MASTER THESIS

DESIGN AND EVALUATION OF A

TECHNOLOGY SUPPORTED MENTAL RESILIENCE TRAINING

Joris Favié, B.Sc.

Media and Knowledge Engineering
Faculty of Electrical Engineering, Mathematics & Computer Science
Delft University of Technology
The Netherlands

Title: Design and evaluation of a technology supported mental resilience training.

Study: Computer Science, Delft University of Technology, the Netherlands

Author: Joris Favié, 9237241, joris@favie.com

Committee: Prof.dr. M.A. Neerincx, Dr.ir. W.P. Brinkman, Dr. V.T. Visch

Date: March 2017

Abstract:

Certain professionals are inherently at risk of experiencing traumatic events, and developing the anxiety disorder Post-Traumatic Stress Disorder (PTSD) as a result. The cost of PTSD is huge, both in terms of human pain and economics. Effective treatments such as exposure therapy do exist, but obviously prevention would be preferable. There is some research into training people to become more psychologically resilient, but this is still in an early stage. Supporting resilience training with technology may offer advantages such as reduced cost and logistics as compared to in vivo training, and being able to present stressors in a very controlled manner. Measuring effectiveness with technology may be more objective than the commonly used self-reporting questionnaires.

In this thesis one possible technology supported resilience training system was designed and evaluated. The system involved video annotation and the cognitive reappraisal paradigm, delivered through three different versions of a computer application that were developed for this purpose. Two novel ways aiming to measure resilience using a pictorial emotional Stroop task and the fear conditioned startle paradigm were evaluated.

Acceptance of the video annotation tool by the target group (military personnel and firefighters) was positive. Due to feasibility considerations its effectiveness could not be measured. Empirical validation of the measurement methods as a measurement of resilience was not successful, possibly as a result of selection bias in the participating sample and use of video material that was not sufficiently aversive.

ACKNOWLEDGEMENTS

I would like to express my thanks to Willem-Paul Brinkman for his guidance, expertise and time, and for his encouragement to write a chapter for a book on a related subject that has now been published (Favié, Vakili, Brinkman, Morina & Neericx, 2016). I would also like to thank Nexhmedin Morina for his feedback on certain issues related to the experimental materials as a psychology expert. Last but not least I am very grateful for the support of my family and girlfriend during the entire study.

TABLE OF CONTENTS

1. INTRODUCTION 1
1.1 Problem definition 1
1.2 Research question 2
1.3 Research approach 2

2. BACKGROUND 5
2.1 Resilience 5
2.2 Measuring resilience 5
2.2.1 Fear potentiated startle reflex 6
2.2.2 Emotional Stroop Task 7
2.3 Resilience training paradigms 8
2.3.1 Stress Inoculation Training 8
2.3.2 Cognitive reappraisal 9
2.4 Requirements 11

3. DESIGN AND DEVELOPMENT OF THE TRAINING TOOL: LABELING VIDEO WITH COGNITIVE REAPPRAISAL THEMES 15
3.1 Design 15
3.1.1 Requirements addressed by core functions of all versions 16
3.1.2 Tool version A: full video annotation 18
3.1.3 Tool version B: label while watching video 20
3.1.4 Tool version C: label after watching video 21
3.2 Development 21

4. DESIGN OF MEASUREMENT INSTRUMENT 23
4.1 Pictorial Emotional Stroop (PEST) 23
4.2 Fear potentiated startle (FPS) blink magnitude 25

5. EVALUATION - VIDEO LABELING TOOL 29
5.1 Introduction 29
5.1.1 Research questions 29
5.2 Methods 29
5.2.1 Experiment design 29
5.2.2 Participants 29
5.2.3 Procedure & materials 30
5.2.4 Measurements 30
5.2.5 Data preparation & analysis 31
5.3 Results 31
Q1. Do participants generally find the tool and training useful? 31
Q2. How do participants perceive ease of use of the tool? 32
Q3. Which version of the tool did the participants find easiest to use? 34
Q4. How did the participants use the tools? (i.e. behavior) 35
5.4 Conclusions 37
5.4.1 Answers to research questions 37
5.4.2 Limitations 37
5.4.3 Future improvements 38

6. EVALUATION - MEASUREMENT INSTRUMENT 39
6.1 Introduction 39
6.1.1 Research questions 39
6.2 Method 39
6.2.1 Measurements 39

6.2.2 Participants 40
6.2.3 Design 41
6.2.4 Procedure and apparatus + materials 41
6.2.5 Data preparation & analysis 44
6.3 Results 45
6.3.1 Pictorial emotional stroop 45
6.3.2 Fear potentiated startle blink 45
6.4 Conclusions 47
6.4.1 Answers to research questions 47
6.4.2 Limitations 48

7. CONCLUSION 49
7.1 Contribution 50
7.2 Limitations 50
7.3 Future Work 50

REFERENCES 53

APPENDIX A VIDEO ANNOTATION TOOLS 57
A.1 Scenario for tool version A: full annotation 57
A.2 Tool version A: full annotation – UI sections explained 59
A.3 Tool version A: full annotation – configuration & log file 60
A.4 Tool version B: label while watching – results & log files 62

APPENDIX B MATLAB SCRIPT TO REDUCE RAW PORTILAB EMG DATA INTO Z-SCORED BLINK MAGNITUDES PER IMAGE STIMULUS TYPE 63

APPENDIX C MATLAB SCRIPT FOR PEST, PRODUCING FAST & SLOW EFFECT& T-TEST PER PARTICIPANT 70

APPENDIX D PEST FAST & SLOW EFFECT REACTION TIMES AND ERROR RATES 72

APPENDIX E FORMS USED IN THE EXPERIMENT 73
E.1: Consent form 73
E.2: Self report scales 74
E.2.1: Brief Resilience Scale (BRS) 74
E.2.2: The core self-evaluations scale (CSES) 74
E.2.2: The International Positive and Negative Affect Schedule Short Form (I-PANAS-SF) .75

APPENDIX F STILLS FROM VIDEOS SHOWN IN THE EXPERIMENT 76
F.1 Stills from negative videos 76
F.1.1 Murder scene from the feature film "American History X" 76
F.1.2 Clip from a documentary about the Rwandan genocide 77
F.1.3 Clip from a news report on the abuse and murder of a baby 78
F.2 Stills from neutral videos 79
F.2.1 News item on a new terminal for Heathrow airport 79
F.2.2 News item on the building of Beijing airport 80
F.2.3 Weather report 1 80
F.2.4 Weather report 2 81
F.3 Still from American History X murder scene, used in FCS+ condition of the FPS phase of the experiment 82

APPENDIX G EVALUATION FORM FOR VIDEO ANNOTATION TOOLS 83

1. INTRODUCTION

Certain professionals such as soldiers, police officers, firefighters, medical personnel and even train drivers are inherently at risk of experiencing work-related traumatic incidents, which may lead to life-debilitating disorders such as posttraumatic stress disorder (PTSD).

PTSD is an anxiety disorder that is known to develop in some individuals after facing extremely traumatic events. To be diagnosed with PTSD in accordance to the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013, p. 271), an individual must have been exposed to actual or threatened death, serious injury, or sexual violence. Other necessary criteria include that, associated with the traumatic event, there is the presence of intrusion symptoms, persistent avoidance of stimuli, negative alterations in cognitions and mood, and marked alterations in arousal and reactivity. These symptoms need to persist for more than one month, and need to cause clinically significant distress or impairment to the functioning of the individual.

Military professionals may face a host of potentially traumatic events: seeing or handling human remains, knowing someone being killed or seriously injured; being under fire, being unable to help wounded civilians because of the rules of engagement, seeing destroyed homes and villages; clearing and searching homes, caves or bunkers, patrolling in dangerous conditions (Bouchard, Baus, Bernier, & McCreary, 2010). Police officers may experience similar acute traumatic incidents, the following of which were rated as highest work stressors in studies by Violanti & Aron (1994) and Spielberger et al (1981): killing or seeing someone being killed in the line of duty, exposure to battered or dead children, and attending to disasters. Firefighters may be risking their own lives when entering burning buildings, and they may witness the suffering of others. Medical staff such as paramedics are frequently exposed to critical incidents, and they are often exposed to suffering and death of patients. Train drivers may experience railway suicides.

1.1 Problem definition

Lifetime prevalence of PTSD in the general population of the USA has been reported to be around 8% (Skogstad, 2013). Prevalence estimates for PTSD in US military populations reach up to 20% (Institute of Medicine, 2012), and similar numbers have been estimated for firefighters, police and ambulance personnel (Skogstad et al, 2013). However, it should be noted that the actual statistics might be much higher, since perceived stigma by peers and leaders and fear of negative career consequences may cause PTSD symptoms to be underreported (Hoge et al., 2005).

The cost of PTSD is huge, both in terms of human pain and economics. In the U.S.A. 8000 veterans commit suicide each year (Kemp & Bossarte, 2012). Comorbidity and substance abuse as an attempt at self-medication is common, and PTSD sufferers struggle with employment and relationships. Annual economic costs of anxiety disorders are difficult to quantify, but have been estimated as $43.2 billion in the U.S. (Greenberg et al, 1999) and €74.4 billion in Europe (Olesen et al, 2012).

PTSD can be effectively treated using e.g. cognitive behavioral therapy (CBT) strategies (Wright, 2006) such as exposure therapy, but obviously it would be preferable if disorders could be prevented by training individuals to become more psychologically resilient, before potentially exposing them to traumatic events.

Research in the area of resilience training is in still an early stage. Typically it involves elements of stress inoculation training (SIT), which is based on the idea that arousal levels of individuals in response to powerful stressors may be reduced by "inoculating" them to potentially traumatizing stressors.

Supporting resilience training with technology may be advantageous for a number of reasons:

- Stressors may be presented in a very controlled manner,
- Cost and logistics factors: using e.g. virtual reality may eliminate the need for hiring real actors and arranging a training zone with actual tanks and explosions,
- Measuring effectiveness with some kind of technology may be more objective than the common practice of using self reporting questionnaires.

1.2 Research question

The following research question was elicited from the problem definition for this thesis:

> *What would be a possible technology supported resilience training environment with high user acceptance, and how could its effectiveness be measured?*

Split up into sub-questions:

1. *What would a possible technology supported resilience training system look like?*
2. *What is the user acceptance of the system?*
3. *How can its effectiveness be measured?*

The "Holy Grail" would be a training that proves to be effective in longitudinal studies for "inoculating" individuals against psychological disorders after encountering traumatic events. Given feasibility considerations and resource constraints for this thesis, the decision was made to stage it as a pilot experiment, with the goal to set in place a platform that may be used for a larger experiment at a later phase.

1.3 Research approach

First a literature study was done into the relevant psychological background, methods to measure resilience, and existing technology-supported resilience training systems (Favié, Vakili, Brinkman, Morina & Neericx, 2016; Favié, 2014). This revealed the potential for a novel approach using cognitive reappraisal (chapter 2). For the design of the training tool and measurement instrument, specific project requirements were created by addressing existing general guidelines and requirements for the development of technology supported resilience training in a military context. A scenario was developed for using a video annotation tool as part of a resilience training for military professionals. All requirements were incorporated into the scenario, and core functions were distilled from this. A tool was built based on these requirements and core functions (chapter 3). This training tool using video annotation and cognitive reappraisal was then tested and evaluated in an experiment involving Dutch military officers in training (chapter 5). Furthermore, two novel approaches to measure resilience using technology were implemented and tested in an integrated experiment, using the emotional pictorial Stroop paradigm, and the fear potentiated startle reflex (chapter 4 & 6). A description of the contributions, limitations and future work concludes this thesis (chapter 7).

2. BACKGROUND

This chapter provides some background and describes how the requirements for a possible technology supported resilience training system were gathered. The goal was to find a promising novel approach that could be applied within the constraints of the thesis project. To this end, first a literature study was done into relevant psychological background, methods to measure resilience, and existing technology-supported resilience training systems (Favié, 2014). This revealed some "gaps" that were considered worth exploring, and that could be used in a novel approach specifically catered to military professionals: for a training tool this was cognitive reappraisal, and for a measurement instrument the pictorial emotional Stroop test and fear potentiated startle reflex paradigms were identified. Using an existing set of general requirements and guidelines for the development of resilience training technology in a military context, a list of specific requirements was distilled for a possible training tool and measurement instrument applying the identified "gaps".

2.1 Resilience

When developing a resilience training tool and measurement instrument there obviously needs to be clarity about what resilience itself is exactly. Unfortunately the empirical construct of resilience is not yet very well defined, consistent and scientifically grounded, but at the most basic level it has been described as the ability to adapt and cope effectively despite threatening or challenging situations (Agaibi & Wilson, 2005; Wilson & Drozdek, 2004). Whether someone is resilient may therefore be inferred post hoc: after being exposed to traumatic events that could likely cause PTSD, is he or she still psychologically healthy?

Studying such notably resilient individuals for commonalities is one means through which researchers are attempting to get a better grasp on the construct of resilience. If such commonalities can be successfully identified, it may be possible for less stress resistant individuals to adopt certain resilience-enhancing attributes by e.g. following a resilience training.

Despite the lack of a clear and consistent definition, the presence of protective factors that help individuals resist life stress is a commonly used indicator of resilience (Kaplan, Turner, Norman, & Stillson, 1996). For instance, Feder, Nestler, Westphal, & Charney (2010) name a range of psychosocial factors that promote successful adaptation to stress, including active coping strategies & facing fears, positive emotionality, cognitive reappraisal, the presence of social support, a sense of purpose in life and spirituality. Adaptive coping, personal control, hardiness and social support have been identified by Simmons and Yoder (2013) as psychological attributes that are evidently related to the literature about resilience in a military context.

2.2 Measuring resilience

Since the aim was not only to develop a training environment, but also to measure its effectiveness, knowledge of resilience measurement instruments is required. As the concept of resilience itself is not that well defined and scientifically grounded yet, it should be no surprise that there is no one standard way of measuring it either. Instead of trying to measure resilience directly, many researchers opt instead to evaluate it indirectly as a function of other constructs (e.g. hardiness), clinical outcomes (e.g., depression, PTSD), or behavioral measures.

Self-report instruments (i.e. questionnaires) are widely used in studies related to resilience training. Despite their common use, there are known disadvantages to self-reporting due to the way that subjects generally behave, such as various kinds of biases. There may also be incentives for participants to not self-report truthfully, and especially within a military context there is a perceived stigma by peers and leaders and fear of negative career consequences for reporting PTSD symptoms (Kim, Britt, Klocko et al., 2011).

Since there is a general sense of professionals in this area of research that the currently used self-report questionnaires are not reliable enough (Rizzo et al., 2011), and obviously effectiveness of a training cannot be validated without a reliable measurement, there is a clear demand for a novel reliable resilience measurement instrument. There is some research into biomarkers (objectively measured indicators of a biological state) related to resilience, but this is still in an early stage. Some biomarkers have been linked to related constructs such as stress or arousal, e.g.: salivary cortisol, skin conductivity, heart rate and heart rate variability (HRV), but as of yet no validated resilience measurement instrument using biomarkers exists.

While studying the literature, two concepts were identified as promising yet underused within the context of measuring resilience with technological support, and therefore deemed interesting for further investigation: the fear potentiated startle (FPS) reflex and the emotional Stroop task.

2.2.1 Fear potentiated startle reflex

Since FPS is widely used in experimental research on conditioned fear, it seems likely to be useful in (indirectly) measuring resilience. And as the author is not aware of research where this potential is investigated, this may be a “gap” worth exploring.

The startle reflex (SR) is a natural physiological reaction to sudden, strong stimuli such as a bright flash of light or a loud sound. It involves muscle responses such as jumping up, contraction of the neck muscles and blinking of the eyes. The function of the SR is believed to be protection of the eye (eye blink) or back of the neck (whole-body startle) and to help escape from sudden stimuli.

Emotional states are known to influence the SR. When a startle probe (e.g. a loud noise) is presented together with an unpleasant stimulus such as a picture of a mutilated body, a scary sound or foul smell, individuals without psychological disorders respond with a larger SR when compared to neutral stimuli (e.g. a picture of a table), while the reflex is smaller when presented with pleasant stimuli such as a picture of a smiling child (Vaidyanathan, Patrick, & Cuthbert, 2009; Lang, Bradley, & Cuthbert, 1990; Vrana, Spence, & Lang, 1988).

When fear is elicited before presenting the startle probe, the amplitude of the reflex is also larger. This phenomenon is called Fear Potentiated Startle (FPS) (Brown, Kalish, & Farber, 1951). The FPS response can be elicited using a threatening stimulus (i.e. some object or situation that would cause someone to experience fear), but it can also be elicited by a neutral stimulus as a result of fear conditioning. Individuals can be conditioned for this purpose by e.g. repeatedly showing a colored object on a screen (CS, conditioned stimulus) followed by an unpleasant stimulus such as a small electric shock or air-blast to the larynx (US, unconditioned stimulus). If the startle probe that triggers the startle response is now presented with or shortly after the colored object on the screen (CS), the startle response will be significantly larger in amplitude (e.g. the eye will blink more). Furthermore, research has shown that the FPS response also takes place after seeing videos or images containing unpleasant or frightening scenes (Vrana et al., 1988).

Individuals suffering from PTSD are known to display an increased FPS response, and under stress their response becomes even larger (Lissek et al., 2005). Moreover, it does not seem to matter whether the stimuli they are presented with are threatening or neutral – their response will be similar either way. This seems to indicate that unlike "normal" people, they can not differentiate a genuinely threatening stimulus from a harmless one (Pole, Neylan, Best, Orr, & Marmar, 2003).

Eye blink is the most common measure of FPS response. This is recorded using a technology called electromyographic recording (EMG), which uses electrodes to measure movement of the eyelid muscles: two directly below the eye, about 2 centimeters apart, and one to ground the signal, e.g. on the forehead.

2.2.2 Emotional Stroop Task

This section describes why the emotional Stroop task may be another candidate for a novel (indirect) resilience measurement approach. The original Stroop effect demonstrates the presence of interference in the reaction time of individuals when naming the color of a word printed in a font other than its name (Stroop, 1935). For instance, the word "blue" printed in red will bring about delayed color-naming. The emotional Stroop paradigm is an adaptation of the original Stroop paradigm. Individuals are still asked to name the colors of words, but here the point of interest is interference caused by words that carry an emotional charge for the participant. For example, studies have linked drug-related words to a delayed response time by substance addicts, and depressing words have been shown to have a similar effect on individuals suffering from depression. Since "a picture is worth a thousand words", some researchers opt to use images instead of lexical representations in a Pictorial Emotional Stroop Task (PEST) (Constantine, McNally, & Hornig, 2001; Lavy & Van den Hout, 1993). In this adaptation of the emotional Stroop task, a color is added to images either by color-filtering them into a single tint, or by surrounding them with a colored border.

The tendency of participants to be slower at color-naming a word or image when it has an emotional charge for them has been interpreted as selective allocation of attentional processes (Williams, Mathews, MacLeod, 1996). A stimulus that a phobic individual associates with fear is automatically and subconsciously processed, and activates pre-existing associations (Foa & Kozak, 1986). This takes time and competes with other cognitive processes such as recognizing the color, which will thus be performed slower as compared to a condition without stimuli perceived as threatening by the participant.

For the emotional Stroop effect the influence of emotional stimuli was long considered to be relevant on only the *current* trial, but some more recent findings show an influence of the valence of *previous* events (Frings, Englert, Wentura et al., 2010). McKenna and Sharma

(2004) in fact showed that the emotional Stroop effect hinges mostly on the valence of the *previous* trial. Therefore some researchers have opted to decompose the emotional Stroop effect into a so-called “fast” and “slow” effect. The fast effect is commonly interpreted as reflecting fast and automatic allocation of attention in the *current* trial, while it is assumed the slow effect results from a general slowdown after the processing of negative stimuli in the *previous* trial. The fast effect is then measured by considering only trials following neutral stimuli, while for the slow effect the after-effects of negative versus neutral stimuli in trial *n*-1 are considered only on the responses to neutral stimuli in trial *n*.

Research by e.g. MacLeod & Hagan (1992) using the emotional Stroop task suggests that adjustment to stress can be predicted by preexisting differences in attention control. Moreover, a meta-analysis of 26 studies by Cisler et al. (2011) concludes that EST performance of PTSD groups was affected by PTSD-relevant words and generally threatening words compared to neutral words, and only PTSD-relevant words impaired the performance of trauma exposed control groups. Non-trauma exposed control groups did not display these effects.

The EST has not been used as a direct measure of resilience, but one can imagine its potential as a novel application for indirect measurement of resilience. Consider for instance a study where participants receive a resilience training, then are exposed to traumatic videos, and finally do a PEST using stills taken from the traumatic videos. We might then perhaps use the response time as an indirect measure of resilience and the effectiveness of the training.

2.3 Resilience training paradigms

As part of the literature study (Favié, 2014) existing resilience training paradigms were researched, with extra attention to those used in current technology supported resilience training studies. Although resilience training per se turned out to be still an active area of research, the related concept of training individuals to be more stress-resistant is not new. Most current resilience training systems use some kind of cognitive training and incorporate breathing and relaxation techniques.

This section starts with a brief description of the commonly used paradigm called Stress Inoculation Training (SIT), and is followed by a description of the cognitive re-appraisal paradigm, which was identified as a promising “gap” in the area of resilience training while researching the literature.

2.3.1 Stress Inoculation Training

SIT aims to reduce and prevent stress with a tailored form of cognitive behavioral therapy (CBT) (Meichenbaum & Deffenbacher, 1988). It stems from models that posit that stress arises when an individual perceives his or her ability to cope with the (perceived) demands of a situation as insufficient.

Inoculation in the medical sense of course induces immunity to a disease by deliberate introduction to a small dose of a pathogenic agent. In a similar vein, “inoculation” in the context of SIT refers to exposing individuals to minor stressors in a controlled setting, where the idea is that learning to manage these minor stressors will help them handle major stressors with reduced arousal levels and boost their confidence in their coping skills.

Meichenbaum discusses three stages to SIT:

> *1. Psychoeducation (learning about stress responses and the need to control them),*
> *2. Training (learning a skill, such as arousal control, in order to tone down the harmful effects of stress)*
> *3. Implementation (utilizing these skills in the stressful context).*

SIT is used in technology supported stress resilience training studies by e.g. Hourani, Kizakevich, & Hubal (2011), Bouchard, Bernier, Boivin, Morin, & Robillard (2012), and Cohn, Weltman, Ratwani, Chartrand, & McCraty (2010).

2.3.2 COGNITIVE REAPPRAISAL

Cognitive reappraisal (CR) is an emotion regulation strategy that involves changing the way one thinks about an event to change its emotional impact. Despite the apparent effectiveness as a coping mechanism, the author is not aware of much research into resilience training focused on CR, and therefore it was considered worth investigating.

