Simulated (Un-)armed Confrontations and Police Decision Making: 
Examining Influencing Factors on Tactical Decision Making

Thesis submitted in accordance with the requirements of the 
University of Liverpool for the degree of Doctor of Philosophy

by

Mario S. Staller

October 2016
Simulated (Un-)armed Confrontations and Police Decision Making:
Examining Influencing Factors on Tactical Decision Making

Mario S. Staller

Abstract

This thesis examines how the tactical decision making of (mainly) German Police Officers was influenced by several factors that are salient in training and operational environments. First, the effects of conventional ammunition (CA) versus non-lethal training ammunition (NLTA) in training settings on psychophysiological parameters and working memory were compared. It was found, that there is no difference in psychophysiological response to a demanding training exercise with regards to different ammunition used. This indicates that there are no differences between CA and NLTA with regards to representativeness of training. Second, a tactical scenario and a physical exercise were compared with regards of the effects on executive functioning. The findings yielded that executive functioning was equally enhanced due to a physical exercise compared to tactical scenario. This leads to the conclusion that cognitive adaptions are probably caused by physical demand. Third, the effects of previous experience in (simulated) violent confrontations on threat-related attentional biases and risk taking were investigated. The results showed, that previous experience as a police officer or a martial artist had no effect on threat-related attentional biases or risk taking. Fourth, the impact of ego depletion on police officers when provoked by a role player in a scenario were quantitatively assessed. The findings revealed that a state of ego depletion shortened the time when police officers displayed offensive aggression towards a provocative role player compared to non-depleted officers. Taken together the current work provided evidence, that: (a) tactical decision making of officers is influenced by physiological load and ego depletion, and (b) the use of CA and NLTA does not influence psychophysiological demand.
In line with previous research on human defensive behaviour, the observed cognitive shifts under conditions of threat can be interpreted as an adaptive behaviour in order to cope with the demand at hand. However, the current results indicate that these cognitive shifts may be mediated by physiological arousal. Further research is needed to further clarify this relationship.

With regards to threat-related attentional bias, the current work indicates that current paradigms probably are not capable of capturing functional threat-related attentional bias. Furthermore, it could be argued, that current training settings do not provide enough valid cues in order to learn functional threat-related attentional bias. Therefore, future work should employ more valid cues in the context of police use of force to further investigate the development of functional threat-related attentional bias.

Concerning ego depletion and self-regulation, the current work for the first time demonstrates that depleted self-control resources transfer to observable physical aggression. Since there is evidence that self-control performance varies across contexts the obtained results are important for both the aggression and the police use of force domain. As such the results are in line with current theories of self-control.

From a practical perspective the results shed light on the design of representative learning and testing environment in the police of force domain: The use of NLTA in the police use of force training should be broadened at the expense of CA, allowing for safer and more representative training settings. Furthermore, police training should emphasize enhancing physical fitness and self-control. However, further research aiming at developing self-control in the police use of force context is clearly needed.
Acknowledgements

Since this PhD journey comes to an end, there are a few individuals I would like to thank in particular.

To Jon – for providing me with excellent supervision and giving me the opportunity to pursue this PhD. I really loved the discussions we had. Your comments on my arguments always hit the weak spots. Only through this precise and direct feedback I could become better in my arguments. In the martial arts we call this process “sparring”. So thanks for sparring with me whenever I wanted / needed / had to spar. Good training partners are rare to find – and you were one of the best.

To Benny – for being there, whenever a helping hand was needed. For bouncing ideas back and forth, for taking the time to revise, comment and argue with me on my ideas and written work. And finally and foremost for being such an excellent friend. Now it’s your turn 😊.

To OBi – for supporting me always and everywhere, as a colleague and a friend. Without the vast amounts of cappuccinos (and discussions) we had, I would have not been this far. Thank you, my friend.

To OWi – for making this PhD project possible. Thank you for your dedication, time and curiosity surrounding this project.

To Torsten, Isabel and Valentina – for helping collecting the data and reminding me about things, whenever I was close to forgetting.

To all the cops from the Hessian Police - for participating in this work. And a big thank you to all the superiors for letting them participate.

To all train2survive members – for always being ready to take part in experiments and for being interested in optimizing the training for conflict situations. You are a bunch of amazing people.

To Barbara – for supporting me in the beginning of this journey. Thank you.
And finally to my personal superheroes: Batman, Spiderman and Wonder Women. Thank you for letting me pursue my dreams. As easy as it sounds – I know how hard this sometimes is. I express my deepest gratiitiude for that.

“I heard about you. Didn’t think you were real.”

“I’m real when it’s useful”

(Justice League)
## Contents

Abstract .......................................................................................................................... 2  
Acknowledgements ........................................................................................................ 4  
Contents .......................................................................................................................... 6  
List of Tables .................................................................................................................. 15  
List of Figures .................................................................................................................. 17  
Glossary of Abbreviations .............................................................................................. 20

### Chapter 1:

**Introduction** ................................................................................................................. 23  
1.1 Background ................................................................................................................. 24  
1.2 Statement of the Problem ............................................................................................ 28  
1.3 Significance of the Thesis ............................................................................................ 28  
1.4 General Aims ............................................................................................................... 29  
1.5 Structure of the Thesis ............................................................................................... 30  
1.6 Summary ..................................................................................................................... 31

### Chapter 2:

**Background Literature** ................................................................................................. 32  
2.1 Human Defensive Behaviour ....................................................................................... 33  
2.1.1 Complexity of Defensive Behaviours ...................................................................... 37  
2.1.2 Challenges in Studying Human Defensive Behaviour ........................................... 39  
2.2 Police Use of Force ..................................................................................................... 40  
2.2.1 Police Use of Force Models – Training Models ....................................................... 41  
2.2.2 Police Use of Force Models – Explaining Models ................................................... 45  
2.2.2.1 The Transactional Model of Police Use of Force by Binder and Scharf (1980) ................................................................. 45  
2.2.2.2 The Model of Use of Control Tactics and Technologies in Police-Citizen Encounters by Amendola (1995) ......................... 47  
2.2.2.3 The Model of an Armed Confrontation by Boulton (2014) .............................. 49  
2.2.3 German Police Use of Force Framework ............................................................... 50  
2.3 Stress and Policing ..................................................................................................... 52  
2.3.1 Stress, Homeostasis and Allostasis ......................................................................... 52
4.3.2.1 Cognitive Measures ......................................................... 135
4.3.2.2 Physiological Measures .................................................... 135
4.3.2.3 Weapons and Safety Equipment ........................................ 135
4.3.3 Procedure ............................................................................. 135
4.3.4 Analysis of HRV data ............................................................. 137
4.3.5 Statistical Analysis ............................................................... 138
4.4 Results ..................................................................................... 139
4.4.1 Phonemic Fluency (Study 1.1) ............................................... 139
4.4.2 Digits Span (Study 1.2) ........................................................ 141
4.4.3 Corsi Block-Tapping Task (Study 1.3) ...................................... 143
4.4.4 Heart Rate Variability .......................................................... 144
4.5 Discussion ................................................................................. 152
4.5.1 Differences Between the Use of CA and NLTA in Training and Testing Environments ......................................................... 152
4.5.2 Psychophysiological Demand Under Physical Pressure .......... 153
4.5.3 Cognitive Functioning ............................................................ 156
4.5.4 Weaknesses and Limitations .................................................. 160
4.5.5 Summary of Chapter 4 .......................................................... 161

Chapter 5:
Scenario vs. Exercise ..................................................................... 162
5.1 Introduction ................................................................................. 163
5.2 Executive Functioning in Police Use of Force Events ............... 164
  5.2.1.1 Exploratory Work of A. P. J. Roberts (2012) on Use of Force Situations and Executive Functioning ......................................................... 165
    5.2.1.1.1 Effects on Verbal Fluency Tasks ............................................... 165
    5.2.1.1.2 Effects on Stroop Task ........................................................ 166
    5.2.1.1.3 Effects on Digit Span Task, Letter Number Sequencing Task, Symbol Search Task and Trail Making Test ............... 166
    5.2.1.1.4 Effects of Corsi Block-Tapping Task and Self Ordering Pointing Task .......................................................... 167
  5.2.1.2 Subsequent Work of Boulton (2014) on Use of Force Situations and Different Components of WM and the Influence of Experience ........................................ 168
5.2.1.3 Work of Condon (2015) on Physical Combat and Executive Functioning .......................................................... 170
5.2.1.4 Work of Taverniers et al. (2011) on the Effects of representative Testing Environments in Police Use of ForceTraining .......................................................... 171
5.2.1.5 Summary of Findings and Rationale for Study 2 .......... 171
5.2.2 Aim of Study 2 ................................................................. 172
5.3 Hypotheses ................................................................. 177
5.4 Methods ........................................................................ 177
5.4.1 Participants .................................................................. 177
5.4.2 Materials and Apparatus .................................................. 178
5.4.2.1 Phonemic Fluency as a Measure of Executive Functioning of WM .................................................................. 178
5.4.2.2 Physiological Arousal .................................................. 178
5.4.2.3 Safety Equipment ......................................................... 178
5.4.3 Procedure ..................................................................... 178
5.4.3.1 Treatment for Scenario and Exercise Group .................. 179
5.4.3.2 General Procedure ...................................................... 179
5.4.4 Analytic Strategy ............................................................ 180
5.5 Results ............................................................................ 181
5.5.1 Sample Characteristics .................................................... 181
5.5.2 Phonemic Fluency Task .................................................... 181
5.5.3 Physiological Arousal ...................................................... 184
5.5.3.1 Systolic Blood Pressure ............................................... 185
5.5.3.2 Diastolic Blood Pressure .............................................. 186
5.5.3.3 Heart Rate ................................................................. 186
5.5.3.4 Core Temperature ....................................................... 187
5.6 Discussion ....................................................................... 188
5.6.1 Physiological Arousal Under Physical and Tactical Demand ...... 188
5.6.2 Executive Functioning Under Physical and Tactical Demand ..... 190
5.6.3 Limitations ................................................................. 194
5.6.4 Summary ................................................................. 196
## Chapter 6:

**Threat Attentional Bias and Risk Taking**

### 6.1 Introduction

6.1.1 Threat-Related Attentional Bias

6.1.1.1 Attentional Bias and Fear and Anxiety

6.1.1.2 Temporal Structure of Attentional Bias

6.1.1.3 Measure of Attentional Bias

6.1.1.3.1 The Dot Probe Task

6.1.1.3.2 The Emotional Stroop Task

### 6.2 Risk Taking

### 6.3 Hypotheses

### 6.4 Methods

6.4.1 Participants

6.4.2 Materials

6.4.2.1 State-Trait Anxiety Inventory

6.4.2.2 Dot Probe Task

6.4.2.3 Emotional Stroop Task

6.4.2.4 Balloon Analogue Risk Task

6.4.3 Procedure

### 6.5 Results

6.5.1 Sample Characteristics

6.5.2 State-Trait Anxiety Inventory

6.5.3 Dot Probe Task

6.5.4 Emotional Stroop Task

6.5.5 Balloon Analogue Risk Task

### 6.6 Discussion

6.6.1 Threat-Related Attentional Biases in Police Officers

6.6.2 Risk Taking in Police Officers

6.6.3 Differences in State and Trait Anxiety

6.6.4 Limitations

### 6.7 Summary of Chapter 6
Chapter 7:

Ego Depletion ................................................................. 235

7.1 Introduction .................................................................. 236
  7.1.1 Aim of the Study ...................................................... 237
  7.1.2 Hypotheses ............................................................ 237

7.2 Methods ..................................................................... 238
  7.2.1 Participants ............................................................ 238
  7.2.2 Materials ............................................................... 239
  7.2.2.1 Brief Mood Introspection Scale ......................... 239
  7.2.2.2 Cardiovascular Measures .................................. 239
  7.2.2.3 Video Stimuli for Ego Depletion Manipulation (Video Task) in Study 4.1 .................................................. 239
  7.2.2.4 Cold Pressor Task for Ego Depletion Manipulation in Study 4.2 .......................................................... 240
  7.2.2.5 Employed Scenario in Study 4.1 and 4.2 .......... 240

7.2.3 Procedure .............................................................. 244
  7.2.3.1 General Procedure of Study 4.1 and Study 4.2 .... 244
  7.2.3.2 Experimental vs. Control Condition in Study 4.1 (Video Task) ............................................................ 246
  7.2.3.3 Experimental vs. Control Condition in Study 4.2 (Cold Pressor Task) ........................................................ 246

7.2.4 Analysis of HRV data ............................................... 247

7.2.5 Statistical Analysis ................................................ 247

7.3 Results of Study 4.1 – Video Task .................................. 249
  7.3.1 Effects of Manipulation on Manipulation Check Questions and Mood .......................................................... 249
  7.3.2 Effects of Manipulation on Offensive Aggression .... 250
  7.3.3 Phonemic Fluency Task .......................................... 250
  7.3.4 Cardiovascular Measures ....................................... 254

7.4 Results of Study 4.2 – Cold Pressor Task ...................... 257
  7.4.1 Effects of Manipulation on Manipulation Check Questions and Mood .......................................................... 257
  7.4.2 Effects of Ego Depletion on Offensive Aggression .... 258
  7.4.3 Phonemic Fluency Task .......................................... 259
Chapter 8:
General Discussion ................................................................. 280
8.1 Key Findings ........................................................................... 281
  8.1.1 Cognitive Shifts Under Threat ........................................... 283
  8.1.2 Representative Learning Environments for Police Use of Force Training ........................................ 288
  8.1.3 Threat-Related Attentional Bias ....................................... 290
  8.1.4 Ego Depletion and Self-Regulation .................................. 291
8.2 Limitations ............................................................................. 292
  8.2.1 Human Defensive Behaviour ........................................... 292
  8.2.2 Physiological Measures ................................................... 293
  8.2.3 Cognitive Measures ......................................................... 293
  8.2.4 Police Data Challenges ..................................................... 294
8.3 Theoretical Implications .......................................................... 295
  8.3.1 Implications for Theories of Human Defensive Behaviour and Executive Functioning ................................. 295
  8.3.2 Implication for Theories of Self-Regulation and Ego Depletion ................................................................. 295
  8.3.3 Implications for Theories of Aggression ............................ 296
  8.3.4 Applications to Methodology ............................................. 297
  8.3.5 Future Research ............................................................... 297
8.4 Practical Implications ................................................................. 298
  8.4.1 Applications to Police Policy .............................................. 299
  8.4.2 Applications for Police Use of Force Training .................... 300
    8.4.2.1 Representative Learning and Testing Environments ....... 300
    8.4.2.2 Self-Control Training ................................................... 301
    8.4.2.3 Physical Fitness of Police Officers ............................... 302
8.5 Conclusion ......................................................................................... 302

References .......................................................................................... 304

Appendices ........................................................................................ 420
Appendix A: Ethical Approval ............................................................... 421
Appendix B: Consent Form – Study 1 and 2 ......................................... 422
Appendix C: Online Consent Form - Study 3 ....................................... 424
Appendix D: Stait-Trait Anxiety Inventory ........................................... 426
Appendix E: Words Used in the Dot Probe Task ................................. 430
Appendix F: Words Used in the Emotional Stroop Task ....................... 432
Appendix G: Consent Form - Study 4 ................................................ 433
Appendix H: Information Sheet - Study 4 ........................................... 435
Appendix I: Debrief Sheet - Study 4 ................................................... 437
Appendix J: Brief Mood Introspection Scale ........................................ 438
List of Tables

