sensitivity to Punishment (SP; example item “Are you often afraid of new or unexpected situations?”). All items are answered on dichotomous scale of Yes/No format. In this study, Cronbach alphas for SP and SR were .85 and .77, respectively.

Both, the BIS/BAS scales and SPSRQ are translated and validated in Croatian (Krupič, Trižanić, Ručević, Gračanin, & Corr, 2014).

2.2.3. Positive Affect and Negative Affect Schedule (PANAS)

PANAS (Watson, Clark, & Tellegen, 1988) consists of 20 items: 10 measuring Positive Affect (PA) and 10 measuring Negative Affect (NA). All items are answered on 5 point Likert scale. In this
study. Cronbach alphas for PA and NA ranged from .78 to .84, and .80 to .87, respectively. The questionnaire was translated and validated in the Croatian language (Knezović & Križanić, 2011).

### 2.2.4. Evaluation list (EL)

EL consists of 14 questions of different contexts, and served generally as a method to mask the hypothesis of the study. From the list, we used three items referring to: (a) the level of effort invested in examination preparation; (b) level of interest for the test; and (c) level of satisfaction with received grade. Since they were not related to the topic of this study, other items were not used in data analysis (the list can be found in Appendix A).

### 2.3. Design and procedure

In order to measure emotional states in a real life situation, a quasi-experimental study was conducted during an obligatory university course. The course included seven tests during one semester. No intervention or manipulations were employed. Participants completed the PANAS: just before they started their first examination (baseline); after they received their test results; and then before each of the subsequent examinations.

Conditions for positive and negative test feedback were operationalized by: (a) the relationship between received and expected grade; and (b) the level of satisfaction with the received grade obtained as assessed by the self-report evaluation list. If the received grade was lower than expected, then this condition was classified as negative test feedback. If the received grade was higher than expected, then this condition was categorized as a positive test feedback. In the case where there was no difference between received and expected grade, an additional criterion was used (e.g., if a student was expecting to receive 4 and she received a 4, then the variable of satisfaction with the received grade was used to determine measurement condition). On the 5-point scale of the variable, we interpreted answers 1 or 2 as dissatisfaction with the received grade, indicating negative test outcome, while answers 4 and 5 were interpreted as satisfaction with a grade, indicating positive test outcome (see Appendix B).

Emotional states for all conditions were calculated as arithmetic means for PA and NA. In this way, there are six dependent variables: PA and NA taken just before the first examination (baseline); PA and NA just after students received their examination feedback; and, then, PA and NA on each of the subsequent testing sessions. For subsequent examinations, mean PA and NA were subtracted from baseline PA/NA to derive a change score.

Procedures to avoid the problem of demand characteristics were employed (Klein et al., 2012). First, students participated under an anonymous password in order to ensure they more freely invested in examination preparation; (b) level of interest for the topic of the course. After negative test outcome, individuals higher on BAS Total (but especially Reward Responsiveness) experienced lower levels of NA. There were no statistically significant correlations between personality scales and changes of emotional states after positive test feedback.

Finally, on a motivational level, there were no individual differences in effort after negative test feedback, while after positive test feedback there was decreased level of effort in examination preparation for individuals higher on SR.

### 3. Results

Table 1 presents the descriptive statistics for the two emotional states before the first examination (baseline) and then after receiving feedback. These data were subjected to repeated-measures ANOVA on total of 80 participants who had complete data. First, we tested requirements for repeated ANOVA by Mauchly’s test of sphericity, which was not significant \((p < 0.05)\) for both PA and NA. Both independent variables and their interaction were statistically significant at the \(p < 0.001\) level: \(F(2,158) = 27.85\) with \(η^2 = .261\) for Condition; \(F(1,77) = 15.53\) and \(η^2 = .164\) for Affect; and finally \(F(2,158) = 206.98\) and \(η^2 = .724\) for the Condition \(\times\) Affect interaction. \(η^2\) indicates strong effects of conditions on students’ emotional states. In the positive test feedback group, PA was much higher than NA; and this pattern was reversed in the negative test feedback group; and both feedback-related emotional states differed from baseline.

#### 3.1. Emotional states

Correlations revealed that students higher on BIS, SP and SR experienced a higher level of NA. Students higher on Drive experienced a higher level of PA in positive test feedback, while in the same group students higher on Reward Responsiveness experienced a lower level of NA. Finally, students higher on SP and SR experienced higher level of NA following negative test feedback (see Table 2).

#### 3.2. Motivational variables

Students higher on BIS and SP reported higher levels of effort invested in examination preparation; and students higher on SR showed lower, while those higher on Drive, showed higher levels of interest for the topic of the course.

Correlations revealed relations between RST scales and changes in level of PA and NA following positive and negative test feedback (Table 3). After negative test outcome, individuals higher on BAS Total (but especially Reward Responsiveness) experienced lower levels of PA. There were no statistically significant correlations between personality scales and changes of emotional states after positive test feedback.

Finally, on a motivational level, there were no individual differences in effort after negative test feedback, while after positive test feedback there was decreased level of effort in examination preparation for individuals higher on SR.

