Eye Movement and test anxiety:

Reducing the detrimental effects of test-related mental images

Antoniëtte Geerman

Begeleider Prof. dr. Iris M. Engelhard

Universiteit van Utrecht, 28 november 2011

A B S T R A C T

Eye movements (EM) during recall of emotional mental images have proven to be successful in degrading these images and reducing their emotional intensity. This holds for images related to past autobiographical events and for future-oriented images: flashforwards. Individuals with test anxiety are troubled by flashforwards, and these may, paradoxically, decrease their actual performance on a test by rendering less working memory resources available for test performance. The aim of the current experiment was to examine whether EM during recall of flashforwards related to test anxiety increases later performance on a test. A sample of 47 students was pre-selected based on high test anxiety levels, and was randomly assigned to one of three conditions: recall+EM (N= 15), recall only (N= 16), or a no intervention control condition (N= 16). Before and after the intervention, they were asked to resolve arithmetic, matrix reasoning and digit span equations. We predicted that recall+EM would improve test performance by reducing vividness and emotionality of flashforwards, compared to recall only and no intervention. Unexpectedly, there was no significant difference between recall+EM and recall only on vividness and emotionality ratings. Findings are discussed and recommendations for future research are made.

Keywords: eye movement, test anxiety, performance, worry, cognition, flashforwards

1. Introduction

The importance to perform and excel academically is felt throughout a student's life. As Seymour B. Sarason, an early pioneering researcher of test anxiety observed, "We live in a test-conscious, test-giving culture in which the lives of people are in part determined by their test performance" (Zeidner, 1998, p. 4). Consequently it is no surprise that test anxiety is a widespread phenomenon and appears to become more prevalent in modern society. What constitutes test anxiety precisely? The term test anxiety refers to "the set of phenomenological, physiological and behavioral responses that accompany concern about possible negative consequences or failure of an exam or a similar evaluative situation" (Zeidner, 1998, p. 17). Test anxiety is a type of performance anxiety, a kind of anxiety where performance is deteriorated and dominated by thoughts of failure and consequences of failing. Other domains affected by performance anxiety include math anxiety (Hopko et al., 2003), music performance anxiety (Yoshi et al., 2009), male sexual performance anxiety (McCabe, 2005) and sport performance anxiety (Salminen et al., 1995).

No large scale research has been done on the epidemiology of test anxiety. Some studies report the estimated prevalence rate of test anxiety ranging from 10 to 33% among American elementary, high school and college students (Hill & Wigfield, 1984; McDonald, 2001; Methia, 2004), about 41% in 8 to 12 year old African Americans (Turner et al., 1993), about 22% for Canadian elementary and high school students (McGuire et al., 1987) and about 42% in Turkish students preparing for high school and university exams (Gündoğdu, as cited in Boyacioglu & Kucuk, 2001, p. 447).

Among students who experience test anxiety, females seem to report higher levels than males (Hembree, 1988; Zeidner, 1998). Besides gender differences some research has reported ethnic and Socio-Economic Background (SEB) differences in test anxiety (Hembree,

1988; Zeidner & Safir, 1989) suggesting higher levels among non-white American students and students from lower socio-economic backgrounds (Turner, 1993; Putwain, 2007).

Various studies report the detrimental effect of test anxiety on academic performance. For example, high test anxiety students have significant but modest lower grade point averages (GPAs) (Culler & Holahan, 1980; McDonald, 2001; Chapell et al., 2005). There is considerable debate concerning the nature of this association: does poor performance result from test anxiety, or vice versa? Or is another third variable involved? To test the relationship between test anxiety and test performance, Lang and Lang (2010) conducted a study alternating the level of perceived competence in students by means of priming. Higher self-perceived competence, activated by priming, resulted in improved performance among test anxious students. According to these results, the skill-deficit theory (Musch & Bröder, 1999) is a highly unlikely explanation for the correlation between test anxiety and test performance. Another way to explain the relationship between test anxiety and test performance is through the work of Yerkes and Dodson (1908), which has become known as the Yerkes & Dodson Law. They describe the relationship of anxiety on performance as an inverted U-shape, indicating that optimal performance is best achievable under a modest degree of anxiety. This means that academic performance suffers when anxiety is too low or too high.