Table 1: Characteristics of Ill-Structured, Complex Problems (Groenendaal & Helsloot, 2015) ........................................59
Table 2: Functionality and Action Fidelity in Self-Defence Situations in Police Use of Force Simulations ................................85
Table 3: Total Number of Participants per Study ................................106
Table 4: Design, IVs and DVs of Studies ........................................107
Table 5: Psychological tests employed in the thesis .............................108
Table 6: Physiological measures employed in the thesis .....................116
Table 7: Time Domain HRV Parameters .......................................119
Table 8: Frequency Domain HRV Parameters ..................................120
Table 9: Means, Standard Deviations and Confidence Intervals of Phonemic Fluency Scores of Study 1.1 ...............................148
Table 10: Means, Standard Deviations and Confidence Intervals of Digit Span Scores of Study 1.2 .................................149
Table 11: Means, Standard Deviations and Confidence Intervals of Corsi Block-Tapping Task Scores of Study 1.3 .......................150
Table 12: Comparison between CA and NLTA Condition of Cardiovascular Variables of Study 1 (n=25) .........................151
Table 13: Studies Investigating Executive Functioning in use of force Training .................................................................174
Table 14: Descriptive Statistics of Phonemic Fluency Scores ..............182
Table 15: Descriptive Statistics of Physiological Arousal Measures ......184
Table 16: Demographical Data of Participants of Study 3 ....................211
Table 17: Means, Standard Deviations and 95% Confidence Intervals around the Mean for State and Trait Anxiety Scores of Study 3 ...........................................................................217
Table 18: Summary Data for Dot Probe Task 50 msec and 200 msec Across All Groups of Study 3 .........................................222
Table 19: Means, Standard Deviations and 95% Confidence Intervals of Emotional Stroop Scores of Study 3 ......................224
Table 20: Means, Standard Deviations and 95% Confidence Intervals of the Adjusted Average Pump Count across all Groups of Study 3 ......................................................226

Table 21: Characteristics of Level in the Scenario ..................................................242

Table 22: Means, Standard Deviations, and 95% Confidence Intervals around the Mean of Phonemic Fluency Task Scores (Study 4.1) ........................................................................................................251

Table 23: Means, Standard Deviations, and 95% Confidence Intervals Around the Mean of Phonemic Fluency Task Scores of Study 4.2 ........................................................................................................260

Table 24: Group Comparisons of Manipulation Check and Mood Variables of Study 4.1 ........................................................................................................266

Table 25: Group Comparisons of Manipulation Check and Mood Variables of Study 4.2 ........................................................................................................267

Table 26: Group Comparisons of Cardiovascular Data of Study 4.1 ....268

Table 27: Group Comparisons of Cardiovascular Data of Study 4.2 ....269
List of Figures

Figure 1: Example of a linear use of force continuum (adapted from National Institute of Justice, 2009) ........................................... 42

Figure 2: Example of a linear - staircase use of force continuum (R. Hoffmann et al., 2004) .............................................................. 43

Figure 3: Example of a circular use of force continuum (R. Hoffmann et al., 2004) ................................................................. 43

Figure 4: Example of a use of force matrix (Desmedt, 2009) ............. 44

Figure 5: The Transactional Model of Police Use of Force (adapted from Binder & Scharf, 1980) .......................................................... 46

Figure 6: The Model of Officer Use of Control Tactics and Technologies (adapted from Amendola, 1995) .......................... 48

Figure 7: A model of an armed confrontation: Six phases towards the achievement of situational control through dominance (Boulton, 2014) ................................................................. 49

Figure 8: The Deescalating Model of Police Operations (adapted from Bernt & Kuhleber, 1991) .......................................................... 51

Figure 9: Relationship between the controllability and predictability of environmental challenges and the life threatening nature of these challenges (adapted from Koolhaas et al., 2011) ............ 55

Figure 10: Recognition-primed decision model (J. K. Phillips, Klein, & Sieck, 2004) ................................................................. 64

Figure 11: Model of situation awareness in dynamic decision making (Endsley, 2000) ................................................................. 65

Figure 12: The multi-component WM model (Baddeley, 2000) ........... 69

Figure 13: A model of the mechanism proposed by the Dual Component Theory of Inhibition Regulation (J. J. Reynolds & McCrea, 2016) ........................................................................ 77

Figure 14: Representativeness in police use of force ................................................................. 80

Figure 15: The testing of generated solutions for police use of force Problems ........................................................................ 83

Figure 16: The Trade-Off Model of Police Use of Force Simulation Design (Staller & Zaiser, 2016) .................................................. 88
Figure 17: Weapon systems and taxonomies in police use of force training.................................................................91
Figure 18: Design of the fighting and shooting exercise of Study 1 ..........137
Figure 19: Means and 95% Confidence Intervals of Phonemic Fluency Scores.........................................................................................................................140
Figure 20: Means and 95% confidence intervals of Digits Span Task scores..............................................................................................142
Figure 21: Means and 95% confidence intervals of block span and total score.......................................................................................144
Figure 22: Mean HR and 95% confidence intervals across all time points for the CA and NLTA condition (n = 25) .............................145
Figure 23: RMSSD and 95% confidence intervals across all time points for the CA and NLTA condition (n = 25).................................147
Figure 24: Means and 95% confidence intervals of Phonemic Fluency Task scores.................................................................183
Figure 25: Means and 95% confidence intervals of systolic blood pressure .......................................................................................185
Figure 26: Means and 95% confidence intervals of diastolic blood pressure.......................................................................................186
Figure 27: Means and 95% confidence intervals of mean HR .................187
Figure 28: Means and 95% confidence intervals of core temperature .....187
Figure 29: Cognitive mechanisms underlying threat processing according to Bar-Haim et al. (2007) ......................................................203
Figure 30: On-Screen Presentation of the Balloon Analogue Risk Task...215
Figure 31: Means and 95% confidence intervals of state and trait anxiety scores across all groups......................................................218
Figure 32: Means and 95% confidence intervals of 50 msec Dot Probe Task scores across all groups..................................................219
Figure 33: Means and 95% confidence intervals of 200 msec Dot Probe Task scores across all groups..............................................220
Figure 34: Means and 95% confidence intervals of mean bias score of 50 msec Dot Probe Task.....................................................220
Figure 35: Means and 95% confidence intervals of mean bias score of 200 msec Dot Probe Task.....................................................221
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>Mean and 95% confidence intervals of RTs of word classes across all groups</td>
<td>225</td>
</tr>
<tr>
<td>37</td>
<td>Mean and 95% confidence intervals of bias scores across all groups</td>
<td>226</td>
</tr>
<tr>
<td>38</td>
<td>Means and 95% confidence intervals of Adjusted Average Pump Count of all groups</td>
<td>227</td>
</tr>
<tr>
<td>39</td>
<td>Levels of the scenario design</td>
<td>243</td>
</tr>
<tr>
<td>40</td>
<td>Schematic overview of experimental procedure in Study 4.1</td>
<td>245</td>
</tr>
<tr>
<td>41</td>
<td>Schematic overview of experimental procedure in Study 4.2</td>
<td>245</td>
</tr>
<tr>
<td>42</td>
<td>Means and 95% confidence intervals of the time (in seconds) when participants displayed offensive aggression</td>
<td>250</td>
</tr>
<tr>
<td>43</td>
<td>Means and 95% confidence intervals of Phonemic Fluency Task scores of Study 4.1</td>
<td>253</td>
</tr>
<tr>
<td>44</td>
<td>Mean HR at the four time points across experimental and control condition (T1: baseline; T2: Video Task; T3: scenario; T4: recovery)</td>
<td>255</td>
</tr>
<tr>
<td>45</td>
<td>RMSSD measures at the four time points across experimental and control condition</td>
<td>256</td>
</tr>
<tr>
<td>46</td>
<td>Means and 95% confidence intervals of self-reports of behavioral and emotional suppressing</td>
<td>258</td>
</tr>
<tr>
<td>47</td>
<td>Means and 95% confidence intervals of the time (in seconds) when participants displayed offensive aggression in Study 4.2</td>
<td>259</td>
</tr>
<tr>
<td>48</td>
<td>Means and 95% confidence intervals of Phonemic Fluency Task scores of Study 4.2</td>
<td>262</td>
</tr>
<tr>
<td>49</td>
<td>Means and 95% confidence intervals of RMSSD for the experimental and control condition across all time points</td>
<td>263</td>
</tr>
<tr>
<td>50</td>
<td>Means and 95% confidence intervals of mean HR for the experimental and control condition across all time points</td>
<td>265</td>
</tr>
<tr>
<td>51</td>
<td>The hypothesised relationship between executive functioning and physiological arousal</td>
<td>285</td>
</tr>
<tr>
<td>52</td>
<td>A model linking the relationship of physiological arousal and executive functioning capacity to defensive responses</td>
<td>287</td>
</tr>
</tbody>
</table>
### Glossary of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFO</td>
<td>authorized firearms officer</td>
</tr>
<tr>
<td>ANS</td>
<td>autonomic nervous system</td>
</tr>
<tr>
<td>AR</td>
<td>autoregressive</td>
</tr>
<tr>
<td>BART</td>
<td>Balloon Analogue Risk Task</td>
</tr>
<tr>
<td>BMIS</td>
<td>Brief Mood Inspection Scale</td>
</tr>
<tr>
<td>bpm</td>
<td>beats per minute</td>
</tr>
<tr>
<td>CA</td>
<td>conventional Ammunition</td>
</tr>
<tr>
<td>CT</td>
<td>congruent trial</td>
</tr>
<tr>
<td>DB</td>
<td>digit span backward</td>
</tr>
<tr>
<td>DF</td>
<td>digit span forward</td>
</tr>
<tr>
<td>HF</td>
<td>high frequency</td>
</tr>
<tr>
<td>HR</td>
<td>heart rate</td>
</tr>
<tr>
<td>HRV</td>
<td>heart rate variability</td>
</tr>
<tr>
<td>IBI</td>
<td>interbeat interval</td>
</tr>
<tr>
<td>IT</td>
<td>incongruent trial</td>
</tr>
<tr>
<td>LF</td>
<td>low frequency</td>
</tr>
<tr>
<td>LSD</td>
<td>least significant differences</td>
</tr>
<tr>
<td>MBS</td>
<td>mean bias score</td>
</tr>
<tr>
<td>MCS</td>
<td>mean cluster size</td>
</tr>
<tr>
<td>MRT</td>
<td>mean reaction time</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>NDM</td>
<td>naturalistic decision making</td>
</tr>
<tr>
<td>NHST</td>
<td>null hypothesis significance testing</td>
</tr>
<tr>
<td>NLT</td>
<td>non-lethal training</td>
</tr>
<tr>
<td>NLTA</td>
<td>non-lethal training ammunition</td>
</tr>
<tr>
<td>NN50</td>
<td>numbers of pairs of adjacent RR intervals differing by more than 50 ms in the entire recording</td>
</tr>
<tr>
<td>NS</td>
<td>number of switches</td>
</tr>
<tr>
<td>pNN50</td>
<td>NN50 count divided by the total number of all RR intervals</td>
</tr>
<tr>
<td>PPE</td>
<td>personal protection equipment</td>
</tr>
<tr>
<td>PT</td>
<td>police trainer</td>
</tr>
<tr>
<td>RMSSD</td>
<td>root mean square of successive RR intervals</td>
</tr>
<tr>
<td>RPO</td>
<td>regular police officer</td>
</tr>
<tr>
<td>RR</td>
<td>interval between two consecutive R waves</td>
</tr>
<tr>
<td>RR mean</td>
<td>mean of RR intervals</td>
</tr>
<tr>
<td>RT</td>
<td>reaction time</td>
</tr>
<tr>
<td>SA</td>
<td>situation awareness</td>
</tr>
<tr>
<td>SDNN</td>
<td>standard deviation of RR intervals</td>
</tr>
<tr>
<td>SOP</td>
<td>standard operating procedure</td>
</tr>
<tr>
<td>SOU</td>
<td>Special Operation Unit</td>
</tr>
<tr>
<td>STAI</td>
<td>Stait-Trait Anxiety Inventory</td>
</tr>
<tr>
<td>STD HR</td>
<td>standard deviation of instantaneous heart rate values</td>
</tr>
<tr>
<td>TDS</td>
<td>total digit span</td>
</tr>
<tr>
<td>TW</td>
<td>total number of words generated</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>VLF</td>
<td>very low frequency</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction
1 Introduction

1.1 Background

The use of potentially lethal force by police officers lies at the extreme edge of policing activity (Burrows, 2007; Kargin, 2016). In this context, the performance of police officers must be efficient and legitimate, in order to ensure safety of the public (E. L. Kavanagh, 2006). Mistakes can have devastating consequences for the resolution of the incident and can impact long-term public perceptions of policing (MacDonald, Manz, Alpert, & Dunham, 2003). Tactical decision making is salient in incidents that attract widespread public attention, evaluation and criticism. For example, in the US the shooting of Alton Sterling in Baton Rouge, Louisiana (Hauslohner & Cusick, 2016), and Philando Castile in Falcon Heights, Minnesota (T. R. Shapiro, Bever, Wesley, & Miller, 2016), preceded the shooting of 5 Police officers in Dallas (New York Times, 2016) and led to widespread discussions in the USA about the decisions made by police officers. In Germany, the shooting of a young axe-attacker by a German Police Special Unit (Focus, 2016; Spiegel, 2016) attracted widespread attention and evaluation by the public: the special unit was called, when a young Afghan citizen started to randomly attack people in a community train near Würzburg in Bavaria. The attacker left the train and hid near the waterfront before the special unit arrived. While looking for the suspect in a dark and unclear area, two officers were attacked at very close distance. The operators neutralized the threat by shooting the suspect in the body and in the head. Shortly after the incident, the officers were publically criticised for applying lethal force and not shooting to the legs, which is common after police shootings. However, it has to be emphasised that these events represent an extremely small number of incidents compared to routine police work.

In these contexts the performance of police officers is impacted by a variety of cognitive, affective, and physiological factors. Many of them are often not taken into account while planning tactical operations or during the subsequent investigations (E. L. Kavanagh, 2006). External factors, such as the suspects’ aggression, and internal factors, such as physiological arousal
or previous experience, interact with the police officers’ perception and appraisal of the environment, which in turn determines their tactical decision making (Knutsson & Strype, 2003; Vrij, 1995; Vrij & Dingemans, 1996; Vrij, van der Steen, & Koppelaar, 1995; M. D. White, 2001; 2003). For example, situations involving acute physical, cognitive and emotional demand induce physiological arousal, which has been found to impact cognitive and behavioural performance in an inverted U-shape (Yerkes & Dodson, 1908): some functions are enhanced up to a certain point, after which continued increases can significantly reduce performance (C. A. Anderson, 2004; Barton, Vrij, & Bull, 2000; 2002; Brisswalter, Collardeau, & René, 2002; Honig & Lewinski, 2008; Hüttermann & Memmert, 2014; McMorris & Graydon, 2000). This change reflects normal neurological and physiological adaptations, that lead to an optimal state, in which an organism can deal with a perceived situational and/or environmental threat (D. M. Clark, 1988; 1989; Lewinski, 2008; Pellow, Chopin, File, & Briley, 1985). These adaptive defensive responses in humans are underpinned by an interaction of genetically hardwired evolutionary responses and more recently evolved cognitive processes (Lewinski, 2008). This may include the redirection of limited resources toward survival mechanism at the expense of secondary functions, such as higher cognition (Dienstbier, 1989; Driskell & Salas, 2006; D. Grossman & Christensen, 2007; Herman & Cullinan, 1997; Korte, Koolhaas, Wingfield, & McEwen, 2005; Selye, 1956).