The core tenet of CR stems from appraisal theories of emotion that posit it is not an event itself, but rather the subjective appraisal (i.e. the meaning and significance an individual attaches to the event) that determines its affective impact (Lazarus & Folkman, 1984). In his theory of psychological stress and coping Lazarus indentifies two processes, cognitive appraisal and coping, as critical mediators of stressful person-environment relations and their outcomes. During cognitive appraisal a person assesses whether an event is relevant to his or her well-being, and if so, to what extent and in what ways. Lazarus makes a distinction between primary and secondary appraisal. In primary appraisal the person evaluates the significance of the event to him or herself (for instance: is there potential harm or benefit regarding commitments, values, goals or self-esteem?). In secondary appraisal the person evaluates what (if anything) can be done to overcome or prevent harm caused by the event: various coping options are evaluated such as changing the situation, acceptance, looking for more information, or suppressing counterproductive impulsive reactions. Lazarus divided the coping mechanisms up into direct actions addressing the problem and cognitive reappraisal processes that aim to dampen the emotional state itself. It is the latter processes that are addressed in this section as offering potential for an effective emotion regulation technique.

(Re)appraising continues beyond the initial event itself. In fact the first event-driven appraisal is thought to be at the beginning of iterative cycles of reappraising that will have emotional, cognitive and practical implications (Gross & Thompson, 2007). These (re)appraisal cycles seem to impact not only our feelings, but also other kinds of cognitive processing such as how well and how often we remember events (Gross, 2003; Schartau, Dalgleish & Dunn, 2009).

The appraisal theories posit that when faced with similar stressful life events, people respond quite differently largely due to the way they appraise them. Therefore a fundamental aspect of many psychological interventions such as CBT is learning to change ones own appraisal process (Samoilov & Goldfield, 2000). In this context, CR can be described as a technique that aims to achieve this through cognitive-linguistic emotion regulation (Gross, 1998). The main idea of CR as an intervention is to implement cognitive strategies to help an individual change the intensity of his or her reaction to an event (Ochsner & Gross, 2005). Participants may be given reappraisal tasks, for which they are instructed to consciously increase or decrease their negative affect in response to aversive stimuli (Ochsner & Gross, 2004; Eippert et al., 2006). For example, participants may be asked to systematically apply certain appraisal themes to reduce negative emotions.

Many of the dysfunctional appraisal themes mentioned in the CBT literature reflect a difficulty to see the bigger picture (Schartau, Dalgleish, & Dunn, 2009). Therefore, negative affect experienced from an event may likely be reduced if participants adopt a broader perspective that integrates positive or adaptive information. For this purpose Schartau et al. selected four functional appraisal themes directed at perspective broadening:

> *"1. Bad things happen — bad things happen in the world, and I need to put them behind me and move on.*
> *2. Silver lining — there are usually some good aspects to every situation, and it is important to focus on these.*
> *3. Broader perspective — bad events are rare overall, and lots of good things are happening all of the time.*
> *4. Time heals — in the (near) future this will not seem anywhere near as bad as it does now."*

In their study, participants were asked to employ one or more of these themes to reduce their emotional response to videos containing aversive events. Participants were presented with examples and were allowed to practice the task on a distressing practice video, before proceeding to the actual test. Participants who practiced appraisal reported less negative emotional responses to a final distressing video compared with controls. Moreover, those who practiced appraisal reported less intrusion and avoidance of the target memories in the week after the study relative to controls.

Several other studies (Gross, 1998; Dandoy & Goldstein, 1990; Lazarus, Opton, Nomikos, & Rankin, 1965) found that participants instructed to use reappraisal reported less negative affect after watching a distressing video compared to those in the "just watch" condition. Furthermore, Gross & John (2003) found that individuals who often use reappraisal experience less negative emotions in emotionally charged events and exhibit positive outcomes over time. CR is also named as one of the psychosocial factors that promote successful adaptation to stress by Feder et al. (2010). Furthermore, Denny & Ochser (2014) found evidence for the longitudinal trainability of reappraisal using short courses of reappraisal practice.

Besides using cognitive appraisal by itself, one could also imagine it to be integrated within an SIT program as one of the skills in the training phase. Existing research on CR is mostly outside the context of stress and depression (Troy et al., 2010), but one resilience training program developed by Rizzo et al. (2011) does include CR. As part of this program, participants are exposed to an emotionally challenging event in a virtual world, after which a virtual "mentor" guides him or her through rational restructuring exercises for appraising the event. The program has not yet been empirically validated.

2.4 Requirements

As mentioned in the previous sections, studying the literature revealed cognitive reappraisal as a paradigm with both novelty and promise for a technology supported resilience training tool. For a resilience measurement instrument, the emotional Stroop task and fear potentiated startle reflex were identified as “gaps” worth exploring: they seemed to show both promise and novelty, while also addressing the need for objective resilience measurement tools related to the stigma on honesty about emotions in the military, which is a known issue with regards to reliability of the current self-reporting instruments.

To come up with specific requirements for applying these paradigms to a training tool and measurement instrument within the context of this thesis, the approach was to make use of lessons learned by developers and researchers within this area of research. Vakili, Brinkman, Morina, & Neerincx (2015) describe a set of general requirements that resilience researchers involved in the development of technological solutions in a military context can use to structure their efforts (table 2.1), along with a set of guidelines to help mitigate risk (table 2.2). These were the result of a qualitative analysis of interviews with developers of technology assisted resilience programs in a military context regarding human-computer interaction and system development, and discussions with clinicians using technology-centered concept storyboards to gather feedback, and refine, validate and extend the initial concepts.

General guidelines		
1.	Beware constraints	Be conscious of project-limiting factors Culture, Effectiveness, Engineering and Resource
2.	Keep operational priority	Training should not negatively affect operational performance
3.	Capitalize on high-tech	Capitalize on the positive regard for high-tech approaches
4.	Get stakeholder buy-in	Ensure stakeholder buy-in and employ user-centered design
5.	Prepare for measurability challenge	Prepare for the scientific challenges of using stress and emotion measures
6.	Acknowledge stigma	Acknowledge the stigma toward psychological and emotional topics

Table 2.1 Guidelines for developers of technology-assisted mental resilience programs (Vakili et al, 2015).

General requirements		
1.	Change	Training should achieve cognitive, affective or behavioral change that enhances resilience
2.	Personalization	Training should be personalized to individual needs and differences
3.	Transferability	Training effects should be transferrable to operational domain and not interfere with operational performance
4.	Durability	Training effects should be durable
5.	Measurability	Training effects should be measurable
6.	Cultural relevance	Training should be relevant to its cultural context
7.	Economy	Training should be economical
8.	Safety	Training should be safe, and should be safe to use in the operational domain
9.	Engagement	Training should be engaging and motivating
10.	Life-cycle	Training should address the entire soldier life cycle

Table 2.2 Technology supported resilience training requirements gathered from interviews with experts (Vakili et al., 2015).

Addressing these general guidelines and requirements for the development of technology supported resilience training in a military context, specific requirements for a training tool and measurement instrument applying the identified "gaps" were summarized in table 2.3. This table provides rationales referencing relevant general requirements from table 2.2.

	PROJECT REQUIREMENT	GENERAL REQUIREMENT	RATIONALE
Training	1. Participants should label events with CR themes	*Change* *Durability* *Economy* *Safety* *Novelty*	CR appears to be an effective coping and emotion regulation strategy. CR appears to have longitudinal effects. Developing a tool for this purpose should be possible with moderate efforts. No risk for physical harm. Little risk for psychological harm if the contents are checked by expert. CR has not often been used in the context of this project.
	2. Content should be easily adaptable	*Cultural relevance* *Personalization* *Economy* *Engagement* *Life-cycle*	Easy customization facilitates picking events and themes that are relevant for specific target groups. Ease of adapting events and themes facilitates personalization. Ease of adapting content reduces costs for customization. When content is highly relevant, the participant will likely be more engaged and motivated. Events and themes can be chosen that are relevant for different phases of the soldier life cycle.
	3. Ease of use for participants	*Engagement*	Participants need to be immersed by the presented contents, while also performing the CR task, keeping the themes in mind at the same time. More overhead caused by a complicated interface may result in mental overload and/or interfere with the level of immersion and performing the CR task. Usability is also related to user satisfaction, which affects engagement.
	4. Ease of deployment	*Economy*	Simple, quick deployment will limit related costs for (customization of) specific equipment and spaces.
	5. Usable in group setting	*Cultural relevance* *Economy*	Troops are used to work in units, which allows for social support. Furthermore, deployment and training schedules are usually set for units rather than individuals in the military. Training groups of participants at once is efficient.
	6. Face validity	*Engagement* *Transferability*	When content is highly relevant, the participant will likely be more engaged and motivated. Face validity of training contents will likely increase transferability to operational domain.
Measurement	7. Measurements must be objective	*Measurability & stigma*	Current self-report measurement tools have known reliability issues related to stigma for honesty about emotions in the military.
	8. Must be grounded in and build on existing studies	*Measurability challenge*	Since measuring resilience is scientifically challenging and access to relevant experience is limited, building a totally novel measurement instrument from the ground up would be overly ambitious.

Table 2.3 Project requirements & rationales.

The project requirements are matched against the general requirements for technology supported resilience training in a military context in figure 2.1. It shows that the project requirements for the training tool address all general requirements, while this is not the case for the measurement tool. Since the primary use of a measurement tool in this pilot project was validation of the training tool using students rather than soldiers as participants it was not deemed necessary for the measurement tool to address all general requirements since the latter are specific to a military context.

	PROJECT REQUIREMENT / GENERAL REQUIREMENT	1. Change	2. Personalization	3. Transferability	4. Durability	5. Measurability	6. Cultural relevance	7. Economy	8. Safety	9. Engagement	10. Life-cycle
Training	1. Label events with CR themes	■			■			■	■		
	2. Easily adaptable events & themes		■				■	■		■	■
	3. Ease of use for participants									■	
	4. Ease of deployment							■			
	5. Usable in group setting						■	■		■	
	6. Face validity			■			■			■	
Measurement	7. Objective measurements					■	■				
	8. Build on existing studies					■		■			

Figure 2.1 Project requirements vs. general requirements for resilience training technology.

3. DESIGN AND DEVELOPMENT OF THE TRAINING TOOL: LABELING VIDEO WITH COGNITIVE REAPPRAISAL THEMES

As mentioned in section 2.3.2, cognitive reappraisal was identified as a promising paradigm for a technology supported training tool. A summary of a study by Schartau et al. (2009) was provided, which used functional CR themes aimed at perspective broadening, asking participants to employ these themes to reduce their emotional response to videos containing aversive events.

Results of this study were promising, and this sparked the idea to implement a similar approach using technology in the form of a video annotation tool for applying CR themes to video material. It seemed feasible to develop such a tool and use it in a pilot experiment within the constraints of the thesis. Requirements for such a tool were specified in figure 2.3. This chapter describes the design and development of the training tool.

3.1 Design

A scenario was developed for using a video annotation tool as part of a resilience training for military professionals (appendix A.1). All requirements were incorporated into the scenario, and core functions were then distilled from the scenario (figure 3.1). For instance, the following part of the scenario describes the use of video material as easily adaptable content (requirement 2):

> *"The subject of the video is the genocide that took place in Rwanda in 1994 and the role of the UN peacekeeping troops. This video was picked beforehand by the trainer, as he found it a relevant scenario for military service members. It was relatively easy for him to find a documentary about this subject online, edit the video material to contain on the most relevant segment, and use it in the tool by just pasting the video file name and folder into an XML configuration file."*

Three core (sub)functions were distilled from this (listed next to requirement 2 in figure 3.1):

1. *Custom videos and themes can be used in the tool through easily adaptable XML configuration files.*
2. *Tool uses video content, which is easier and cheaper to find/produce than e.g. virtual worlds.*
3. *Tool provides support for most common video codecs through Windows Media Player API.*

It also addresses requirement 6 (face validity), with a corresponding core function:

> *"Trainers can pick relevant videos for the participants, who will then recognize that stressful situations in the video are relevant for them personally, which will make the exercise feel more like an actual resilience training for them than some abstract cognitive task."*

	REQUIREMENT	COVERAGE / CORE FUNCTIONS
1	Participants should label events with CR themes	1. Tool facilitates tagging and/or annotating videos containing aversive events with CR themes.
2	Content should be easily adaptable	1. Custom videos and themes can be used in the tool through easily adaptable XML configuration files. 2. Tool uses video content, which is easier and cheaper to find/produce than e.g. virtual worlds. 3. Tool provides support for most common video codecs through Windows Media Player API.
3	Ease of use for participants	1. Tool facilitates task with as few distractions as possible. 2. Design is simple, minimalist, and intuitive. 3. Usability heuristics are applied: o Consistency with similar applications such as video players, Youtube, etc. o Avoiding irrelevant information & elements, hiding elements until they are needed.
4	Ease of deployment	1. Application is simple and light, uses only standard Windows libraries. 2. Does not require admin privileges, can be run from USB stick
5	Usable in group setting	1. Tool will be used individually, but introduction and practice should be done in a group setting facilitated by an instructor. 2. Distraction from other participants during individual use should be limited by providing headphones.
6	Face validity	1. Trainers can pick relevant videos for the participants, who will then recognize that stressful situations in the video are relevant for them personally, which will make the exercise feel more like an actual resilience training for them than some abstract cognitive task.

Figure 3.1 Satisfying the project requirements

Three different applications were designed, which all involve applying predefined functional CR themes to video material, but each with a different approach:

A. Full video annotation (described in the scenario, appendix A.1)
B. Label while watching video
C. Label after watching video

3.1.1 REQUIREMENTS ADDRESSED BY CORE FUNCTIONS OF ALL VERSIONS

Ease of use for participants

The participants will be performing a dual task: watching a video containing aversive events while applying CR themes at the same time. The participants probably do not have any experience with the latter technique yet, and the dual task will likely place considerable cognitive demands on the participants, therefore any cognitive overhead from the application user interface should be minimal.

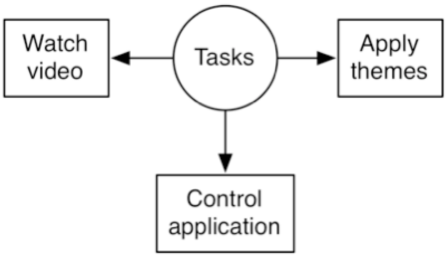


Figure 3.1 Tasks performed using the training tool

For that reason, and in accordance with requirement 3 – "ease of use for participants", the applications were designed to be simple and easy to use, the specifics of which are discussed in the respective sections for the individual versions of the tool. In general the idea was to make the applications as self-explicatory and intuitive as possible, making use of usability heuristics (Nielsen, 1994) and "Minimal Manual" guidelines (Carroll, 1987), avoiding irrelevant or rarely needed information and user interface elements, or hiding them until they become relevant for the user. Furthermore, the applications are all rendered full screen, to avoid distractions and to ensure that the participant will only be using the training tool on the computer.

Ease of adapting content

Addressing requirement 2 – "easily adaptable events & themes", the tools use video material as main content, which is much easier and cheaper to find or produce and edit than e.g. virtual worlds. The latter are used in e.g. the STRIVE system (Rizzo, John, Newman, et. al., 2013), but require a complete design studio for producing and editing the virtual reality content, which obviously involves substantial cost and complexity. Furthermore, all versions of the tool allow for the videos and themes to be easily adapted through simple configuration files (appendix A.3).

Face validity

Allowing trainers to easily pick relevant videos for the participants will likely improve acceptance of the tool: when participants recognize stressful situations in the video that are relevant for them personally while exercising their profession, this will feel more like an actual resilience training for them than some abstract cognitive task for example, thus addressing requirement 6 - "face validity". Furthermore, actual video footage from relevant events that took place in reality (e.g. from documentaries, news, field footage, Youtube etc.) are likely to be more realistic than e.g. scripted virtual reality material produced by a design studio with voice actors etc.

Logging input from participants

Input from participants is logged, so that researchers may later analyze this, for instance as part of evaluating acceptance of the tool.

3.1.2 TOOL VERSION A: FULL VIDEO ANNOTATION

The main idea here was to facilitate deep learning (Marton & Saljo, 1976; Entwistle, Hanley, & Hounsell, 1979) by actively engaging the participant in not only applying CR theme labels to events in context relevant video material, but also requiring the participant to enter a rationale for why an applied theme is relevant for a particular moment in the video. When a participant recognizes that an event in the video is suitable for applying a certain CR theme, a "theme label" can be attached to the corresponding playback location.

The interface was designed in such a way as to prevent cognitive overload and distractions during the dual task process of both watching a video with aversive events and applying CR themes. First, it is consistent with existing relevant systems such as media players, film editors and Youtube, and the video itself is displayed as largely as possible. Secondly, the design is minimalist (providing only necessary, relevant elements and information) and interface elements are hidden until they are actually needed. For example, instead of constantly showing a list of all themes that can be applied, a single "+" button is displayed. Only after the participant has decided to add a theme by pressing this button, a list of themes is displayed along with a textbox for entering a rationale. Furthermore, functionality related to editing existing annotations is only displayed when the participant clicks on an actual annotation. Thirdly, video is paused automatically when the participant adds a CR theme to the timeline of the video, thereby momentarily directing full focus away from the video and onto the task of selecting a relevant CR theme and providing a rationale.

Annotations may be edited (i.e. the CR theme, its location on the timeline and/or rationale) or even deleted, and the video may be paused or rewound at will by the participant by changing the playback position on the timeline until he or she is satisfied that the entire video has been properly annotated with suitable CR themes, allowing for more reflection than by forcing the participant to make instant, final decisions.

The video, available themes (typically consisting of a few words) and markers (typically the first letter of a theme in a particular color) for annotations may be configured through an XML file (appendix A.3). This facilitates easily adaptable events & themes (requirement 2).

Once the participant is satisfied with all his or her annotations, the application may be exited by pressing the button in the lower right. This button is only enabled if at least one annotation has been entered, pops up a dialog to confirm the action (preventing accidental termination), and its location is consistent with a top to bottom, left to right workflow. At this point the user input from the session is written to a log file (appendix A.3) for later analysis by researchers.

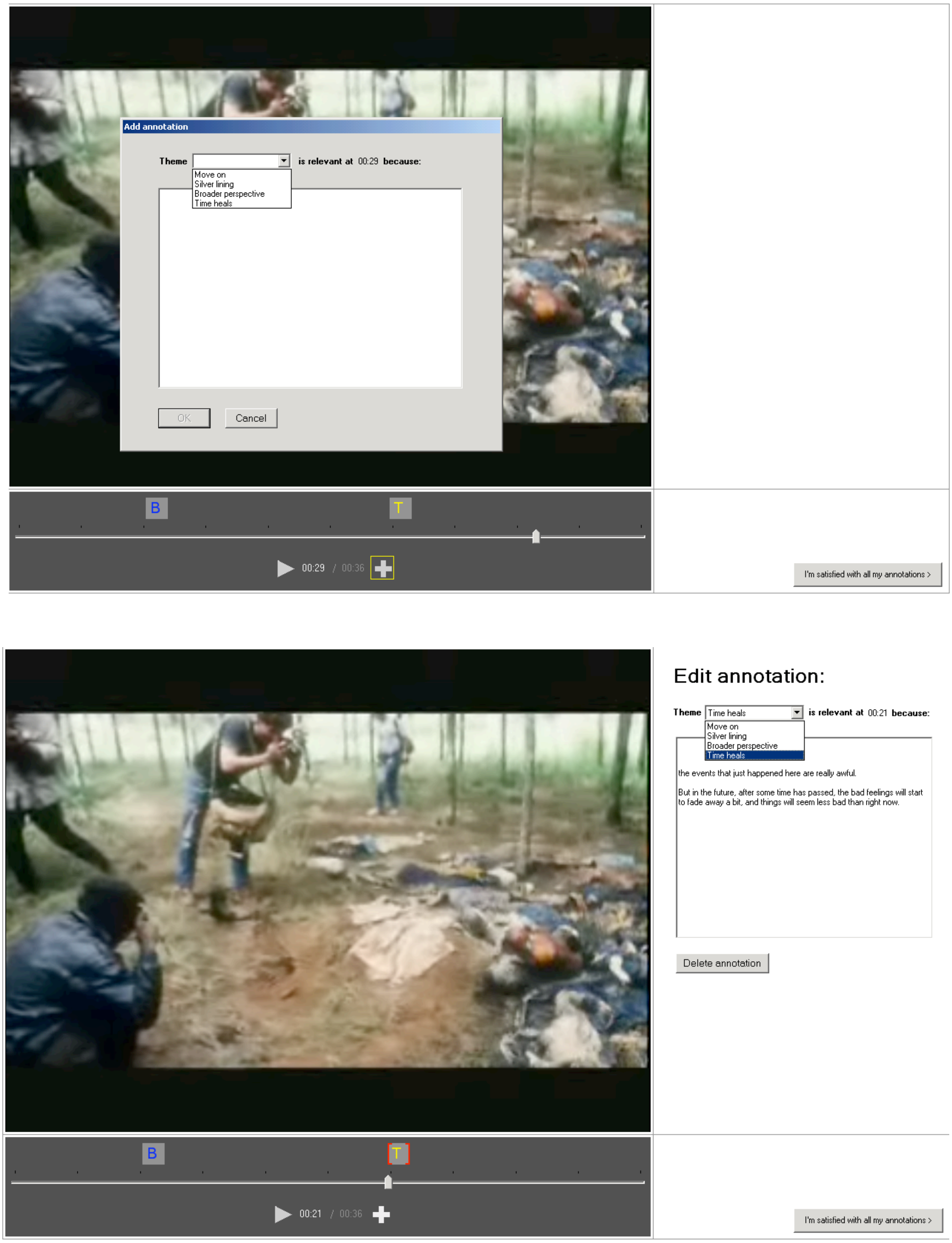


Figure 3.2 Screenshots of tool version A - full video annotation

3.1.3 TOOL VERSION B: LABEL WHILE WATCHING VIDEO

This version is much simpler than the full annotation tool. The idea is for the participant to practice applying CR themes "in the moment", by clicking a suitable CR theme instantly as an aversive event occurs in the video. In this version of the tool video cannot be paused or rewound, so there is less room for reflection than with the full annotation version. This "in the moment" approach is more congruent with arousal suppression skills often taught as part of SIT that aim to facilitate better task performance under stress (e.g. during a mission on the battlefield), while the full annotation tool may be more useful to teach functional reappraisal after the event has already passed (e.g. back in the barracks, still thinking about aversive events that happened during battle). So in a sense, this version of the tool is more about appraisal rather than re-appraisal.

One could also imagine a training where skills are acquired in different phases. In the first phase the full annotation tool could be used, which gives the participant more time to reflect and allows for correcting mistakes, thereby facilitating internalization of the application of CR themes through deep learning. Once the participant has acquired the skill of applying CR themes in the first phase, the second version of the tool could then be used to practice applying themes "in the moment".