It has often been proposed that arousal and worry are the major components of test anxiety causing performance reduction. Early cognitive interference models have suggested poor performance to be a consequence of worry cognitions occupying limited processing resources (Wine, 1971). Later models, such as the processing efficiency theory (Eysenck & Calvo, 1992), suggest that worry cognitions may reduce processing efficiency. Currently dominating the research field is the working memory theory, suggesting that worry has detrimental effects on performance, because it taxes working memory, leaving less resources available for the task at hand (Richards et al., 2000).

Ergene (2003) reported a meta-analysis of 21 intervention techniques aimed at reducing test anxiety and its detrimental effects on performance. The best treatment so far seems to be a combined therapy, including skill focused approach, behavioral (e.g. desensitization) and cognitive approach. From all 21 interventions, cognitive restructuring had better effect as an intervention technique in reducing test anxiety targeting worry. In a recent experimental laboratory and real-life study, Ramirez and Beilock (2011) found that writing about worries related to test anxiety improved test performance for college students who were anxious about test taking. Research has also found reduced physiological arousal, worry, and fears of negative evaluation with a single session of Eye Movement Desensitization Reprocessing (Maxfield & Melnyk, 2000).

EMDR involves many components, but the most crucial one is EM during recall of an upsetting mental image. Clinical studies so far have used mental images of past negative events. However, laboratory studies have shown that EM during recall reduces vividness and emotionality of mental images of past events, but also of mental images of feared future events: flashforwards (e.g., Engelhard et al., 2010). This has been found in analogue samples and in students with high OCD-like symptoms (Engelhard et al., 2011a) or test anxiety (Engelhard et al., 2011b). In the latter study, 29 participants were asked to describe two feared test-anxiety related flashforward and were randomly assigned to recall+EM or recall only condition. Results showed that recall+EM reduced vividness and emotionality of flashforwards, relative to recall only.

This suggests that the beneficial effects of EM on flashforwards among students with test anxiety may also increase test performance. The aim of this study was to examine if EM while recalling test related mental images would result in better test performance, compared to recall only and no intervention. Students with high test anxiety scores were pre-selected and were randomly assigned to one of the three conditions. Before and after the intervention,

they took an IQ test. After completing the pre-IQ test, a pressure speech was given to increase pressure and anxiety related to test taking, creating a high stake testing environment. Recall+EM and recall only were then requested to elaborate and write down their negative cognitions regarding the upcoming post-IQ test (flashforwards) and to visualize these test-related flashforwards as vividly as possible. Recall+EM were instructed to recall their flashforward during EM, participants in the recall only group were to recall their flashforwards while focusing on a dot, while the control group had no intervention. We hypothesized that (1) participants in recall+EM would show larger decreases in vividness and emotionality of test-related flash forwards compared to recall only, and (2) participants in recall+EM would show increased performance on the IQ test, compared to recall only and the no intervention control group.

2. Method

2.1 Participants and general procedure

A total of 345 undergraduates from Utrecht University were approached during lectures or online (net-questionnaire) to complete a short test anxiety survey, which included four questions that were rated on a scale from 1 (never) to 4 (always). The questions were (1) During tests, I feel very anxious, (2) During tests, I sometimes get so nervous that I forget facts I really know, (3) During test, I often think about the consequences of failure, (4) I tend to worry allot before taking an important test. These items were taken from Test Anxiety Inventory (TAI) Spielberger et al. (1980). Students were invited to participate in a laboratory study if they scored at least three times a score of 3 or twice a score of 4 on this scale. From the 345 student, 26% met the criteria for test anxiety and 54 students agreed to participate for course credit or financial award. Six of them were excluded, resulting in a final sample of N= 48. Reasons for exclusion were the following. Four students failed to meet the required

vividness and emotionality ratings (>40 on 0-100 scale as used in Engelhard et al., 2011a). One case was excluded due to current psychological treatment. Another case was excluded because of non-compliance with test requirements (not completing the questionnaires and not believing the pressure speech manipulation). One outlier on the SCL-90-R was changed into M ± 2.5*SD.

