Effects of Virtual Training on Emotional Response

A Comparison between Different Emotion Regulation Strategies

Tibor Bosse[1], Charlotte Gerritsen[1,2], Jeroen de Man[1,2], and Jan Treur[1]

[1] Vrije Universiteit Amsterdam, Agent Systems Research Group
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
{t.bosse,c2.gerritsen,j.de.man,j.treur}@vu.nl
[2] Netherlands Institute for the Study of Crime and Law Enforcement
De Boelelaan 1077a, 1081 HV Amsterdam, The Netherlands
{cgerritsen,jdeman}@nscr.nl

Abstract. Learning to regulate one's emotions under threatening circumstances is important, among others, for professionals like police officers and military personnel. To explore the opportunities of Virtual Reality-based training for such professionals, this paper describes an experiment performed to investigate the impact of virtual training on participants' experienced emotional responses in threatening situations. A set of 15 participants was asked to rate the subjective emotional intensity of a set of affective pictures at two different time points, separated by six hours. The participants were divided into three groups: the first group performed a session of virtual training in between, in which they received a choice-reaction task, the second group performed a session of virtual training, in which they had to apply reappraisal strategies, and a control group did not have any training session. The results indicate that the reappraisal-based training caused the participants in that group to give significantly lower ratings for the emotional intensity of the negative pictures, whereas the content-based training resulted in significantly higher ratings compared to the group without training. Moreover, a second experiment, performed with the same participants six months later, indicated that these effects are fairly persistent over time, and that they transfer to different pictures with similar characteristics.

Keywords: Virtual training, emotion regulation, emotional response.

1 Introduction

The ability to cope with negative stimuli from the environment is a useful characteristic of human beings. Almost on a daily basis, we are confronted with situations that in one way or the other invoke negative emotions. A particular type of negative emotion, which is typically induced by perceived threats, is *fear* [19]. Depending on the person, different types of stimuli that may trigger fear vary from horror movies and scary animals to enclosed spaces and public speaking. The probability of being confronted with such stimuli depends, among others, on the person's profession. On average, professionals in domains such as the police, military and public transport are more

K. Imamura et al. (Eds.): BHI 2013, LNAI 8211, pp. 21–32, 2013.

likely to be confronted with fear-inducing stimuli than people with an office job. It is therefore not surprising that these types of job are usually more appropriate for people that are strong at regulating their levels of fear.

Nevertheless, even the 'coolest' of individuals may have difficulties to function adequately in case the stimuli are extreme, such as in cases of military missions or terrorist attacks. First, the extreme emotions experienced in these situations may impair their cognitive processes like attention and decision making [14, 20]. And second, even if they make optimal decisions from an external perspective, they have an increased risk of developing anxiety related disorders such as Post-Traumatic Stress Disorder (PTSD) [4]. For these reasons, much time and money is spent on developing appropriate training in these domains. Increasingly often, virtual environments are successfully used to train performance and decision making of professionals under more realistic and stressful situations (see for example [3, 9]). Furthermore, methods to prevent or treat PTSD after a traumatic event are costly and may even have negative effects [7]. *Primary prevention*, before any traumatic event has occurred, has been proposed as a promising alternative [5, 6]. A promising technique for primary prevention, which has recently received much attention, is 'stress inoculation training' based on Virtual Reality (VR). The assumption behind this approach is that, by gradually exposing a trainee to fear-provoking stimuli, a VR system is able to increase her 'mental readiness' [21, 23]. In that sense, this approach has similarities with exposure therapy [8, 12]. VR-based stress training has proved to be successful, among others, for bank employees [16] and airline crew [24], to increase preparation for hostage situations.

Despite these developments, the domain of VR-based training is still in its infancy and many questions remain. Three of such questions that are of interest for this paper are the following:

1) What type of training should be provided in order to maximise training effectiveness in reducing negative emotional effects?

2) What are the long-term effects of such types of training?

3) To what extent is there transfer of training to different, but comparable stimuli?