The cognition of police officers in real world settings is located in an environment that is defined by ill structured problems, uncertainty, time pressure and dynamic environments with shifting or competing goals and high stakes (Cannon-Bowers & Salas, 1998a; 1998b; Endsley, Hoffman, Kaber, & Roth, 2007; Flin, Salas, Strub, & Martin, 1997; Orasanu & Connolly, 1993; Orasanu & Martin, 1998; Parent & Verdun Jones, 1998; Zsambok & Klein, 2014). In these real world situations naturalistic decision making (NDM) models may be most applicable to the tactical decision making of police officers. NDM models suggest, that experienced decision makers have superior declarative and procedural domain knowledge, the ability to construct more complete and accurate mental models, and to use efficient
information processing and decision strategies (L. Alison et al., 2013; Lipshitz & Shaul, 2014). They also are thought to rely on intuitive decision making based on the recognition of situational cues (G. A. Klein, 2008; 2009; 2015), which flexibly shunts from one process to the next depending on the situational assessment (Rasmussen, 1976). In contrast, inexperienced decision-makers rely more on slower analytical decision-making models (Kahneman & Klein, 2009; G. A. Klein, 2008; Lintern, 2010; Lipshitz, Klein, Orasanu, & Salas, 2001).

Police officers rely on situational cues to ascertain critical features of dynamic situations in order to determine a course of action (Endsley, 1995). In complex situations, like police-citizen encounters, situation awareness (SA) may be incomplete or uncertain. In that case, decision makers have to employ meta-cognitive control in order to critique, correct, and re-interpret information (J. D. Cohen et al., 1997; Lintern, 2010). These processes are thought to rely on working memory (WM) and executive functioning (Norman & Shallice, 1986; Shallice & Burgees, 1993).

WM enables the encoding, consolidation, active maintenance, manipulation and regulation of information, which are necessary for the completion of complex cognitive tasks (Frank, Loughry, & O'Reilly, 2001; Jonides et al., 1993; Miyake & Shah, 1999; Rosenthal, Riccio, Gsanger, & Jarratt, 2006). It has a limited storage capacity and as such is sensitive to cognitive load, which may interfere with performance. WM was conceptualized by Baddeley and Hitch (1974) as a theoretical multicomponent model that includes distinguishable systems (Baddeley, 1992; 2001; 2007; 2010; 2012; Baddeley & Hitch, 1974). Among these are the central executive and three temporary storage systems, (a) the visuospatial sketchpad, (b) the phonological loop, and (c) the episodic buffer as a more general integrated storage system. According to Baddeley (2007, 2012) the central executive is the most complex and most important subsystem of this multicomponent model. However, due to the variety of tasks and their complexity, it is the subcomponent that is least understood (Baddeley, 1996; 2000; Collette & Van der Linden, 2002). Baddeley (2012) does not claim to fully understand how the central executive functions. The central executive is responsible for
(a) monitoring and processing temporarily stored information, (b) planning and attentional switching, and (c) inhibition (Baddeley, 1996; J. Holmes, Adams, & Hamilton, 2008; Miyake et al., 2000). As such, the central executive controls and moderates the operation of various cognitive sub-processes which thereby determine decision making and the subsequent implementation of action (Baddeley, 2007; Doerner & Ho, 1994; N. P. Friedman et al., 2006; Lezak, 1982; Miyake et al., 2000). The ability to initiate, self-monitor, and correct actions based on environmental feedback are the basis for these processes. Hence, in situations where the reliance on learned procedures does not suffice, the cognitive flexibility to adapt to new or complex situations as a part of executive functioning is crucial (Ardila, 2008; L. H. Goldstein & McNeil, 2004; A. R. Luria, 1973; Stuss & Benson, 1984). Furthermore, there are situations in policing where impulses to act must be inhibited for the sake of long-term goals. For example, in situations where the use of force may be legitimate, a focus on communication can be beneficial to avoid the use of that force (Zaiser & Staller, 2015). Also, after use of force has been applied, police officers must demonstrate self-regulation in order to not use disproportionate or excessive force, which may result in fatal outcomes, for example due to prolonged restraining positions, which make breathing more difficult (C. Hall et al., 2015; Pinizzotto, Davis, Bohrer, & Infanti, 2012). Therefore, the application of self-control in policing is crucial. It is known that exercising self-control leads to difficulties in a subsequent task, where self-control is needed (Barlett, Oliphant, Gregory, & Jones, 2016; Baumeister, 2014; Englert, 2016; Stucke & Baumeister, 2006).

Besides the physiological and psychological demands on WM and previous tasks that affect self-control performance in the task at hand, the direction of attention also plays a crucial role with regards to tactical actions. The direction of attention to relevant cues impact the decision making in the specific domain (Bruya, 2010). These processes are guided by conscious and unconscious processing (Morsella & Poehlman, 2013). In policing contexts, the attention to threat is of particular interest, since earlier attention to threat may contribute to faster decision-making in this domain.
1.2 Statement of the Problem

Tactical decision making is a crucial aspect of effective police work. However, little research has been conducted in this context so far (Amendola, 1996; Barton et al., 2002; Boulton & Cole, 2016; Flin, Pender, Wujec, Grant, & Stewart, 2007; Knutsson & Strype, 2003; Nieuwenhuys, Cañal-Bruland, & Oudejans, 2012; Nieuwenhuys, Savelsbergh, & Oudejans, 2015; Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b; Renden, Savelsbergh, & Oudejans, 2016). Possible reasons for the relatively small amount of research conducted in this area are difficulties in gaining access to police samples and challenges in the collection and analysis of police data that is both ecologically valid and scientifically objective. Hence, the majority of police use of force research to date has applied either computer-based simulated scenarios or qualitative analysis of post incident reports and judicial accounts from real incidents (Amendola, 1996; Barrett, 2008; Binder & Scharf, 1980; 1982; M. D. White, 2006). Recent approaches also employed experimental paradigms with standardised measures whilst accounting for the associated influences of operational demands on tactical decision making (Barton et al., 2000; 2002; Boulton, 2014; Condon, 2015; Doerner, 1991; Doerner & Ho, 1994; Renden, Landman, Savelsbergh, & Oudejans, 2015a; A. P. J. Roberts, 2012; Vrij, Van der Steen, & Koppelaar, 1994).

1.3 Significance of the Thesis

The current thesis has a practical and theoretical significance. From a practical point of view, the thesis will help to better understand the use of non-lethal training ammunition (NLTA) and conventional ammunition (CA) in the context of police training. Since accidents in training environments are a life threatening problem (Griffith, 2003), the current study will afford safer yet representative police use of force training involving the use of firearms. Furthermore, shedding further light on to the existing knowledge on how tactical decisions are made (Boulton, 2014; Boulton & Cole, 2016; Condon, 2015; A. P. J. Roberts, 2012), is crucial for the reduction of risk in the context of complex, highly dynamic, and pressured environments of policing. By
knowing the influential factors underlying tactical decision making, policing agencies and coaches in these contexts are able to optimize training and operational procedures. Furthermore, the current research may have an impact on people’s attitudes and behaviours toward the legitimacy of the police and justice system (Friedrich, 1980). For example, the previously described shooting of an axe-wielding attacker near Würzburg, led to widespread discussions about the use of force in policing (Focus, 2016; Spiegel, 2016). In the light of current reports and discussions about excessive use of force (Amnesty International, 2010; Friedrichs, 2010), understanding the processes that result in such actions by police officers are crucial.

From a theoretical point of view, the current thesis will broaden the theoretical understanding of various psychological concepts. First, the research will help to better understand changes in WM functioning in the context of human defensive behaviour. Second, the knowledge with regards to the development of threat-related attentional biases will be broadened. Third, since research regarding the interaction between ego depletion and real aggressive behaviour is limited (as there is much debate over the construct validity of laboratory models of human aggression), the current thesis will enhance our understanding of the effects of depleted self-control resources on actual aggressive behaviour.

1.4 General Aims

The work in the current thesis examines several factors that possibly influence the decision making process of police officers with regards to tactical actions. As such, it builds on previous work of the research group (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012), that has shown that cognitive processing during simulated armed confrontations adaptively changes in order to cope with the demand/threat at hand.

The thesis first focuses on possible differences between CA and NLTA in demanding training sessions in order to inform the methodology of studying armed confrontations. Building on this knowledge subsequent studies employed NLTA for the sake of safety while being representative at the same
time. Second, the thesis investigates weather adaptive cognitive shifts during simulated confrontations are caused by the cognitive demands of tactical training or by the sheer physical activity involved with this training. Third, it is examined if previous experience in regular conflict training or experience as a police officer may contribute to the development of threat-related attentional biases. Fourth, the thesis investigates the effects of ego depletion on subsequent performance in a use of force scenario.

With regards to methodology, the current thesis employed a quantitative approach in mainly naturalistic settings (except for Study 3) in order to provide a body of research work with high reliability and ecological validity. Physiological arousal was measured in three studies in order to facilitate discussion of physiological and cognitive adaptations to the demands at hand. Overall, the current thesis produces a model for understanding the relationship of various influential factors in the context of police tactical decision making. The aim is to establish a valid theoretical framework for future research into the cognitive processes involved in the police use of force context.

1.5 Structure of the Thesis

The thesis is divided into three main sections; (a) literature review and general introduction to the empirical research and methodology, (b) empirical research and (c) general discussion and applications of the work.

Section I consists of two chapters which introduce the key topics of concern. Chapter 2 provides a general overview of the relevant literature and theories including police use of force, stress, defensive behaviours, decision making, WM, executive functioning, ego depletion and representative learning environments. This is followed by identification of the specific aims of the empirical work with its rationale. Chapter 3 describes the methods used, including a detailed rationale for their use.

In Section II, the empirical research is organised in four chapters consisting of seven studies. Chapter 4 examines differences in the use of CA and NLTA with regards on its effects on different aspects of WM. Chapter 5 focuses on
the differences between a simulated armed confrontation and physiologically demanding exercise with regards to physiological arousal and executive functioning. Chapter 6 examines threat-related attentional biases and risk taking in police officers and martial artists comparing them to a control group. Chapter 7 investigates the effects of ego depletion on offensive aggression of police officers.

Section II consists of one chapter. Chapter 8 summarises and discusses the general findings of the work, its limitations and suggestions for further research. This section draws some final conclusions and discusses the general implications of this research in terms of psychological methodology and to police policy and training.

1.6 Summary

The thesis focuses on various factors that influence tactical decision making within the policing context. Since research is limited in this area, the current work integrates theories from topics such as defensive behaviour, WM, decision making paradigms, representative learning environments and police use of force research. A quantitative approach is taken in order to examine the effects of distinct factors on tactical decision making.
Chapter 2: Background Literature
2 Background Literature

2.1 Human Defensive Behaviour

In almost every species situations of perceived threat induce various responses on a neurological and physiological level. These adaptations place the organism in a state, where it is able to deal with the stimuli, that is perceived as the source of the threat (D. M. Clark, 1988; Lewinski, 2008; Mobbs, Hagan, Dalgleish, Silston, & Prévost, 2015; Pellow et al., 1985). Quickly and effectively choosing among a limited set of defensive behaviours is a vital capacity for an organism, to ensure survivability (Löw, Weymar, & Hamm, 2015).

The perception of threat leads to patterns of species specific and non-specific defensive behaviours, that aim at optimizing survival of the own species (Mobbs et al., 2015). These behaviours, ranging from flight to attack, have evolved to deal with environmental challenges that show a common structure within all animals (including humans). Prototypical situations are the attacking of an aggressor, fleeing from a predator or the hiding from an inescapable threat (Harrison, Ahn, & Adolphs, 2015). Several years of theoretical and empirical studies put forward the idea, that a common structure in defensive behaviours across species exist (D. C. Blanchard, Hynd, Minke, Minemoto, & Blanchard, 2001). Various schemes have been proposed, answering the question of how defensive behaviours are organised (D. C. Blanchard, Griebel, Pobbe, & Blanchard, 2011; R. J. Blanchard, Blanchard, & Hori, 1989; Fanselow & Lester, 1988; Kozlowska, Walker, McLean, & Carrive, 2015; P. J. Lang, Bradley, & Cuthbert, 1998; Mobbs et al., 2015). For example, Kozlowska et al. (2015) term these hard-wired, automatically activated defence behaviours, “the defense cascade”. They conceptualize (a) arousal as the first step in activating this cascade, (b) flight or fight as an active defensive response to threat, (c) freezing as a fight-or-flight response, that is put on hold, (d) tonic immobility and collapsed immobility as options of a last resort to inescapable threats, and (e) quiescent immobility as a state of quiescence promoting healing and rest (Kozlowska et al., 2015). Recently, Mobbs et al. (2015) propose the Survival
Optimization System to account for the strategies that humans and other animals use to defend against recurring and novel threats. This behavioural repertoire involves (a) prevention (niche constructing/herding), (b) prediction (passive avoidance), (c) threat orienting (attention/freezing), (d) threat assessment (freezing/directed escape) and (e) defence strategies (indirect escape) (Mobbs et al., 2015). Fighting in this context occurs as an active coping strategy, when escape is not directly possible. In animal models, the distance between the threat (the predator) and the prey, the type of predator and the predator’s size as attack strategy influence the prey’s response (R. J. Blanchard & Blanchard, 1990; R. J. Blanchard, Flannelly, & Blanchard, 1986; Fanselow & Lester, 1988; Mobbs et al., 2015; Rodgers & Randall, 1987).

A common thread in these non-species specific models is that aggressive behaviour is a positive coping strategy for life threatening encounters. Human aggression, which has been defined as any overt behaviour produced with the intent to harm another who wishes to avoid that harm (Baron & Richardson, 1994; Berkowitz, 1989; J. Bushman & Anderson, 2001) or as the intent to increase one’s own position in a dominance hierarchy at the expense of another (Ferguson & Beaver, 2009), is typically framed as negative, unwanted facet of human behaviour (Hawley & Vaughn, 2003; Shaver & Mikulincer, 2011; P. K. Smith, 2007). Given this moralistic view of aggression, research aiming at understanding human aggression has mainly focused on understanding its origins and how to prevent the excessive use of violence. (C. A. Anderson, Benjamin, & Bartholow, 1998; C. Anderson & Bushman, 2002; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Berkowitz, 1990).

However, as outlined above, aggression as an integral part of defensive behaviour is innate to both humans and animals alike. (D. C. Blanchard et al., 2001; P. J. Lang, Davis, & Öhman, 2000). The type of demonstrated behaviour often varies depending on the motivation. For example, in animal models of defensive behaviour, offensive aggression is characterized by an initiation by the aggressor, whereas defensive aggression is initiated in response to the aggressor (although it is extremely difficult to differentiate between the two forms of aggression as it is often said that “offence is
sometimes the best defence”). Predatory aggression is displayed when animals are seeking food and is possibly best considered as feeding-related behaviour rather defensive behaviour (Sukikara, Mota-Ortiz, Baldo, Felício, & Canteras, 2006).

In humans the defensive response system seems to be activated by fear and anxiety stemming from perceived threat to the sustainability of the individual (P. J. Lang et al., 2000). The perceived threats can have many different forms: perceptible and overt or indistinguishable and hidden. Threats can involve challenges to health, safety or life. Threats activate fear or a state of anxiety, which in turn trigger defence responses (Darwin, 1872; Eilam, Izhar, & Mort, 2011; P. J. Lang et al., 2000).