Participants were tested in a soundproof cabin room that contained a one way mirror window. First, participants received written and oral information about the study. After providing written informed consent, participants were randomly assigned to one of the three conditions. Participants in the recall+EM or recall only condition were administered: four questionnaires (see below), the pre-IQ test, the pressure speech instructions, the mental image manipulation, and post-IQ test. Participants in the control condition followed the same procedure except that the mental image manipulation was replaced with no intervention.

The pre-IQ test included 3 subtasks of the Wechsler Adult Intelligence Scale third edition (WAIS-III; Harcourt Test Publisher, 2004). The first subtask arithmetic (13 items) demands concentration while manipulating mental mathematical problems (e.g., "How many 20-cent eggs can you buy for two euro's"). The second subtask matrix reasoning (13 items) requires nonverbal abstract problem solving capacities and spatial reasoning. The third subtask digit span (14 items) demands attention, concentration and mental control (e.g., "Repeat the numbers 1-2-3 in reverse sequence"). Matrix reasoning is part of the perceptual reasoning index, arithmetic and digit span are part of the working memory index, representing two of the four major components of intelligence.

The pressure speech instructions were modified from Ramirez and Beilock (2011) to create a high stake testing environment. We told participants that an improvement of 20% (on score and time) on the post-IQ test compared to the pre-IQ test would yield an additional financial reward of 10 euro. Participants were told they have been coupled to a buddy and in

order to receive the additional financial reward both the participant and their buddy would have to meet the required 20% improvement on the post-IQ test. Each participant was then informed that their buddy had already completed the post-IQ test and had shown the required 20% improvement, and that the participant's performance would determine whether an additional financial reward would be given to both the participant and buddy. Participants were also told that two professors would be observing their behavior behind a one way mirror window.

After pressure speech, participants in the recall+EM and recall only were asked to elaborate on negative cognitions regarding the upcoming post-IQ test (e.g., "I will fail to improve on the post-IQ test and my buddy will not receive the financial reward because of me"). Participants were instructed to write these cognitions down on a paper and to visualize them as clearly and vividly as possible. During recall+EM or recall only, participants were told to recall the visual image related to their negative cognition, and at the same time fixate on a white dot on a computer screen, moving from left to right (EM) or stationary (recall only). Following the procedure done by Engelhard et al. (2010b) the duration of recall+EM or recall only was six times 24 seconds with 10 seconds breaks in between. Participants in the no intervention control group were not given any task and were instructed to wait (8 minutes) for the experimenter.

The post-IQ test was identical to the pre-IQ test. The total score was computed (i.e., correct number of items) for pre-IQ test and post-IQ test separately (arithmetic + matrix reasoning + digit span), with a potential score range from 0-40.

State anxiety was measured three times: (1) after the questionnaires (baseline), (2) right after the pressure speech instructions, and (3) right before the post-IQ test. Participants also rated emotionality and vividness of the visual images about the negative cognitions related to the post-IQ test, before and after the conditions recall+EM, recall only and control.

Participants also rated the difficulty experienced in retaining these visual images. State anxiety, difficulty, emotionality and vividness were rated using 0 (‘not at all) to 100 (‘extremely’) VAS presented on a computer using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

The experiment ended with a relaxation exercise and debriefing that involved some questions (e.g. what do you think the aim of the research was?), the credibility of manipulation (e.g. did you believe you were coupled to a buddy?).

1.2 Questionnaires

The 20-item Test Anxiety Inventory (TAI; Spielberger et al., 1980) was used to measure anxiety related to test situations. Items (e.g., During tests, I sometimes get so nervous that I forget facts I really know) were rated on a 4 point scale (1= almost always, 4= almost never) for the past week. It measures the two major components of test anxiety; worry and emotionality. Potential total score (worry + emotionality component) was computed (score range from 20-80).