The current paper makes some steps towards the investigation of these research questions by means of an experiment where participants were exposed to negative stimuli via a computer screen. The paper is organised as follows. First, an experiment is introduced that was used to assess the impact of different types of virtual training on the experienced emotional intensity towards the stimuli presented. The experiment involves a first part that was mainly designed to investigate research question (1), and a second part (performed after six months) to address research question (2) and (3). After that, the results of the experiment are discussed, and the paper is concluded with a discussion.

2 Experimental Design

The first research question addressed is what type of VR-based training is appropriate, in order to obtain a successful decrease of emotional responses towards negative stimuli. Previous research outside the VR domain (e.g., [17]) suggests that the effectiveness of exposure therapy is partly determined by the way the person deals with the negative stimuli. Within the context of virtual training, a number of strategies can be used, varying from just looking at the stimuli to performing different emotion regulation strategies such as 'attentional deployment', 'cognitive change', and 'suppression' [10]. As a first step to investigate and compare the effects of different strategies, an experiment was performed in which participants' reactions to viewing negative pictures from the IAPS picture set, 'developed to provide a set of normative emotional stimuli' [13], were assessed for two different types of training.

The setup of this experiment, which is described in the next section, was inspired by an experiment by Helm et al. [11] in the context of REM sleep. In the experiment by Helm et al. [11], participants' reactions to emotional pictures were assessed at two different time points. One group of participants was allowed to sleep in between the sessions, whilst the other group was not allowed to sleep. In the experiment reported in the current paper, we mostly re-used the setup of that experiment, but we replaced the task of sleeping by the task to perform (different variants of) virtual training. Note that the initial phase of our experiment has been reported in [2] and that by copying the setup from [11] it has already been established that no effects of circadian rhythm play a role.

2.1 Participants

Fifteen healthy adults (of age between 26 and 32 years, with a mean of 28.2) participated in the experiment, and were randomly assigned to one of three groups (to which we will refer as the 'training 1' group, the 'training 2' group, and the 'no training' group), in such a way that each group consisted of 5 participants. Six of the participants were female and nine were male.

2.2 Setup

In the first part of the experiment, the participants in the control (no training) group participated in two rounds, separated by a pause of six hours (see Figure 1, lowest line, test 1 and 2). In these rounds they were presented 5 times 30 pictures from the IAPS picture set [13]. The sets of pictures used were identical to the sets used within [11], covering pictures with valence scores ranging from 1.45 (negative) to 8.28 (positive) and arousal scores from 2.28 (low) to 7.12 (high), according to the standard IAPS classification (ranging from 1 to 9). The participants were first shown a black fixation mark for 500ms, after which the image was shown for 2000ms. For 2500ms after the image, a question was shown asking the participants to rate the emotional intensity of the picture on a scale from 1 to 5 (1 being non emotional and 5 being very emotional)[1]. Finally, for another 2700ms a grey fixation mark was shown, followed by the same sequence for the next image, and so on. The images were shown in 5 blocks of 30 images with small breaks in between. Image order was fixed for each block of images, but the order of blocks was randomised between participants. For a total of 150 images, the participants needed approximately 20 to 25 minutes to complete the test. Furthermore, the heart rate and the skin conductance of the participants were measured with the PLUX wireless sensor device (http://www.plux.info/)[2].

The participants in both training groups also participated in these rounds, just like the control group. However, in between these two rounds they performed a virtual training session. This training occurred three hours after the first round and three hours before the second round (see Figure 1, upper and middle line). The virtual training made use of the same pictures used in the other rounds. Within the training 1 group, the participants were given a choice-reaction task in which they had to assess the valence of the picture as quickly as possible while the image increased in size up to double the size of the original picture (i.e., whether it gave them a positive or a negative emotion). They could make this distinction by either clicking the mouse or pressing the spacebar. Within the training 2 group, the participants were asked to view them while actively reducing their emotional response until they felt comfortable looking at the picture (e.g., by assuring themselves that the pictures were not real). The motivation for using these two types of training is that we wanted to investigate the impact of type of training and task instructions (i.e., learning to work with potentially negative stimuli in a dynamic context vs. learning to cope with negative stimuli directly by performing reappraisal) on subjective emotional response. Although a large number of emotion regulation strategies exist, we selected reappraisal because this strategy has proven successful in a number of related domains, e.g., [11].