Fear is defined as the response to a perceptible threat, like fire or an attacking predator, whereas anxiety is the response to an abstract danger, like the potential risk of earthquakes or olfactory signals of a predator (Eilam, 2005; Eilam et al., 2011; Öhman, 2000). It is noteworthy in this context, that contrary to fear, there is no external termination signal in anxiety. Relief from this state is highly subjective and therefore varies among individuals (Eilam et al., 2011).

Upon encountering a perceptible threat to life, the fear response takes the form of freezing, fleeing, fighting back or tonic immobilisation (D. C. Blanchard, 1997; R. J. Blanchard & Blanchard, 1989; Eilam, 2005). The Distance-Dependent-Defence-Hierarchy states that these defensive reactions are a function of the distance between a prey species and the source of a danger, which is the predator (Rodgers & Randall, 1987). These behavioural patterns of response are a reflection of a predator entering the animal’s environment and continuing to approach. Freezing refers to the state of “attentive immobility”, where the organism is passive, but attentive and engages in selective scanning (Marks, 1987; Volchan et al., 2011). With being motionless and vigilant the chances of not being noticed by the predator increases. However, the organism is automatically mobilized and primed to respond, but not yet active (P. J. Lang et al., 2000). An exaggerated startle reflex to any suddenly imposed stimulus displays an
increase on this state of fear, when the threat cannot be avoided (P. J. Lang et al., 2000; Löw et al., 2015). Löw et al. (2015) showed, that with the approaching imminence of the potential threat, sympathetic activation and autonomic arousal to the approaching threat is exhibited in humans. However, if the organism has the opportunity to avoid the approaching threat, startle responses are inhibited in preparation for active avoidance (Löw et al., 2015). If the attack of the predator starts, overt defensive behaviour (fight or flight) takes place. Tonic immobility is considered as the last line of defence against entrapment by a predator within the sequence of defensive behaviours, aiming at reducing the probability of a continued attack (Volchan et al., 2011).

In cases where identifiable triggering threats are missing, humans and animals alike are capable of risk assessment, followed by precautionary measures in order to avoid the potential danger (D. C. Blanchard et al., 2011; Eilam et al., 2011; Hinds et al., 2010; Woody & Szechtman, 2011). Risk assessment involves the gathering of information with regards to a potential threat in order to produce an optimal response (D. C. Blanchard, Blanchard, & Rogers, 1991; Eilam et al., 2011). In this context, the Security Motivation System (Hinds et al., 2010; Szechtman & Woody, 2004; Woody & Szechtman, 2011) has been proposed as a distinct module in the brain, which evolved to be specially adapted for handling these potential threats (Pinker, 1997; Tooby & Cosmides, 1990; 2006). As such it features three characteristics: First, it is devoted to the detection of relatively specific classes of stimuli, facilitating the rapid processing of information and potential relevance for survival, and functions in a relatively encapsulated and automatic way (Hinds et al., 2010). Second, it operates as a motivational system driving relevant responses, once it is activated (Kavaliers & Choleris, 2001). Third, the output of the security motivational system consists of a species-specific set of behaviours, that play an essential role in the termination of the activation of the module once the organism engages in these behaviours (Hinds et al., 2010; Woody & Szechtman, 2011; 2013).

The output of behaviour with regards to identified and unidentified threats is usually effective and thus adaptable, allowing the organism to escape or to
avoid life-threatening events like a predatory attack. Also the display of aggression may be useful in humans (Ferguson & Beaver, 2009) and may be regarded as an adaptive response when avoidance and escape are not viable response options (D. C. Blanchard et al., 2001; 2011; Eilam et al., 2011; Parker, 1974; Ydenberg & Dill, 1986). However, there are instances, where defensive responses are maladaptive and inappropriately activated due to an underlying pathology. For example, mistaken appraisal of threat such as misinterpretations of facial and social cues may result in inappropriate defensive actions to ambiguous social situations (Akhtar & Bradley, 1991; B. M. Curtis & O'Keefe, 2002; Dodge & Schwartz, 1997) or a dysregulation of the security motivation system may result in maladaptive behaviours of handling potential threats (Boyer & Bergstrom, 2011; Flannelly, Galek, Tannenbaum, & Handzo, 2007; Hinds, Woody, Schmidt, Van Ameringen, & Szechtmn, 2015; Neuberg, Kenrick, & Schaller, 2011). It is noteworthy, that pathological disorders like obsessive-compulsive behaviours demonstrate, that an intuitive, affective-based module like the security motivation system is very powerful and is able to override rational control of behaviour (Woody & Szechtmn, 2013).

2.1.1 Complexity of Defensive Behaviours

The understanding of human defensive behaviour is built upon the examination of aggression and defensive actions in animals (Darwin, 1872; Eilam, 2005; Eilam et al., 2011). Based on this understanding, models of human defensive behaviours have been developed (Harrison et al., 2015; Kozlowska et al., 2015; Löw et al., 2015; Mobbs et al., 2015).

However, there are phylogenetic differences in brain structure between animals and humans, that are reflected by cognitive capacity of humans and the complexity of cognitive processes, like the use of language and abstract thinking (A. G. Carter, Creedy, & Sidebotham, 2016; M. H. Christiansen & Kirby, 2003a; M. Christiansen & Kirby, 2003b; Corballis, 2003; Dumontheil, 2014; P. K. Murphy, Rowe, Ramani, & Silverman, 2014; Tattersall, 2016). The prefrontal cortex plays a dominant role in this context as it is the neocortical region that is most highly developed in primates (E. K. Miller &
The key function of the prefrontal cortex is cognitive control (E. K. Miller & Cohen, 2001). It allows for inherently more complex defensive behaviours and greater flexibility in the range of responses due to the neural connections with evolutionarily older, genetically hard wired systems (e.g. the limbic system) and more recently evolved cognitive ones (Craig, 2003; E. K. Miller & Cohen, 2001).

The prefrontal cortex is the primary centre for WM (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; D'Esposito, Postle, & Rypma, 2000; 2002), which is the main cognitive system that has underpinned the development of more complex behaviours and actions (M. Christiansen & Kirby, 2003b). WM is defined as a limited capacity system, that is responsible for the temporary storage and manipulation of information required for the completion of cognitive tasks and physical action (Baddeley, 1992; 2000; 2010; Miyake & Shah, 1999). It acts as an interface between the external environment, internal memory constructs, and action in order to facilitate the completion of a task specific output (Andrade, Kemps, Werniers, May, & Szmalec, 2002; Conway et al., 2005; Miyake & Shah, 1999). WM contains information that can be processed and acted upon, guiding behaviour in the absence of external cues (Goldman-Rakic, Chafee, & Friedman, 1993). As such it coordinates the resources, such as verbal and spatial memory, regulates and distributes attention and behavioural outputs (Baddeley, 1992).

In the context of performing simple, reflexive behaviours, such as the automatic orientation to an unexpected sound, WM does not play a critical role (E. K. Miller & Cohen, 2001; Schweizer, 2001). Those responses are assumed to be genetically hard wired deriving from long established evolutionary systems. Likewise, as a task is practiced more and more, its completion demands less attentional control and utilizes less WM capacity (C. Blais, Harris, Guerrero, & Bunge, 2012; Fuster, 2001; Jackson, Ashford, & Norsworthy, 2006; I. H. Jenkins, Brooks, Nixon, Frackowiak, & Passingham, 1994; B. P. Lewis & Linder, 1997; Norman & Shallice, 1986; Petersen, van Mier, Fiez, & Raichle, 1998; R. A. Schmidt & Wrisberg, 2008). However, in the context of novel or fast changing, threatening environments, where there is no well-established behavioural pattern, rigid behaviours are
maladaptive. Hence, complex behaviours in general, including defensive ones, must be processed via WM (J. D. Cohen, Dunbar, & McClelland, 1990; J. B. Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000).

Taken together, human defensive behaviour is far more complex than in any other species. The execution of an adaptive responses to perceived threats in a complex environment involves the interaction of genetically hard wired evolutionary responses in conjunction with more recently evolved cognitive processes (Craig, 2003; Lewinski, 2008; Schweizer, 2001).

However, it has to be emphasized that defensive behaviours in humans are dynamic structures, where certain behavioural changes are compensated for by changes in other competing behaviours (A. K. Dixon, Fisch, Huber, & Walser, 1989; Krsiak, 1991). As such those behaviours can only be understood with regards to the context in which the behaviour occurs. With regards to focus of this thesis, the environmental context is the police use of force against offenders.

2.1.2 Challenges in Studying Human Defensive Behaviour

The study of human defensive behaviour is linked to a major empirical problem: It is impossible to ethically harm people for research purposes so laboratory models of human aggression have been criticised for a lack of validity (Giancola & Chermack, 1998; Ritter & Eslea, 2005; Tedeschi & Quigley, 1996; 2000). In laboratory aggression paradigms (Bandura, 1983; Berkowitz, Corwin, & Heironimus, 1963; J. D. Lieberman, Solomon, Greenberg, & McGregor, 1999; Norlander, Nordmarker, & Archer, 1998; Russell, Arms, Loof, & Dwyer, 1996; S. P. Taylor, 1967) problems occur with regards to the supposed target of the aggression, either because of the distance involved (e.g. the target is in another room) or because the behaviour does not seem likely to cause real harm (e.g. the target is a doll). Furthermore, paradigms are also criticised for aggression being the only way to interact with the opponent and not allowing participants to choose from a range of available response options, including prosocial response options or communication (Ritter & Eslea, 2005).
This problem in the study of human defensive behavior has lead to (a) a heavy reliance on animal studies (D. C. Blanchard et al., 2001; D. C. Blanchard & Blanchard, 2008; R. J. Blanchard & Blanchard, 1989; Eilam, 2005; Eilam et al., 2011; P. J. Lang et al., 2000) and (b) approaches that investigate humans that are already in threatening situations (J. Kavanagh, 2005; MacDonald, Kaminski, & Smith, 2009; T. D. Smith & DeJoy, 2014; I. Y. Sun, Payne, & Wu, 2008; Vrij et al., 1994; M. D. White, 2006). In contrast, the underpinning mechanism of human defensive behaviour have rarely been directly investigated (A. P. J. Roberts, 2012; Staller & Cole, 2016).

However, the legitimate use of force by military and law enforcement personnel creates a unique platform for studying human defensive behavior. Police officers, in particular, are required to use a proportionate amount of force to neutralise the (potentially lethal) threat posed by an individual during a violent confrontation in order to remain with existing human rights legislation. Consequently, during training military and law enforcement personnel are routinely exposed to simulated (un-)armed confrontations from which data can be collected in an ethical way (Staller & Cole, 2016). As such, several researchers have investigated physiological and cognitive responses within this framework (Nieuwenhuys et al., 2015; Nieuwenhuys & Oudejans, 2010; Nieuwenhuys, Weber, van der Hoeve, & Oudejans, 2016; Oudejans, 2008; Renden et al., 2016; 2014).

### 2.2 Police Use of Force

Police officers are regularly tasked with the resolution of conflicts (Amendola, 1996; G. S. Anderson, Litzenberger, & Plecas, 2002; Anshel, 2000). As executive organs of the state the use of proportionate use of force is a legitimate part of their role (Amendola, 1996; Terrill, Paoline, & Manning, 2003). However, conflict resolution and achieving compliance is not limited to the use of force: Communication and negotiation skills are a valuable tool in the officer’s toolbox as well (Zaiser & Staller, 2015).

The profession of a police officer involves coping with a broad range of situations and how these situations develop and unfold cannot be foreseen. On the one hand, even routine deployments can result in life threatening
situations for the officer (e.g. terror attack while on patrol). On the other hand, ‘high risk’ deployments can turn out to be non-threatening once the police officer arrives at the scene (e.g. martial arts class practicing in the park). Additionally, the types of threats posed to police officers can vary from verbal to deadly assaults. Even unarmed encounters pose a serious risk of injury to the police officer and the other party involved (Bochenek, 2014; Ellrich, Baier, & Pfeiffer, 2010a; 2010b; 2011; Ellrich, Pfeiffer, & Baier, 2010c; Jager, Klatt, & Bliesener, 2013; M. R. Smith & Alpert, 2000). Law enforcement personal have numerous technologies and a variety of tactical options at their disposal in order to deal with these situations. The precise options available differ depending on the country, the local jurisdiction, and the unit the police officer is assigned to.

Practitioners and scholars alike have focused on models depicting the decision making process in police use of force incidents. Even though regularly referred to as “use of force models”, there are two broad types of models: (a) models that aim to depict the different kind of options a police officer can choose in a police – citizen encounter and (b) models that try to explain how decisions are made in such incidents. The first type of model are used as a training tool in order to help the officer to develop the decision makings skills needed in an encounter and are mainly based on the agency’s policies about use of force (training models). The second type of model aims to reveal the influencing factors that impact decision making in such situations (explaining models).

2.2.1 Police Use of Force Models – Training Models

Police agencies establish parameters for the use of force in their use of force policies in order to help the officer in their decision making process in the field (Terrill & Paoline, 2012). Although these models aim at meeting the legal criteria for the use of force in a police-citizen encounter, they are not intended to serve as a justification for officers to prescribe specific response options appropriate to a situation (R. Hoffman, Lawrence, & Brown, 2004). Their main purpose is to train officers to employ reasonable responses to the actions of a citizen (R. Miller, 2012; Peters & Brave, 2006).
Throughout the last 50 years a variety of police use of force models have been proposed (Seagrave, 1997; G. Williams, 2002). There are several structural permutations regarding the presentation of reasonable police use of force. These model are often referred to as “force continuums” (R. Hoffman et al., 2004; R. Miller, 2012; Peters & Brave, 2006; Terrill & Paoline, 2012; G. Williams, 2002), which refers to a list of actions in the escalation of force in response to changes in the perceived threat (National Institute of Justice, 2009; M. D. White, 2003).

Agencies use different types of structural designs to depict the use of force options in any given event: linear designs (Figure 1 and Figure 2), circular/wheel designs (Figure 3) and matrix/box designs (Figure 4).

Figure 1: Example of a linear use of force continuum (adapted from National Institute of Justice, 2009)
Models of Police Use of Force

In the United States and Canada, models describing police use of force for more than three decades. The development of models supporting the training and articulation of police expertise to explain their reactions to use of force is one firm ground in police use of force. Many of the models were overly complex and difficult for training purposes. However, a significant concern arose from their adoption by police agencies. Many of the models lacked a common language and shared standard. At the same time, there were too many models with different terminology.

The old adage that a picture is worth a thousand words has been applied to the different models. Many of the models were overly complex and difficult to understand. In addition, many of the models relied on the model and the research and development of the Ontario Use-of-Force Model, 1993.

Figure 2: Example of a linear - staircase use of force continuum (R. Hoffmann et al., 2004)

Figure 3: Example of a circular use of force continuum (R. Hoffmann et al., 2004)
In a survey of 1083 police agencies in the United States (response rate: 662; 61.1%) Terrill and Paoline (2012) found, that more than 80% of the agencies utilize a use of force continuum, of which the linear model is the most popular one (over 70% of agencies using a version of this model). The placement of various force tactics and considerations regarding the suspect’s level of
resistance varied widely across the departments, leading to the conclusion that there is no consensus about a use of force continuum by practitioners.