The 12-item Brief Fear of Negative Evaluation (B-FNE; Leary, 1983) measures social anxiety. Each item contains a statement (e.g., I often worry that I will say or do the wrong things) rated for the past week. Items were rated from 0 (not at all characteristic of me) to 4 (extremely characteristic of me). B-FNE total score range from 0-48.

The 90-item Symptom Check List (SCL-90-R; Arrindell & Ettema, 2004) was used to measure a broad range of psychological problems. Items (e.g., Feeling so restless you couldn’t sit still) were rated on a 5 point scale (1= not at all, 5= extremely) for the past week. The total score (range from 90-450) was computed.

The Neuroticism and Extraversion subscales of the Eysenck Personality Questionnaire (EPQ; Sanderman et al., 1991) were administered. Only the 22-item neuroticism subscale will

be reported (N-EPQ). Items were rated on a 'yes or no' format (e.g., Does your mood often go up and down?). N-EPQ total score range from 0-22.

2. Results

2.1 Analysis of questionnaires

Table 1

Mean and Standard Deviation for Questionnaires per Condition

Condition		SCL-90-R	N-EPQ	TAI	B-FNE
Recall+EM	M	149.53	10.33	45.80	26.80
	SD	35.30	4.87	8.55	8.35
Recall only	M	191.00	13.13	47.19	25.31
	SD	63.89	5.93	9.25	10.12
Control	M	147.38	10.31	48.94	29.75
	SD	43.97	5.21	12.15	8.11

Table 1 shows the means and standard deviations for the questionnaires for the three conditions. We examined whether the scores on the questionnaires differed between the conditions. ANOVAs showed recall+EM, recall only and no intervention control did not significantly differ on levels of test anxiety (TAI; $F(2) = .37$, $p = .69$), fear of being negatively evaluated (Brief-FNE; $F(2) = 1.03$, $p = .37$), and neuroticism (N-EPQ; $F(2) = 1.44$, $p = .25$). Conditions did significantly differ on the SCL-90-R questionnaire; $F(2) = 3.93$, $p = .03$. Independent t-tests showed that recall+EM and recall only differed significantly, $t(29) = -2.22$, $p = .04$, recall only and control condition differed, $t(30) = 2.25$, $p = .032$, with participants in recall only showing higher SCL-90-R scores than recall+EM and control condition. No significant difference is found between recall+EM and control condition for SCL-90 scores, $t(29) = .15$, $p = .88$.

Table 2

Mean and Standard Deviation for Questionnaires

	SCL-90-R	N-EPQ	TAI	B-FNE
M	162.99	11.28	47.34	27.30
SD	52.71	5.42	10.00	8.92

Table 2 shows total mean and standard deviation on the questionnaires. Overall, participants had higher SCL-90 scores than a sample of non-clinical adults (M = 118.28, SD = 32.38; Arrindell & Ettema, 2004), and lower scores compared to a sample of adult psychiatric patients (M = 203.55, SD = 61.60; Arrindell & Ettema, 2004). Neuroticism scores were higher for participants compared to non-clinical adults (M = 8.6, SD = 5.2; Sanderman et al., 1980) and lower when compared to a sample of outpatients with anxiety disorder (M = 13.2, DS = 5.4; Sanderman et al., 1980). Participants had higher levels of test anxiety (TAI) when compared to the norm of American undergraduates (male: M = 46.48, SD = 12.43, female: M = 42.79, SD = 13.70; Spielberger et al., 1980). Participants mean score on the B-FNE was somewhat lower when compared to undergraduate students (M = 35.7, SD = 8.10; Leary, 1983).

2.2 Analysis of baseline difficulty recalling the image, vividness, and emotionality

Table 3 shows the mean and standard deviation on baseline difficulty, vividness and emotionality ratings between recall+EM and recall only group. With respect to difficulty visualizing negative cognitions, an independent t-test (two-tailed) showed no significant difference between recall+EM and recall only, $t(29) = .19$, $p = .85$. There were no differences for baseline vividness between recall+EM and recall only, $t(29) = .43$, $p = .67$, and for emotionality ratings between recall+EM and recall only, $t(29) = -.44$, $p = .66$.