### 4. Discussion

As expected, punishment sensitivity, as measured by BIS and SP, correlated positively with NA just before the first test, at baseline. Additionally, we examined the role of individual differences in motivational aspects defined as effort invested in examination preparation and level of interest to study. Individuals higher on BIS and SP invested more effort. This was expected on the basis of the RST operationalization of BIS behavioral repertoire when an individual is approaching aversive situation with caution (Corr, 2008). An alternative explanation of these results can be related to fear of failure construct. Further studies are needed to tease apart these two different possibilities.

Consistent with expectation, individuals higher on the BAS experienced higher PA after positive test feedback. The same result was expected for SR, but not observed. Furthermore, SR, but not BAS, correlated positively with NA just before the test and then again after negative test feedback. Similar findings can be found in previous studies (Dufey, Fernández, & Mourgues, 2011; Smillie et al., 2012).
Was this test hard or complicated for you? How much effort in preparation have you invested for the next examination? Are you satisfied with how your test has turned out?

Correlation matrix for personality scales and changes (compared to baseline) of PA and NA after positive and negative test feedback.

Table 2
Correlations between PA and NA in all three situations with SPSRQ and BIS/BAS scales.

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Positive test feedback</th>
<th>Negative test feedback</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PA</td>
<td>NA</td>
<td>PA</td>
<td>NA</td>
</tr>
<tr>
<td>BAS Total</td>
<td>.01</td>
<td>.06</td>
<td>.21</td>
<td>.21</td>
</tr>
<tr>
<td>BAS Drive</td>
<td>.07</td>
<td>-.07</td>
<td>.20</td>
<td>-.18</td>
</tr>
<tr>
<td>BAS Fun seeking</td>
<td>-.07</td>
<td>.11</td>
<td>.10</td>
<td>-.08</td>
</tr>
<tr>
<td>BAS Reward Responsiveness</td>
<td>.03</td>
<td>.10</td>
<td>.16</td>
<td>-.24</td>
</tr>
<tr>
<td>BIS</td>
<td>-.02</td>
<td>.23</td>
<td>.05</td>
<td>.07</td>
</tr>
<tr>
<td>SR</td>
<td>.12</td>
<td>.33</td>
<td>.18</td>
<td>.14</td>
</tr>
<tr>
<td>SP</td>
<td>-.13</td>
<td>.35</td>
<td>-.04</td>
<td>.22</td>
</tr>
</tbody>
</table>

GLE – general level of effort; GLI – general level of interest.

*p < 0.05.

**p < 0.01.

& Jackson, 2005), where BAS but not SR correlated with PA, and SR but not BAS correlated with NA.

These differences can be explained by different views of the underlying personality trait of the BAS. Torrubia et al. (2001) suggest impulsivity, while Depue and Collins (1999) propose extraversion, as underlying personality trait. Based on our results, the Carver and White (1994) BAS scale seems conceptually more similar to extraversion, since extraversion correlates positively only with PA (Larsen & Ketelaar, 1991). Additional analyses related to motivational variables of effort also confirmed differences between these SR and BAS scales.

It was expected that, over the series of examinations, negative test feedback would increase the level of NA, especially for individuals higher on BIS or SP, but that was not observed. Table 3 shows only a near significant positive correlation with PA. One explanation can be that high BIS individuals after negative test feedback invest more effort in examination preparation, and in that way they feel more secure or prepared for the next examination. Additional analysis supported this interpretation: there was a significant correlation between effort invested in examination preparation after negative test feedback.

A somewhat unexpected finding was that individuals higher on Reward Responsiveness experienced lower levels of PA after negative test feedback. Finally, additional motivational variables indicated another conceptual difference between BAS and SR scales. With regard to SR, there was a negative correlation with effort change after receiving positive test feedback while BAS individuals did not show this tendency.

4.1. Limitation of the study

In real life studies, effects sizes are usually smaller than in true experiments, which is a result of uncontrolled factors influencing the dependent variables. Hence, in future studies with more controlled conditions and with more participants, even stronger effects can be expected. Furthermore, as single items can reduce variability and, therefore, correlations with external variables, in future studies it would be appropriate to use an approximate measure of effort such as time spent in examination preparation (e.g. in hours) instead of estimation of invested effort on 5-point Likert-type response scale.

We did not differentiate the FFFS and BIS as currently there are no agreed scales for these constructs within RST. An important focus of future research should be to compare the roles played by these two major defensive systems in emotion elicitation in educational settings. Appropriate scales are starting to emerge to inform the interaction of sensitivity to reward and punishment. They show significant effects of examination feedback on emotional elicitation and reveal that personality differences predict these states. On the basis of our findings, the interaction of sensitivity to reinforcement and personality deserves much more attention than it has hitherto received in the higher educational literature.

Appendix A. Evaluation list administered immediately after finishing the test

1. Was this test hard or complicated for you? How much effort in preparation have you been invested for this test? Are you satisfied with how your test has turned out?

(continued on next page)
Appendix B. Evaluation list administered immediately after receiving evaluated report

1. What mark did you receive ______
2. Are you satisfied with the received mark ______
3. Are you satisfied with the quality of the corrector’s feedback ______
4. Are you satisfied with your report in global ______
5. Do you believe that you could do better in next report ______

1 – Not at all; 2 – A little bit; 3 – Moderate; 4 – Quite yes; 5 – Definitely yes.

References