What the linear models have in common is that actions are modelled chronologically, which leads to the impression that later options can only be applied if earlier ones have been tried and did not work. This led to discussions about potentially negative outcomes of these models (Aveni, 2003; R. Hoffman et al., 2004; Peters & Brave, 2006; Petrowski, 2002; G. Williams, 2002), although it is regularly acknowledged within these models that encounters can be highly dynamic and thus demand quick changes with regards to an appropriate use of force option in the prevailing circumstances, such as the sudden and unexpected production of a firearm (National Institute of Justice, 2009).

2.2.2 Police Use of Force Models – Explaining Models

Several authors have put forward police use of force models, that articulate distinct variables that impact the decision making process in police use of force events. Unlike the training models, the explaining models are grounded on empirical data from police use of force events.

2.2.2.1 The Transactional Model of Police Use of Force by Binder and Scharf (1980)

Binder and Scharf (1980) proposed the Transactional Model of Police Use of Force (see Figure 5), which was grounded on qualitative analysis of police statements and court records of encounters between citizens and police officers in the United States involving the use of force. The main difference between the transactional model and continuum models is the recognition that force decisions are multifaceted. The final decision is a culmination of numerous decisions combined with internal and external factors. Together with the understanding that police use of force encounters are dynamic, rapidly evolving, and potentially life threatening (Binder & Scharf, 1980; Fyfe, 1981; Hontz, 1999), the Transactional Model shows the limitations of the rigid, linear approach to police use of force models (Binder & Scharf, 1980; 1982; M. D. White, 2015)
The model itself comprises of four phases: (a) anticipation, (b) entry, (c) information exchange and (d) final decision. During the first phase (anticipation) the police officer’s decision making is influenced by the manner in which the officer is deployed. For example, in a planned event (SWAT operation, dispatch through communications centre, etc.) the police officer has time to gather information about the event and the parties involved. In contrast, in a spontaneous event (assault on the officer, officer generated occurrence, etc.) the officer has very little time to receive further intelligence. Physiological, cognitive and affective demands that are placed upon the officer during this phase can impact decision making in later stages of the unfolding event (G. S. Anderson et al., 2002; Barton, Vrij, & Bull, 1998; Fridell & Binder, 1992; Fyfe, 1981).

Frequently this impact is based upon the perceived level of threat (Alpert & Dunham, 1997; Barton et al., 2000; Correll, Park, Judd, & Wittenbrink, 2002; Fridell & Binder, 1992; Fyfe, 1980; R. Wolf, Mesloh, Henych, & Thompson, 2009). Furthermore, officers anticipate the demand that will be placed upon them and experience anticipatory physiological reactions to incidents or shifts (G. S. Anderson et al., 2002).

During the phase of entry the police officer has to process vast amounts of perceptual information, which impacts the decision making about the best course of action in the particular event (Honig & Lewinski, 2008; Vrij et al., 1994; Vrij & Dingemans, 1996). The assessment of the level of threat, cover options, distance, and positioning towards the offender are of particular
importance with regards to officer safety and situational control (Bochenek, 2014; Füllgrabe, 2014; Schmalzl, 2008).

The phase of information exchange typically involves non-verbal communication (e.g. subtle facial expressions and body language) and verbal exchanges between the police officer and the citizens present at the scene. Officers are more likely to use (lethal) force after a period of negative exchange with a suspect (Fridell & Binder, 1992). Furthermore, the officer’s heart rate (HR) increases at this phase when communicating with a suspect (G. S. Anderson et al., 2002). During this phase un-holstering behaviour becomes a crucial aspect: while un-holstering the weapon increases the likelihood of a weapon discharge, both intentionally (Doerner, 1991; Doerner & Ho, 1994) and unintentionally (Heim, 2009; Heim, Schmidtbleicher, & Niebergall, 2006a; 2006b), it is often used as a protective action by the officer in order to gain compliance.

During the final decision phase the police officer decides to shoot or not to shoot based upon the decisions that were made earlier in the encounter. Earlier tactical decisions such as positioning, keeping distance, focus on the hands, and tactical planning during the encounter can impact later decisions in a transactional manner (Binder & Scharf, 1982; Fridell & Binder, 1992; Honig & Lewinski, 2008).

It is important to note that the phases of the model are not necessarily present in all potentially violent confrontations.

2.2.2.2 The Model of Use of Control Tactics and Technologies in Police-Citizen Encounters by Amendola (1995)

Amendola (1995, 1996) proposed an elaborated descriptive model to help understand the dimensions and specific variables that affect the need for, or likelihood of, a police officer using force. Building on the limitations of previous models the Model of Use of Control Tactics and Technologies captures the situational dynamics within a broader framework (Figure 6). Numerous internal and external (environmental) factors that influence use of force decision making have been included. It describes specific decision
points including plan formation prior to attending the encounter, establishing a goal once the situation had been assessed, selection of the appropriate technologies, and the tactical implementation from the available options (Amendola, 1995; 1996). Within this model the officer’s behaviour is the culmination of a variety of influences and others’ actions. As such, the model takes account of antecedent events, officer characteristics, citizen characteristics, situational characteristics, the range of options, and the associated constraints upon using any particular tool, technology, tactic, or behaviour as an intervention.

![Diagram](Model_of_Officer_Use_of_Control_Tactics_and_Technologies_in_Police-Citizen_Encounters.png)

*Figure 6: The Model of Officer Use of Control Tactics and Technologies (adapted from Amendola, 1995)*

Like the Transactional Model Amendola’s model is based on police witness statements from court proceedings. Although the model has not been empirical validated, its descriptive nature provides a theoretical framework that enables researchers to “conduct comprehensive, multi-variable studies designed to more thoroughly understand the multiple influences on officer use of force” (Amendola, 1996, p. 2).
2.2.2.3 The Model of an Armed Confrontation by Boulton (2014)

The model by Boulton (2014) conceptualizes tactical decision making of authorized firearms officers during an armed confrontation as a process of continual reassessment of the police officers ability to achieve dominance as a means of controlling the situation (Figure 7). Supporting previous models (Amendola, 1996; Binder & Scharf, 1980) the model depicts an armed confrontation as a contingent sequence of decisions and the resulting behaviours. Hence, the model is structured as a process diagram representing the associations between tasks, events, and behaviour, in order to provide insight into the multifaceted layers of activities that are involved in an armed confrontation.

![Diagram of Armed Confrontation Model]

**Figure 7:** A model of an armed confrontation: Six phases towards the achievement of situational control through dominance (Boulton, 2014)

Building on cognitive task analysis of authorised and specialized firearms officers Boulton’s model puts emphasis on the impact of expertise on the decision-making process. Hence, the model sees adaptive expertise as the determining factor of the situational (re)assessment process throughout an armed confrontation, that influences the processes of predictive mental
modeling and adaptive response to change under increased demand (Boulton, 2014; Boulton & Cole, 2016).

### 2.2.3 German Police Use of Force Framework

In Germany, police training requirements are determined separately in each police agency (16 state police agencies, 2 federal agencies). Hence, training is not standardized across the nation. However, recruits have to go through three years of training in all agencies, which comprises academic and physical education. German police officers are trained in legislative requirements, policy and procedure, use of force training, and have to maintain a certain fitness level, which is tested throughout the three years of education. After graduation, requalification of training is set by the agency and according to the unit the new police officers are assigned to.

When police officers take on their duty of preserving peace, ensuring safety, and resolving conflict in may become necessary to use force. Officer’s safety and the use of force by the officer have to be seen in a contextual framework. The German framework is based on the balance between the potential risk to officer’s safety and the use of force options that are reasonably available to the officer and proportionate to the situation. The nationwide framework “Leitfaden 371 - Eigensicherung” has been developed in order to assist in training officers on how to assess risks and engage in reasonable force and is built on the “Deescalating Model of Police Operations” (Bernt & Kuhleber, 1991; Kraft, 2004), which is depicted in Figure 8.
The model represents the encounter as a three-phase process with the phase preparation (plan), operation (act) and debriefing (assess). Each phase comprises of several options the police officer should engage in in order to optimise acceptable and safe policing. The Deescalating Model of Police Operations is circular to allow for continuous updating as situational, environmental, and internal cues change. Therefore, it can be applied to larger operations with long periods of preparation, operation, and debriefing but also on a micro level in every encounter with multiple cycles of planning, acting, and assessing. Like the previous discussed models, the Deescalating Modell of Police Operations has no empirical validation supporting its use. In addition to the model, Füllgrabe (2014) emphasised the development of a “threat radar” (“Gefahrenradar”) for police officers in order to optimise the efficiency of the described model.

*Figure 8: The Deescalating Model of Police Operations (adapted from Bernt & Kuhleber, 1991)*
2.3 Stress and Policing

Policing is generally seen as a stressful occupation (Anshel, 2000; Greenwood-Ericksen, Oron-Gilad, Szalma, Stafford, & Hancock, 2004; Lewinski, 2006; 2008). However, it is not clearly defined to what “stress” in this context refers to and how it affects the tactical decision making of police officers. When police officers report for duty they do not know what will happen during their shifts. Even though police work is largely sedentary (G. S. Anderson & Plecas, 2000), officers have to deal with violence on an occasional basis (Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b). Police officers may have to resolve conflicts and cope with life-threatening situations without knowing when this will happen. As such the phenomenon of violence against police officers is a pertinent problem for police forces around the world (Bochenek & Staller, 2014; Ellrich et al., 2011; Ellrich & Baier, 2016; Ellrich, Baier, & Pfeiffer, 2010a; 2010b; Ellrich, Pfeiffer, & Baier, 2010c; Fachner & Carter, 2015; Fachner & Thorkildsen, 2015; Jager et al., 2013; C. Pfeiffer, Ellrich, & Baier, 2010; J. K. Stewart, Fachner, King, & Rickman, 2012; Timmer & Pronk, 2011). For example, officers may be confronted with suspects, who threaten them, actively resist arrest, punch, kick, or shoot at them. These situations represent the climax of a dangerous situation, that often started to unfold in the interaction before the violence occurred (Binder & Scharf, 1982; Pinizzotto et al., 2012). Such situations may be described as acutely “stressful” compared to chronically “stressful”, with the latter studied more intensively in the policing context than the former (G. S. Anderson et al., 2002; Anshel, 2000; Burke, 1994; C. L. Cooper, 2003; Deschamps, Paganon-Badinier, Marchand, & Merle, 2003; Hem, 2004; Kirkaldy, Cooper, & Ruffalo, 1995; Vena, Violanti, Marshall, & Fiedler, 1986; Violanti & Aron, 1995).

2.3.1 Stress, Homeostasis and Allostasis

It is widely agreed that stress is an experience which is common to all living organisms (Haber, 2009), however, there are numerous different definitions (Staal, 2004). Stress was originally defined as the non-specific response of a body to any noxious stimulus (Selye, 1950). Later on, the concept was...
refined to distinguish between “stressor” and “stress response” (Koolhaas et al., 2011). A stressor in this context is defined as a stimulus that threatens homeostasis (Chrousos, 2009), which is defined as the active maintenance of a steady internal state, in the face of an ever-changing external environment (Cannon, 1929; Levenhagen et al., 2001; McEwen & Wingfield, 2003). The stress response is the reaction of the organism to regain homeostasis.

When homeostasis is compromised, the organism changes the parameters of its internal milieu by matching them appropriately to the demand of the environment; a process which is termed allostasis (Juster, McEwen, & Lupien, 2010; McEwen, 2000; 2004; 2007; McEwen & Seeman, 1999; McEwen & Wingfield, 2003; Sterling, 2004). Allostasis prioritises vital homeostasis maintaining systems at the cost of non-vital functions (D. S. Goldstein & McEwen, 2002). It is mediated by the complex interaction of various systems and hormones; the two main systems being the hypothalamic pituitary adrenocortical axis and the sympathetic adrenomedullary system (Koolhaas et al., 2011; McEwen, 2000; 2007).

During allostasis change takes place in the central nervous system and in various peripheral organs and tissues (Chrousos, 1992b; 2009). Concerning the central nervous system this includes (a) the facilitation of arousal, alertness, vigilance, cognition, attention and aggression (Chrousos, 2007; Haller & Kruk, 2001), (b) the inhibition of vegetative functions, like reproduction, feeding and growth (Chrousos, 1992b; 1997; 1998; 2007) and (c) the activation of counter-regulatory feedback loops (Chrousos, 1992b). Changes in the periphery include (a) increase of oxygenation (Chrousos, 1992b; Paisansathan et al., 2007; Palm & Nordquist, 2011), (b) nutrition of the brain, heart and skeletal muscles (Chrousos, 1992b; A. D. M. Greenfield, 1962), (c) increase in cardiovascular tone and respiration (Chrousos, 1992b; P. Grossman, 1983), (d) increase of metabolism (Chrousos, 2007; Taché & Bonaz, 2007), (e) increase of detoxification of metabolic products and foreign substances (Chrousos, 1992b) and (f) activation of counter-regulatory feedback loops including immunosuppression (Chrousos, 1995; 2000; Elenkov & Chrousos, 1999; Elenkov et al., 2008; Elenkov, Ppanicolaou,
Wilder, & Chrousos, 1996; Karalis et al., 1991). These changes serve as time-limited adaptive functions with regards to preparing the body for action (Berthoud, Kressel, & Neuhuber, 1995; 1992; Cunningham & Sawchenko, 1989; Tsigos & Chrousos, 2002).

Revising the concept of stress Koolhaas et al. (2011) consider stress as a cognitive perception of uncontrollability and/or unpredictability that is expressed in a physiological and behavioural response. In this definition of stress, the term “is restricted to conditions, where environmental demands exceeds the natural regulatory capacity of an organism, in particular situations that include unpredictability and uncontrollability” (Koolhaas et al., 2011, p. 1291). Unpredictability is elicited by the absence of an anticipatory response, whereas loss of control is characterized by a delayed recovery response and the presence of an adrenaline response (Koolhaas et al., 2011). There is a lot of evidence that these two situational features are central to the concept of stress (Ely & Henry, 1978; Fokkema, Koolhaas, & van der Gugten, 1995; Manuck, Kaplan, & Clarkson, 1983; Salvador, 2005; Sapolsky, 1995).

The more narrow definition of Koolhaas et al. (2011) avoids confusion with normal physical reactions that are mandatory to support behaviour, and thus are adaptive. With repeated exposure to a stimulus, predictability and controllability may increase. Hence environmental challenges may initially be perceived as stressful, but due to habituation these situations will still be demanding, but not termed as stressful. However, the intensity of the stressor has also to be considered: while a stressor may be mild in terms of its potential consequences, other stressors may be life-threatening. The relationship between the degree of uncontrollability/unpredictability and the intensity (safe vs. life-threatening) of the stressor is depicted in Figure 9.

The systematic habituation to a stressor is an important aspect of police training and is also known as stress inoculation training or stress exposure training in the literature (Artwohl, 2003; Baumann, Gohm, & Bonner, 2011; Driskell & Johnston, 1998; Driskell, Salas, Johnston, & Wollert, 2008; Driskell, Salas, Saunders, & Hall, 1996; McClernon, McCauley, O’Connor, &
According to the conceptualisation of Koolhaas et al. (2011), life-threatening situations like firearms incidents for patrol officers would always be stressful events due to the unpredictability and uncontrollability of the environmental challenge. However, for specialised units (such as SWAT teams) that routinely engage in armed confrontations such situations become more controlled and predictable.

Regardless of the definition of the term “stress”, the demands of police-citizen encounters (stressor) place load on the officers that has to be coped with. Such situations only become stressful, if the load is acutely or chronically unsustainable.
2.3.2 Demand

Allostasis can be induced by physical, cognitive and emotional demands (Backs & Seljos, 1994; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Veltman & Gaillard, 1998). Even though placing load on different structures, the neural activation pattern involved in these demands are similar (Critchley, Corfield, Chandler, Mathias, & Dolan, 2000a; Critchley, Elliott, Mathias, & Dolan, 2000b; Critchley, Mathias, & Dolan, 2001a).