In addition to the above, the participants in all three groups participated in a second part of the experiment, which was performed in the afternoon six months after the first part.[3] This second part was the same for all three groups (see Figure 1, test 3). In this part, participants were presented the same 5x30 pictures from the IAPS picture set as presented in the first part, and were again asked to rate the emotional intensity of the pictures on a scale from 1 to 5. This was done to address the second research question mentioned above, i.e., studying the long term effects of the different types of training. Next, after a short break, they were again presented 5x30 pictures from the IAPS picture set, which were new pictures, but with similar valence and arousal scores as the first 5x30 pictures. This was done to find out whether there was any transfer of the training to different pictures, thereby addressing the third research question.

---

[1] Note that this intensity was independent of the 'valence' of the picture (i.e., pos. vs. neg).

[2] The results of these measurements are not further discussed in this paper. They were collected to gain more insight in the relation between the presented stimuli and physiological states, which will be further explored in a follow-up experiment.

[3] In both the control and training 2 group one participant dropped out.

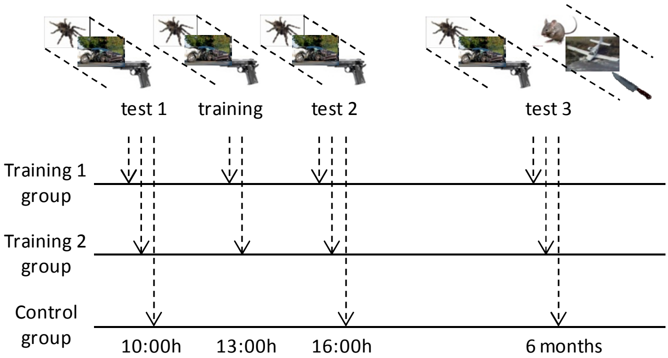

Fig. 1. Experimental design

2.3 Implementation

Both the test environment and the two training environments were implemented using the PsychoPy software (http://www.psychopy.org/). This package provides an API for creating psychological experiments using the programming language Python. In combination with the Python API provided by PLUX, all ingredients for implementing both environments were available. The implementation itself is relatively straightforward, looping through the different images in fixed intervals and recording both the physiological measurements from the PLUX device as well as the manual responses from the participants.

3 Results

The results of the experiment will be described in three separate sub-sections, addressing the three respective research questions. First, for the first part of the experiment, the two types of training are compared with the control group. Recall that the difference between the three groups was the following:

- control group: no training
- training 1 group: choice reaction task
- training 2 group: reappraisal task

After this comparison, we investigate whether these types of training have lasting effects, based on the second part of the experiment. Finally, data gathered with a new set of images is compared to the original set to find out if transfer of learning takes place between similar images.

3.1 Type of Training

Figure 2 shows for all 150 pictures (horizontal axis) the absolute change in emotional ratings (averaged over all participants, vertical axis) between the first and the second test, both for the control group and the two training groups. Pictures are sorted as follows: pictures with a negative valence (i.e., a value smaller than 4.5 according to the IAPS classification) are shown on the left and those with a positive valence on the right hand side. Note that we are particularly interested in the negative images, since emotion regulation training usually aims at decreasing emotional response to negative stimuli. Moreover, both the negative and the positive pictures are sorted with respect to the change in emotional rating of the control group (the solid red line). As can be seen, the curve for the control group is situated around 0 (mean value -0.05), whereas the curves of both training groups are lower (mean value ‘training 1’ -0.11; ‘training 2’ -0.4). A paired t-test confirmed that training 2 significantly lowered the emotional ratings more than the control group (t(149) = -8.15, p < 0.001). However, this change was not significant for the training 1 group (t(149) = -1.34, p = 0.18). This indicates that, for this set of participants[4], training 2 resulted in significantly lower ratings of the images in the second test.