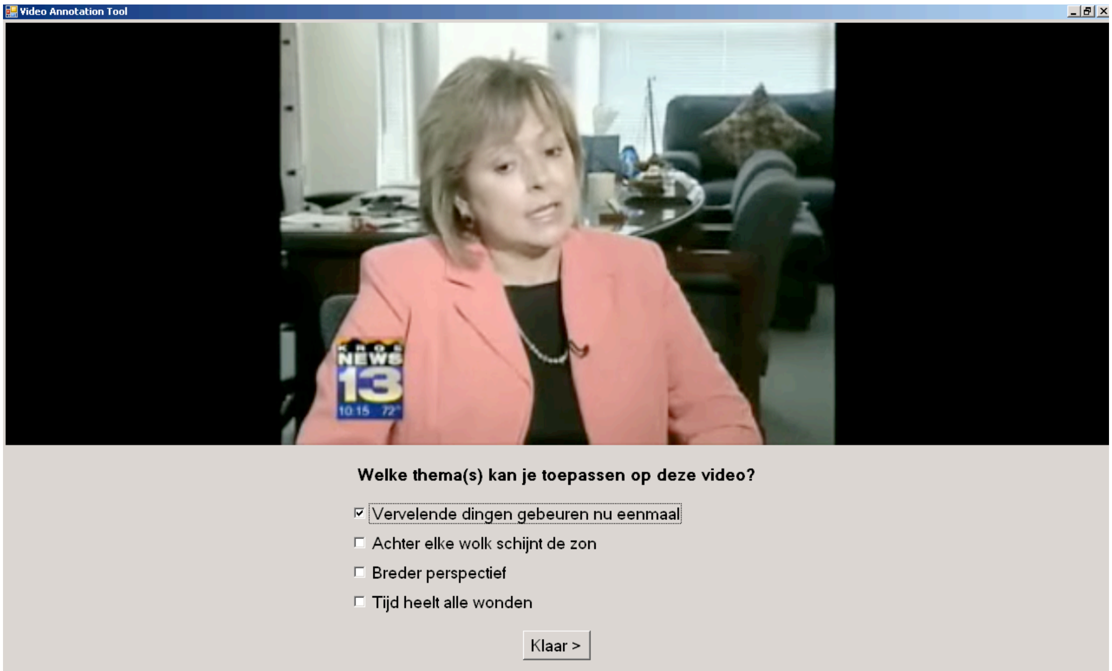


Figure 3.3 Screenshot of Tool version B - label while watching video

The screen is split into two sections: 1) video and 2) a short list of CR themes that can be "checked" by the participant when he or she feels they are relevant to the video.

In fact, the only interaction that is possible with this tool is the selection of one or more themes, and pressing the "Ready >" button, which is only enabled once the video has completed, and which will store the results and exit the tool.

3.1.4 TOOL VERSION C: LABEL AFTER WATCHING VIDEO

This is the simplest version of the tool: the video plays full screen, and when it has finished, the participant is asked to indicate which CR themes can be applied to the video.

The rationale for the design of this version was to stay as closely as possible to the study by Schartau et al. (2009), which would then provide a means to compare results.

3.2 Development

Since Windows is the ubiquitous software platform within the (Dutch) military, the decision was made to build an application for Windows 7. Once this platform was determined, the Microsoft Visual Studio integrated development environment was an obvious choice for building the application. It is known to be a mature development platform for Windows applications, and its tight integration with Windows through the .NET framework would likely facilitate easy deployment of the developed application onto a variety of Windows machines. Furthermore, video playback would likely be smooth through the use of the Windows Media Player API and codecs. These points are relevant for requirement 4 - “ease of deployment”.

4. DESIGN OF MEASUREMENT INSTRUMENT

Ideally an evaluation of a resilience training tool for military professionals would involve a longitudinal study with actual soldiers who are sent to war and followed for several years, but obviously there are practical and ethical issues with this, and it would be outside the scope of this project. To keep the evaluation feasible it was therefore decided to design an objective measurement instrument that builds on existing studies.

Two promising paradigms were identified for this purpose after studying the literature: the pictorial emotional Stroop task, and fear potentiated startle eye blink magnitude.

4.1 Pictorial Emotional Stroop (PEST)

The general principles of the PEST were covered in section 2.2.2. For building a PEST instrument, the relevant practicalities are that a sequence of color-filtered images is shown on a monitor to the participant, who needs to identify the color of each image as quickly as possible by pressing the button of the same color on a response input box. The control of timing and synchronization of stimulus delivery with recordings needs to be highly accurate for such experiments. This is typically done on fast computers running specialist software packages such as E-prime (Schneider, Eschman, & Zuccolotto, 2012), of which a license (version 2.0.10.242, Standard) was available for the project. E-prime is claimed to have built in timing systems to deliver high timing accuracy by coordinating with the operating system to have as much possible control over the computer's execution.

Time-critical responses from the participant, such as the button presses identifying the color of the image, are typically collected via an external button box, which is designed to provide higher accuracy and precision in timing than a standard PC keyboard. The serial response input box offered by Psychology Software Tools was deemed prohibitively expensive for this project, and no similar, compatible products appeared to be commercially available from other manufacturers at a substantially lower cost either. Literature research revealed an affordable and relatively straightforward DIY alternative: building a key input box that interfaces with the PC via the parallel port (Alrymple-Alford, 1992). The parallel port is a simple and affordable interface that provides 4 input bits with accurate timing. Although the parallel interface can now almost be considered legacy and has been largely phased out by USB, the latter is less suited for time-critical responses by default due to its higher complexity, and a USB based solution therefore requires extra features such as a built in hardware clock for time-stamping key presses and a more complex communication protocol. It seemed feasible to build the key input box for the parallel port within the constraints of the project, so it was constructed from off-the-shelf electronic parts (see figure 4.1) and code was written for it to operate as an input device within E-prime. The colors of the buttons were chosen accordingly to the colors used in the PEST, but if needed they can be changed for future projects by fitting other caps. The buttons were positioned ergonomically according to the natural differences in length of the fingers of the human hand.

Figure 4.1 Response box for parallel interface

Stimuli images were color filtered into red, blue, green and yellow by a Photoshop script written by the author, which first converts images into grey-scale and then adjusts the color balance accordingly (see figure 4.2).

Figure 4.2 Color filtered negative stimulus image

The experiment was programmed in E-Studio, the sub-program of E-Prime dedicated to creating experiments. E-Studio features a GUI for structuring the experiment by dragging and dropping components similar to Microsoft Visual Studio, but practical scenarios usually involve writing code in the E-Basic language. For this experiment code was written for e.g. pseudo-randomization of the stimuli images and interfacing with the parallel interface input box.

All responses along with their timing are outputted to a file that can be analyzed to some basic extent with the E-DataAid program, and can be exported to formats compatible with other data analysis programs such as SPSS.

4.2 Fear potentiated startle (FPS) blink magnitude

An FPS blink magnitude measurement tool was built consisting of several hardware components including an air blast device (figure 4.4) and a bioamplifier, which were integrated into a single experiment (figure 4.3) that was scripted on a computer running E-prime. Scientific background regarding the FPS reflex paradigm was already provided in section 2.2.1, but in brief the purpose of the tool was to condition the participant with fear to certain image stimuli linked with aversive airblasts to the larynx, and to record the amplitude of the eye blink triggered by a noise startle probe after viewing a stimulus image.

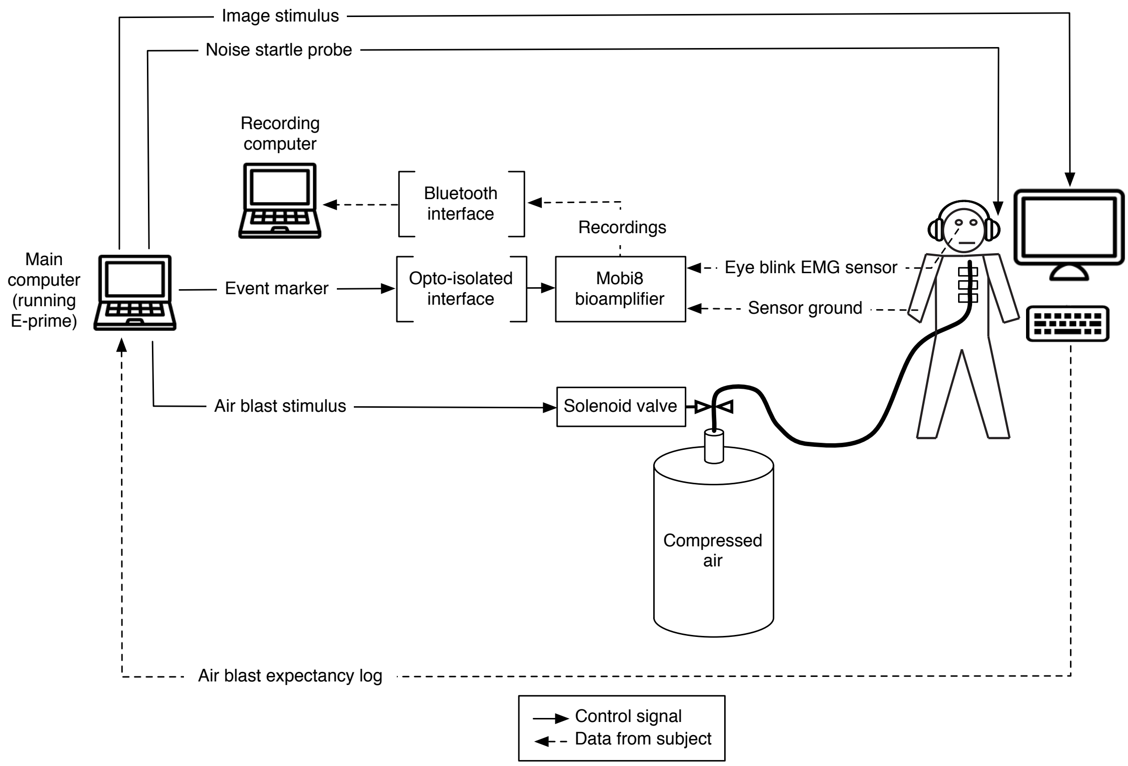


Figure 4.3 Fear potentiated startle blink measurement setup

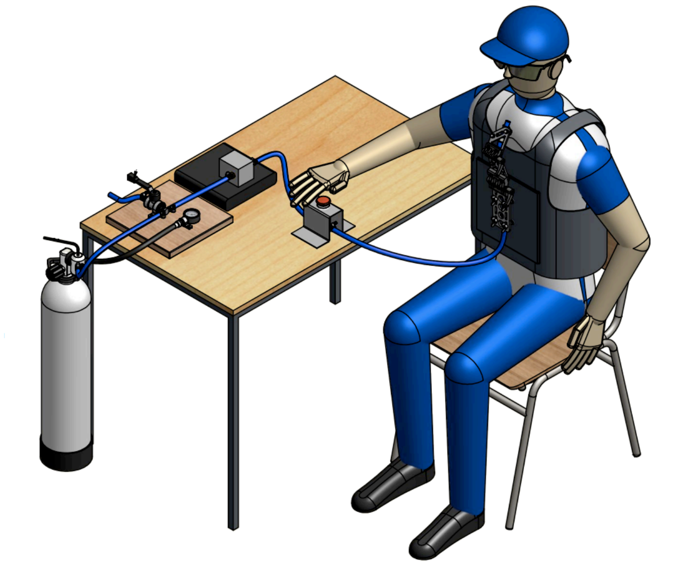


Figure 4.4 Air blast device consisting of solenoid valve with USB interface, compressed air bottle, tubing with manual pressure regulator, pressure meter and release valve running to the larynx of the participant via a vest, and a red emergency stop button.

The E-prime computer was scripted to display image stimuli to the participant on a 22-inch monitor, and to play white noise startle probe sounds from an M-audio Transit USB sound interface over a Superlux HD-661 headphone. The white noise sample was created with the noise generator of Audacity software (Version 1.3.7 for OSX; Audacity, 2015). The volume of the white noise played over this headphone was calibrated to 108 dB(A), which is loud enough to induce startle eye blink for most individuals but still within safety standards (Blumenthal, Cuthbert, Filion, et al., 2005).

Calibration of the level of the noise startle probe on the headphone was performed using a KEMAR head and ear simulator for acoustic measurements. This KEMAR dummy has microphones built into its artificial ears, the audio signal from which can be externally analyzed, thus enabling accurate acoustic measurements taken from headphones placed on the head. A calibrated setup consisting of a sound level meter directly plugged into the audio signal from the ear microphones of the KEMAR would have been ideal for the purpose of calibrating the sound level of the noise probe, but such a setup was not available. Therefore an alternative calibration procedure was performed with an available calibrated Bruel & Kjaer 2236 SPL sound level meter (which cannot be directly plugged into the audio signal from the in-ear microphones, but instead relies on its own built in microphone), a laptop with Audacity software for recording and analyzing the audio signal of the in-ear microphones, and an active monitor speaker. The white noise sample was played over the speaker from the USB sound interface, and the volume was adjusted with the level knob on the monitor speaker until the sound level meter, which was held directly next to the KEMAR ear, indicated 108 dB(A). At that point the output of the in-ear microphones was recorded on the laptop, and analyzed with a live sound level meter built into the Audacity software. The next phase consisted of matching this 108 dB(A) level of noise just played over the speaker with a similar level of noise played over the headphones. The headphones were now placed on the KEMAR, the noise probe was played through the headphones instead of the speaker, and the audio signal from the in-ear microphones was again analyzed on the laptop with the live sound level meter built into Audacity. The volume of the noise probe played over the headphones was adjusted with a Behringer MA400 headphone amplifier until the Audacity sound meter on the laptop indicated the same level as it was at the 108 dB(A) measurement taken from the speaker. At this point a marker was placed at the position of the level knob of the headphone amplifier, indicating the position for the calibrated 108 dB(A) noise probe for the headphones.

The air blast device can administer blasts of compressed air to the larynx by opening/closing a solenoid valve through commands sent from E-prime over the USB interface. Compressed air is stored in a 7 liter diving bottle at 200 bar, and a pressure valve with manually adjustable regulator reduces this to a maximum of 10 bar. When the solenoid valve is opened for delivering an air blast (typically for 250 ms), air from the bottle flows through a plastic hose to the nozzle with a diameter of 6 mm, which is typically placed about 10 cm from the larynx of the participant through an adjustable vest. When the regulator is set to the maximum 10 bar, the measured centerline velocity of the air flow at 10 cm from the nozzle is 156 m/s.

A separate computer running Portilab software (version 2.2.12; TMSi, 2015) was used to record the electromyographic (EMG) signal from two 2mm Ag/AgCl EMG electrodes connected on the right orbicularis muscle below the pupil of the participant, using a Mobi8 bioamplifier from Twente Medical Systems International (TMSI) linked to the recording computer via Bluetooth.

Besides the monitor, keyboard, and headphones, the E-prime experiment script interfaced with two other external devices: 1) a custom built solenoid valve connected via USB, used to trigger aversive blasts of compressed air to the larynx of the participant, and 2) an opto-isolated digital interface connected via the parallel port to the Mobi8 bioamplifier, used for adding event markers to the EMG recordings of the participant's eye blinks triggered by the noise probe regarding the onset of noise startle probe and the type of image stimulus that was displayed. These event markers were required for later analysis of the relevant blink recordings. Since only a single bit was available on the Mobi8 for adding such digital markers, a simple kind of pulse width encoding was scripted for this purpose. Furthermore, the participant was asked to log his or her expectancy of air blasts using the standard PC keyboard.

The raw EMG recordings from the Portilab software were later analyzed off-line with Matlab code written by the author (appendix B), in order to reduce the dataset to z-scored magnitudes per image stimulus. The script first reads the data from the proprietary Portilab format into Matlab using code provided by TMSi. Next, potential noise is removed from the EMG signal with a 4th order Butterworth bandpass (28-500 Hz) filter, after which the resulting signal is rectified and smoothed with a 4th order Butterworth lowpass filter at 40 Hz. The sections within in the recording containing possible blinks triggered by the noise probe are identified by decoding the digital pulse markers inserted by the E-prime script, which convey the timing of each noise probe and the type of image stimulus that was presented at that moment. Blink magnitude is then scored as the peak absolute amplitude in the period from 20 ms to 200 ms after the noise probe, minus the mean of the absolute amplitude during the baseline period from 50 ms prior to 20 ms after the noise probe. A blink is rejected and replaced by the mean of the previous and next blink if the signal during the baseline period is very noisy, or to be more exact if the variance of the signal during the baseline period is larger than the median of the variances of all baseline periods of the participant plus 3 times the median absolute deviation of the baseline periods. Finally, to compensate for individual differences in absolute blink magnitude between participants, the blink magnitudes are standardized into z scores: the mean of all the participant's blink magnitudes is subtracted from the current blink magnitude, and divided by the standard deviation of all the participant's blink magnitudes.

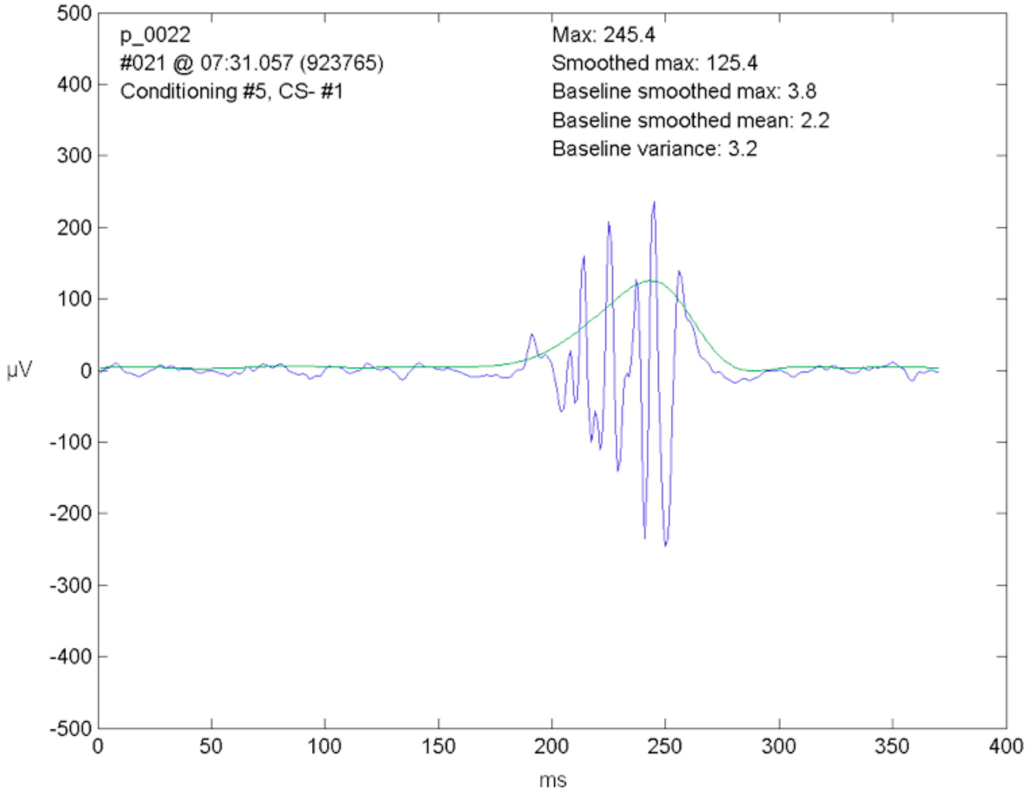


Figure 4.5 Raw (blue) and smoothed rectified (green) eye blink EMG, generated by Matlab script.

5. EVALUATION - VIDEO LABELING TOOL

5.1 Introduction

The three versions of the video labeling tool were evaluated as part of a resilience training for two groups of professionals: non-commissioned officers (NCOs) in training of the Dutch army and firefighters. Obviously the ideal outcome of such an evaluation would be an indication of the effectiveness in terms of enhancing resilience, but this was beyond the scope of this thesis study. The focus of this evaluation was therefore mainly to get an understanding of the level of acceptance and engagement of the users and how they perceived the tool in the context of an actual training program. In particular, this chapter will answer the following research questions:

5.1.1 RESEARCH QUESTIONS

Q1. Do participants generally find the tool and training useful?
Q2. How do participants perceive the ease of use of the tool?
Q3. Which version of the tool did the participants find easiest to use?
Q4. How did the participants use the tools? (i.e. behavior)

5.2 Methods

5.2.1 EXPERIMENT DESIGN

Options for experiment design were limited since the experiment had to be integrated into an existing military training schedule and all participants were required to follow the same training. Therefore a within subject design was used. The experiment was approved by the TU Delft ethics committee.

5.2.2 PARTICIPANTS

In total 76 participants used the video annotation tools as part of a training: 59 non-commissioned officers (NCOs) in training of the Dutch army and 17 firefighters. There was a significant difference between the age of NCOs ($M=19.70$, $SD=1.79$) and firefighters ($M=40.94$, $SD=9.33$) as revealed by an independent samples t-test, $t(16.35)= -9.33$, $p < .001$. Only one of the participants was female.

For the NCOs the reappraisal training was integrated into their regular training program as a separate module, while the firefighters were presented this training as a workshop. Participation was voluntary.

5.2.3 PROCEDURE & MATERIALS

The regular instructors of the NCOs were provided with a “train the trainers” session, so that they could deliver the training to the NCOs themselves, in order to prevent an anomaly in the regular military training by allowing civilians to act as instructors. The workshop for the firefighters was handled by a TU Delft instructor with a degree in psychology.

Before the video labeling tools were actually used, a short presentation was given to the participants explaining the concept of reappraisal and four different reappraisal themes, which were called “kijkmanieren” (which roughly translates to “ways of looking at situations”) in staying with the language of the target groups. This presentation also included a video that explained how one could apply the themes to an example adversary event.

The three versions of the tool (as described in section 3.2) were used in sequence, each version containing a different video. The order of the sequence was the same for all participants, again due to restrictions beyond our control: prior to using each individual version, there was a group video presentation demonstrating its use and briefly introducing the content of each video. This also required the tool versions to be linked to fixed videos, therefore there was no randomization of videos between tool versions.

The reappraisal themes were picked by domain experts from TNO and the military. Videos used in the session for the NCOs were approved by a psychologist from the military and TNO. Videos for the firefighters were approved by a social worker employed by the fire department. The videos for the NCOs were about hurricane Katrina in New Orleans, the Rwandan genocide and a prominent child abuse case in the Netherlands. The videos for the firemen were about a passenger flight that crashed during landing at Amsterdam, a fire during a soccer match in a stadium, and again the Rwandan genocide. All footage was from current affairs TV broadcasts.

5.2.4 MEASUREMENTS

The labeling tools recorded the input of participants, such as which themes they selected at what points in a video, and in case of the full annotation tool the texts they entered to explain why they found the selected theme relevant at the point in the video where they chose to make the annotation. These data were later analyzed as an indication of user engagement.