Physical demand is placed upon the individual, when tasks of physical nature (e.g. standing, walking, climbing stairs, running, and fighting) are performed. In a laboratory setting, HR will go up in relation to the intensity of the performed physical activity (G. S. Anderson et al., 2002). Law enforcement personal, that operate in the public are subject to these basic physiological demands. However, in tasks that involve fighting or shooting (e.g. fight for survival or armed confrontations) physiological demand is much higher (Bouhlel et al., 2006; M. R. Davis et al., 2016; Joel, Paiva, & Andreato, 2014). This is supported by various studies that have shown, that regular police officers encounter many forms of physiological demand throughout their work (G. S. Anderson et al., 2002; G. S. Anderson, Plecas, & Segger, 2001; Bonneau & Brown, 1995). Physiological demand in law enforcement personal has been identified through changes in cardiovascular function and the release of corticosteroids (Barton et al., 2000; Colin, Nieuwenhuys, Visser, & Oudejans, 2013; Meyerhoff, 2004; Nieuwenhuys & Oudejans, 2011b; Nieuwenhuys, Savelsbergh, & Oudejans, 2011; Taverniers, Smeets, Van Ruysseveldt, Syroit, & Grumbkow, 2011). Obviously, specialist operators, that are specially trained for armed encounters, have to cope with high physiological demand on a regular basis (Abelbeck, Lipsey, & Reiser, 2011; Andersen & Papazoglou, 2016; Galyean, Wherry, & Young, 2009; Garbarino, Cuomo, Chiorri, & Magnavita, 2013; Taverniers, Van Ruysseveldt, Smeets, & Grumbkow, 2010). It is noteworthy in this context, that physiological arousal can contribute to the generation of perceived affective demand via visceral sensory feedback in the insular cortex (Critchley, 2005; Critchley & Seth, 2013).
Cognitive demand can be described as the mental effort that is required for task completion. Tasks involving executive cognitive demand have been shown to be associated with activation of the anterior cingulate, amygdala, and prefrontal cortex (Collette, Hogge, Salmon, & Van der Linden, 2006; Garavan, Ross, Murphy, Roche, & Stein, 2002; J. G. Kerns et al., 2004; A. M. Owen, Evans, & Petrides, 1996; E. E. Smith & Jonides, 1999). Cognitive demand like complex decision-making under time pressure has been shown to increase physiological arousal (Fairclough & Houston, 2004; Veltman & Gaillard, 1998). Conversely aerobic exercise has been shown to lead to increases in cognitive performance (Hillman, Belopolsky, Snook, Kramer, & McAuley, 2004; Hillman, Weiss, Hagberg, & Hatfield, 2002; Lambourne, Audiffren, & Tomporowski, 2010; Tomporowski, 2003; Tomporowski & Ellis, 1986). In police officers cognitive demand has been identified through changes in mental effort and WM function (Colin et al., 2013; Nieuwenhuys et al., 2011; 2012; Nieuwenhuys & Oudejans, 2010; 2011b; Taverniers et al., 2011).

Affective demands refer to demands that; (a) are associated with the experience and expression of emotions in oneself or in others, (b) stem from the regulation of one’s emotions, and (c) involve meeting other people’s emotional needs (Tuxford & Bradley, 2013). In the context of policing all three aspects regularly occur: police officers are confronted with hostility and threat (Jager et al., 2013). In (un)armed confrontations the fear of injury and pain is constantly present (D. Grossman & Christensen, 2007; Nieuwenhuys et al., 2012; Renden, Landman, Savelsbergh, & Oudejans, 2015a). Nevertheless, in order to ensure individual safety, the officer has to deliver optimal performance in emotionally demanding situations (Staller, 2011). Emotional labour has also to be displayed by police officers in police-citizen encounters. Various studies have shown that negative emotions can impact policing performance negatively (Klukkert, Ohlemacher, & Feltes, 2009; van Reemst, Fischer, & Zwirs, 2015). Therefore, the regulation of emotions can also be considered as part of the professional practice of police officers (Bochenek, 2014). Finally, officers have to engage in emotion work in order to resolve conflict in a non-violent way. Understanding, dealing with, and
actively influencing other people’s emotions is a crucial feature of the negotiation process with aroused individuals (Matusitz, 2013; McMains, 2002; Mullins, 2002; Vecchi, 2009; Vecchi, Van Hasselt, & Romano, 2005; Zaiser & Staller, 2015). Various studies have identified emotional demands in law enforcement personal through changes in self-reported anxiety, aggression, and distress (Barton et al., 2000; Colin et al., 2013; Meyerhoff, 2004; Nieuwenhuys et al., 2011; 2012; Nieuwenhuys & Oudejans, 2010; 2011b; Taverniers et al., 2011; Vrij et al., 1994; 1995; Vrij & Dingemans, 1996).

2.4 Decision Making

In natural environments there is the problem of potentially catastrophic events that are of low or unknown probability (Hinds et al., 2010; Woody & Szechtmman, 2011). In the context of police work, the occurrence of violence against officers and the use of deadly force can be considered as such events. The decisions made in these situations have to be as optimal as possible, since they could have severe consequences. In police-citizen encounters officers are required to make decisions under various environmental constraints: First, information about what is happening is rarely fully available and any information is likely to be out of date by the time the officer arrives on scene (R. Miller, 2012; Sharps, 2010). Second, such situations involve the risk of physical harm to themselves, the public, and/or the suspect. Whilst the threat of harm is obvious in armed encounters (J. J. Williams & Westall, 2003), the same threat is often not obvious upon contacting the citizen (Pinizzotto et al., 2012). Third, the temporal dynamics of encounters are fast, leading to fast changing situations, where decisions have to be made rapidly and thus are time sensitive (Artwohl & Christensen, 1997; FitzGerald & Bromley, 1998; Jensen & Wrisberg, 2014; L. Miller, 2015). However, it has to be noted, that not all police-citizen encounters, even armed confrontations, are “split-second decisions” (Terrill, 2016). Even though the decision to use force may be made within a short time-frame, it may be the end result of a series of decisions that were made from the start of the encounter (Fridell & Binder, 1992; Klukkert et al., 2009). This interaction, that finally leads to the use of force, may also be influenced by
other events, that predispose the use of force. For example, Chapter 7 investigates the effects of ego depletion caused by previous events on the decision to use force.

A (un-)armed confrontation in the context of policing can be described as a critical, ill-structured and complex problem. Based on the work of several scholars (Funke, 2001; Griffin & Bernard, 2003; Rittel & Webber, 1973; Simon, 1973; Zsambok & Klein, 2014), Groenendaal and Helsloot (2015) compiled a list of the characteristics of such problems (see Table 1).

Table 1: Characteristics of Ill-Structured, Complex Problems (Groenendaal & Helsloot, 2015)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambiguous</td>
<td>It is unclear as to what the problem is</td>
</tr>
<tr>
<td>Opaque</td>
<td>The problem’s cause and effect is elusive</td>
</tr>
<tr>
<td>Unpredictable</td>
<td>It is unclear how the problem will develop</td>
</tr>
<tr>
<td>Uncertain</td>
<td>It is unclear if information is complete, accurate and still relevant</td>
</tr>
<tr>
<td>Unstable</td>
<td>The problem is susceptible to change</td>
</tr>
<tr>
<td>Fragmented</td>
<td>Information concerning the problem is available from a number of sources</td>
</tr>
<tr>
<td>Multiple interests</td>
<td>A problem involves a variety of stakeholders with sometimes conflicting interests</td>
</tr>
<tr>
<td>Limited Feedback</td>
<td>The consequences of actions taken are not always immediately evident</td>
</tr>
</tbody>
</table>

In order to account for how decisions are made within these contexts, the following sections briefly outline decision making theories, that could contribute to the understanding of how police officers decide under these high-pressured situations.

2.4.1 **Classical Decision Making**

The majority of classical decision making theories are normative and prescriptive and thus aim at identifying the best decision to take (Camerer & Fehr, 2006; S. H. Kaplan, Greenfield, Gandek, Rogers, & Ware, 1996). This approach included the assumption, that the decision maker is fully informed,
purely rational and able to compute accurately (Beach, 1997; M. S. Cohen, Freeman, & Thompson, 1998; Dawes, 1979; 1989; W. Edwards, 1954; Lipshitz et al., 2001). Hence, decisions in this context are made by following prescriptive rules, evaluating options, and seeking maximum output for minimal costs (Green, 2006; Kahneman, 2003; Tversky & Kahneman, 1986). Classical decision making theories assess decision making according to its utility relative to the cost-benefit ratio of the implemented actions (Green, 2006). In order to do so, classical decision making paradigms typically consist of well-defined problems, stable environments, a decision maker, who is not invested in the outcome and a single well understood goal. Hence, within this framework situational factors are removed that are usually present in real world decision tasks (G. A. Klein, 2014; Orasanu & Connolly, 1993; Orasanu & Martin, 1998).

The classical decision paradigms are primarily applied to fields such as economics, business, and politics (Cabantous & Gond, 2011; Gabor, 1976; Green, 2006; Simon, 1979; Tversky & Kahneman, 1986). However, it has to be noted that the ecological validity of such work outside these areas is weak (Brehmer, 1992). Furthermore, this analytical or rational decision making style can lead to negative utility under conditions of uncertainty and time pressure (W. Edwards, 1954; Lipshitz & Strauss, 1997). In this context Simon (1955) introduced the concept of satisficing in order to point out, that decision makers typically do not do all the work to ensure optimal decisions: Rather than maximising utility, people choose a good enough option, which is not necessarily the best one. By using an adequacy criterion a decision is made whether an option is satisfactory or not. Then, the first option meeting this criterion is chosen.

Taken together, traditional decision making with its analytical approach is based on evaluating options and selecting statistically optimal solutions (T. Elliott, 2005). However, naturalistic environments, with constrains like time-pressure, high stakes and uncertainty, create contexts, where the evolution of options is not possible.
Naturalistic Decision Making (NDM) focuses on how individuals make decisions in applied “real world” settings (G. A. Klein, 2008; 2015). The emphasis in this context is placed upon how decisions are made rather than how decisions should be made (Lipshitz et al., 2001; Zsambok & Klein, 2014). The field of NDM has defined an area of research studying expert decision makers in real time, dynamic, high stakes events through the use of a number of methodologies (Baumann, Sniezek, & Buerkle, 2001; McAndrew & Gore, 2013).

By locating decision making in context, NDM models are applicable to real world situations that are defined by complex and ill-structured problems as described in Table 1 (Cannon-Bowers & Salas, 1998a; Endsley et al., 2007; Flin et al., 1997; G. A. Klein, 2008; 2009; 2014; Lipshitz et al., 2001; Orasanu & Connolly, 1993; Orasanu & Martin, 1998; Parent & Verdun Jones, 1998; Saus et al., 2006; Zsambok & Klein, 2014).

NDM studies have investigated decision making in a variety of contexts, such as the police (Boulton & Cole, 2016), the military (Kaempf, Klein, Thorsden, & Wolf, 1996; G. A. Klein, Klein, Lande, Borders, & Whitacre, 2015), hospital intensive care (Crandall & Getchell-Reiter, 1993; Fackler et al., 2009), and the aviation domain (Orasanu, 1995). The results have highlighted several important factors within these contexts, including SA, situation assessment, mental modelling, pattern matching, and recognition (G. A. Klein, Calderwood, & Clinton-Cirocco, 1986).

Decisions in emergency situations have to be often made under conditions of acute demand, due to the existing constraints in these contexts (e.g. high risk, little time, and incomplete information). These constraints impact professional judgements in such situations (Greenwood-Ericksen et al., 2004; Kowalski-Trakofler, Vaught, & Scharf, 2003; Mitchell & Flin, 2007). However, it seems that professional decision making is not always compromised under demand: Since different situations call for different kinds of cognitive activity (Kowalski-Trakofler et al., 2003), this may represent an adaptive response to cope with the increased cognitive load of such
emergency situations (J. D. Cohen et al., 1997; Endsley, 1995; Flin et al., 2007; G. A. Klein et al., 1986; G. A. Klein, Calderwood, & MacGregor, 1989; Kowalski-Trakofler et al., 2003; Lipshitz et al., 2001).

Therefore, in the “real world” experts use their experience to identify meaningful cues, generate reasonable options, using simple heuristics to select a course of action and then implement the first satisfying solution (G. A. Klein, 2009), rather than methodically calculate utility. However, classical and NDM are complementary: in operational settings individuals employ both styles of decision making (Kahneman & Klein, 2009; Killion, 2000).

Two systems (or dual process) models (J. S. B. T. Evans, 2007; J. S. B. T. Evans & Frankish, 2009; Kahneman, 2011; Kahneman & Klein, 2009) offer further insight, about which type of decision making is used and when. In the terminology of these models, System 1 processes refer to intuitive, fast decision making (i.e. NDM), whereas System 2 operations are controlled, voluntary and effortful. The vast majority of decisions are made through System 1 processes since this is typically the most efficient in terms of using mental and time resources to solve problems and achieve goals (Kahneman & Klein, 2009).

2.4.2.1 Recognition and Pattern Matching

In contrast to Kahneman (2011), the work of Klein and colleagues (G. A. Klein, 2009; 2015; K. G. Ross, Klein, Thunholm, Schmitt, & Baxter, 2004; Zsambok & Klein, 2014) has focused on examining how practitioners make professional fast System 1 decisions in pressurised circumstances. Early NDM research with firefighters indicated that decision makers in naturalistic settings make decisions in a seeming less effortless, fast paced, automatic manner (G. A. Klein et al., 1986). Investigating the contradictory findings compared to traditional decision making approaches, Klein et al. (1986) found that commanders were not comparing optimal choices but relied on the first plausible option they considered. Based on this work, G. A. Klein et al. (1989) proposed the recognition-primed decision making model, which describes the decision making process in three distinct patterns (see Figure 10).
In the first variation, the decision maker examines the situation and responds with the initial option that was identified. In the second variation, commanders would “story build” (mentally simulate) events that would have lead up to the observed features and make decisions based on these simulations (G. A. Klein et al., 1989). In the third variation, the decision maker mentally simulates a course of action, evaluating if this option would work. The simulations aim at looking for unintended consequences as part of the process (G. A. Klein et al., 1989; Lipshitz, Klein, & Carroll, 2006). In order to identify cues, to generate the required mental maps, and to make projections of outcomes (G. A. Klein & Calderwood, 1991), the decision maker needs SA and strong pattern matching abilities (G. A. Klein et al., 1986).
### Figure 10: Recognition-primed decision model (J. K. Phillips, Klein, & Sieck, 2004)

**Variation 1: Simple Match**
- Experience the Situation in a Changing Context
- Perceived as typical [Prototype or analog]
- Recognition has four byproducts
  - Expectancies
  - Relevant Cues
  - Plausible Goals
  - Typical Action
- Implement Course of Action

**Variation 2: Diagnose the Situation**
- Experience the Situation in a Changing Context
- Diagnose [Feature Matching, Story Building]
  - Inference
- Is Situation Typical? [Prototype or Analog]
- More data
- Recognition has four byproducts
  - Expectancies
  - Relevant Cues
  - Plausible Goals
  - Typical Action
- Clarify
- Anomaly
- Evaluate Action (n) [Mental Simulation]
  - Will it Work?
  - Modify yes, but no
- Yes
- Implement Course of Action

**Variation 3: Evaluate Course of Action**
- Experience the Situation in a Changing Context
- Perceived as typical [Prototype or analog]
- Recognition has four byproducts
  - Expectancies
  - Relevant Cues
  - Plausible Goals
  - Action 1...n
- Recognition has four byproducts
  - Expectancies
  - Relevant Cues
  - Plausible Goals
  - Typical Action
- Clarify
- Anomaly
- Evaluate Action (n) [Mental Simulation]
  - Will it Work?
  - Modify yes, but no
- Yes
- Implement Course of Action
2.4.2.2 Situation Awareness

Even though recognition of cues is crucial for proficient decision making, other processes, like SA, are also fundamental for success (M. S. Cohen et al., 1998). SA is regarded as one of the primary factors that influence NDM (Pfaff et al., 2013) and is simply defined as “knowing what is going on around you” (Endsley, 2000, p. 2). It is the perception of elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status into the near future (Endsley, 1988).