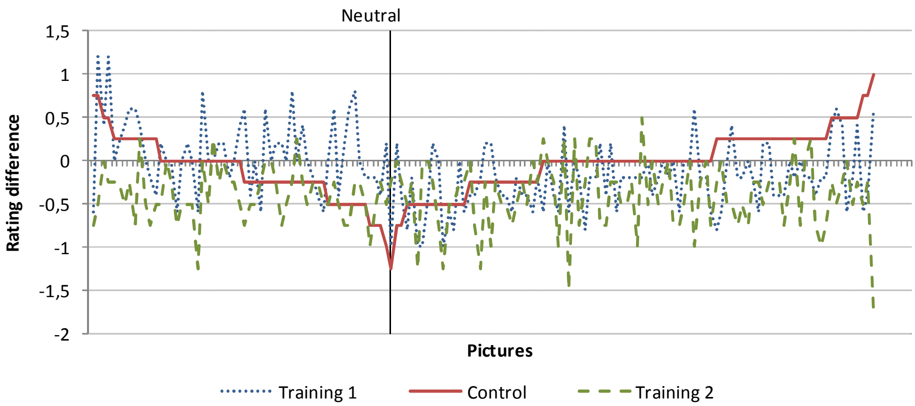

Fig. 2. Absolute change in emotional ratings for 150 pictures (averaged over all participants)

If we focus on only those images that are negatively valenced (left part of the graph in Figure 2), the results are different. The mean change for the control group is slightly lower at -0.11. Training 1 resulted in an increase of emotional ratings with a mean change of 0.11, while training 2 still has a mean change of -0.39. Both these changes were confirmed to be significant with t(56) = 3.32, p = 0.0016 for training 1 and t(56) = -4,65, p < 0.001 for training 2. From this we conclude that, again for this

[4] This test took, for each of the 150 pictures, the average change in rating given by the participants in the training group, and compared this with the average change in rating given to the same picture by the participants in the control group. Since this way of testing takes the pictures (instead of the participants) as a basis, the results cannot be generalised for the population as a whole. However, an additional (unpaired two sample) t-test has been performed in which for each participant the average change in rating over all 150 pictures was calculated, and these averages were used to compare the training 2 with the control group. Due to the low number of participants, these results were not statistically significant on the p<0.05 level, but a clear trend was found in the results (with t(8) = 2.05, p = 0.07).

set of participants, training 1 resulted in a significant increase of emotional ratings for negative images, while training 2 significantly decreased emotional ratings for those images.

In addition, Figure 3 shows for both training groups and the control group, the relative change in emotional ratings between the first and the second round. On the horizontal axis ratings from 1 to 5 are shown. Each bar represents the percentage change between the first and second round for one particular group for that rating. Thus a positive change represents an increase in the absolute number of pictures that were given that rating in the afternoon.

Looking at the left graph, showing the results for all images, it can be seen that there are small differences between the control group and training 1, in line with the findings above. For training 2, it is clear that there is a decrease in the pictures with a high rating and a corresponding increase of pictures with a low rating.

Focussing on only those images with a negative valence as shown in Figure 3 on the right, the results for both the control group and training 2 show a similar pattern compared with all images. However, for training 1, a trend can be seen towards the higher emotional ratings. This is in compliance with the results above, where we found a significant increase of the mean ratings for this group.