After using the tools, the participants were asked to complete a custom questionnaire (see appendix G) consisting mostly of 5-point Likert scale questions (1: strongly disagree, 2: disagree, 3: neutral, 4: agree, 5 strongly agree). The questions were related to the usefulness of the reappraisal themes, in how far the video labeling exercise helped the participants master the reappraisal themes, and if the participants felt the labeling exercises would help them see possible future stressful events in a different perspective. Furthermore, participants were asked to rate how disturbing they found each video that was used, how easy to use they found each version of the tool, and which version they felt was best to learn to apply the reappraisal themes.

5.2.5 DATA PREPARATION & ANALYSIS

Log-files generated by the tools were analyzed, in particular the quantity of themes used (for all versions) and the length of the textual inputs (for the full annotation version). This data was entered into SPSS, along with the responses from the questionnaires. Additionally, two overall index values were added: "overall ease of use" was created by summing the "ease of use" ratings of the three versions, and "overall disturbingness of videos" by summing the rated levels of the versions. Due to an error in two questions on the questionnaires handed out to one group of 30 NCOs, responses to the concerning questions were excluded from the analysis.

The Wilcoxon signed rank test was used to indicate whether opinions of participants were significantly above "neutral". The Friedman test was used to indicate if there was a significant difference between how the different versions were rated in terms of ease of use. A nonparametric Spearman rank-order correlation was run to determine if there was a correlation between the video content that was used in each version and its ease of use.

5.3 Results

Q1. DO PARTICIPANTS GENERALLY FIND THE TOOL AND TRAINING USEFUL?

Participants were asked to respond to three questions related to usefulness on a 5 point Likert scale (1: strongly disagree, 2: disagree, 3: neutral, 4: agree, 5 strongly agree). The results are displayed in figure 5.1. With regards to the first two questions, most participants (N=45, 60%, and N=45, 59% resp.) expressed agreement, but for the third question most responses (N=23, 30%) were neutral. This tendency of agreement with the first two questions but neutrality to the third is further supported by the corresponding medians and inter quartile ranges. Furthermore, a Wilcoxon signed-rank test revealed that the first two questions were answered significantly higher than neutral, but that this was not the case for the third question (table 5.1).

	Survey question	***Mdn (IQR)***	***Wilcoxon***	
			Z	***p***
1	I found the four "ways of looking at situations" useful.	4 (1)	-5.13	<.001
2	I found labeling the videos a useful exercise to master the four "ways of looking at situations".	4 (1)	-4.24	<.001
3	The video labeling exercises will help me see potential future disturbing situations in a different perspective.	3 (2)	-.45	.651

Table 5.1 Results of Wilcoxon signed rank test comparing responses with a constant of 3 (*neutral*), and median of responses to the questions related to usefulness.

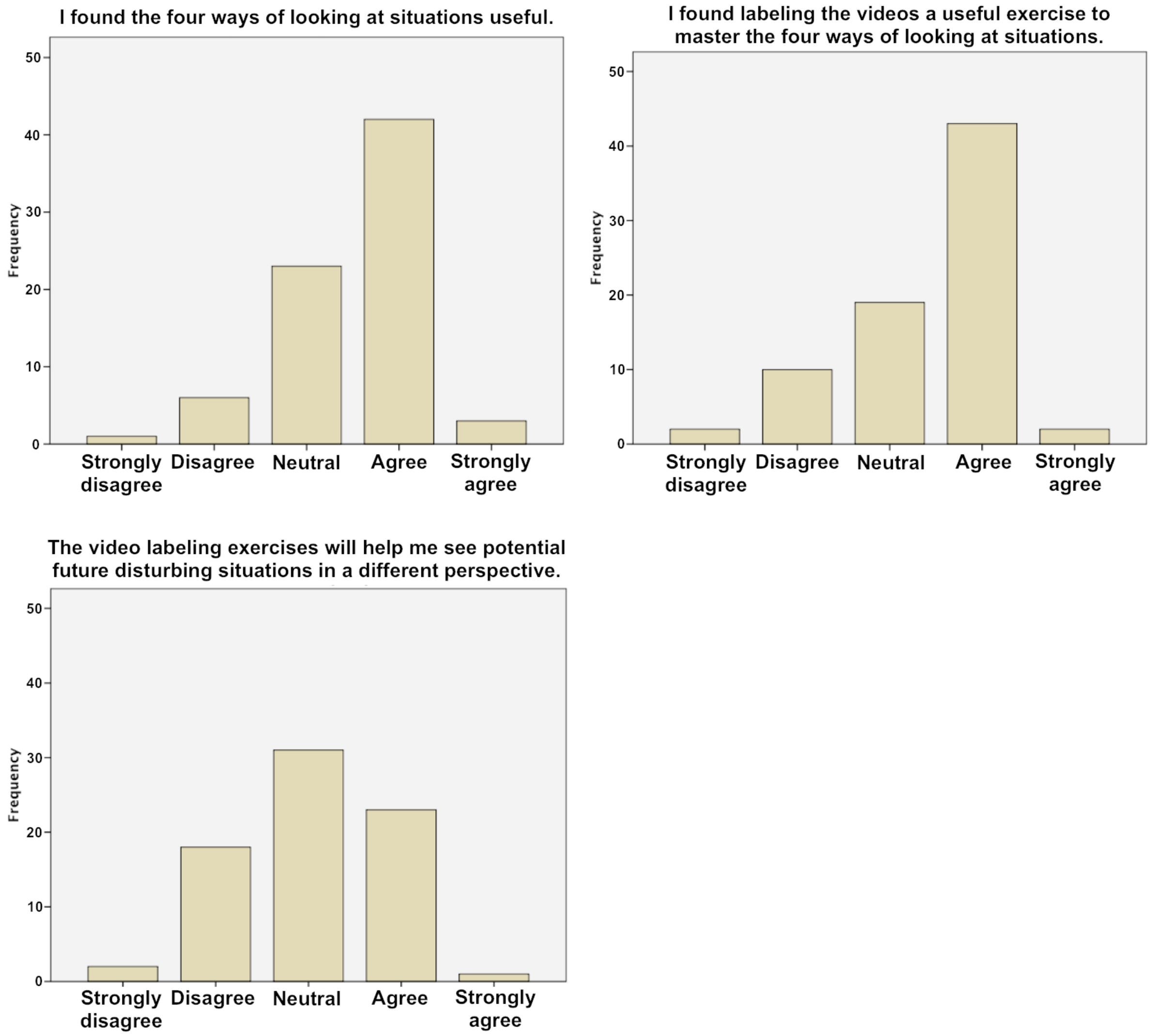


Figure 5.1 Frequencies of responses to questions related to usefulness.

Q2. How do participants perceive ease of use of the tool?

For all versions of the tool, most participants expressed agreement with the statement "I found this version easy to use". The median was 4 (*agree*), with a small inter quartile range, and a Wilcoxon signed rank test comparing the responses with a constant of 3 (*neutral*) confirmed that the responses were significantly higher than neutral (table 5.2).

			Wilcoxon	
	"I found this version of the tool easy to use"	***Mdn (IQR)***	***Z***	***p***
1	Full annotation	4 (0)	-6.85	<.001
2	Label afterwards	4 (0)	-5.34	<.001
3	Realtime labeling	4 (1)	-4.04	<.001

Table 5.2 Results of Wilcoxon signed rank test comparing responses with a constant of 3 (*neutral*), and median of responses to the survey questions related to ease of use of the different versions of the tool.

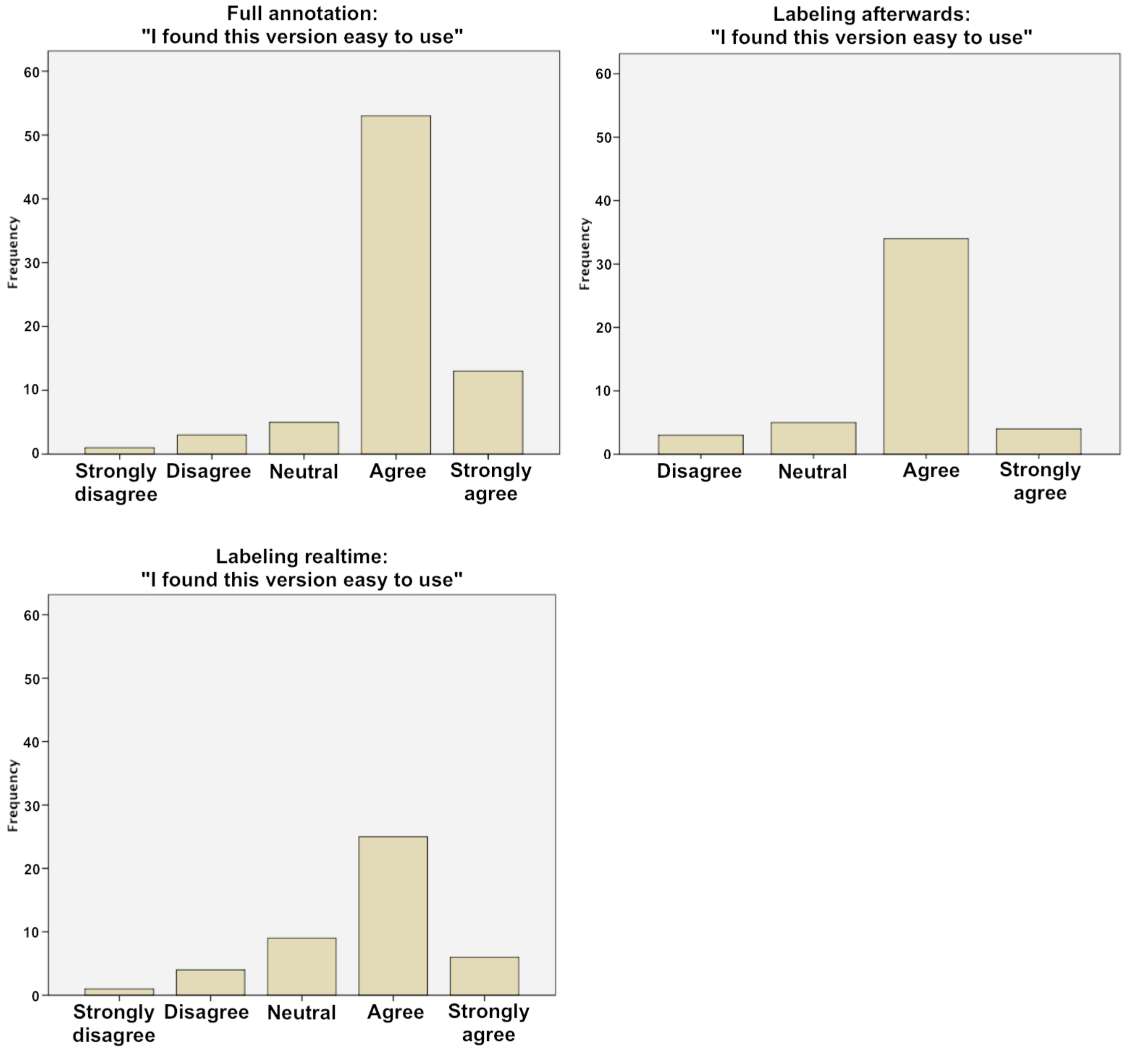


Figure 5.2 Frequencies of responses to questions related to ease of use of the versions of the tools.

A Spearman's rank-order correlation was run to determine if there was a relationship between how disturbing the participants found the video used in each version of the tool and its reported ease of use. No significant correlation was found (see table 5.3). There was also no significant correlation between a sum of the responses of the different versions and a sum of the reported distress caused by the video contents.

Version	***Spearman's rho coefficient***	***N***	***p***
Full annotation	.042	74	.719
Label afterwards	.003	46	.983
Realtime labeling	.080	45	.603
Sum of all versions	-.049	75	.676

Table 5.3 Spearman correlation coefficient with self reported distress caused by video content

Q3. WHICH VERSION OF THE TOOL DID THE PARTICIPANTS FIND EASIEST TO USE?

With regards to the statement "I found this version easy to use" (as reported in figure 5.2), a Friedman test revealed no significant difference between the versions, $\chi^2(2, N=44) = 3.493$, $p = 0.174$.

Participants were also asked to rank the labeling exercises in ascending order of difficulty. The easiest exercise was reported to be the one that involved the "full annotation" version by 50% of the participants, the "label after watching" version by 38%, and "label while watching" by 12%, from a total of 72 valid responses. A Chi-squared test for goodness of fit revealed a significantly different distribution than would be expected if the answers had been purely random, $\chi^2(2, N=71) = 20.64$, $p < .001$. When leaving out "label while watching", and comparing just the "full annotation" and "label after watching" versions, there was no significant difference from equal distribution however, $\chi^2(1, N=64) = 2.25$, $p = .134$.

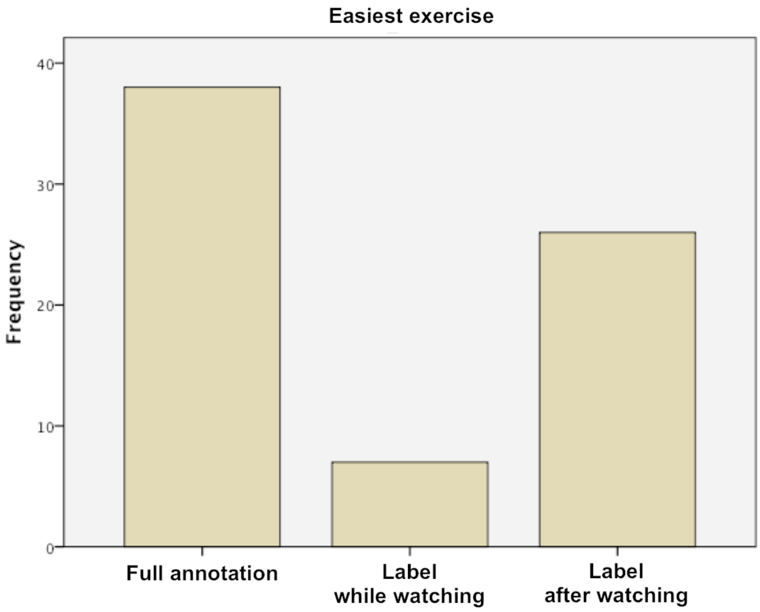


Figure 5.3 Easiest exercise

Furthermore, participants were also asked which version of the tool they find the best to help them learn to use the reappraisal themes.

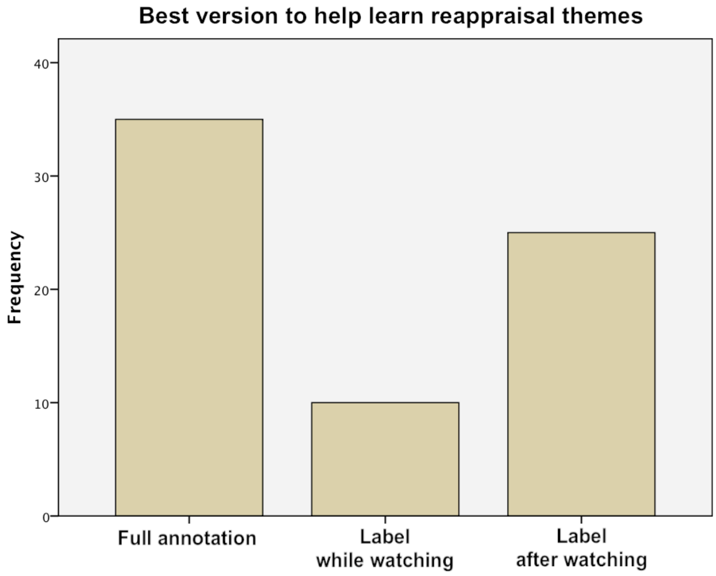


Figure 5.4 Best version to help learn reappraisal themes

Again, “full annotation” was chosen the most, and again a Chi-squared test for goodness of fit revealed that distribution was significantly different than would be expected if the answers had been purely random, $\chi^2(2, N=70) = 13.57, p = .001$. When leaving out “label while watching” and comparing just the remaining two versions, the difference from equal distribution is not significant, $\chi^2(1, N=60) = 1.667, p = .197$.

Q4. How did the participants use the tools? (i.e. behavior)

Log files generated by the tools were analyzed, and this revealed no apparent anomalies with regards to the relevance of themes, since they were all widely used (see figure 5.5). Also on average each participant selected more than one theme in each of the versions of the tool (see table 5.4).

Tool version	**M (SD)**
Full annotation	1.96 (1.22)
Real-time labeling	1.43 (0.96)
Label afterwards	1.72 (0.75)

Table 5.4 Number of themes used per participant

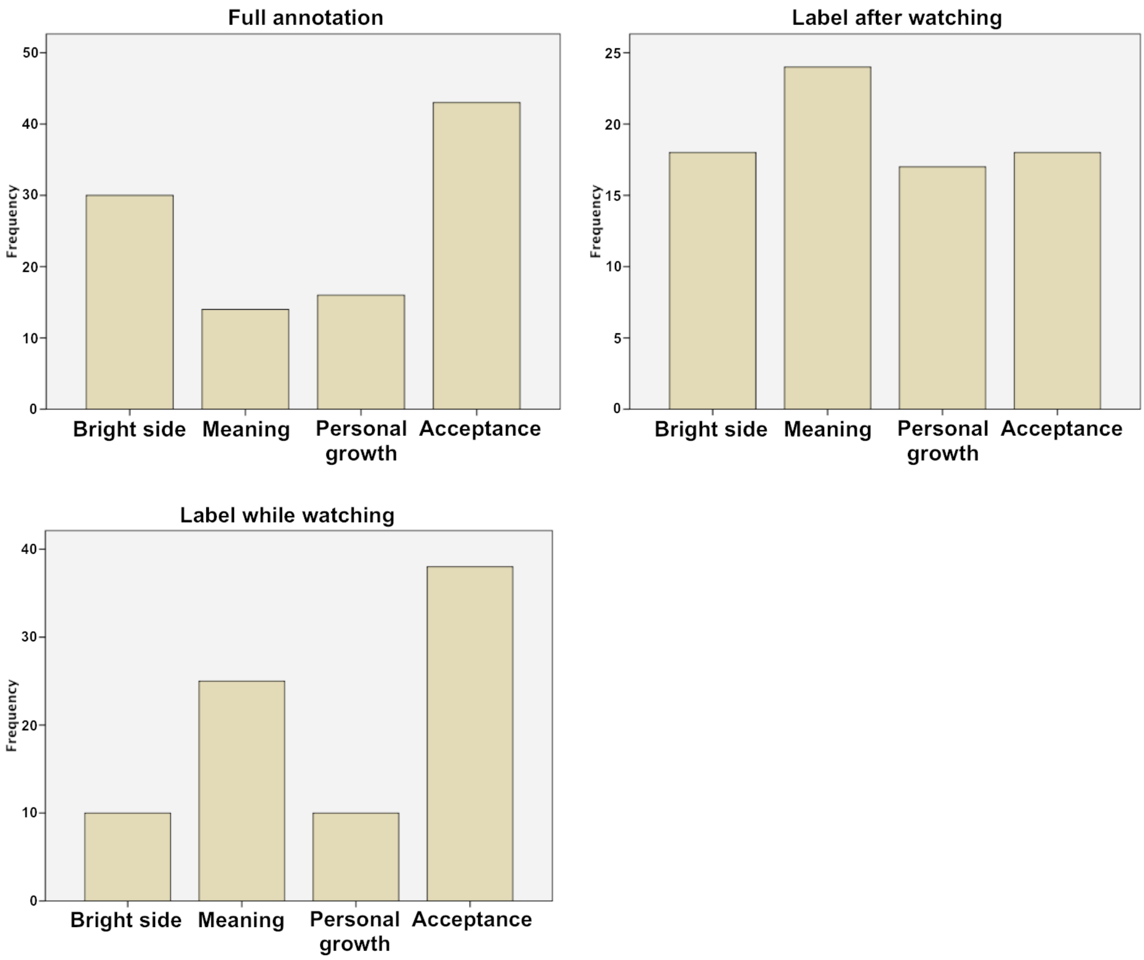


Figure 5.5 Frequencies of selected themes per version of the tool

For the full annotation version, the length of the entered responses (i.e. why the participants felt the selected theme was relevant at the point in the video) was also analyzed: M = 99 characters (SD = 56).

5.4 Conclusions

5.4.1 Answers to research questions

Q1. Do participants generally find the tool and training useful?

The survey results suggest that the participants generally felt the reappraisal themes were useful, and that the labeling tools helped them learn to use these themes. However, it seems that they could not really say if the video labeling exercises would help them see potential future disturbing situations in a different perspective. This seems understandable, given the fact that most participants had not actually experienced much hardship themselves, and the time they spent with the tools was limited.

Q2. How do participants perceive the ease of use of the tool?

Both the survey and the analyzed log files of the tools suggest that participants did not experience difficulties using the tools. They generally reported that all versions of the tool were easy to use, though there was a detail of the "real-time labeling" tool that some participants remarked could be improved. The video contents also did not appear to have any confounding effects on perceived ease of use of the tools.

Q3. Which version of the tool did the participants find easiest to use?

The "real-time labeling" version was rated significantly lower than the other two versions. The "full annotation" version was rated higher than the "label afterwards" version, but not significantly so.

Q4. How did the participants use the tools?

Log files generated by the tools revealed that all themes were applied, and on average each participant selected more than one theme in each exercise. In the full annotation tool, the textual input was on average 99 characters (SD = 56).

5.4.2 Limitations

The experiment had to be run within a predetermined, fixed setting. Therefore ideal control conditions were not possible. This resulted in some unavoidable potentially confounding factors: the distinction between the tools and the entire training may not have been clear to the participants, the different videos were linked to particular versions of the tools, and the order of the exercises was not randomized. Time spent with the tools was limited to one session. This obviously limits the potential for evaluation, and may also not be enough to have any lasting effect.

Participants of two different professions were used, and only one participant was female, so it is hard to say if the results can be generalized beyond male military personnel and firefighters. Generalizing to different age groups may be possible however: the military personnel were young, while the firefighters were middle-aged, and both groups found the tools easy to use.

Furthermore, participants may have felt obliged to give socially acceptable answers in the survey. Especially the NCOs were in a setting where it is expected to agree with higher-ranking officers such as their trainers. However this was a deliberate trade off: NCOs were likely to take the training more seriously when this was handled by their usual trainers.

The open-ended comments section of the survey revealed some further potential limitations. Two participants mentioned that they found it hard to put themselves into the situations depicted in the video, suggesting that perhaps the video contents was not relevant enough to the personal life experiences of all participants. Two other participants mentioned that watching videos did not cause them any real stress, and therefore they did not feel it prepared them properly for actual events. And two participants felt they could only apply this training once they actually experienced something bad.