Table 3

Mean and Standard Deviation for Baseline Ratings Between Conditions

	Recall+EM		Recall only	
Baseline Ratings	M	SD	M	SD
Difficulty	39.70	22.87	38.22	20.68
Vividness	67.88	14.03	65.70	13.99
Emotionality	70.90	15.55	73.99	16.19

2.3 Analysis of Pre-Post differences on vividness and emotionality ratings

Figure 1 shows vividness and emotionality ratings for the two recall conditions. To examine whether vividness and emotionality decreased for the recall+EM condition, relative to recall only, two 2 x 2 repeated measures of analyses were conducted with condition (recall+EM, recall only) and time (pre-IQ test, post-IQ test). For vividness ratings, there were no main effects for condition, $F(1,14) = .43$, $p = .52$, and time, $F(1,14) = .18$, $p = .68$, and the crucial condition x time interaction was also not significant, $F(1,14) = 1.54$, $p = .25$. Exploratory t-tests showed no significant change in pre-post vividness scores for recall+EM, $t(14) = .41$ $p = .34$, and recall only, $t(15) = -1.53$, $p = .08$.

For emotionality ratings, there were no main effects for condition, $F(1,14) = 2.54$, $p = .13$. Significant effect was found for time, $F(1,14) = 5.68$, $p = .03$. The crucial Condition x Time interaction was not significant, $F(1,14) = 2.30$, $p = .15$. Exploratory t-tests showed a significant decrease in pre-post emotionality scores for the recall+EM condition, $t(14) = 3.307$, $p = <.005$, but not for the recall only condition, ($t(15) = .21$ $p = .42$).

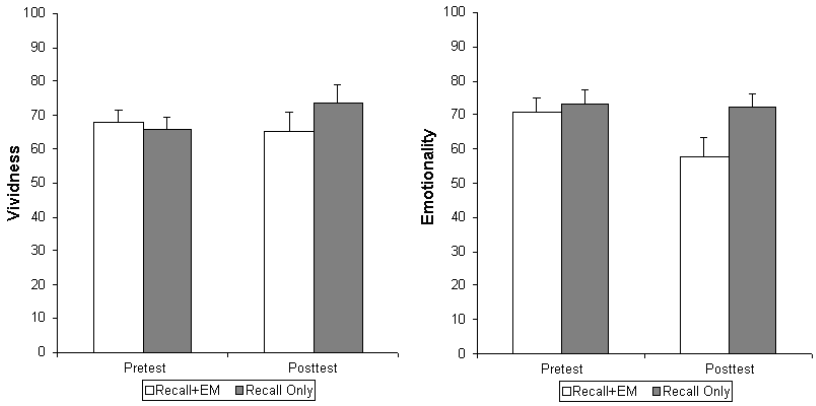

Figure 1. Pre and post vividness and emotionality ratings for recall+EM and recall only.

2.4 Analysis of state anxiety

To examine whether state anxiety differed between recall+EM, recall only and control condition, a 3 x 3 repeated measures of analyses were conducted with condition (recall+EM, recall only and control) and time (baseline, after the pressure speech instructions and right before the post-IQ test). There was a significant main effect for time, $F(1,14) = 85.517$, $p = <.01$ but not for condition, $F(1,14) = .11$, $p = .75$. The crucial Condition x Time interaction was significant, $F(1,14) = 7.63$, $p = < .05$.

Figure 2 shows the pressure ratings for each condition on pressure moment (1) baseline, (2) after the pressure speech instructions and (3) right before the post-IQ test. Exploratory t-test showed difference in state anxiety ratings between measurement (2) after the pressure speech and (3) right before the post IQ-test significantly differed between recall+EM and control condition $t(29) = .77$ $p = <.05$, with recall+EM showing larger increase in state anxiety rating (see figure 2). Other t-tests between measurement moments and conditions showed no significant results.