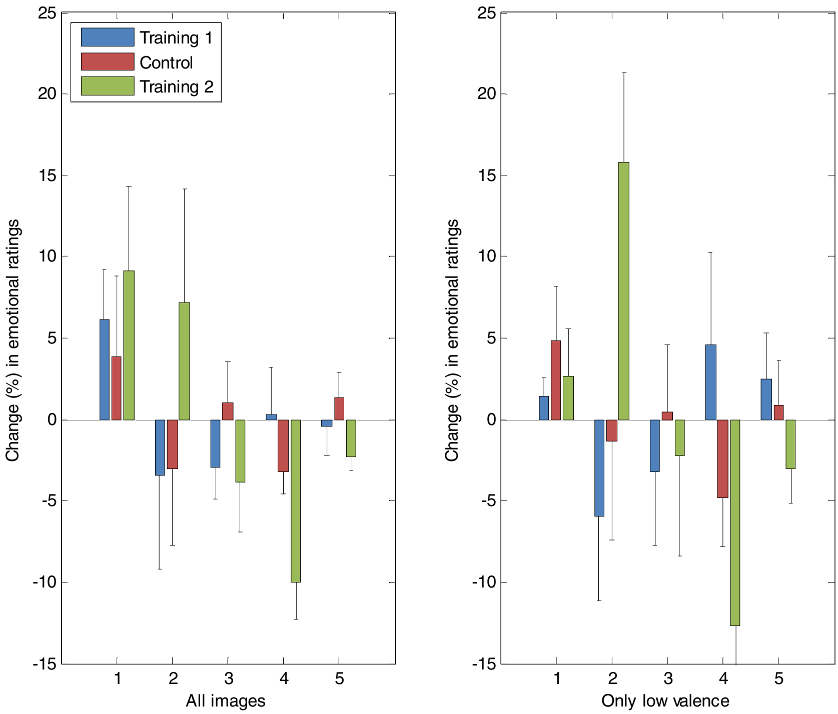

Fig. 3. Relative change in emotional ratings

Finally, in addition to comparing the absolute (Figure 2) as well as the relative (Figure 3) change in emotional ratings among the different groups, we explicitly compared the ratings given in the afternoon with those given in the morning for each group individually (no figures shown). We found that training 2 is the only group in which a significant drop of emotional ratings has occurred for all images ($t(3) = 7.57$, $p = 0.048$) as well as only the negative images ($t(3) = 9.69$, $p = 0.023$). For both the control group and training 1 the differences between the morning and afternoon measurements were not significant. This provides strong evidence for the hypothesis that reappraisal-based virtual training can be used to reduce subjects' emotional responses to negative stimuli at later times.

3.2 Six Months Later

Figure 4 shows the subjective emotional ratings for the two measurements (test 1 and 2, i.e., morning and afternoon) in the first part of the experiment as well as the same measurement (test 3) using the same participants 6 months later. Pictures are sorted on the emotional rating given during the first measurement, such that those pictures with a negative valence are shown on the left and those with a positive valence on the right hand side. For the control group (middle graph), it can be seen that both later measurements vary around the initial ratings given for those images and even so no significant differences were found.