5.4.3 FUTURE IMPROVEMENTS

There were many indications that the "full annotation" tool was the best (also qualitative in the comments section of the survey), so this suggests proceeding with this version instead of the others in potential future research. And instead of having a real live person introduce the tools, possibly this could be handled by a virtual "tutor". A benefit would be that the difference between live trainers could be eliminated as potential confounding factor.

6. EVALUATION - MEASUREMENT INSTRUMENT

6.1 Introduction

The measurement instrument, as described in chapter 4, included a Pictorial Emotional Stroop Test (PEST) and a Fear Potentiated Startle (FPS) blink magnitude measurement tool. This chapter describes the experiment that was designed to assess the predictive validity of this instrument, i.e. whether the designed measurement instrument could in fact measure a difference between neutral versus negative affective stimuli material presented to subjects. In this experiment, seven short negative and neutral affective videos were shown to the participants. The PEST phase of the experiment used stills from these videos, and the working hypothesis was that there would be a difference in response times between the stills from the negative and neutral videos. In the FPS blink measurement part of the experiment, participants completed a fear conditioning procedure with air-blasts to the larynx and measurements of their blink magnitude as triggered by a noise probe. The working hypothesis here was that participants would come to expect an aversive air-blast when they were shown either a fear conditioned still from a negative video or a fear conditioned geometric shape, and in addition that they would blink "harder" in response to the noise probe when they were shown the still from the negative video when compared to being shown the neutral geometric shape. More concisely formulated, this chapter then aims to answer the following research questions:

6.1.1 RESEARCH QUESTIONS

Q1. Does the implementation of the PEST, with stimuli conditioned by negative affective videos, produce a difference in reaction time between negative emotional and non-emotional stimuli?

Q2. Is the startle-blink amplitude higher for an air blast fear conditioned "neutral" still image from the negative affective video, compared with a neutral geometric shape with no negative association from the video that has also been fear conditioned via air blast?

6.2 Method

6.2.1 MEASUREMENTS

Three questionnaires were used to potentially identify participants with extraordinary personality traits (see appendix E.2 for the complete forms):

- The Brief Resilience Scale (BRS) aims to assess the ability to bounce back.
- The Core Self-Evaluation Scale (CSES) measures personality traits on four dimensions: locus of control, neuroticism, generalized self-efficacy, and self-esteem.
- The International Positive and Negative Affect Schedule Short Form (I-PANAS-SF) measures cross-cultural positive or negative affectivity.

Pictorial emotional Stroop task

During the PEST phase of the experiment, response time and accuracy was measured. The response time was later split out in the so-called fast and slow effect, as some studies suggest that the emotional Stroop task hinges mostly on the slow effect (see also section 2.2.2). To recap, the slow effect refers to the effect of a negative affective stimulus in the previous trial still affecting the response time in the current trial, while the fast effect only relates to the current trial. The fast effect is then measured by considering only trials following neutral stimuli, while for the slow effect the after-effects of negative versus neutral stimuli in trial *n*-1 are considered only on the responses to neutral stimuli in trial *n*.

Fear potentiated startle blink magnitude

At the beginning of the FPS phase of the experiment, 2mm Ag/AgCl EMG electrodes were applied to the face of the subjects on the right orbicularis muscle, below the pupil, and no more than 20mm toward the outer eye (Blumenthal et al., 2005). These electrodes were connected to a Mobi8 bioamplifier from Twente Medical Systems International (TMSI) linked to a recording computer via Bluetooth. The raw EMG recordings were later analyzed off-line with Matlab code written by the author, in order to reduce the dataset to z-scored magnitudes per image stimulus (see section 4.2 for full details).

6.2.2 PARTICIPANTS

In total 24 participants participated in the experiment in exchange for a small gift (USB stick). They were all students from Delft University of Technology, except for one from Rotterdam University of Applied Sciences. 3 participants failed to complete the experiment due to the image-rating task not resulting in enough stimuli fitting the criteria. According to the consent form that was signed by each participant, nobody was affected by color or face blindness, anxiety disorders, depression, startle disorders, severe psychological trauma, memory or learning disorders or hearing problems.

The scores of the self-report forms filled in by the participants are shown in table 6.1; there were no obvious outliers.

Table 6.1: Participant responses to self report scales

Scale	**Score** *M* (*SD*)
Brief Resilience Scale	3.38 (0.42)
Core Self-Evaluations Scale	3.42 (0.51)
I-PANAS-SF - positive	17.54 (2.25)
I-PANAS-SF - negative	10.75 (2.97)

6.2.3 DESIGN

A within subject design was used with two main parts run consecutively: a pictorial emotional Stroop task (PEST) and a fear potentiated startle (FPS) blink magnitude measurement.

The conditions for the PEST were:

- *Negative emotion*: image stimulus was one of the five most arousing images as rated by the participant with a valence of 3 or below on the SAM rating scale.
- *Neutral emotion*: image stimulus was one of the five least arousing images as rated by the participant with a valence within 1 point of the midpoint on the 9-point SAM rating scale.

The conditions for the FPS part:

- *NA*: Noise probe Alone, no image was shown (i.e. the screen remained black while the noise probe was triggered).
- *CS+*: Neutral image of a geometric shape that was fear-conditioned via air-puff in the conditioning phase.
- *CS-*: Neutral image of another geometric shape, not fear-conditioned.
- *FCS+*: Negative-association image that was fear-conditioned via air-puff in the conditioning phase. This image was a still displays the head of the killer from the American History X murder scene that was shown to the participant during the video watching phase (see appendix F.3). It was not used before in the Stroop task.

6.2.4 PROCEDURE AND APPARATUS + MATERIALS

A simplified sequence of the experimental setup is provided in figure 6.1, and a more detailed version in figure 6.2.

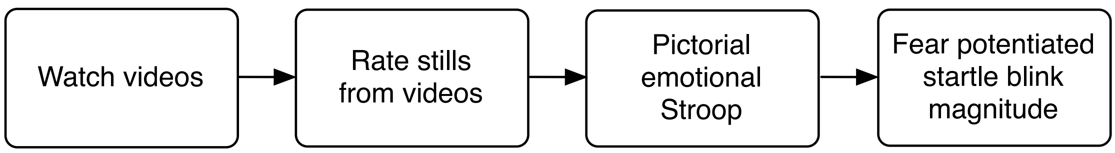
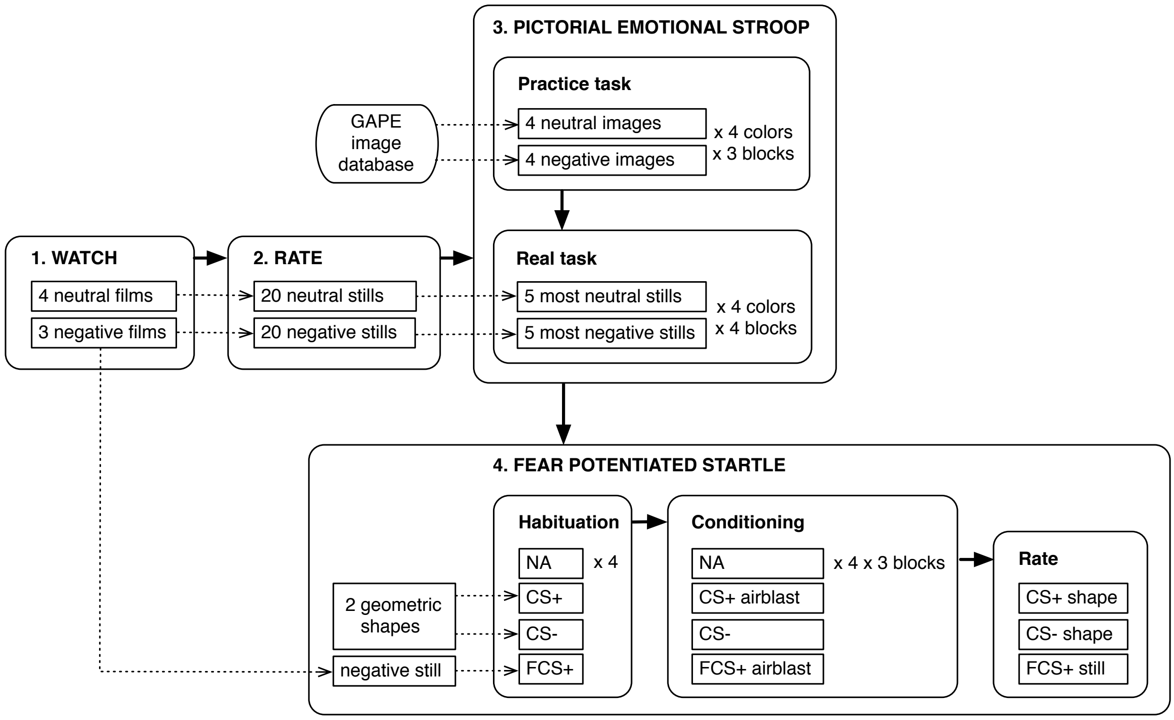


Figure 6.1: Simplified sequence of experiment

The experiment was implemented using E-prime software version 2.0.10.242, Standard (Schneider, Eschman, & Zuccolotto, 2012).



NA: Noise probe Alone, no image is shown.
CS+: Neutral image of a geometric shape that will be fear-conditioned via air-puff in the conditioning phase.
CS-: Other neutral image of a geometric shape, not fear-conditioned.
FCS+: Negative-association image (the same for all subjects and not used before in the Stroop task), that will be fear-conditioned via air-puff in the conditioning phase

Figure 6.2: Detailed schematic of experiment

Phase 1: "Watch videos"

Participants were first asked to watch seven short videos. Four odd-numbered videos were emotionally neutral in content. Three even-numbered videos contained content that was intended to illicit horror and distress, as applied by Schartau, Dalgliesh (2011). Stills from the videos are provided in appendix F. Between the videos there was a pause of 3 seconds with a black screen. Within these constraints, the order of the videos was randomized.

The neutral videos were as follows (duration in parentheses):

- Weather report 1 (1m 20s).
- Weather report 2 (1m 45s).
- News item on a new terminal for Heathrow airport (1m 30s).
- News item on the building of Beijing airport (1m 34s).

The negative videos were as follows:

- Clip from a news report on the abuse and murder of a baby (35s).
- Murder scene from the feature film "American History X" (36s).
- Clip from a documentary about the Rwandan genocide (35s).

Phase 2: “Rate stills from videos”

A total of 40 still images were selected from these videos for possible use in the pictorial emotional Stroop task (PEST): 20 from the neutral and 20 from the negative videos (see appendix F). Participants were asked to rate all these images on valence and arousal using a Self Assessment Manikin (SAM; Lang, 1980) with 9-point Likert scales for valence and arousal. This was to select five neutral, low-arousing images and five negative, high-arousing images for each participant on the basis of the individual ratings. The selection criteria were that the five images rated lowest by the participant on arousal of those rated with a valence within 1 point of the midpoint were used as the neutral stimuli, while the five images rated highest on arousal of those rated with a valence of 3 or below were used as the negative stimuli. The images were color filtered into four colors (red, green, blue, yellow) for use in the PEST.

Phase 3: “Pictorial Emotional Stroop task”

Participants were asked to identify the color of the pictures shown to them using a custom-built button box with four colored buttons connected to the parallel port of the PC (see also section 4.1). The selected images were presented at the center of a black screen on a 22-inch TFT monitor. Instructions were provided on screen before each segment.

Before proceeding to the real experimental trials of the PEST, there was a practice task, for which four neutral, low-arousing and four negative, high-arousing images were selected from the Geneva affective picture database (GAPE; Dan-Glauser & Scherer, 2011). The GAPE database contains 730 images rated on valence and arousal on a scale between 0 and 100 by 60 psychology students. The selected neutral images had ratings between 40 and 60 for valence and below 20 for arousal. The selected negative images had a valence rating below 30 and an arousal rating between 70 and 80.

The practice task contained 3 blocks of 32 practice trials. In every block the negative and neutral images were presented once in each color with no repeating colors or pictures. For the practice run, participants received feedback concerning the accuracy of their reaction after each trial. Each trial stayed on screen until a response was made. There was a pause between the blocks, where participants could press a button to proceed to the next block. After the practice blocks, there was a pause with a screen that informed the participant it was now time for the real task, where there would be no feedback. The participant was asked to press the spacebar to proceed.

For the real experimental trials, the inter-trial interval was set to 30ms, allowing for the measurement of fast as well as slow effects. 4 blocks were run, each containing 40 trials. In each block the previously selected five negative and five neutral images were presented once in each color, with no repeating colors or pictures. After the last block, the participant was informed that it was time for the next phase. The experiment supervisor would then attach electrodes below the eye and provide a headphone. After this a screen provided instructions for the next phase.

Phase 4: “Fear potentiated startle blink measurement”

The fear conditioning procedure was based on Norrholm et al. (2011) but with slightly different conditions and without an extinction phase. The exhaust tube of the air-blast device was attached to the participant pointing at the larynx. The participant was asked to put on a

Superlux HD-661 headphone, which was used to administer 40ms, 108dB noise probes of broadband white noise between 20 Hz to 20 kHz (Blumenthal et al., 2005) to induce blink. The apparatus is described in more detail in section 4.2.

Participants underwent a fear-potentiated startle protocol with Acquisition only. The Acquisition phase contains first one Habituation block, then 3 Conditioning blocks. The purpose of the Habituation block is for the participant to "get used" to the setting first without the aversive air-blast, before proceeding to the actual fear conditioning with air-puff in the Conditioning block. The Habituation block contains randomized trials where 4 trials are of the noise-probe alone, 4 are of the CS+, 4 of the CS-, and 4 are of the FCS+. In the Conditioning blocks, there are randomized trials where 4 are of the noise-probe alone, 4 are of the CS+ with air-blast, 4 are of the FCS+ with air-blast, and 4 of the CS-. For both phases, the intertrial intervals were randomized to be between 9 and 22 seconds. Participants were asked to log their expectancy of receiving the US with the keyboard on each trial.

6.2.5 Data preparation & analysis

The timing and accuracy of the colored buttons that were pressed during the Stroop task were recorded in a data file by the E-prime script. During the FPS phase, electromyographic recordings were made to analyze the magnitude of the eye blinks using a Mobi8 device and Portilab software, which also recorded pulse coded trigger signals generated by the E-prime script to mark the timing of the noise probes and types of stimulus. Expectancy of air blast as indicated by the participants through pressing PC keyboard buttons was recorded by the E-prime script.

Pictorial emotional Stroop

The E-prime data files of the PEST were processed with a custom Matlab script (see appendix C), which identified outliers for reaction times based on an exclusion criterion of the median plus or minus 3 times the Median of Absolute Deviations (as recommended by Leys, Ley, Klein, Bernard & Licata; 2013), and excluded incorrect responses from the reaction times for further analysis. Since the experiment involved repeated measures where data cannot assumed to be independent, a multilevel approach was taken for the analysis using R version 3.2.5. Two models were created for both the fast and slow effect: Model 0 was a null model predicting reaction time based on a random intercept per participant, and Model 1 added a fixed effect for the emotion of the stimulus. Log likelihoods of these models were compared to determine if Model 1 had a better fit.

Fear potentiated startle blink

A custom Matlab script was written to reduce raw Portilab EMG data into z-scored (per participant) blink magnitudes per image stimulus type (see section 4.2 for details and appendix B for complete script). This data was exported into SPSS and R for statistical analysis. Again a multilevel approach was taken for the analysis using R version 3.2.5. A null model was fitted for z-scored blink amplitude allowing intercepts to vary by participant. Model 1 added a fixed effect for the four conditions. Log likelihoods of these models were compared to determine if Model 1 had a better fit, i.e. if the conditions had an effect on the blink amplitude. Another null model was created from only the blinks in the CS+ and CS- conditions allowing intercepts to vary by participant. This was compared to a Model 1 with the added fixed effect for these conditions.

6.3 Results

6.3.1 PICTORIAL EMOTIONAL STROOP

Adding a fixed effect for the valence (negative vs. neutral) of the image stimulus in Model 1 did not result in a significantly better fit than Model 0 when comparing log likelihoods, neither for the fast nor the slow effect (see tables 6.2 & 6.3). In other words: the valence of the image stimuli did not serve as a predictor of the reaction time.

Table 6.2
Fast effect - comparing the null model with a model including an emotion effect.

Model	df	AIC	logLik	χ^2	Pr (>χ^2)
0: Null model	3	22165.35	-11079.68		
1: + valence	4	22166.88	-11079.44	0.48	0.49

Table 6.3
Slow effect - comparing the null model with a model including an emotion effect.

Model	df	AIC	logLik	χ^2	Pr (>χ^2)
0: Null model	3	22206.85	-11100.42		
1: + valence	4	22207.78	-11099.89	1.07	0.30

For reference, appendix D displays the mean reaction times, error rates and the fast and slow effect of the entire group of participants according to the protocol used by Frings, Englert, Wentura and Bermeitinger (2010).

6.3.2 FEAR POTENTIATED STARTLE BLINK

To get a quick overview, the z-scored blink amplitudes (*M* and *SD*) are displayed in table 6.4, and the densities are plotted in figure 6.1.

Table 6.4: Z-scored amplitudes of blinks per condition

Condition	Amplitude *M* (*SD*)
NA	-0.18 (0.83)
CS-	0.06 (0.90)
CS+	0.10 (0.88)
FCS+	0.12 (0.84)

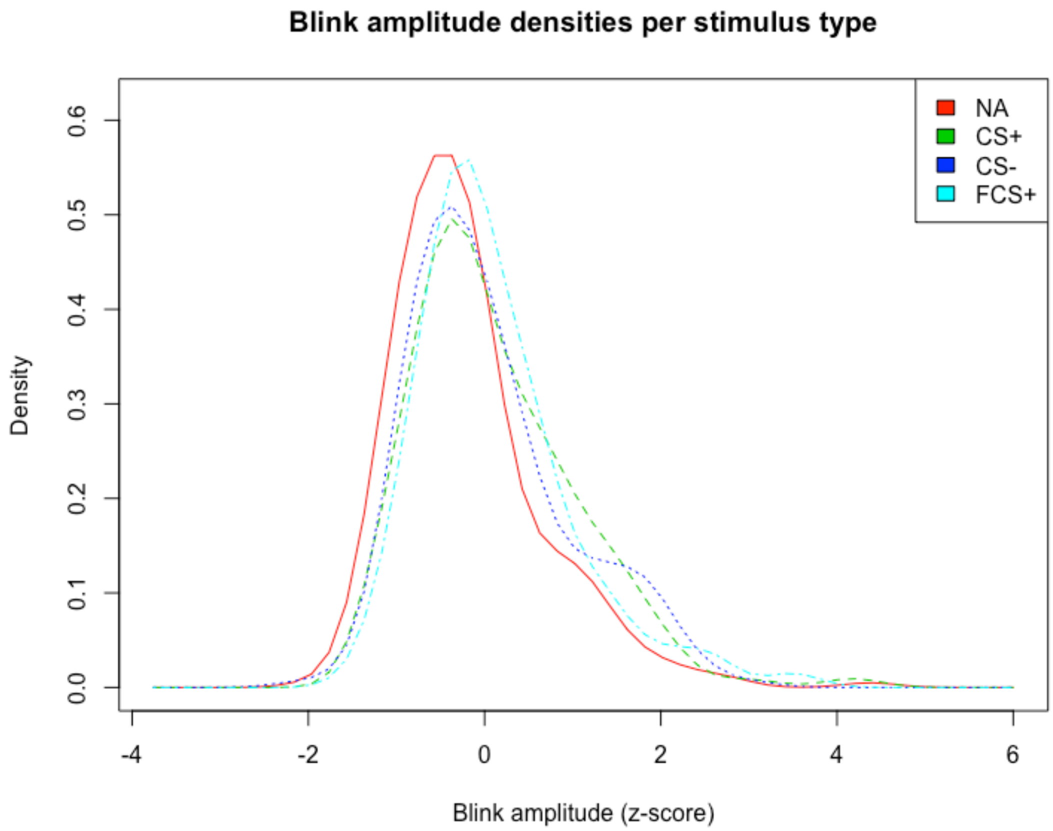


Figure 6.1: Blink amplitude (z-score) densities of the stimulus types

NA:	Noise probe Alone, no image is shown.
CS+:	Neutral image of a geometric shape that was fear-conditioned via air-puff in the conditioning phase.
CS-:	Other neutral image of a geometric shape, not fear-conditioned.
FCS+:	Negative-association image (the same for all subjects and not used before in the Stroop task), that was fear-conditioned via air-puff in the conditioning phase

The addition of the fixed effect of the four conditions to the null model resulted in a significantly better fit (see table 6.5). For the condition NA (noise probe alone), the estimate for the z-scored blink amplitude was significantly different from 0, and it was significantly smaller than for that of the other conditions (see table 6.6). The same multi-level analysis was then performed for all blinks without those of the condition NA. In this case Model 1 did not result in a better fit (see table 6.7), so in fact from all four conditions NA could be isolated as the only predictor for the blink amplitude. In other words: when participants were presented with just a noise probe without any image stimulus, the amplitude of their blink was significantly negatively affected compared to the conditions when they were also presented with an image stimulus prior to the noise probe. But there was no significant predictive effect for any of the other conditions, not even for the fear conditioning of the neutral geometric shape.

Table 6.5
Comparing the null model with a model including an effect for the condition.

Model	df	AIC	logLik	χ^2	Pr (>χ^2)
0: Null model	3	2759.70	-1376.85		
1: + condition	6	2742.80	-1365.40	22.89	<.01

Table 6.6
Model 1 – overview of the fixed effects

Parameter	Estimate	Std. Error	t-value	p
Intercept (Condition NA)	-0.196	0.0564	-3.476	<.01
Condition CS+	0.3163	0.0767	4.126	<.01
Condition CS-	0.2538	0.0767	3.299	<.01
Condition FCS+	0.3149	0.0767	4.108	<.01

Table 6.7
Comparing the null model with a model including an effect for the condition excluding NA.

Model	df	AIC	logLik	χ^2	Pr (>χ^2)
0: Null model	3	2078.35	-1036.17		
1: + condition	5	2081.47	-1035.74	0.88	0.65

6.4 Conclusions

6.4.1 ANSWERS TO RESEARCH QUESTIONS

Q1. Does this implementation of the PEST, with stimuli conditioned by a negative affective video, produce a significant difference in reaction time between negative emotional and non-emotional stimuli?

It cannot be concluded that the PEST was effective in producing a significant difference in reaction time between negative emotional and non-emotional stimuli. This can possibly be explained by a combination of selection bias and video material that was not actually perceived as very shocking: during the debriefing many of the participants stated that the recruitment text and consent form had set their expectations for the video material to be very graphic and shocking, some said that they in fact had been looking forward to the challenge of facing such material, but that the actual videos and images used in the experiment were not that shocking to them. Conversely, a number of people who had initially signed up for the experiment later declined to participate after reading the consent form, stating that they were afraid that they would find the material too shocking.