![Model of situation awareness in dynamic decision making](Endsley, 2000)

SA comprises of three different levels (see Figure 11): First, the perception of cues in the current situation, which is a fundamental aspect, since without the basic perception of important information, the odds of forming an incorrect picture of the situation increases dramatically. Second, it encompasses how people combine, interpret, store, and retain information. The integration of
multiple pieces of information and the determination of the relevance to the individual's goals forms the second level of SA. Third, the highest level of SA is the ability to forecast future situation events and dynamics (Endsley, 2000). In Endsley's model SA is conceptualized as a stage separate from decision making and performance. The reason being, that it is possible to have perfect SA without making correct decisions. Vice versa, good decisions can be made (by luck) with poor SA (Endsley, 2000).

In naturalistic conditions, cognitive resources are allocated to gather and update SA, in order to prepare for rapid changes in the situation. Conversely, the reasons for incorrect decision making frequently can be found in incorrect SA rather than in a flaws in the decision making process (Endsley, 1995; Endsley et al., 2007). SA involve schemata that direct information searching and that actively build mental models of the situation (Lipshitz & Shaul, 2014; E. M. Roth et al., 2010). These models help expert decision makers to quickly focus on critical aspects of the situation through recognition and pattern matching (Lintern, 2010; Lipshitz & Shaul, 2014).

In complex situations, such as police-citizen encounters or (un)armed confrontations, SA may be incomplete or uncertain. Therefore, meta-cognitive control is needed in order to allow the decision maker to critique, correct, and re-interpret information (J. D. Cohen et al., 1997; Lintern, 2010). Police and military personnel rely on SA in varying and complex situations in order to determine a course of action (Endsley, 1995; Füllgrabe, 2014; Matthews & Beal, 2002; P. M. Salmon, Stanton, Walker, & Jenkins, 2009; Saus et al., 2006; Strater, Jones, & Endsley, 2001). SA becomes more proficient the more an individual gathers experience in a particular environment. For example, novices may be aware of perceptual cues, but do not comprehend them in order to influence future actions (Fleischmann, 1995). In these situations, cognitive load is the result of task relevant maintenance and manipulation of large sets of information, which relies heavily on WM and executive functioning (Norman & Shallice, 1986; Shallice & Burgees, 1993).
2.5 Working Memory

Memory consists of three different systems with differing neuronal activation and neuropsychological functioning (J. Brown, 1958; Milner, 1963; L. Peterson & Peterson, 1959; Shallice & Warrington, 1970): (a) sensory stores, that register external environmental cues, (b) short-term memory, that stores and represents this information, and (c) long-term memory, that 'permanently' stores processed information as knowledge (R. C. Atkinson & Shiffrin, 1968).

Initially short-term memory was proposed to act as a gate for the transference of information to long-term memory. However, Baddeley and Hitch (1974) proposed their seminal concept of WM, as an active system that processes information (Baddeley, 1992; 2000; 2007; Engle, 2002). The concept of WM became a major topic of cognitive research with researchers proposing various kind of models (Fuster, 2001; Goldman-Rakic et al., 1993; Werheid et al., 2002).

WM enables the encoding, consolidation, active maintenance, manipulation and regulation of information that is necessary for the completion of complex cognitive tasks (Frank et al., 2001; Jonides et al., 1993; Miyake & Shah, 1999; Rosenthal et al., 2006; Werheid et al., 2002). The information stems from sensory stores and interacts with long term memory in order to guide behaviour (Baddeley, 2001; Baddeley, Gathercole, & Papagno, 1998; A. M. Owen, McMillan, Laird, & Bullmore, 2005). In contrast to long-term memory, WM is a limited capacity system. It is responsible for the temporary storage and manipulation of information that is required for the completion of cognitive and/or physical tasks (Baddeley, 2000; Conway & Engle, 1994; Miyake et al., 2000). As such it is also involved in defensive behaviours (Miyake et al., 2000). WM is sensitive to cognitive load, which may affect performance (Chase & Simon, 1973; Just & Carpenter, 1992; Shachak, Hadas-Dayagi, Ziv, & Reis, 2009).
2.5.1 The Multi-Component Model of WM

Baddeley and Hitch (1974) conceptualized WM as a theoretical multicomponent model that includes distinguishable processes consisting of temporary memory stores and the regulatory mechanism of executive attention. Among these are an attentional control system, the central executive, and three temporary storage systems: (a) the visuospatial sketchpad, (b) the phonological loop, and (c) the episodic buffer as a more general integrated storage system (Baddeley, 1996; 2007; 2010; 2012). The model is depicted in Figure 12. Initially the model consisted of three subsystems; the episodic buffer was added to the model upon later review (Baddeley, 2000).

The central executive is the most complex and most important subsystem of this multicomponent model (Baddeley, 2012). It maintains attentional control over the functioning of the subsidiary slave systems, which preserve incoming information from the immediate environment or the long-term memory stores for the completion of a current task (Baddeley, 2007; Pechmann & Mohr, 1992; Rosenthal et al., 2006). The phonological loop comprises a temporary phonological store in which auditory memory traces decay over a period of few seconds, unless they get revived by articulatory rehearsal (Baddeley, 2000). The visuospatial sketchpad integrates, stores and manipulates visual and spatial information on a temporarily basis (Baddeley, 2003b). Finally, the episodic buffer is a limited-capacity temporary storage system that is capable of integrating information from a variety of sources (Baddeley, 2000). It binds information from the phonological loop, the visuospatial sketchpad, and the long-term memory into a unitary, episodic representation. This enables the individual to consider multiple sources of information simultaneously, creating an model of the environment that may be manipulated to solve problems and to plan behaviour (Baddeley, 2001).
Functional neuroimaging, lesion, and behavioural studies suggest that the three basic components of WM (central executive, phonological loop, and visuospatial sketchpad) are located in different brain regions and that the functioning of each one can be individually and separately impaired (Baddeley, 2003a; Basso, Spinnler, Vallar, & Zanobio, 1982; De Renzi & Nichelli, 1975; Hanley, Young, & Pearson, 1991; Logie, Venneri, Sala, Redpath, & Marshall, 2003; Salla & Logie, 2002; Guiseppe Vallar & Papagno, 2002; P. P. Wang & Bellugi, 1994). This supports the idea of the multi-component structure of the model.

2.5.1.1 Phonological Loop

The phonological loop holds phonological information in a limited capacity storage for unfamiliar sound patterns and verbal information (Baddeley, 2000; Baddeley et al., 1998; D. M. Jones, Macken, & Nicholls, 2004; Saito, 2001). It comprises a phonological store, which holds phonological information, and an articulatory control process, which refreshes and maintains decaying representations via rehearsal (Baddeley, 2000; 2001;
Baddeley & Wilson, 1985; Caplan & Waters, 1995; Logie et al., 2003; Saito, 2001; Giuseppe Vallar & Baddeley, 1984). Even though the precise neuroanatomical structures and mechanics are still unclear, the phonological loop is accountable for the majority of effects concerning the phonological component of WM (Baddeley, 1968; 2012; Baddeley & Larsen, 2007; Baddeley, Hitch, & Allen, 2009; Baddeley, Lewis, & Vallar, 1984; Baddeley, Thomson, & Buchanan, 1975; Colle & Welsh, 1976; Conrad, 1964; Conrad & Hull, 1964; Logie et al., 2003; Salamé & Baddeley, 1986). It is assumed to be the best developed component of the multi-component model of WM (Baddeley, 2000; 2012).

2.5.1.2 **Visuospatial Sketchpad**

The visuospatial sketchpad is the visual equivalent of the phonological loop. It is limited in capacity (typically to about three or four objects) and it operates in a similar fashion as the phonological loop for the temporary storage and manipulation of visual and spatial information (Baddeley, 2003b; 2007; Sims & Hegarty, 1997). It is the visual equivalent of the phonological loop (Baddeley, 2007; Bruyer & Scailquin, 1998) and is comprised of two components: The visual cache is the passive storage component of the system and the inner-scribe as an active system for visuospatial information rehearsal (Baddeley, 2001; Bruyer & Scailquin, 1998; Logie, 1986).

Since compared to the phonological loop, experiments for investigating the visuospatial part of WM are difficult to design, the sketchpad is poorly understood (Baddeley, 2003b; Salla & Logie, 2002). Hence, the precise nature of function and the neuroanatomical structure of the visual component of WM is still debateable (Baddeley, 2003b). Compared to phonological processing, visual processing is more complex, resulting in increased executive demand, requiring attentional and executive control (Baddeley, 1996; 2007)

2.5.1.3 **Central Executive**

The central executive is the most complex and most extensively investigated theoretical construct of WM (Baddeley, 2012; Collette & Van der Linden,
2002). The central executive has the capacity to; (a) focus attention, (b) divide attention, (c) switch attention, and (d) provide a link between WM and long-term memory (Baddeley, 1996; J. Holmes et al., 2008; Miyake et al., 2000). It is assumed to be an attentional control system that is responsible for strategy selection and for control and coordination of various processes that are involved in short-term storage and in more general processing tasks (Baddeley, 2007; Collette & Van der Linden, 2002). The central executive is required in situations; (a) involving planning or decision making, (b) involving error correction or trouble-shooting, (c) which contain novel sequences of action, (d) which are judged dangerous or technically difficult, and/or (e) require the inhibition of a habitual response (Baddeley, 1996; 2000; D. E. Meyer & Kieras, 1997; Norman & Shallice, 1986; Perner & Lang, 1999; Shallice & Burgees, 1993). These processes are generally described as executive functions (Norman & Shallice, 1986).

Executive functions control and moderate the operation of various cognitive sub-processes, determining decision making and the implementation of action (N. P. Friedman et al., 2006; Miyake et al., 2000). They enable planning, forethought and goal-directed action, thus regulating decision making (A. Diamond, 2013; Lezak, 1982; Shields, Sazma, & Yonelinas, 2016b; Suchy, 2009; P. G. Williams, Suchy, & Rau, 2009). This process requires the ability to initiate, self-monitor, and correct action based on environmental feedback (Collette & Van der Linden, 2002; Collette et al., 2005).

The frontal lobes, in particular the prefrontal cortex have been identified as crucial to executive function (Baddeley, 1996; Collette et al., 2005; D'Esposito et al., 1998; D'Esposito, Zarahn, Aguirre, & Rypma, 1999; M. J. Kane & Engle, 2002; Lezak, 1982; A. M. Owen, Roberts, Polkey, Sahakian, & Robbins, 1991; Postle, Berger, & D'Esposito, 1999; Stuss & Benson, 1984; Werheid et al., 2002). It is activated during tasks that involve the shifting of attention, updating and/or inhibition (Ardila, 2008; D'Esposito et al., 1999; De Linden et al., 1999). It has been shown, that lesions in the prefrontal cortex are associated with difficulties in the completion of executive tasks, especially when these involve novel or complex stimuli (Fuster, 2001).
Impairments in executive functioning can result in several performance decrements, namely a reduced ability to update task relevant information, shift attention and inhibit irrelevant cues. In the context of police-citizen encounters, all of these could lead to potentially severe errors in tactical decision making (Boulton, 2014). Furthermore, executive cognitive functioning deficits are hypothesised to be an aetiological factor in aggression and violent behaviour in general (Paschall & Fishbein, 2002). However, it is not always beneficial to exert executive (top-down) control. When performing a new task, individuals who recruit the lateral prefrontal cortex most often perform best (Duncan & Owen, 2000; Poldrack et al., 2005). In contrast, this region is recruited least, when something is performed that is no longer new (Chein & Schneider, 2005; Garavan, Kelley, Rosen, Rao, & Stein, 2000; Landau, Garavan, Schumacher, & Esposito, 2007; Milham, Banich, Claus, & Cohen, 2003).

In the context of simulated (un-)armed confrontations, this has been shown for police officers (Condon, 2015; A. P. J. Roberts, 2012): executive functioning performance increased, when officers had to tactically adapt to a novel problem. In contrast executive functioning decreases as soon as officers could rely on well-trained standard operating procedures (SOP).

2.5.1.4 Episodic Buffer

The episodic buffer is the latest introduced component in the multi-component WM model (Baddeley, 2000). It is assumed to be a limited capacity store that binds together information to form integrated episodes (Baddeley, 2003b). It can be understood as a buffer store between the components of WM, that also links WM to perception and long-term memory (Baddeley, 2012). Furthermore, it is assumed to be attentionally controlled by the central executive and to be accessible to conscious awareness (Baddeley, 2003b). Hence, the episodic buffer is regarded as a fundamental feature of the capacity of WM.
2.5.2 Alternative Models of WM

There are number of alternative conceptualisations of WM. However, these models similarly propose WM to be a multi-faceted construct (Conway et al., 2005; M. J. Kane, Conway, Miura, & Colflesh, 2007; Miyake et al., 2000; Miyake & Shah, 1999).

The embedded processes model (Cowan, 1995; 1999) views WM as an active part of the long-term memory that is currently the focus of attention. Although the model does not account for abstract subsystems, it accepts the possibility of domain specific representations (Baddeley, 2007). According to the embedded processes model limitations in WM capacity stem from: (a) a lack of active maintenance as information decays across time, and (b) a limited-capacity attention system.

The biologically-based computational model maps WM function to specific brain structures (O'Reilly & Frank, 2006; O'Reilly, Braver, & Cohen, 1999). In this model WM is viewed to be domain general without being a unitary construct. The prefrontal cortex and the hippocampus are hypothesised to actively maintain information while the posterior perceptual and motor cortex is used to process information and complete actions. The model acknowledges the existence of domain specific storage systems; however, these are not considered to be part of what the model defines as WM. The biologically-based computational model could be considered to explain executive processes rather than WM in general (O'Reilly et al., 1999), which demonstrates that a differences in the emphasis and in definitions of WM can be accounted for the differences between models (Baddeley, 2003a; Miyake et al., 2000).