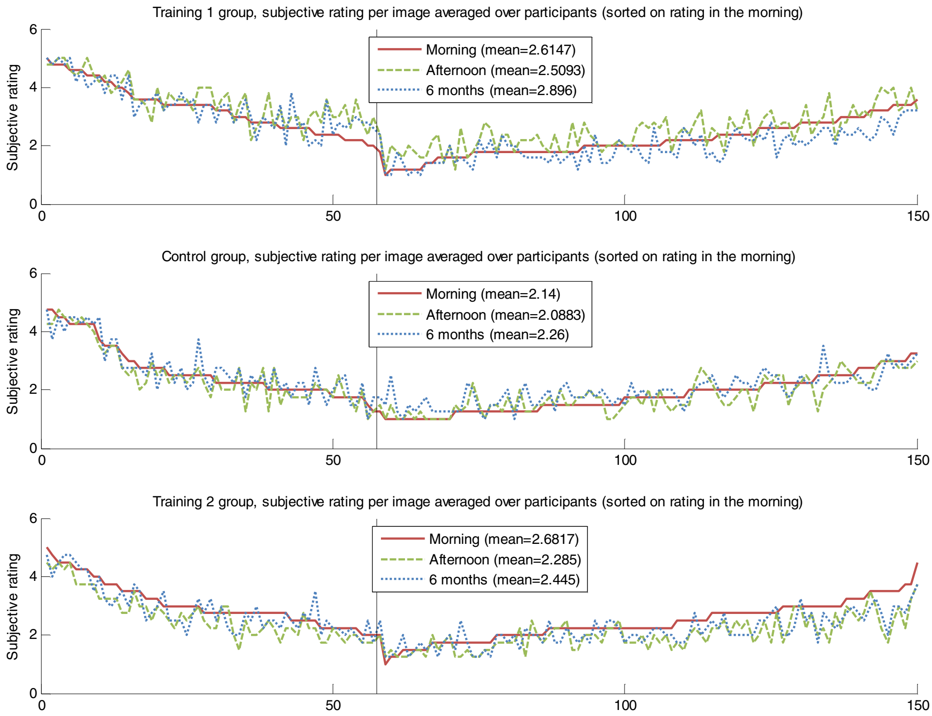

Fig. 4. Emotional ratings for 3 different points in time

For training 1 (top graph) the difference is harder to see. However, statistically, the measurement in the afternoon gave slightly lower ratings overall ($t(149) = 3.0478$, $p = 0.0027$), while six months later the ratings had increased compared to the first measurement ($t(149) = -8.1722$, $p < 0.0001$). Taking into account only the negative images, the mean rating increased from 3.29 in the morning, via 3.40 in the afternoon, to 3.56 after six months, with both the difference between the morning and the six months as well as that between the afternoon and six months measurement ($t(56) = -4.4742$, $p < 0.0001$ and $t(56) = -2.6364$, $p = 0.011$ respectively).

The lower graph in Figure 4 (training 2) shows that both the measurement taken in the afternoon and the one taken six months later resulted on average in lower ratings, whereby the ratings after six months roughly lie between the ratings of the other two time points measured. Furthermore, all these differences are statistically significant with $p < 0.0001$. Regarding the negative images, similar results are found with statistical significance of $p < 0.001$.

Thus, after six months, the control group still had a similar response towards the images as they had initially. The participants taking part in training 1 already showed an increased response towards the negative images in the afternoon and after six months this had increased even more for the negative images as well as the complete set of images. The lowered emotional response caused by training 2 was still present (and significant) after 6 months, albeit less pronounced.

3.3 Transfer of Training

Six months after the initial experiment, each participant also rated a second set of images on the emotional intensity. Those images were randomly selected from the remaining IAPS pictures, with the constraint that the set as a whole matched to the first set on both valence and arousal. Figure 5 shows the differences in ratings between the initial measurement of the first set of pictures and the ratings for the new images after six months. If transfer of training would take place, or in case of the control group with no training at all, similar results would be expected for both sets of stimuli. This can also be seen in the graph, were the curve for the control group is situated very closely around 0. As expected, for this group no significant differences were found using an unpaired t-test for either all images or only the negative ones ($t(298) = -0.3462$, $p = 0.73$; $t(113) = -0.0933$, $p = 0.93$).

Furthermore, Figure 5 shows very few differences between the control group and training 1, indicating little transfer regarding this type of training. An unpaired t-test confirms this: $t(298) = -0.62703$, $p = 0.73$ for all images and $t(113) = -1.0934$, $p = 0.14$ for the negative ones. For training 2, the mean rating has dropped, which can also been seen in Figure 5. These differences were significant for all images ($t(298) = 4.342$, $p < 0.0001$) as well as only the negative pictures ($t(113) = 1.7808$, $p = 0.039$). Thus, although no transfer can be shown for training 1, transfer does seem to take place for training 2.