Q2. Is the startle-blink amplitude higher for an air blast fear conditioned “neutral” still image from the negative affective video (FCS+), compared with a neutral geometric shape with no negative association from the video that has also been fear conditioned via air blast (CS+)?

Although the means of the blink amplitudes are ordered in the way that one would expect (FCS+ > CS+ > CS- > NA), the experiment did not produce significantly different eye blink magnitudes between the FCS+ and CS+ conditions. Only the condition where no image was shown at all (NA) revealed significantly different blink magnitudes from the other conditions. This suggests that subjects were only affected by whether there was a picture shown or not, and therefore it cannot be ruled out that the setup was ineffective at the most basic level: fear conditioning subjects to certain images via aversive air blasts, and therefore that it also failed at the higher level measurement of the differences in emotional responses of subjects between neutral and negative affective content that is shown to them. Perhaps the air-blast was simply not aversive enough. Additionally, there is again the possible explanation of selection bias and use of video material that was not perceived as very shocking by the participants.

6.4.2 LIMITATIONS

Due to ethical considerations and limited experience with these kinds of experiments, a lot of care was taken to not "damage" participants by using a strong warning in the recruitment texts and video material reportedly was perceived as not incredibly shocking by many of the participants.

7. CONCLUSION

The following research question was elicited from the problem definition for this thesis:

> *What would be a possible technology supported resilience training environment with high user acceptance, and how could its effectiveness be measured?*

With the following sub-questions:

1. *What would a possible technology supported resilience training system look like?*
2. *What is the user acceptance of the system?*
3. *How can its effectiveness be measured?*

A direct answer to the main research question would be lengthy, so it is instead addressed in sections via the sub-questions.

What would a possible technology supported resilience training system look like?

There are many possible ways to support resilience training with technology - see e.g. Favié, Vakili, Brinkman, Morina et al. (2016) for a state of the art review. In this thesis, appraisal was chosen to be the focus. The project requirements were that participants should label events in videos with cognitive reappraisal themes, the labeling tool should be easily adaptable to other content and themes, it should be simple to use, easy to deploy, usable in a group setting, and face validity was deemed important. Requirements for the measurement tool were that it should be grounded in and build on existing studies, and measurements should be objective.

What is the user acceptance of the system?

User acceptance is key in getting a training environment for military professionals off the ground (see Vakili et al; 2015). In order to study user acceptance of the developed system, it was implemented in a training environment for non-commissioned officers of the Dutch army and firefighters. Usage data of the tools was analyzed, along with questionnaire responses regarding use of the tool. Analysis of the log files of the tools and survey results suggest that user acceptance was not an issue: the training was perceived as useful, and the tool as easy to use. This thesis did not formally research the acceptance of the measurement apparatus, which included a pictorial emotional Stroop task (PEST) and a fear potentiated startle (FPS) blink magnitude measurement. The debriefing after the experiment did not reveal significant issues however.

How can its effectiveness be measured?

The quest was for a concrete measure with a causal link to resilience, like a physiological response to negative affective material. As there is a whole sequence of events leading to possible symptoms of PTSD, that process could be influenced by “better” handling of traumatic events/materials. Therefore a relevant measurement could be related to how one reacts to traumatic material. For this thesis it was decided to design a measurement instrument with a PEST and FPS blink magnitude measurement. However an analysis of the data from the evaluation experiment of the instrument did not reveal a difference between neutral and negative material. So this approach does not appear to be an easy way to measure resilience.

7.1 Contribution

A flexible video annotation tool was built that can be applied for resilience training, and has already been used in a study by Beer, Neerincx, Morina & Brinkman (2017). Its video content and annotation themes can be easily changed and therefore the tool can also be used in other domains such as social phobia or even marketing training. Furthermore, a practical setup with FPS measurement with air blast and emotional Stroop task was realized, which can be used for other experiments. The concept of using appraisal in resilience training is somewhat of a novelty, and therefore this thesis could serve as a stepping stone in this area. Acceptance of the annotation tool by the target domain (military personnel & fire fighters) can be concluded, which is an important factor for any tool or intervention to be effective within this target domain. Furthermore, this thesis provides some insight into the complexities of measuring resilience with an emotional Stroop task and fear potentiated startle blink. These two paradigms appear to not be easily applicable in measuring resilience in individuals without disorders.

7.2 Limitations

It was not deemed feasible to study the effectiveness of the video annotation tool, and the measurement tool made use of only two out of many possible paradigms for measuring resilience. Furthermore, the target group for the video annotation tool was limited to military personnel and fire fighters, although results may reasonably be expected to extrapolate to similar professions such as the police. Finally, the measurement tool was evaluated with university students only, and no significant effect was found, possibly due to selection bias and use of video material that may not have been sufficiently aversive. However with regards to a possible selection bias towards “tough” individuals for the participants: the same can be said of the target group of military personnel, fire fighters etc.

7.3 Future Work

An effect study would be a first obvious choice for future work related to this thesis, ideally longitudinal: is the training indeed effective in improving appraisals, and how long do the training effects last – can people still apply the techniques a year later? Also, do people actually use the techniques when they are confronted with traumatic events in real life? And if so, does this actually help reduce the PTSD symptoms?

In this study the video annotation tool was used as part of a one-hour workshop, but one could consider other options and investigate what would be the best possible way to implement it in a training. Perhaps it is more effective when split up in separate sessions on multiple days for example. It could also be turned into a more standalone solution, e.g. by including a virtual coach instead of requiring a live instructor for explaining the concept and themes. A virtual coach/avatar could also provide realtime feedback. Gamification or scaffolding are other possible approaches worth investigating. To help with this, attention could be focused on the trainers in further studies, while this thesis focused mainly on the experience of the students.

As mentioned before, the video material and themes can easily be adapted to suit many other paradigms besides resilience alone – there are plenty of scenarios where one might benefit from being able to see things in a different view. One could imagine it being used for other cognitive therapies that involve reflection (e.g. social phobia) or various other kinds of educational purposes (e.g. marketing trainings).

As for the designed measurement tool, there may have been issues with the experimental setup that could be investigated: the image and video material may not have been shocking enough, the air blast may not have been aversive enough, and there may have been selection bias in the participants. Besides investigating these issues, more physiological markers could be used. The entire setup could also be made more portable and easier to use by non-experts.

In conclusion, this thesis showed that the novel application of appraisal in technology supported resilience training may be worth researching further. The experimental evaluation of the implemented measurement tool appears to indicate that PEST and FPS are not promising paradigms for easily measuring resilience.

REFERENCES

Agaibi, C. E., & Wilson, J. P. (2005). Trauma, PTSD, and Resilience A Review of the Literature. *Trauma, Violence, & Abuse*, *6*(3), 195–216.

Alrymple-Alford, E. C. (1992). Response-key input via the IBM PC/XT/AT's parallel printer port. *Behavior Research Methods, Instruments & Computers*, *24*(1), 78–79.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Audacity [Computer software]. (2015). Retrieved from http://audacityteam.org/

Beer, U.M., Neerincx, N.A., Morina, N., & Brinkman, W. (2017), Technological Support of Appraisal Training in Perspective Broadening. *Health and Technology*. Manuscript submitted for publication.

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V, & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42*(1), 1–15.

Bouchard, S., Bernier, F., Boivin, E., Morin, B., & Robillard, G. (2012). Using biofeedback while immersed in a stressful videogame increases the effectiveness of stress management skills in soldiers. *PloS One*, *7*(4), e36169.

Brown, J. S., Kalish, H. I., & Farber, I. E. (1951). Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *Journal of Experimental Psychology*, *41*(5), 317.

Carroll, J. M. (1987). Minimalist design for active users. In R. M. Baecker & W. A. S. Buxton (Eds.), *Human-computer Interaction* (pp. 621–626). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Cisler, J. M., Wolitzky-Taylor, K. B., Adams, T. G., Babson, K. A., Badour, C. L., & Willems, J. L. (2011). The emotional Stroop task and posttraumatic stress disorder: a meta-analysis. *Clinical Psychology Review*, *31*(5), 817–28.

Cohn, J., Weltman, G., Ratwani, R., Chartrand, D., & McCraty, R. (2010). Stress Inoculation through Cognitive and Biofeedback Training. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)* (Vol. 2010, pp. 1–11).

Constantine, R., McNally, R. J., & Hornig, C. D. (2001). Snake Fear and the Pictorial Emotional Stroop Paradigm. *Cognitive Therapy and Research*, *25*(6), 757–764

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, *43*(2), 468–77.

Dandoy, A. C., & Goldstein, A. G. (1990). The use of cognitive appraisal to reduce stress reactions: A replication. *Journal of Social Behavior & Personality; Journal of Social Behavior & Personality*.

Denny, B. T., & Ochsner, K. N. (2014). Behavioral effects of longitudinal training in cognitive reappraisal. *Emotion*, *14*(2), 425–33.

Eippert, F., Veit, R., Weiskopf, N., Erb, M., Birbaumer, N., & Anders, S. (2006). Regulation of emotional responses elicited by threat-related stimuli. *Human Brain Mapping*, *28*(5), 409–423.

Entwistle, N., Hanley, M., & Hounsell, D. (1979). Identifying Distinctive Approaches to Studying. *Higher Education*, *8*, 365–380.

Feder, A., Nestler, E. J., Westphal, M., & Charney, D. S. (2010). Psychobiological mechanisms of resilience to stress. In J. W. Reich, A. J. Zautra, & J. S. Hall (Eds.), *Handbook of Adult Resilience* (pp. 35–54). Guilford Publications.

Favié, J. (2014). *Research assignment: technology supported resilience training – a state of the art*. TU Delft.

Favie, J., Vakili, V. G., Brinkman, W.-P., Morina, N., & Neericx, M. A. (2016). State of the art in technology supported resilience training for military professionals. In A. Gaggioli, A. Ferscha, G. Riva, S. Dunne, & I. Viaud-Delmon, I. (eds.). Human Computer Confluence: Transforming human experience through symbiotic technologies. Warsaw: De Gruyter Open.

Foa, E. B., Steketee, G., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy*, *20*(2), 155–176.

Frings, C., Englert, J., Wentura, D., & Bermeitinger, C. (2010). Decomposing the emotional Stroop effect. *The Quarterly Journal of Experimental Psychology*, *63*(1), 42–49.

Greenberg, P. E., Sisitsky, T., Kessler, R. C., Finkelstein, S. N., Berndt, E. R., Davidson, J. R. T., … Fyer, A. J. (1999). The economic burden of anxiety disorders in the 1990s. Journal of Clinical Psychiatry, 60(7), 427–435.

Gross, J. J. (1998). Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362.

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (Vol. 3, pp. 3–26). New York: Guilford Press.

Hoge, C., & Castro, C. (2005). Impact of Combat Duty in Iraq and Afghanistan on the Mental Health of U.S. Soldiers : Findings from the Walter Reed Army Institute of Research Land Combat Study. Strategies to Maintain Combat Readiness During …, (2005), 1–6.

Hourani, L., Kizakevich, P., & Hubal, R. (2011). Predeployment stress inoculation training for primary prevention of combat-related stress disorders. *Journal of Cybertherapy & Rehabilitation*, *4*(1).

Institute of Medicine. (2012). Treatment for Posttraumatic Stress Disorder in Military and Veteran Populations : Initial Assessment. Washington, DC: National Academy of Sciences.

Kaplan, C. P., Turner, S., Norman, E., & Stillson, K. (1996). Promoting resilience strategies: A modified consultation model. *Children & Schools*, *18*(3), 158–168.

Kemp, J., & Bossarte, R. (2012). Suicide Data Report 2012. Department of Veterans Affairs, 1–59.

Olesen, J., Gustavsson, A., Svensson, M., Wittchen, H. U., & Jönsson, B. (2012). The economic cost of brain disorders in Europe. European Journal of Neurology, 19(1), 155–162.

Kim, P. Y., Britt, T. W., Klocko, R. P., Riviere, L. A., & Adler, A. B. (2011). Stigma, Negative Attitudes About Treatment, and Utilization of Mental Health Care Among Soldiers. *Military Psychology*.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, *97*(3), 377–395.

Lavy, E., & van den Hout, M. (1993). Selective attention evidenced by pictorial and linguistic stroop tasks. *Behavior Therapy*, *24*(4), 645–657.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer Publishing Company.

Lazarus, R. S., Opton Jr, E., Nomikos, M. S., & Rankin, N. O. (1965). The principle of short-circuiting of threat: Further evidence. *Journal of Personality*.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). *Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median*. Journal of Experimental Social Psychology, 49(4), 764–766.

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour Research and Therapy*, *43*(11), 1391–424.

MacLeod, C., & Hagan, R. (1992). Individual difference in the selective processing of threatening information, and emotional responses to a stressful life event. *Behaviour Research and Therapy*, *30*(2), 151–161.

Marton, F., & Saljo, R. (1976). On qualitative differences in learning: I - Outcome and process. *British Journal of Educational Psychology*, *46*(1), 4–11.

Marton, F., & Saljo, R. (1976). On qualitative differences in learning: I - Outcome and process. *British Journal of Educational Psychology*, *46*(1), 4–11.

McKenna FP, and Sharma D (2004) “Reversing the emotional stroop effect reveals that it is not what it seems: the role of fast and slow components” *J Exp Psychol Learn Mem Cogn,* 30(2):382-92

Meichenbaum, D. H., & Deffenbacher, J. L. (1988). Stress inoculation training. *The Counseling Psychologist*, *16*(1), 69–90.

Nielsen, J. (1994). Heuristic Evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods* (pp. 25–62). John Wiley & Sons, New York, NY.

Ochsner, K. N., & Gross, J. J. (2004). Thinking makes it so: A social cognitive neuroscience approach to emotion regulation. *Handbook of Self-Regulation: Research, Theory, and Applications*, 229–255.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–9.

Pole, N., Neylan, T. C., Best, S. R., Orr, S. P., & Marmar, C. R. (2003). Fear-potentiated startle and posttraumatic stress symptoms in urban police officers. *Journal of Traumatic Stress*, *16*(5), 471–479.

Psychology Software Tools (2015). E-Prime 2.0 Timing Data. Retrieved October 12, 2015, from http://www.pstnet.com/support/kb.asp?TopicID=5628

Rizzo, A., John, B., Newman, B., Williams, J., Hartholt, A., Lethin, C., & Buckwalter, J. G. (2013). Virtual Reality as a Tool for Delivering PTSD Exposure Therapy and Stress Resilience Training. *Military Behavioral Health*, *1*(1), 48–54.

Rizzo, A., Parsons, T. D., Lange, B., Kenny, P., Buckwalter, J. G., Rothbaum, B., … Reger, G. (2011). Virtual reality goes to war: a brief review of the future of military behavioral healthcare. *Journal of Clinical Psychology in Medical Settings*, *18*(2), 176–87.

Samoilov, A., & Goldfried, M. R. (2000). Role of Emotion in Cognitive-Behavior Therapy. *Clinical Psychology: Science and Practice*, *7*(4), 373–385.

Schartau, P. E. S., Dalgleish, T., & Dunn, B. D. (2009). Seeing the bigger picture: training in perspective broadening reduces self-reported affect and psychophysiological response to distressing films and autobiographical memories. *Journal of Abnormal Psychology*, *118*(1), 15.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.

Shurick, A. a., Hamilton, J. R., Harris, L. T., Roy, A. K., Gross, J. J., & Phelps, E. a. (2012). Durable Effects of Cognitive Restructuring on Conditioned Fear. *Emotion*, *12*(6), 1393–1397.

Simmons, A., & Yoder, L. (2013). Military Resilience: A Concept Analysis. *Nursing Forum*, *48*(1), 17–25.

Skogstad, M., Skorstad, M., Lie, A., Conradi, H. S., Heir, T., & Weisæth, L. (2013). Work-related post-traumatic stress disorder. Occupational Medicine.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Troy, A. S., Wilhelm, F. H., Shallcross, A. J., & Mauss, I. B. (2010). Seeing the silver lining: Cognitive reappraisal ability moderates the relationship between stress and depressive symptoms. *Emotion*, *10*(6), 783.

Vaidyanathan, U., Patrick, C. J., & Cuthbert, B. N. (2009). Linking dimensional models of internalizing psychopathology to neurobiological systems: affect-modulated startle as an indicator of fear and distress disorders and affiliated traits. *Psychological Bulletin*, *135*(6), 909.

Vakili, V., Brinkman, W., & Neerincx, M. (2012). Lessons learned from the development of technological support for PTSD prevention: a review. In B. K. Wiederhold & G. Riva (Eds.), *Annual Review of Cybertherapy and Telemedicine 2012: Advanced Technologies in the Behavioral, Social and Neurosciences* (pp. 22–26). IOS Press, Incorporated.

Vakili, V., Brinkman, W.-P., Morina, N., & Neerincx, M. A. (2014). Characteristics of successful technological interventions in mental resilience training. *Journal of Medical Systems*, *38*(9), 1–14.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: A new measure of emotion? *Journal of Abnormal Psychology*, *97*(4), 487.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3–24.

Wilson, J. P. (2004). Broken spirits. In J. P. Wilson & B. Drozdek (Eds.), *Broken spirits: the treatment of traumatized asylum seekers, refugees and war and torture victims* (pp. 141–173). New York: Brunner/Routledge.

Wright, Jesse H. (2006). Cognitive behavior therapy: Basic principles and recent advances. Focus, 4(2), 173.

APPENDIX A VIDEO ANNOTATION TOOLS

A.1 Scenario for tool version A: full annotation

Scenario: Use of full annotation tool as part of a resilience training by a group of military professionals

Introduction by instructor

Participants are situated in a classroom with computers that have the tool installed and headphones attached.

Since applying cognitive reappraisal themes will be outside the comfort zone of military professionals, an instruction and group discussion are deemed helpful before proceeding to use the actual tool. The instructor (ideally a military professional with combat experience, who the participants are more likely to accept than e.g. a civilian psychologist) gives a general introduction about reappraisal. He explains that how one feels about an event is not directly determined by the event itself, but rather by the way one looks at or thinks about it, and that it is possible to change these "ways of thinking about a situation". The instructor names four CR themes, and explains what they mean:

- *Bright side: I can also see positive aspects of this situation*
- *Purpose: I can see this serves a greater purpose*
- *Personal growth: This will help me grow as a person*
- *Acceptance: I accept the situation as it is*

The instructor then describes an example aversive event, and possible ways of how to apply the CR themes: participants are asked to imagine that just as they were about to go on a holiday, they broke a leg in a traffic accident. They are now stuck to a hospital bed, and face a lengthy period of recuperation. The instructor explains how each of the CR themes can be applied to help them look at the situation in a different light.

Participants then get to practice by labeling videos with these "ways of thinking about a situation" using the annotation tool. A short instruction video explains how to use the tool. The instructor then gives a brief lead-in about the video that will be used in the session.

The subject of the video is the genocide that took place in Rwanda in 1994 and the role of the UN peacekeeping troops. This video was picked beforehand by the trainer, as he found it a relevant scenario for military service members. It was relatively easy for him to find a documentary about this subject online, edit the video material to contain only the most relevant segment, and use it in the tool by just pasting the video file name and folder into an XML configuration file.

The participants are asked to imagine that they were part of the troops on this mission while watching the video. After using the tool, the instructor asks the group to discuss which themes they chose, and why.

Finally, the group is told to proceed to the next phase, where they will all individually use the full annotation tool. As the participants now already have some experience with applying the themes and using the tool, their cognitive demands will be reduced during this actual task.

Use of the video annotation tool

Each participant starts the tool, and enters his or her assigned ID (which is used in the output file that stores the results for later analysis by researchers).

The tool switches to full screen mode (preventing distractions and using other programs) and a video starts playing in an interface with a timeline and play/pause controls that is similar to popular media players such as VLC-player and Youtube (consistency with familiar applications adds to the usability of the tool). There is also a "+" button to add a theme-label. This minimalistic, uncluttered interface helps prevent distraction and cognitive overload: the participant needs to be sufficiently "immersed" in the video for his or her emotions to be affected, while at the same time keeping the CR themes in mind and deciding when to apply which theme. This is probably already demanding enough that any "cognitive overhead" by using the program must be reduced to a minimum.

At some point in the video the narrator explains that the UN troops had an insufficient mandate and capacity to prevent the genocide. The participant recognizes that he could apply the theme "Acceptance" here. He clicks the "+" button to add a theme-label. Playback is automatically paused, and a popup appears, where the participant selects the "Acceptance" theme from a dropdown containing the four CR themes previously explained by the instructor. He also explains why he feels the theme is relevant at this point in the video in a textbox: "There is nothing I could have done about this horrible situation since the mandate and capacity of the mission were insufficient, and then the best you can do is accept this and try and move on." After clicking "OK" the popup closes, a theme-label with the letter "A" (for Acceptance) appears on the current position in the timeline, and playback of the video is automatically resumed.

Later on, a segment in the video shows that the troops evacuated foreign civilians. The participant feels that he can apply the theme "Bright side" here. He presses the "+" button again, selects the theme and explains in the textbox: "We did do something positive here - without us these people may also have been murdered". Again a theme-label is added to the timeline, this time with the letter "B" (for Bright side).

The video resumes playback, and after a while the participant realizes that he could also see the whole experience as a potential for personal growth, so he adds a label for this theme. A marker with the letter "P" (for Personal growth) is added to the timeline, though not at a position that directly refers to any particular event in the video. Towards the end of a the video a former service member that took part in the mission explains that although he found much of the experience quite horrific, he now feels that he has moved on and that having experienced the mission helps him put other issues in perspective. The participant realizes that this part of the video better suits the "Personal growth" label that he previously added to a somewhat random position, so he pauses the video and drags the corresponding "P" label along the timeline towards the current playing position. Selecting the label also brings up an editing subwindow at the right hand side of the video that allows changing the theme, editing the text or deleting the annotation altogether. The participant decides to change the text a bit, referring more directly to the statement of the former service member. This change is automatically stored. The participant presses "play" to resume playing.

After watching the entire video, and feeling satisfied with all his annotations, the participant clicks the button to exit the tool at the lower right hand side. A popup asks for confirmation (preventing accidental termination) and the annotations are stored for later analysis by researchers. Finally, the instructor asks the participants to fill in a form with questions for evaluating acceptance of the tool.