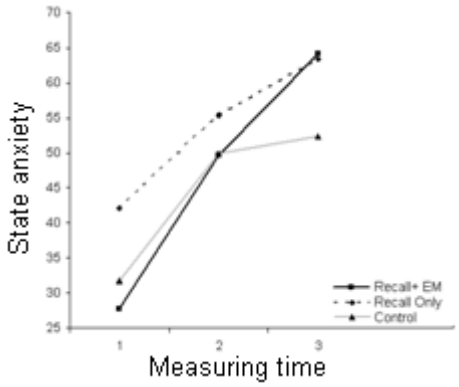

Figure 2. State anxiety ratings for recall+EM, recall only and control condition on (1) baseline, (2) after the pressure speech instructions and (3) right before the post IQ-test).

2.5 Analysis of IQ test scores

Table 4

Mean and Standard Deviation for Test Scores per Condition

	Recall+EM		Recall only		Control	
	M	SD	M	SD	M	SD
Pre-IQ test	24.40	4.03	23.00	3.62	24.19	5.04
Post-IQ test	27.13	3.94	24.88	3.12	25.81	4.86

Table 4 shows descriptives for pre and post IQ test scores per condition. To examine whether test scores differed between recall+EM, recall only and control condition, a 3 x 2 repeated measures of analyses were conducted with condition (recall+EM, recall only and control group) and time (pre-IQ test and post-IQ test). There was a significant main effect for time, $F(1,14) = 32.32$, $p = <.01$ but not for condition, $F(1,14) = .30$, $p = .59$. The crucial Condition x Time interaction was not significant, $F(1,14) = .83$, $p = .38$. This indicates a similar increase in test scores between groups. Figure 3 shows pre and post test scores for recall+EM, recall only and no intervention control group.

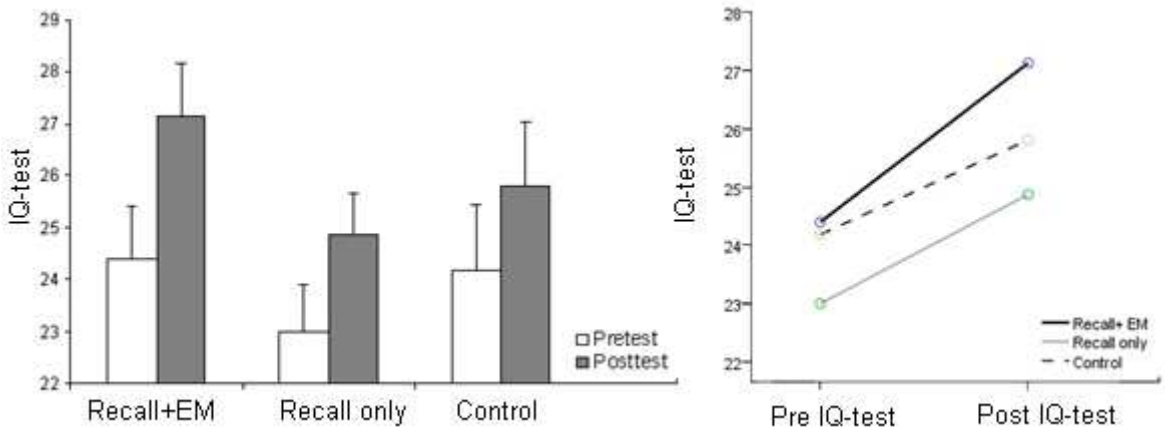

Figure 3. Pre and post test scores for recall+EM, recall only and control group.

3. Discussion

The current study aimed at investigating the relationship between EM and test-related images on the performance of test anxiety students. We predicted that EM would (1) reduce the vividness and emotionality of test-related mental images in students with test anxiety compared to recall only and control group and (2) that this would increase performance on a high stake task. To investigate this, students with test anxiety were pre-selected and randomly assigned to one of three conditions. All groups followed same protocol except for treatment, one group receiving EM, the second had no EM (recall only, stationary) and a third no intervention control condition.