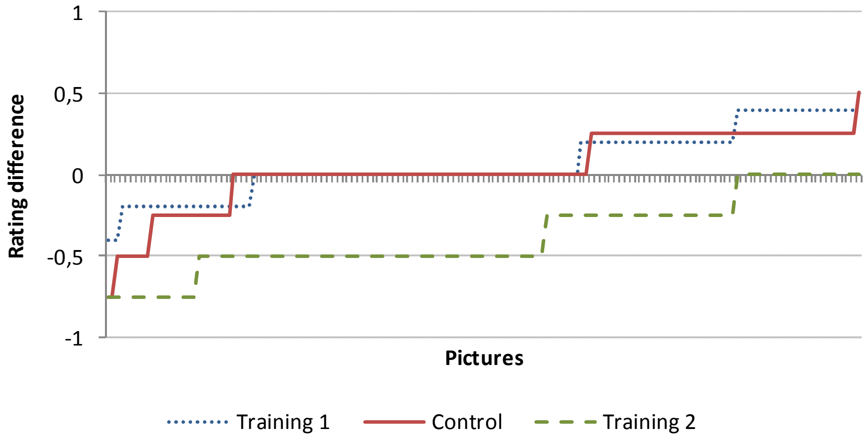

Fig. 5. Differences in emotional ratings between two different sets of pictures

4 Discussion

In this paper an experiment is reported addressing the impact of virtual training on participants' experienced emotional responses towards negative stimuli. Participants were asked to rate the subjective emotional intensity of a set of affective pictures at two different points time. The participants were divided into a first group performing a session of virtual training in between these time points, a second group performing virtual training thereby applying reappraisal strategies, and a control group without any training session. The results are that the reappraisal-based training caused the participants in that group to give significantly lower ratings for the emotional intensity of the negative pictures, whereas the content-based training resulted in significantly higher ratings compared to the group without training. Moreover, a second experiment, performed with the same participants six months later, indicated that these effects are fairly persistent over time, and that transfer to pictures with similar characteristics takes place.

The outcomes of this experiment indicate that, depending on its setup, virtual training may either strengthen the emotional responses to stimuli or weaken them. The fact that the first group shows enhanced responses might be explained by a form of fear conditioning (e.g., [15]) taking place in this setup. By presenting the stimuli in an active, dynamic manner, and by asking the subjects to perform an action as a response, a state of enhanced attention on the stimuli was induced, which may have an opposite effect compared to, for example, the emotion regulation strategy called attention deployment (cf. [10]), and by a process of fear conditioning this may lead to a form of up-regulation as opposed to down-regulation. In contrast to this, in the second group it was explicitly asked to apply an emotion regulation strategy based on reappraisal (cf. [10]). The outcomes indicate that indeed such a setup can strengthen the emotion regulation, which can be explained as inducing a form of fear extinction learning [18, 22].

For further research it is planned to perform more experiments like this, with more participants and a greater focus on interpersonal differences. At the moment, similar experiments are being conducted with different types of stimuli (such as movies and games) to elicit emotional responses. An interesting additional element here is a personality questionnaire to consider individual differences in relation to for example specific personality traits. This opens up the possibility to investigate whether particular personality traits indicate what type of training would be most beneficial for that particular person. Moreover, in future experiments it will be investigated to what extent physiological measurements can provide useful data, in addition to the reported emotion levels. Also different variations in training setups will be explored. Finally, the aim of the project is to build a VR training environment in which the knowledge acquired is incorporated.

Acknowledgements. This research was supported by funding from the National Initiative Brain and Cognition, coordinated by the Netherlands Organisation for Scientific Research (NWO), under grant agreement No. 056-25-013.

References

1. Berking, M., Meier, C., Wupperman, P.: Enhancing Emotion-Regulation Skills in Police Officers: Results of a Pilot Controlled Study. Behavior Therapy 41(3), 329–339 (2010)
2. Bosse, T., Gerritsen, C., de Man, J., Treur, J.: Measuring Stress-Reducing Effects of Virtual Training Based on Subjective Response. In: Huang, T., Zeng, Z., Li, C., Leung, C.S. (eds.) ICONIP 2012, Part I. LNCS, vol. 7663, pp. 322–330. Springer, Heidelberg (2012)
3. Bouchard, S., Guitard, T., Bernier, F., Robillard, G.: Virtual reality and the training of military personnel to cope with acute stressors. In: Brahnam, S., Jain, L.C. (eds.) Advanced Computational Intelligence Paradigms in Healthcare 6. SCI, vol. 337, pp. 109–128. Springer, Heidelberg (2011)
4. Brewin, C.R., Andrews, B., Valentine, J.D.: Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. Journal of Consulting and Clinical Psychology 68(5), 748–766 (2000)
5. Deahl, M.: Traumatic stress-is prevention better than cure? Journal of the Royal Society of Medicine 91, 531–533 (1998)
6. Deahl, M., Srinivasan, M., Jones, N., Thomas, J., Neblett, C., Jolly, A.: Preventing psychological trauma in soldiers: the role of operational stress training and psychological debriefing. British Journal of Medical Psychology 73(1), 77–85 (2000)
7. van Emmerik, A.A., Kamphuis, J.H., Hulsbosch, A.M., Emmelkamp, P.M.: Single session debriefing after psychological trauma: a meta-analysis. The Lancet 360(9335), 766–771 (2002)
8. Foa, E.B., Dancu, C.V., Hembree, E.A., Jaycox, L.H., Meadows, E.A., Street, G.P.: A comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims. J. Consult. Clin. Psychol. 67, 194–200 (1999)
9. Graafland, M., Schraagen, J.M., Schijven, M.P.: Systematic review of serious games for medical education and surgical skills training. British Journal of Surgery 99(10), 1322–1330 (2012)