A.2 Tool version A: full annotation – UI sections explained

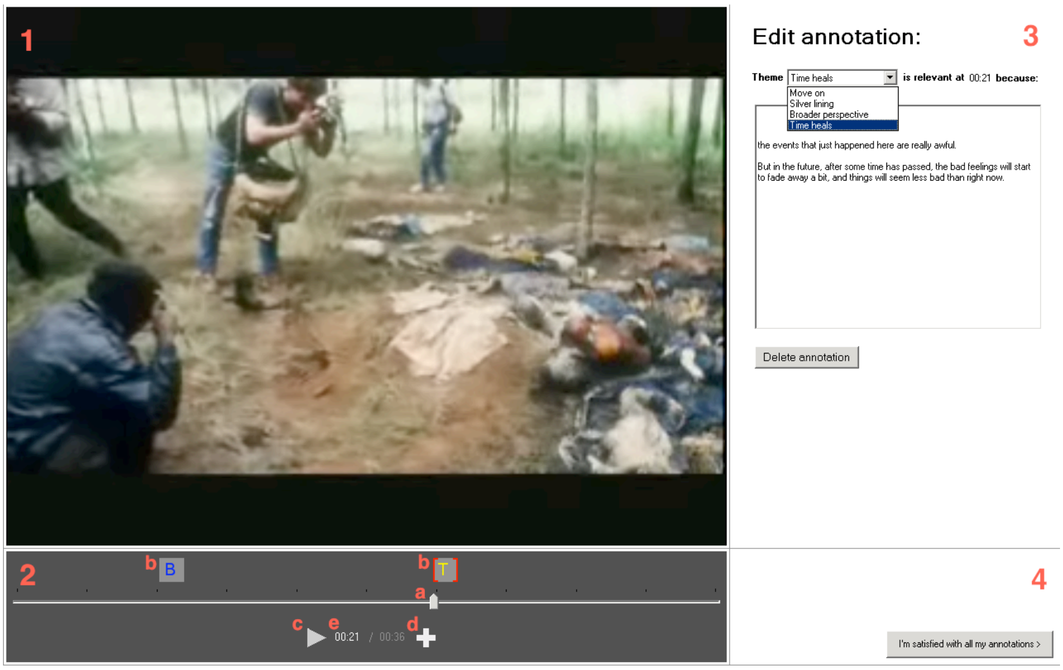


Figure A.1: full annotation tool sections

1. Video. Playback is automatically paused when adding an annotation.
2. Timeline & controls
    a. Current playing position on the timeline. The playing position of the video may be changed by dragging this control left or right.
    b. Labels of CR themes that the participant applied at corresponding positions on the timeline. Clicking a label selects the annotation for editing its details in subwindow 3. Labels may also be dragged to a different position along the timeline.
    c. Pause/play control.
    d. "+" control to add a new annotation at the current playing position. Brings up a popup to choose from a list of CR themes and enter a text for the participant to describe why the selected theme is relevant at the playing position.
    e. Numeric display of current playing position / total time of video.
3. Editing subwindow, where existing annotations may be changed. This area is only functional when a label on the timeline has been selected. When no label is selected, this area is left blank to reduce clutter and distractions.
4. Contains a button to exit the tool once the participant is satisfied with all annotations. This button is only enabled when at least one annotation has been added, and brings up a popup to confirm the action, in order to prevent accidental termination of the tool. The location is at the bottom right, marking the exit as the final part of a top-to-bottom, left-to-right workflow.

A.3 Tool version A: full annotation – configuration & log file

Configuration file

The annotation tool can be configured to use any video and a set of themes by editing a simple XML configuration file, as specified by table A.1 and figure A.1.

XML element		Contents	Example
VideoURI		Path & filename of video file	videos\v_01.avi
Themes	DropdownName	Name of theme displayed in selection dropdown	Personal growth
	MarkerLetter	Letter displayed in marker on timeline	P
	MarkerColor	Color of marker letter	Green

Table A.1: Configuration file for main annotation tool

Configuration file example

```
<?xml version='1.0'?>

<document>

      <videoURI>videos\video.avi</videoURI>


      <themes>

            <theme>
                  <DropdownName>Move on</DropdownName>
                  <MarkerLetter>M</MarkerLetter>
                  <MarkerColor>Red</MarkerColor>
            </theme>

            <theme>
                  <DropdownName>Silver lining</DropdownName>
                  <MarkerLetter>S</MarkerLetter>
                  <MarkerColor>Green</MarkerColor>
            </theme>

            <theme>
                  <DropdownName>Broader perspective</DropdownName>
                  <MarkerLetter>B</MarkerLetter>
                  <MarkerColor>Yellow</MarkerColor>
            </theme>

            <theme>
                  <DropdownName>Time heals</DropdownName>
                  <MarkerLetter>T</MarkerLetter>
                  <MarkerColor>Blue</MarkerColor>
            </theme>

      </themes>

</document>
```

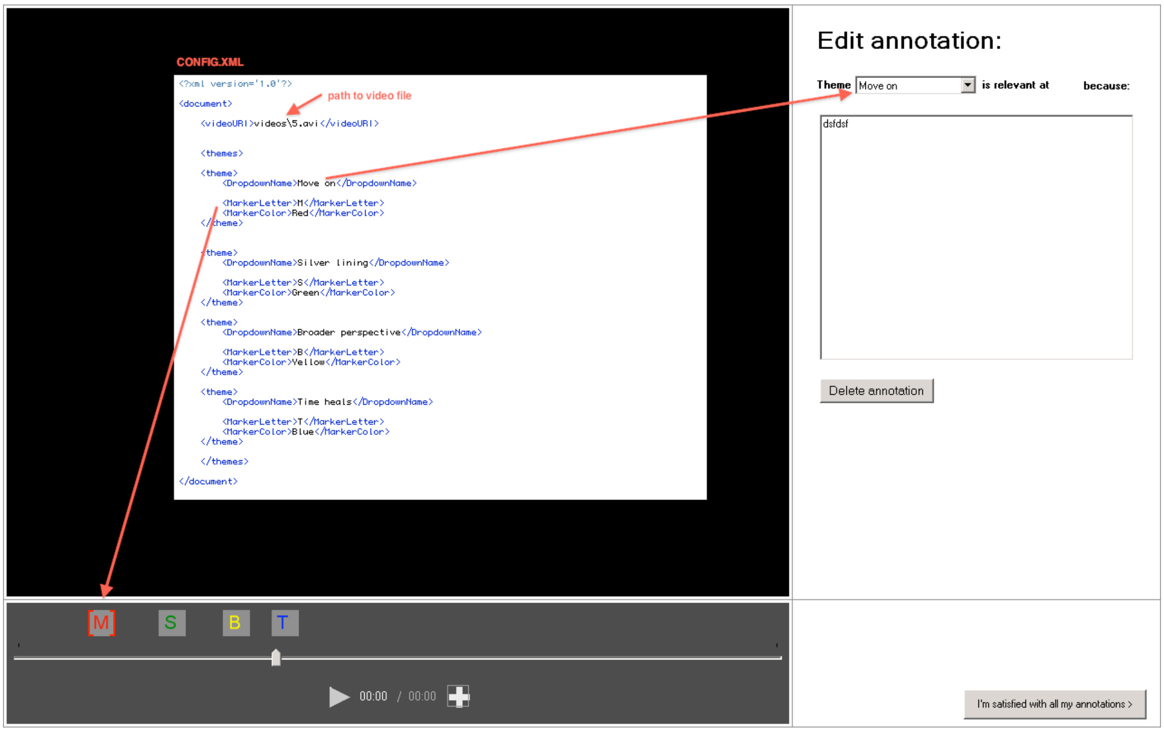


Figure A.1: Configuration file elements affecting the UI elements

Results (log of user input)

The results are stored in an XML file with a unique date/time stamp, and may be used for later analysis by researchers (so this functionality is not intended to be used by the participants).

XML element		Contents	Example
Participant		Unique ID entered by participant upon startup of the tool	0012
DateTime		When was the tool started	2015-05-06 13:52:03
VideoURI		Path & filename of video file	E:\videos\v_01.avi
Annotations	Theme	CR theme chosen by participant at a particular time position in the video	B – personal growth
	Position	Time position in the video that the theme was applied to	02:04
	Text	Text entered by the participant, explaining why he or she felt the selected theme was relevant at the position in the video	Now I know more about this type of criminal behavior, and that may help me improve the way I deal with similar situations in the future

Table A.2: Results from full video annotation tool stored in XML file

A.4 Tool version B: label while watching – results & log files

Results file

XML element		Contents	Example
Participant		Unique ID entered by participant upon startup of the tool	0012
DateTime		When was the tool started	2015-05-06 13:52:03
Video	Filename	Path & filename of video file	E:\videos\v_01.avi
	Theme	CR theme chosen by participant for this video	4

Table A.3: Results file for tool version B

Log file

XML element		Contents	Example
StartSession	Participant	Unique ID entered by participant upon startup of the tool	0012
	DateTime	When was the tool started	2015-05-06 13:52:03
StartVideo	VideoFilename	Path & filename of video file	videos\video3.avi
	DateTime	When was playback of video started	2015-05-06 13:52:04
Click	Theme	Theme that was clicked on	4
	CheckedValue	Value of checkbox of theme after click: "True" or "False"	True
	VideoPosition	Playback position in video when participant clicked	01:39
	DateTime	When was this theme clicked	11-6-2014 10:42:16

Table A.4: Log file for tool version B

APPENDIX B MATLAB SCRIPT TO REDUCE RAW PORTILAB EMG DATA INTO Z-SCORED BLINK MAGNITUDES PER IMAGE STIMULUS TYPE

```
function a = blinks_analyze_v03_plus_baseline (filename_prefix)

file_prefix = filename_prefix;
file_URI = strcat('portilab2_files\',file_prefix,'.S00')

output_graph_filename_prefix = strcat('pics\', file_prefix, '_');
output_graph_without_text_filename_prefix = strcat('pics_without_text\', file_prefix, '_');
csv_filename = strcat('csv\', file_prefix, '.csv');


% import mobi8 file using tms_read script
tms_read (file_URI);

data_blink = ans.data {1,1};
data_trigger = ans.data {2,1};


% remove "battery low" and "battery empty" signals from data channel of Mobi8, so that only the "trigger" signal is left

for i = 1 : size (data_trigger,2)
        if data_trigger(i) == 3
                data_trigger(i) = 1;
        end
        if data_trigger(i) == 5
                data_trigger(i) = 1;
        end
        if data_trigger(i) == 2
                data_trigger(i) = 0;
        end
        if data_trigger(i) == 4
                data_trigger(i) = 0;
        end
end



data_trigger(1) = 0;


sr = 2048; % sample rate

% 4th order Butterworth bandpass filter 28-500 Hz
[b,a]=butter(4,[28,500]/(sr/2));
filtered_data_blink = filtfilt(b,a,data_blink);

% rectify
abs_filtered_data_blink = abs (filtered_data_blink);

% smoothe with 40 Hz lowpass filter
[b,a]=butter(4, 40/(sr/2), 'low' );
lowpass_abs_data_blink = filtfilt(b,a,abs_filtered_data_blink);



% create a vector "ind" for the indices of the (onset of the) trigger
% and a vector "cnt" for the duration of the trigger (in samples)

t = [0,diff([data_trigger])];

ind = find (t==1);

cnt = find (t==-1);

cnt = cnt-ind;

% "ind" iformatted as "MM:SS.FFF"
ind_minutes = datestr(ind/sr/24/3600, 'MM:SS.FFF')

% cnt in milliseconden, i.e. per index the duration of the trigger in milliseconds
cnt_ms = 1000*cnt/sr;
```

```
% create array "amp" with the peak values from 20 ms to 200 ms after the noise startle probes

amp = zeros(1,size(cnt,2));

for i = 1 : size(cnt,2)
  amp(i) = max (abs_filtered_data_blink(ind(i)+cnt(i)+0.020*sr : ind(i)+cnt(i)+0.200*sr));
end



amp_smoothed = zeros(1,size(cnt,2));

for i = 1 : size(cnt,2)
  amp_smoothed(i) = max (lowpass_abs_data_blink(ind(i)+cnt(i)+0.020*sr : ind(i)+cnt(i)+0.200*sr));
end


% baseline period = 50 ms prior to 20 ms after stimulus onset

% baseline max amplitude
baseline_amp_smoothed_max = zeros(1,size(cnt,2));

for i = 1 : size(cnt,2)
  baseline_amp_smoothed_max(i) = max (lowpass_abs_data_blink(ind(i)+cnt(i)-0.050*sr : ind(i)+cnt(i)+0.020*sr));
end

% baseline mean amplitude
baseline_amp_smoothed_mean = zeros(1,size(cnt,2));

for i = 1 : size(cnt,2)
  baseline_amp_smoothed_mean(i) = mean (lowpass_abs_data_blink(ind(i)+cnt(i)-0.050*sr : ind(i)+cnt(i)+0.020*sr));
end

% baseline amplitude variance
baseline_abs_filtered_variance = zeros(1,size(cnt,2));

for i = 1 : size(cnt,2)
  baseline_abs_filtered_variance(i) =  var (abs_filtered_data_blink(ind(i)+cnt(i)-0.050*sr : ind(i)+cnt(i)+0.020*sr));
end


max_amp_minus_baseline_mean = amp_smoothed - baseline_amp_smoothed_mean;

max_amp_minus_baseline_mean_z = ( max_amp_minus_baseline_mean - mean (max_amp_minus_baseline_mean) ) / std(max_amp_minus_baseline_mean);




% 200 ms: start of Habituation phase (1)         [stimulus: 0]
%
% 20 ms: Habituation CS+                         [stimulus: 1]
%
% 40 ms: Habituation CS-                         [stimulus: 2]
%
% 60 ms: Habituation Noise Alone                 [stimulus: 3]
%
% 80 ms: Habituation FCS+                        [stimulus: 4]
%
% The last 4 triggers are immediately followed by a 40 ms white noise startle sound
%
%
%
% 300 ms: start of Conditioning phase (2)        [stimulus: 0]
%
% 100 ms: Conditioning CS+US                     [stimulus: 1]
%
% 120 ms: Conditioning CS-                       [stimulus: 2]
%
% 140 ms: Conditioning Noise Alone               [stimulus: 3]
%
% 160 ms: Conditioning FCS+                      [stimulus: 4]
%
% The last 4 triggers are immediately followed by a 40 ms white noise startle sound
%
% 400 ms: end of startle experiment (3)          [stimulus: 0]
```

```
% define a matrix with: duration of trigger (milliseconds), phase, image stimulus type

trigger_ms_phase_stimulus = [          20      1       1;...
                                       40      1       2;...
                                       60      1       3;...
                                       80      1       4;...
                                       100     2       1;...
                                       120     2       2;...
                                       140     2       3;...
                                       160     2       4;...
                                       200     1       -1;...
                                       300     2       -1;...
                                       400     3       -1;
                             ];


phase = zeros(1,size(cnt,2));
stimulus = zeros (1,size(cnt,2));


% use the duration of the trigger signal to determine the associated phase and image stimulus type
% (pulse code width decoding)
% checking for triggers that are within a range of 5ms of the defined trigger duration values

for i = 1 : size (cnt,2)

        for t = 1 : size (trigger_ms_phase_stimulus,1)

                if abs(cnt(i)-(0.001*trigger_ms_phase_stimulus(t,1)*sr)) < 0.5*0.020*sr
                        phase (i) = trigger_ms_phase_stimulus(t,2);
                        stimulus(i) = trigger_ms_phase_stimulus(t,3);
                end

        end

        if cnt(i) > 0.400*sr
                stimulus (i) = -1;
        end

end


data_matrix = [phase; stimulus; amp; amp_smoothed; cnt_ms];


% data matrix without the "phase" markers

data_matrix_without_phase_triggers = zeros (8, size(cnt,2) - sum (stimulus == 0) );

data_index = 1;

for i = 1 : size (cnt,2)

        % skip the "phase" markers
        if  stimulus(i) ~= -1

                data_matrix_without_phase_triggers (:,data_index) = [phase(i); stimulus(i); max_amp_minus_baseline_mean_z(i); amp(i);
amp_smoothed(i); baseline_amp_smoothed_max(i); baseline_amp_smoothed_mean(i); baseline_abs_filtered_variance(i) ];
                data_index = data_index +1;
        end
end


% save pictures of plots

pic_index = 1;

plot_filenames = cell (1, size(cnt,2) - sum (stimulus == 0) );


count_habituation = 0;
count_conditioning = 0;

count_stim_1_1 = 0;
count_stim_1_2 = 0;
count_stim_1_3 = 0;
count_stim_1_4 = 0;

count_stim_2_1 = 0;
count_stim_2_2 = 0;
count_stim_2_3 = 0;
count_stim_2_4 = 0;

amp_hab_cs_plus = [];
amp_hab_cs_min = [];
amp_hab_na = [];
amp_hab_fcs_plus = [];
amp_hab_cs_plus_timestamp = [];
```

```
amp_hab_cs_min_timestamp = [];
amp_hab_na_timestamp = [];
amp_hab_fcs_plus_timestamp = [];

amp_hab_cs_plus_baseline_max = [];
amp_hab_cs_min_baseline_max = [];
amp_hab_na_baseline_max = [];
amp_hab_fcs_plus_baseline_max = [];

amp_hab_cs_plus_baseline_variance = [];
amp_hab_cs_min_baseline_variance = [];
amp_hab_na_baseline_variance = [];
amp_hab_fcs_plus_baseline_variance = [];


amp_cond_cs_plus_us = [];
amp_cond_cs_min = [];
amp_cond_na = [];
amp_cond_fcs_plus = [];

amp_cond_cs_plus_us_timestamp = [];
amp_cond_cs_min_timestamp = [];
amp_cond_na_timestamp = [];
amp_cond_fcs_plus_timestamp = [];

amp_cond_cs_plus_us_baseline_max = [];
amp_cond_cs_min_baseline_max = [];
amp_cond_na_baseline_max = [];
amp_cond_fcs_plus_baseline_max = [];

amp_cond_cs_plus_us_baseline_variance = [];
amp_cond_cs_min_baseline_variance = [];
amp_cond_na_baseline_variance = [];
amp_cond_fcs_plus_baseline_variance = [];

all_baseline_max = [];
all_baseline_variance = [];


for i = 1 : size (cnt,2)

        % skip the "phase" markers
        if  stimulus(i) ~= -1

                hold off;
                plot (filtered_data_blink(ind(i)+cnt(i)+round(0.020*sr) : ind(i)+cnt(i)+round(0.200*sr)));
                hold all;
                plot (lowpass_abs_data_blink(ind(i)+cnt(i)+round(0.020*sr) : ind(i)+cnt(i)+round(0.200*sr)));

                axis([0,400,-500,500]);

                text_temp = '';


                if phase(i) == 1

                        count_habituation = count_habituation + 1;


                        if stimulus(i) == 1
                                count_stim_1_1 = count_stim_1_1 + 1;
                                text_temp = strcat('Habituation #', num2str(count_habituation), ', CS+ #',num2str(count_stim_1_1));

                                amp_hab_cs_plus = [amp_hab_cs_plus; max_amp_minus_baseline_mean_z(i)];
                                amp_hab_cs_plus_timestamp = [amp_hab_cs_plus_timestamp; ind(i)];
                                amp_hab_cs_plus_baseline_max = [amp_hab_cs_plus_baseline_max; baseline_amp_smoothed_max(i)];
                                amp_hab_cs_plus_baseline_variance          =          [amp_hab_cs_plus_baseline_variance;
baseline_abs_filtered_variance(i)];

                                all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
                                all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];

                        end
                        if stimulus(i) == 2
                                count_stim_1_2 = count_stim_1_2 + 1;
                                text_temp = strcat('Habituation #', num2str(count_habituation), ', CS- #',num2str(count_stim_1_2));
                                %text_temp = 'Habituation, CS-';

                                amp_hab_cs_min = [amp_hab_cs_min; max_amp_minus_baseline_mean_z(i)];
                                amp_hab_cs_min_timestamp = [amp_hab_cs_min_timestamp; ind(i)];
                                amp_hab_cs_min_baseline_max = [amp_hab_cs_min_baseline_max; baseline_amp_smoothed_max(i)];
                                amp_hab_cs_min_baseline_variance          =          [amp_hab_cs_min_baseline_variance;
baseline_abs_filtered_variance(i)];

                                all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
                                all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];

                                end
```

```
if stimulus(i) == 3
    count_stim_1_3 = count_stim_1_3 + 1;
    text_temp = strcat('Habituation #', num2str(count_habituation), ', NA #',num2str(count_stim_1_3));
    %text_temp = 'Habituation, NA';

    amp_hab_na = [amp_hab_na; max_amp_minus_baseline_mean_z(i)];
    amp_hab_na_timestamp = [amp_hab_na_timestamp; ind(i)];
    amp_hab_na_baseline_max = [amp_hab_na_baseline_max; baseline_amp_smoothed_max(i)];
    amp_hab_na_baseline_variance = [amp_hab_na_baseline_variance; baseline_abs_filtered_variance(i)];

    all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
    all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];

end
if stimulus(i) == 4
    count_stim_1_4 = count_stim_1_4 + 1;
    text_temp = strcat('Habituation #', num2str(count_habituation), ', FCS+ #',num2str(count_stim_1_4));
    %text_temp = 'Habituation, FCS+';

    amp_hab_fcs_plus = [amp_hab_fcs_plus; max_amp_minus_baseline_mean_z(i)];
    amp_hab_fcs_plus_timestamp = [amp_hab_fcs_plus_timestamp; ind(i)];
    amp_hab_fcs_plus_baseline_max = [amp_hab_fcs_plus_baseline_max; baseline_amp_smoothed_max(i)];
    amp_hab_fcs_plus_baseline_variance = [amp_hab_fcs_plus_baseline_variance; baseline_abs_filtered_variance(i)];

    all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
    all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];
end

else
if phase(i) == 2

    count_conditioning = count_conditioning + 1;

    if stimulus(i) == 1
        count_stim_2_1 = count_stim_2_1 + 1;
        text_temp = strcat('Conditioning #', num2str(count_conditioning), ', CS+ #',num2str(count_stim_2_1));

        %text_temp = 'Conditioning, CS+US';

        amp_cond_cs_plus_us = [amp_cond_cs_plus_us; max_amp_minus_baseline_mean_z(i)];
        amp_cond_cs_plus_us_timestamp = [amp_cond_cs_plus_us_timestamp; ind(i)];
        amp_cond_cs_plus_us_baseline_max = [amp_cond_cs_plus_us_baseline_max; baseline_amp_smoothed_max(i)];

        amp_cond_cs_plus_us_baseline_variance = [amp_cond_cs_plus_us_baseline_variance; baseline_abs_filtered_variance(i)];

        all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
        all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];
    end
    if stimulus(i) == 2
        count_stim_2_2 = count_stim_2_2 + 1;
        text_temp = strcat('Conditioning #', num2str(count_conditioning), ', CS- #',num2str(count_stim_2_2));

        %text_temp = 'Conditioning, CS-';

        amp_cond_cs_min = [amp_cond_cs_min; max_amp_minus_baseline_mean_z(i)];
        amp_cond_cs_min_timestamp = [amp_cond_cs_min_timestamp; ind(i)];
        amp_cond_cs_min_baseline_max = [amp_cond_cs_min_baseline_max; baseline_amp_smoothed_max(i)];

        amp_cond_cs_min_baseline_variance = [amp_cond_cs_min_baseline_variance; baseline_abs_filtered_variance(i)];

        all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
        all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];
    end
    if stimulus(i) == 3
        count_stim_2_3 = count_stim_2_3 + 1;
        text_temp = strcat('Conditioning #', num2str(count_conditioning), ', NA #',num2str(count_stim_2_3));

        %text_temp = 'Conditioning, NA';

        amp_cond_na = [amp_cond_na; max_amp_minus_baseline_mean_z(i)];
        amp_cond_na_timestamp = [amp_cond_na_timestamp; ind(i)];
        amp_cond_na_baseline_max = [amp_cond_na_baseline_max; baseline_amp_smoothed_max(i)];
        amp_cond_na_baseline_variance = [amp_cond_na_baseline_variance; baseline_abs_filtered_variance(i)];

        all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
        all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];
    end
    if stimulus(i) == 4
        count_stim_2_4 = count_stim_2_4 + 1;
        text_temp = strcat('Conditioning #', num2str(count_conditioning), ', FCS+ #',num2str(count_stim_2_4));

        %text_temp = 'Conditioning, FCS+';

        amp_cond_fcs_plus = [amp_cond_fcs_plus; max_amp_minus_baseline_mean_z(i)];
```

```
                        amp_cond_fcs_plus_timestamp = [amp_cond_fcs_plus_timestamp; ind(i)];
                        amp_cond_fcs_plus_baseline_max = [amp_cond_fcs_plus_baseline_max;
baseline_amp_smoothed_max(i)];
                        amp_cond_fcs_plus_baseline_variance = [amp_cond_fcs_plus_baseline_variance;
baseline_abs_filtered_variance(i)];

                        all_baseline_max = [all_baseline_max; baseline_amp_smoothed_max(i)];
                        all_baseline_variance = [all_baseline_variance; baseline_abs_filtered_variance(i)];
                    end
                end

            end

            %{

            filename = [output_graph_without_text_filename_prefix num2str(pic_index,'%03d')];
            plot_filenames {pic_index} = filename;
            print ('-dpng',  filename);


            text(10 , 470, file_prefix, 'Color', 'k', 'Interpreter','none');

            text_temp_1 = strcat('#', num2str(pic_index,'%03d'), {' @ '}, datestr(ind(i)/sr/24/3600, 'MM:SS.FFF'), ' (', num2str(ind(i)), ')' )
            text(10 , 430, text_temp_1, 'Color', 'k');

            text(10 , 390, text_temp, 'Color', 'k');

            text(200, 470, strcat({'Max: '}, num2str(amp(i),'%0.1f')) );

            text(200, 430, strcat({'Smoothed max: '},  num2str(amp_smoothed(i),'%0.1f')) );

            text(200, 390, strcat({'Baseline smoothed max: '},  num2str(baseline_amp_smoothed_max(i),'%0.1f')) );

            text(200, 350, strcat({'Baseline smoothed mean: '},  num2str(baseline_amp_smoothed_mean(i),'%0.1f')) );

            text(200, 310, strcat({'Baseline variance: '},  num2str(baseline_abs_filtered_variance(i),'%0.1f')) );

            hold off;




            filename = [output_graph_filename_prefix num2str(pic_index,'%03d')];

            plot_filenames {pic_index} = filename;

            print ('-dpng',  filename);

            %}

            pic_index = pic_index +1;

        end

end


data_matrix_without_phase_triggers_transposed =  transpose(data_matrix_without_phase_triggers);

csvwrite (csv_filename, data_matrix_without_phase_triggers_transposed (:,:) );



% replace blinks with very noise baseline by the mean of the previous and next blink

baseline_variance_outlier_threshold = median (all_baseline_variance) + 3 * mad(all_baseline_variance);

non_blinks =  find (all_baseline_variance > baseline_variance_outlier_threshold)

for i = 2 : size(amp_cond_cs_plus_us,1)-1
        if amp_cond_cs_plus_us_baseline_variance(i) > baseline_variance_outlier_threshold
            amp_cond_cs_plus_us(i) = mean([amp_cond_cs_plus_us(i-1) amp_cond_cs_plus_us(i+1)])
        end
end
for i = 2 : size(amp_cond_cs_min,1)-1
        if amp_cond_cs_min_baseline_variance(i) > baseline_variance_outlier_threshold
            amp_cond_cs_min(i) = mean([amp_cond_cs_min(i-1) amp_cond_cs_min(i+1)])
        end
end
for i = 2 : size(amp_cond_fcs_plus,1)-1
        if amp_cond_fcs_plus_baseline_variance(i) > baseline_variance_outlier_threshold
            amp_cond_fcs_plus(i) = mean([amp_cond_fcs_plus(i-1) amp_cond_fcs_plus(i+1)])
        end
end
for i = 2 : size(amp_cond_na,1)-1
        if amp_cond_na_baseline_variance(i) > baseline_variance_outlier_threshold
            amp_cond_na(i) = mean([amp_cond_na(i-1) amp_cond_na(i+1)])
```

```
        end
end


csvwrite (['csv\' file_prefix '_amp_cond_cs_plus_us.csv'], amp_cond_cs_plus_us);
csvwrite (['csv\' file_prefix '_amp_cond_cs_min.csv'], amp_cond_cs_min);
csvwrite (['csv\' file_prefix '_amp_cond_fcs_plus.csv'], amp_cond_fcs_plus);
csvwrite (['csv\' file_prefix '_amp_cond_na.csv'], amp_cond_na);


% plot conditioning amplitudes
hold off;
plot (amp_cond_fcs_plus, 'rx' );
hold all;
plot (amp_cond_cs_plus_us, 'g+' );
hold all;
plot (amp_cond_cs_min, 'b*' );
hold all;
plot (amp_cond_na, 'ko' );
hold all;
legend ('FCS+', 'CS+US', 'CS-', 'NA');

title([file_prefix ' conditioning phase'],'Interpreter','none')
xlabel('stimulus number')
ylabel('amplitude (normalized)')

filename = [output_graph_filename_prefix '_amplitudes_normalized_conditioning_dots'];
plot_filenames {pic_index} = filename;
print ('-dpng',  filename);


% plot conditioning amplitudes
hold off;
plot (amp_cond_fcs_plus, 'r' );
hold all;
plot (amp_cond_cs_plus_us, 'g' );
hold all;
plot (amp_cond_cs_min, 'b' );
hold all;
plot (amp_cond_na, 'k' );
hold all;
legend ('FCS+', 'CS+US', 'CS-', 'NA');

title([file_prefix ' conditioning phase'],'Interpreter','none')
xlabel('stimulus number')
ylabel('amplitude (normalized)')

filename = [output_graph_filename_prefix '_amplitudes_normalized_conditioning_lines'];
plot_filenames {pic_index} = filename;
print ('-dpng',  filename);



% plot habituation amplitudes
hold off;
plot (amp_hab_fcs_plus, 'rx' );
hold all;
plot (amp_hab_cs_plus, 'g+');
hold all;
plot (amp_hab_cs_min, 'b*' );
hold all;
plot (amp_hab_na, 'ko' );
hold all;
legend ('FCS+', 'CS+US', 'CS-', 'NA');

title([file_prefix ' habituation phase'],'Interpreter','none')
xlabel('stimulus number')
ylabel('amplitude (normalized)')


filename = [output_graph_filename_prefix '_amplitudes_normalized_habituation_dots'];
plot_filenames {pic_index} = filename;
print ('-dpng',  filename);
```