With respect to the first hypothesis (larger decrease in emotionality and vividness for EM compared to recall only), repeated measures ANOVA showed that the crucial interaction between groups on vividness and emotionality was not significant, which was unexpected and inconsistent with prior research (Engelhard et al., 2010, 2011a, 2011b). With respect to the second hypothesis (participants in recall+EM show better test performance compared to recall only and control), no differences were found between the conditions.

Why did we fail to replicate earlier findings concerning reduced vividness and emotionality after EM? One possibility might be that the mental images were not vivid or

emotional enough. This, however, does not seem to be the case: Figure 1 shows relatively high pre vividness and emotionality scores, indicating that students reported test-anxiety related mental images.

A more likely explanation for the failure of EM to reduce vividness and emotionality may relate to the type of flashforwards used during the experiment. For this research we deviated from the usual protocol (Engelhard et al., 2011a, 2011b) in which autobiographical recurrent mental images were used. In this study, mental images were specifically evoked and assessed that related to manipulated supposed high stake IQ test. A pressure speech was given to create a high stake testing environment for the post-IQ test (cf. Ramirez & Beilock; 2010). Afterwards we asked participants what might worry and concern them regarding the upcoming post-IQ test (e.g. I'm afraid I will not improve on the post-IQ test and lose the additional reward for me and my buddy). These worries and concerns were further explored and formed into cognitions (e.g. I am not smart enough; I am ashamed and my buddy gets angry because I lost us our monetary reward). Participants were asked to form a vivid image from these cognitions and followed the procedures for recall+EM or recall only (see methods 2.1 participants and general procedure). Thus, participants formed images of a real event that would follow, and it is possible that such images may be affected less by recall+EM compared to images of past events or potential future events. That is, the anticipation of an impending event and impending real doom may be less modifiable by such a brief intervention.

There are a few limitations to the current study. First, the sample size per group was small, which might have impeded the chance to find significant effects. Further research should include larger sample sizes. Finally, participants performed better during post-IQ test, which may have left little room in variance for group differences. Future research may use

other tests and actual flashforwards that the person suffers from, which may increase ecological validity.

References

Arrindell, W. A., Ettema, J. H. M. (2004). *Symptom Checklist SCL-90*. Lisse: Swets & Zeitlinger.

Boyacioglu, N., & Kucuk, L. (2001). Irrational beliefs and test anxiety in Turkish school adolescents. *The Journal of School Nursing December, 27,* 447-45.

Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97,* 268–274.

Culler, R. E., & Holahan, C. J. (1980). Test anxiety and academic performance: The effects of study-related behaviors. *Journal of Educational Psychology, 72,* 16-20.

Engelhard, I. M., van den Hout, M. A., Janssen, W. C., & van der Beek, J. (2010). Eye movements reduce vividness and emotionality of "flashforwards". *Behavior Research and Therapy, 48,* 442-447.

Engelhard, I. M., Van den Hout, M. A., Dek, E. C. P., Giele, C. L., van der Wielen, J., Reijnen, M. J., & van Roij, B. (2011a). Reducing vividness and emotional intensity of recurrent "flashforwards" by taxing working memory: An analogue study. *Journal of Anxiety Disorders, 25,* 599-603.

Engelhard, I. M., Sijbrandij, M., van den Hout, M. A., Rutherford, N. M., Rahim, H. F., & Kocak, F. (2011b). Choking under pressure: Degrading flashforwards related to performance anxiety. *Journal of Experimental Psychopathology, In press.*

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6,* 409-434.

Harcourt Test Publishers. (2004). WAIS-III Nederlandstalige bewerking. *Wechsler Adult Intelligence Scale, derde editie*. Swets Test Publishers, Lisse.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety, *Review of*

Educational Research, 58, 47-77.

Hill, K. T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it. *Elementary School Journal, 85,*105–126.