10. Gross, J.J.: Emotion Regulation in Adulthood: Timing is Everything. Current Directions in Psychological Science 10(6), 214–219 (2001)
11. van der Helm, E., Yao, J., Dutt, S., Rao, V., Saletin, J.M., Walker, M.P.: REM Sleep Depotentiates Amygdala Activity to Previous Emotional Experiences. Current Biology 21(23), 2029–2032 (2011)
12. Krijn, M., Emmelkamp, P.M.G., Olafsson, R.P., Biemond, R.: Virtual Reality Exposure Therapy of Anxiety Disorders: A Review. Clinical Psychology Review 24(3), 259–281 (2004)
13. Lang, P.J., Bradley, M.M., Cuthberth, B.N.: International Affective Picture System (IAPS): Technical Manual and Affective ratings. Gainesville, Fl. The Center for Research in Psychophysiology, University of Florida (1999)
14. Loewenstein, G.F., Lerner, J.S.: The role of affect in decision making. In: Davidson, R., Scherer, K., Goldsmith, H. (eds.) Handbook of Affective Science, pp. 619–642. Oxford University Press, New York (2002)
15. Maren, S.: Neurobiology of Pavlovian fear conditioning. Annu. Rev. Neurosci. 24, 897–931 (2001)
16. Markus, E.: Antwoord op agressie, het effect van training op het omgaan met agressief klantgedrag en bankovervallen. Ph.D. Thesis. Erasmus Universiteit, Rotterdam (2000) (in Dutch)
17. Muris, P., de Jong, P., Merckelbach, H., van Zuuren, F.: Is Exposure Therapy Outcome Affected by a Monitoring Coping Style? Adv. Behav. Res. Ther. 15, 291–300 (1993)
18. Myers, K.M., Davis, M.: Mechanisms of fear extinction. Molecular Psychiatry 12, 120–150 (2007)
19. Öhman, A.: Fear and anxiety: Evolutionary, cognitive, and clinical perspectives. In: Lewis, M., Haviland-Jones, J.M. (eds.) Handbook of Emotions, pp. 573–593. The Guilford Press, New York (2000)
20. Ozel, F.: Time pressure and stress as a factor during emergency egress. Safety Science 38, 95–107 (2001)
21. Popović, S., Horvat, M., Kukolja, D., Dropuljić, B., Ćosić, K.: Stress inoculation training supported by physiology-driven adaptive virtual reality stimulation. Studies in Health Technology and Informatics 144, 50–54 (2009)
22. Quirk, G.J., Mueller, D.: Neural mechanisms of extinction learning and retrieval. Neuropsychopharmacology 33, 56–72 (2008)
23. Rizzo, A.A., Reger, G., Gahm, G., Difede, J., Rothbaum, B.O.: Virtual Reality Exposure Therapy for Combat Related PTSD. In: Shiromani, P., Keane, T., LeDoux, J. (eds.) Post-Traumatic Stress Disorder: Basic Science and Clinical Practice. Springer (2008)
24. Strentz, T., Auerbach, S.M.: Adjustment to the stress of simulated captivity: effects of emotion-focussed versus problem-focused preparation on hostages differing in locus of control. Journal of Personality and Social Psychology 55(4), 652–660 (1988)