APPENDIX C MATLAB SCRIPT FOR PEST, PRODUCING FAST & SLOW EFFECT& T-TEST PER PARTICIPANT

```
% participant#, block#, trial#, emotion, reaction_time_milliseconds, correct_response
% (emotion: 0=neutral, -1=negative)

stroop_M = csvread ('merged_data_files_filtered_v01_excel.csv');

stroop_M_without_outliers = [];

% http://www.sciencedirect.com/science/article/pii/S0022103113000668


trials_per_participant = 160;

participants = size (stroop_M,1) ./ trials_per_participant;


% remove outliers

for i = 1 : participants

	start_trial = 1 + (i-1)*(trials_per_participant);
	end_trial = i*trials_per_participant;

	outlier_min = median( stroop_M (start_trial:end_trial,5) ) - 3 * mad( stroop_M (start_trial:end_trial,5) );
	outlier_max = median( stroop_M (start_trial:end_trial,5) ) + 3 * mad( stroop_M (start_trial:end_trial,5) );

	for ii = start_trial : end_trial

		if (outlier_min < stroop_M (ii,5)) && (stroop_M (ii,5) < outlier_max)
			stroop_M_without_outliers = [stroop_M_without_outliers; stroop_M(ii,:)];
		end

	end

end


neutral_neutral = [];
neutral_negative = [];
negative_neutral = [];
negative_negative = [];


for n = 2 : size (stroop_M_without_outliers,1)


	% if same participant & same block
	if stroop_M_without_outliers(n,1) == stroop_M_without_outliers(n-1,1) && stroop_M_without_outliers(n,2) ==
stroop_M_without_outliers(n-1,2)

		% column 4 is emotion: 0 = neutral, -1 = negative
		if stroop_M_without_outliers(n-1,4) == 0 && stroop_M_without_outliers(n,4) == 0
			% insert: (participant#, RT, accurate)
			neutral_neutral = [neutral_neutral; stroop_M_without_outliers(n,1) stroop_M_without_outliers(n,5)
stroop_M_without_outliers(n,6) ];
		end
		if stroop_M_without_outliers(n-1,4) == 0 && stroop_M_without_outliers(n,4) == -1
			neutral_negative = [neutral_negative; stroop_M_without_outliers(n,1) stroop_M_without_outliers(n,5)
stroop_M_without_outliers(n,6) ];
		end
		if stroop_M_without_outliers(n-1,4) == -1 && stroop_M_without_outliers(n,4) == 0
			negative_neutral = [negative_neutral; stroop_M_without_outliers(n,1) stroop_M_without_outliers(n,5)
stroop_M_without_outliers(n,6) ];
		end
		if stroop_M_without_outliers(n-1,4) == -1 && stroop_M_without_outliers(n,4) == -1
			negative_negative = [negative_negative; stroop_M_without_outliers(n,1) stroop_M_without_outliers(n,5)
stroop_M_without_outliers(n,6) ];
		end

	end

end
```

```
% Stroop fast effect: only if emotion (n-1) == 0 : RT (n, negative) - RT (n, neutral)  / so previous trial can only be neutral
% Stroop slow effect: only if emotion (n) == 0 : RT (n, negative) - RT (n, neutral) / so previous trial can be either negative or neutral, but
current trial must be neutral



neutral_neutral_mean = [];
neutral_negative_mean = [];
negative_neutral_mean = [];
negative_negative_mean = [];

fast_ttest = [];
slow_ttest = [];

participant_nrs = unique (negative_negative (:,1));

idx = [];

for p = 1:size(participant_nrs,1)

	participant_nrs(p)

	holm_M = [];

	idx = (neutral_neutral (:,1) == participant_nrs(p));
	neutral_neutral_participant =  neutral_neutral (idx,:);
	mean_temp = mean( neutral_neutral (idx,2));
	accurate_perc_temp = sum(neutral_neutral(idx,3)) ./ size(neutral_neutral (idx,3));
	neutral_neutral_mean = [neutral_neutral_mean; mean_temp ttest_temp accurate_perc_temp];
	holm_M = [holm_M; neutral_neutral_participant(:,2) 1*ones(size(neutral_neutral_participant, 1), 1)];


	idx = neutral_negative (:,1) == participant_nrs(p);
	neutral_negative_participant =  neutral_negative (idx,:);
	mean_temp = mean( neutral_negative (idx,2));
	accurate_perc_temp = sum(neutral_negative(idx,3)) ./ size(neutral_negative (idx,3));
	neutral_negative_mean = [neutral_negative_mean; mean_temp ttest_temp accurate_perc_temp];
	holm_M = [holm_M; neutral_negative_participant(:,2) 2*ones(size(neutral_negative_participant, 1), 1)];

	idx = negative_neutral (:,1) == participant_nrs(p);
	negative_neutral_participant =  negative_neutral (idx,:);
	mean_temp = mean( negative_neutral (idx,2));
	accurate_perc_temp = sum(negative_neutral(idx,3)) ./ size(negative_neutral (idx,3));
	negative_neutral_mean = [negative_neutral_mean; mean_temp ttest_temp accurate_perc_temp];
	holm_M = [holm_M; negative_neutral_participant(:,2) 3*ones(size(negative_neutral_participant, 1), 1)];

	idx = negative_negative (:,1) == participant_nrs(p);
	negative_negative_participant =  negative_negative (idx,:);
	mean_temp = mean( negative_negative (idx,2));
	accurate_perc_temp = sum(negative_negative(idx,3)) ./ size(negative_negative (idx,3));
	negative_negative_mean = [negative_negative_mean; mean_temp ttest_temp accurate_perc_temp];
	holm_M = [holm_M; negative_negative_participant(:,2) 4*ones(size(negative_negative_participant, 1), 1)];

	[fast_ttest_participant_h, fast_ttest_participant_p] = ttest2(neutral_negative_participant(:,2), neutral_neutral_participant(:,2));
	[slow_ttest_participant_h, slow_ttest_participant_p] = ttest2(negative_neutral_participant(:,2), neutral_neutral_participant(:,2));

	fast_ttest = [fast_ttest; fast_ttest_participant_p];
	slow_ttest = [slow_ttest; slow_ttest_participant_p];

	holm (holm_M);

end



csvwrite ('stroop_met_ttest.csv', [participant_nrs neutral_neutral_mean(:,[1:2]) neutral_negative_mean(:,[1:2]) negative_neutral_mean(:,[1:2])
negative_negative_mean(:,[1:2]) fast_ttest slow_ttest]);
```

APPENDIX D - PEST FAST & SLOW EFFECT REACTION TIMES AND ERROR RATES

Table D.1
PEST Fast & slow effect reaction times and error rates as a function of valence in trial *n* and trial *n*-1
Note: The fast effect is calculated here as the difference between negative and neutral trials after neutral trials. The slow effect is computed as the difference between neutral trials following negative trials and neutral trials following neutral trials. Reaction times in ms, error rates (in parentheses) in percentages.

	Trial *n* – 1					
	Neutral		Negative		Slow effect	
Trial *n*						
Neutral	807	(3.0)	805	(2.3)	-2	(-0.7)
Negative	798	(2.3)	786	(1.7)		
Fast effect	-9	(-0.7)				

T-tests revealed only a single significant result - the fast effect for one participant (p=0.002).

APPENDIX E - FORMS USED IN THE EXPERIMENT

E.1: Consent form

Welcome to the experiment "Sensitivity of Emotional Stroop Test and Startle Paradigm to Negative Film Images". Please read the following carefully.

If you have, or think you have, any of the conditions below, you are exempt from participation. You do not need to tell experiment staff which condition you have. Simply say you are excluded by one of the exemptions. You will still receive the same compensation.

Exclusions:

- Colour Blindness
- Face Blindness
- Anxiety Disorders or Depression
- Startle disorders
- Severe psychological trauma
- Memory or learning disorders
- Hearing problems

You will be asked to watch three videos that contain scenes of horror and distress and may be found disturbing by some viewers. These videos are about 30seconds long and are played twice.

The films used in the experiment depict events from documentary footage or fictional movies. Some participants may experience negative intrusive recollections following the experiment. If it occurs, it should not last more than a week or so beyond the experiment and should not affect participants' daily life or cause more than a minor, momentary disturbance. If after the experiment you experience any problems due to intrusive memories, or otherwise from this experiment, please contact the experiment operators (below).

You will also be shown a series of pictures and asked to identify the colour using a button box. These pictures may include still images from the videos.

Finally, experiment staff will attach two sensors to the skin below one of your eyes, using removable tape. He or she will also ask you to put on a vest or harness, with an air-delivery tube directed at your throat. The following part of the experiment involves seeing and hearing pictures and sounds. These pictures may also include still images from the videos watched previously. The sounds are no more than .5sec in duration and may be as loud as 108dB, which is loud but will not cause any hearing damage. Puffs of air will be administered that may cause discomfort but should not be painful. This part of the experiment is meant to induce a startle response. (Schrikreflex)

If you wish to stop the experiment at any time or to withdraw your participation, please inform the experiment staff. You will receive full compensation regardless of completion.

No identifying data will be collected. The data will be stored at the TU Delft and only accessible to experiment staff.

E.2: Self report scales

E.2.1: BRIEF RESILIENCE SCALE (BRS)

Please indicate the extent to which you agree with each of the following statements by using the following scale: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree.

1. I tend to bounce back quickly after hard times
2. I have a hard time making it through stressful events (R)
3. It does not take me long to recover from a stressful event
4. It is hard for me to snap back when something bad happens (R)
5. I usually come through difficult times with little trouble
6. I tend to take a long time to get over set-backs in my life (R)

(R) = reversed scale

E.2.2: THE CORE SELF-EVALUATIONS SCALE (CSES)

Instructions
Below are several statements about you with which you may agree or disagree. Using the response scale below, indicate your agreement or disagreement with each item by placing the appropriate number on the line preceding that item.

Strongly Disagree = 1 Disagree =2 Neutral =3 Agree =4 Strongly Agree =5

1. ____ I am confident I get the success I deserve in life.
2. ____ Sometimes I feel depressed. (r)
3. ____ When I try, I generally succeed.
4. ____ Sometimes when I fail I feel worthless. (r)
5. ____ I complete tasks successfully.
6. ____ Sometimes, I do not feel in control of my work. (r)
7. ____ Overall, I am satisfied with myself.
8. ____ I am filled with doubts about my competence. (r)
9. ____ I determine what will happen in my life.
10. ____ I do not feel in control of my success in my career. (r)
11. ____ I am capable of coping with most of my problems.
12. ____ There are times when things look pretty bleak and hopeless to me. (r)

Notes: r=reverse-scored. This measure is non-proprietary (free) and may be used without permission.

E.2.2: THE INTERNATIONAL POSITIVE AND NEGATIVE AFFECT SCHEDULE SHORT FORM (I-PANAS-SF)

(Thompson, 2007)

Question: Thinking about yourself and how you normally feel, to what extent do you generally feel:

Interval measure: *never {*1 2 3 4 5} *always*

Items in order:

1) Upset
2) Hostile
3) Alert
4) Ashamed
5) Inspired
6) Nervous
7) Determined
8) Attentive
9) Afraid
10) Active

Scoring: Add the scores for items 1,2,4,6,9 for the negative affect score. Add the scores for items 3,5,7,8, 10 for the positive affect score.

APPENDIX F - STILLS FROM VIDEOS SHOWN IN THE EXPERIMENT

F.1 Stills from negative videos

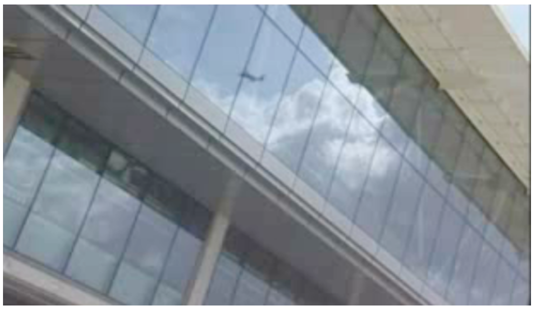
F.1.1 MURDER SCENE FROM THE FEATURE FILM "AMERICAN HISTORY X"

F.1.2 CLIP FROM A DOCUMENTARY ABOUT THE RWANDAN GENOCIDE

F.1.3 CLIP FROM A NEWS REPORT ON THE ABUSE AND MURDER OF A BABY

F.2 Stills from neutral videos

F.2.1 NEWS ITEM ON A NEW TERMINAL FOR HEATHROW AIRPORT



F.2.2 News item on the building of Beijing airport

F.2.3 Weather report 1

F.2.4 Weather report 2

F.3 Still from American History X murder scene, used in FCS+ condition of the FPS phase of the experiment

APPENDIX G – EVALUATION FORM FOR VIDEO ANNOTATION TOOLS

Evaluatie formulier Reflectie applicatie

Hieronder staan een aantal stellingen en vragen. Geef bij de stellingen aan hoe oneens of eens je het met elke stelling bent door een kruis in de cirkel bij je antwoord te zetten. Let op, er zijn geen juiste of onjuiste antwoorden. Het is alleen van belang zo eerlijk mogelijk de vragen te beantwoorden.

Algemeen

Stelling: *Ik vond de 4 kijkmanieren nuttig*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Stelling: *Ik vond het labelen van de video's een nuttige oefening om me de 4 kijkmanieren eigen te maken*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Stelling: *De video-labeling oefeningen zullen me helpen om eventueel toekomstige verontrustende in een ander perspectief te zien*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Video

Kindermisbruik

Stelling: *De inhoud van deze video was verontrustend*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Orkaan Katrina

Stelling: *De inhoud van deze video was verontrustend*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

	Rwanda Genocide

Stelling: *De inhoud van deze video was verontrustend*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Videospeler: Stoppen, labelen & annoteren
(versie waarbij je zelf video kan starten en stoppen, labels toevoegen, en aantekeningen maken)

Vraag: Ik heb eerst de hele video bekeken voordat ik begon met het labelen
O Klopt, eerst afgekeken **O** Nee, gelijk begonnen met labelen

Stelling: *Het labelen leidde me erg af van het kijken naar de video*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Stelling: *Ik vond deze versie van de videospeler makkelijk te gebruiken*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Videospeler: Achteraf labelen
(versie waarbij je eerst de film bekijkt en dan labels moet kiezen, met uitleg waarom)

Vraag: Ik heb eerst de hele video bekeken voordat ik begon met het labelen
O Klopt, eerst afgekeken **O** Nee, gelijk begonnen met labelen

Stelling: *Het labelen leidde me erg af van het kijken naar de video*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Stelling: *Ik vond deze versie van de videospeler makkelijk te gebruiken*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Videospeler: Realtime labelen
(versie waarbij je moet labelen wanneer de video bekijkt, zonder pauze mogelijkheid)

Vraag: Ik heb eerst de hele video bekeken voordat ik begon met het labelen
O Klopt, eerst afgekeken **O** Nee, gelijk begonnen met labelen

Stelling: *Het labelen leidde me erg af van het kijken naar de video*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Stelling: *Ik vond deze versie van de videospeler makkelijk te gebruiken*
O zeer mee oneens **O** mee oneens **O** neutraal **O** mee eens **O** zeer mee eens

Tot slot

Vraag: Welke versie van de videospelers vind je het beste om de kijkmanieren te leren gebruiken

O De **Stoppen, labelen & annoteren** videospeler
O De **Realtime labelen** videospeler
O De **Achteraf labelen** videospeler

Plaats de video-labeling oefening van (1) makkelijkste, (2) tussen in, tot (3) moeilijkste.
(Zet cijfer tussen de bIokhaken)

[ ] De **Stoppen, labelen & annoteren** videospeler
[ ] De **Realtime labelen** videospeler
[ ] De **Achteraf labelen** videospeler

Schrijf je opmerkingen over de videospelers:

BEDANKT!!