Hopko, D. R., McNeil, D. W., Lejuez, C. W., Ashcraft, M. H., Eifert, G. H., & Riel, J. (2003). The effects of anxious responding on arithmetic and lexical decision task performance. *Journal of Anxiety Disorder, 17,* 647-665.

Lang, J. W. B., & Lang, J. (2010). Priming competence diminishes the link between cognitive test anxiety and test performance: Implications for the interpretation of test scores. *Psychological Science, 21,* 811-819.

Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation scale. *Personality and Social Psychology Bulletin, 9*, 371-375.

Maxfield, L., Melnyk, W. T. (2000). Single session treatment of test anxiety with Eye Movement Desensitization and Reprocessing (EMDR). *International Journal of Stress Management, 7 (2),* 87-101.

McCabe, M. P. (2005). The role of performance anxiety in the development and maintenance of sexual dysfunction in men and women. *International journal of stress management, 4,* 379-388.

McDonald, A.S. (2001). The prevalence and effects of test anxiety in schoolchildren. *Educational Psychology, 21* , 89-102.

McGuire, D. P., Mitic, W., & Neumann, B. (1987). Perceived stress in adolescents: What normal teenagers worry about. *Canada's Mental Health, 35,* 2-5.

Methia, R. A. (2004). *Help your child overcome test anxiety & achieve higher test scores.* College Station, TX: Virtual Bookworm.com Publishing.

Musch, J., & Broder, A. (1999). Test anxiety versus academic skills: A comparison of two

alternative models for predicting performance in a statistics exam. *British Journal of Educational Psychology*, *69*, 105–116.

Putwain, D.W. (2007). Test anxiety in UK schoolchildren: Prevalence and demographic patterns. *British Journal of Educational Psychology, 77,* 579–593.

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76,* 92-104.

Ramirez, G., & Beilock, S.L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science, 331,* 211-213.

Richards, A., French, C. C., Keogh, E. & Carter, C. (2000). Test anxiety, inferential reasoning and working memory load. *Anxiety, Stress, and Coping, 13*, 87-109.

Salminen, S., Liukkonen, J., Hanin, Y., & Hyvönen, A. (1995). Anxiety and athletic performance of Finnish athletes: Application of the zone of optimal functioning model. *Personality and Individual Differences*, *19,* 725-729.

Sanderman, R., Arrindell, W.A., Ranchor, A.V., Eysenck H. J. & Eysenck S. B. G. (1991). *Eysenck Personality Questionnaire (EPQ).* Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime Reference Guide. Pittsburgh: Psychology Software Tools, Inc.

Spielberger, C. D. (1980). *Preliminary professional manual for the Test Anxiety Inventory ("Test Anxiety Inventory") TAI.* Palo Alto: Consulting Psychologists Press, Inc.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory (STAI) (Form Y) ("Self-Evaluation Questionnaire")*. Palo Alto: Consulting Psychologists Press, Inc.

Turner, B. G., Beidel, D. C., Hughes, S. & Turner, M. W. (1993). Test anxiety in African American school children. *School Psychology Quarterly, 8*, 140-152.

Van den Hout, Muris, P., Salemink, E., Kindt, M. (2001). Autobiographical memories

become less vivid and emotional after eye movements. *British Journal of Clinical Psychology*, *40*, 121-130.

Yerkes, R.M. & Dodson, J.D. (1908). The relationship of strength of stimulus to rapidity of habit formation. *Journal of Comparative and Neurological Psychology, 18*, 459-482.

Yoshi, M., Kudo, K., Murakoshi, T., & Ohtsuki, T. (2009). Music performance anxiety in skilled pianist: Effects of social-evaluative performance situation on subjective, autonomic, and electromyography reactions. *Experimental Brain Research, 199*, 117-126.

Zeidner, M. (1998). *Test anxiety: The state of the art*. Plenum Publishing Corporation, New York.

Zeidner, M. & Safir, M.P. (1989). Sex, ethnic, and social differences in test anxiety among Israeli adolescents. *Journal of Genetic Psychology, 150*, 175-185.