The current thesis focuses on Baddeley’s multi-component model of WM, since the abstract separateness of the components facilitates experimentation. However, in reality the constructs are more intertwined on a functional and neurobiological level (Baddeley, 2007; Fletcher, 1998; Fuster, 2001; 2002).
2.6 Self-Control

In a variety of domains, humans are able to exert deliberate control over impulses (Ainslie, 1975; Eisenberg et al., 2003; Fujita & Han, 2009) and direct behaviour towards future goals and long-term desirable outcomes (Fishbach & Labroo, 2007; Logue, 2010) at the expense of satisfying immediate urges and temptations (Metcalfe & Mischel, 1999; Mischel, Shoda, & Rodriguez, 1989). This ability to inhibit, override, or otherwise circumvent responses that are motivated by short-term rewards for the sake of long-term benefits is known as self-control (Casey, 2015; Fujita, 2011; Hofmann, Friese, & Strack, 2009; Mischel et al., 1989; J. J. Reynolds & McCrea, 2016; Vohs & Heatherton, 2000). Despite this capacity for self-regulation, there are numerous examples in which self-control fails. For example, risky sexual behaviour, overeating, drug and alcohol abuse, and aggression are consequences of a failure to regulate behaviour (Baumeister, Heatherton, & Tice, 1993; Denson, DeWall, & Finkel, 2012; DeWall, Finkel, & Denson, 2011; Finkel, DeWall, Slotter, Oaten, & Foshee, 2009; Quinn & Fromme, 2010; Slotter et al., 2012; J. D. Stinson, Becker, & Sales, 2008).

2.6.1 Theories of Self-Control

Research into the field of self-regulation has mainly focused on the “strength model” (or resource model) of self-control (Baumeister & Heatherton, 1996; Baumeister, Bratslavsky, Muraven, & Tice, 1998; Heatherton & Baumeister, 1996; Vohs & Heatherton, 2000). The central tenet of this theory is that self-regulation is governed by a limited resource that allows people to control impulses and desires. These self-regulatory resources can be depleted or fatigued by self-regulatory demands. If demands on self-control are very high and/or if demands have been maintained over a prolonged period, self-control resources are depleted and subsequent self-control will be impaired. This state of depleted self-control resources has been termed “ego depletion” (Baumeister et al., 1998; Muraven, Tice, & Baumeister, 1998). Recent investigations have shown that exerting self-control reduces the ability to regulate behaviour in a subsequent task, even when the domains of the self-control tasks are different. For example, it was shown, that controlling the
temptation to eat reduced the ability to regulate aggressive behaviour in responding to negative comments by an experimenter (Stucke & Baumeister, 2006). With regards to aggressive behaviour, a study found that ego depletion resulted in increased aggressive responding, especially after an insulting provocation (DeWall, Baumeister, Stillman, & Gailliot, 2007). In addition, Finkel et al. (2009) investigated self-regulatory failure as a predictor of intimate partner violence perpetration. The results showed, that ego depleted participants were more violent in responses to partner provocations than non-depleted participants. They also reported, that participants whose self-regulatory resources were experimentally bolstered in a 2-week training program exhibited less violent inclinations than did participants whose self-regulatory resources had not been bolstered.

Despite emerging evidence on the effects of ego depletion, the specific underlying mechanisms are less clear (Berkman & Miller-Ziegler, 2013). In a meta-analysis, Hagger et al. (2010) showed, that the effects of ego depletion were associated with the degree of self-reported effort that was exerted during self-control tasks, as predicted by the strength model of self-control. Several studies suggest that instead of the actual depletion itself, the perception of being in an ego depleted state accounts for the ego depletion effect (Ackerman, Goldstein, Shapiro, & Bargh, 2009; Alberts, Martijn, & de Vries, 2011; Alberts, Martijn, Greb, Merckelbach, & Vries, 2010; J. J. Clarkson, Hirt, Jia, & Alexander, 2010; Job, Dweck, & Walton, 2010; Job, Walton, Bernecker, & Dweck, 2013): For example, Ackerman et al. (2009) showed that the simulation of self-control results in the depletion of self-control resources, while observing self-control in others actually increases self-control in the observers. Furthermore, the manipulation of beliefs about the availability of self-control resources can also protect from the effects of ego depletion. For example, it has been shown that individuals are resistant to ego depletion manipulations if they are primed to believe that self-control resources are available (J. J. Clarkson et al., 2010) or that self-control is a unlimited resource (Job et al., 2010). Additionally, other factors can also reduce the effects of ego depletion, such as the unconscious priming of persistence (Alberts et al., 2010) or self-awareness (Alberts et al., 2011) or
representing the task of self-regulatory effort in a less tempting manner (Mischel, 2014).

Another idea that has been put forward is that ego depletion of self-control resources is mediated by fatigue in specific brain areas that are involved in behavioural regulation, most likely the prefrontal cortex (Gailliot & Baumeister, 2007). Using functional magnetic resonance Imaging (D. D. Wagner & Heatherton, 2013) found that consuming self-regulatory resources leads to an exaggerated neural response in the amygdala that appears specific to negatively valenced stimuli. This suggests a failure to recruit top-down prefrontal regions involved in emotional regulation. A recent study found, that exercising self-control temporarily increases approach motivation, as revealed by patterns of electrical activity in the prefrontal cortex (Schmeichel, Crowell, & Harmon-Jones, 2016). This suggests, that exercising self-control causes shifts in motivation and attention.

Recently emerging findings showed, that self-control is not energetically dependent as initially proposed (Beedie & Lane, 2012; E. C. Carter, Kofler, & Forster, 2015; J. J. Clarkson et al., 2010; Dang, 2016b; Job et al., 2010; Kurzban, 2010; Magen & Gross, 2007; Molden et al., 2012; Muraven & Slessareva, 2003). For example, recent meta-analytic evidence (E. C. Carter et al., 2015) does not support the proposition, that self-control relies on a limited resource, when tested in laboratory settings. Likewise, a multilab preregistered replication study of the ego depletion effect (Hagger & Chatzisarantis, 2016) with 23 laboratories (N = 2141) using a sequential task paradigm failed to replicate the ego depletion effect, providing evidence, that if there is any ego depletion effect, it is close to zero. However, a recent complimentary analysis of the data (Dang, 2016a) suggests, that the failure of Hagger et al.’s (2016) replication study may result from the ineffectiveness of their manipulation (E-Crossing Task). Analyses of the data sets, where people consider the manipulation as effortful indicates an ego depletion effect (Dang, 2016a).

The discussion around the emerging findings, that (a) self-control is not energetically dependent and (b) the strength of ego depletion effects may not
be as straightforward as initially proposed (E. C. Carter & McCullough, 2014; Inzlicht & Berkman, 2015), suggest the need for alternative conceptualizations of a theory of self-control. Recently, Reynolds and McCrea (2016) proposed the Dual Component Theory of Inhibition Regulation focusing on the functional aspect of self-control (see Figure 13). As such it assumes, that: (a) it is functional to apply self-control in certain situations, in order to regulate impulsive behaviour, (b) it is undesirable for self-control to be applied indefinitely, and (c) self-control varies across contexts.

The model is composed of two distinct components: a monitor and a threshold component. The first detects that an impulsive behaviour needs to be inhibited, producing a signal of inhibition effort that is fed forward to the threshold component, which in turn processes this signal to determine whether the threshold has been met. Therefore, the threshold component represents the individual’s tolerance for applying inhibitory effort. A central implication of the Dual Component Theory of Inhibition Regulation is that inhibitory self-control is viewed as an information processing mechanism, compared to the view of self-control as energy in the resource model. As such the cessation of inhibition is not considered as “failure” per se, as it depends on the context. For example, it can be highly adaptive to be
aggressive (cessation of inhibition to aggress) if a violent threat is placed upon a police officer.

2.6.2 Ego Depletion and Aggression

With regards to aggressive behaviour, there are only a few studies that have tested the effect of ego depletion on aggression (Barlett et al., 2016; DeWall et al., 2007; Stucke & Baumeister, 2006; Vohs, Glass, Maddox, & Markman, 2011). Stucke and Baumeister (2006) showed that participants that were required to resist tempting food or to concentrate on a boring film while stifling their physical and facial movements (depletion condition), subsequently reacted more aggressive to an insult by giving worse feedback to the sender of the insult. Likewise, Vohs et al. (2011) provided evidence that depleted participants elicited more aggressive responses in a game involving blasting noise at an opponent than non-depleted participants. Depletion was manipulated by tasking participants to watch disgusting footage from two movies, while the depletion group was additionally required to show no facial expression and to neutralize their innermost feelings. The results further showed that sleep deprivation had no effect on the aggressive responses. Confirming the results of Stucke and Baumeister (2006), DeWall et al. (2007) showed that aggression was highest when participants were provoked and ego depleted. However, if participants were not insulted and the urge to aggress was relatively weak, ego depletion did not increase aggressive behaviour. In a recent study, systematically testing aggression change as function of ego depletion and provocation, Barlett et al. (2016) found that aggression was highest for ego depleted participants who were provoked. In this study aggression was measured by asking participants to take raffle tickets from an ostensible partner, while depletion was inferred by remembering a big number throughout the experiment.

However, since previous studies only measured aggression indirectly (i.e. rating by the experimenter) or with non-physical forms of aggression (i.e. noise blast or hot sauce) and in the absence of the receiver of the form of aggression, it cannot be inferred that a state of ego depletion transfers to
physical forms of aggression. Hence, evidence for ego depletion leading to physical forms of aggression is still missing.

As the Dual Component Theory of Inhibition proposes, inhibition plays a crucial role in the display of aggression and depending on the context a cessation of inhibition may sometimes be functional (J. J. Reynolds & McCrea, 2016). With regards to policing, the display of aggression is warranted in situations of immediate threat and in situations where compliance cannot be achieved in any other way. Yet, the interpretation of “cannot be achieved in any other way” varies across individuals. As Klukkert et al. (2009) indicated, emotional cognitions of fear of losing authority combined with fear of escalation may trigger offensive aggression towards initially non-compliant citizens. Even though, still in line with legal and institutional guidelines, the display of aggression too early in an police–citizen interaction process may create later costs (Reisig, McCluskey, Mastrofski, & Terrill, 2004; L. W. Sherman, 1993; Tyler & Huo, 2002; Wolfe, 2011). Regulation of self-control especially in the light of provocation would be beneficial in order to further negotiate non-physically (Zaiser & Staller, 2015). However, it is unclear if ego depletion leads to a cessation of inhibition of offensive aggression in a police-citizen encounter.

2.7 Representative Learning and Testing Environments

The current thesis employs simulated (un)armed confrontations as a naturalistic setting to study the various factors that influence tactical decision making in the use of force domain. Existing laboratory models for studying human aggression are routinely criticised over their validity (see 2.1.2). The simulated (un-) armed confrontation paradigm provides an excellent platform for research within this context. In order to foster ecological validity of the research with regards to applicability of the findings in real world contexts, a high level of representative testing design is paramount. The current section provides an overview of the current discussions around the theme of designing representative testing and learning designs in the police use of force domain.
2.7.1 The Transferability of Use of Force Skills

Recently studies (Jager et al., 2013; Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b) revealed that performance is different in use of force training (the learning environment) compared to a real incident (the criterion environment). This difference between the learning environment and the criterion environment is fundamental to the understanding of the acquisition of police use of force skills. The development of skills that transfer into the real world is the underlying goal of police use of force training. This transfer refers to the dependency of current or future behaviour on prior experience (Thorndike & Woodworth, 1901a; 1901b; 1901c). In the context of perceptual motor skills, including use of force skills, transfer involves the capability to use prior experiences from perceptual motor skill performance and learning trials in use of force situations (training sessions or real incidents) and then to adapt these experiences to similar or dissimilar contexts (Collard, Oboeuf, & Ahmaidi, 2007). Therefore, the effectiveness of training programs refers to the transferability of use of force skills from the learning environment to the criterion environment, where optimal performance is needed (see Figure 14).

Figure 14: Representativeness in police use of force
Transferability of skills to real incidents can only be measured through the analysis of performance in the criterion environment. Corresponding studies focus only on self-reports of participants (Jager et al., 2013; Jensen & Simpson, 2014; Jensen & Wrisberg, 2014; Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b). What are missing are analyses of performance in real incidents based on objective data like video footage (e.g. CCTV and body-cams). A major drawback of analysing performance in the criterion environment is the delayed feedback, since it is legally and ethically impermissible to actively seek violent confrontations in order to capture performance after new skills have been taught. Therefore, the performance of combat skills has to be tested in a testing environment that simulates the criterion environment. Valid results about the transferability of self-defence skills can only be obtained if the testing environment is representative to the criterion environment (red arrow in Figure 14). The same is true for the learning environment: the more representative the learning environment, the better the transfer of skills from that environment to the criterion environment (Broadbent, Causer, Williams, & Ford, 2015).

2.7.2 The Simulation of Reality of Use of Force Tasks

Practitioners and scholars in the police use of force and/or self-defence domain regularly refer to “realistic” or “reality-based” training with regards to the design of corresponding learning or testing environments (Andersen, Pitel, Weerasinghe, & Papazoglou, 2016; J. Armstrong, Clare, & Plecas, 2014; Dzida, Hartunian, & Santiago, 2010; Hoff, 2012; K. R. Murray, 2004; Oudejans, 2008; J. Wagner, 2005; Wollert et al., 2011). Yet, there are various definitions and explanations to what the term “realism” exactly refers to in the context of learning environments. For example, Armstrong et al. (2014) define realistic environments as an environment, that “replicates what an officer would expect to encounter in a real-life situation” (p. 52), whereas Hoff (2012) states that the “more realistic the environment, the greater the benefit” (p. 21) without giving further explanations what “realistic” refers to. In the context of scenario-based training, Wollert et al. (2011) point out that a scenario is a simulation of reality and that in order “to be realistic it must ‘feel right’ to the user” (p. 47). Furthermore, they use the term “scenario fidelity” in
order to describe “how accurately the scenario reflects realistic conditions” (p. 47). To accommodate for the evasive nature of the term, they introduced three dimensions: (a) equipment, (b) sensory, and (c) psychological fidelity. Yet, these dimensions do not emphasize the functional properties of the simulation that align with learning or testing objectives. Scholars in the medical domain also suggest abandoning the mere term of “fidelity” in simulation design, due to its imprecise nature and its lack of emphasis regarding functional task alignment (Hamstra, Brydges, Hatala, Zendejas, & Cook, 2014).

At this point it is worth noticing the skill transfer can be fostered in many activities during a training session and not necessarily through the means of scenario-based training (Staller & Bertram, 2015; Staller & Zaiser, 2015b). Nevertheless, a simulation of reality (via scenario based training) is the only viable way to test the effectiveness of technical and tactical solutions to problems encountered in the field (see Figure 15). Deliberate testing of learned use of force skills in the field is ethically impermissible, whereas the testing in ideal conditions leads to the erroneous assumption that generated (technical and tactical) solutions work in the field. Therefore, the simulation of reality has to include conditions that are prevalent in real life incidents (see section 2.3.2). For example, for the simulation of violent encounters, surprising attacks, aggressiveness and high amounts of pressure have to be present (Jager et al., 2013; Jensen & Wrisberg, 2014; B. S. Miller, 2008).

At the same time the scenario designer has to ensure the safety of the participants by omitting the real-world features that bear the risk of injuring participants (S. A. Murray, Yanagi, Ensign, Clark, & Dast, 2010; Wollert et al., 2011). For example, practicing self-defence techniques in highly dynamic and surprising situations using real guns or knives bears the risk of serious injury if the learner makes a mistake. Another option would be to work with real guns or knives, but to drastically reduce the speed, dynamics and surprising character of the situation (Staller & Bertram, 2015).
2.7.3 The Concept of Realistic Training is Flawed

This example illustrates the imprecise nature of the term “realistic” in training or testing environments. Both situations can be described realistic in one aspect, but unrealistic in another aspect. It seems that in most cases practitioners refer to the physical resemblance of the training setting as being resembling reality or not. Yet, from a learning perspective, the “functional alignment with the learning task, the instructional design, and the instructor is likely have far greater impact on immediate learning, retention and transfer to new settings” (Hamstra et al., 2014, p. 389).

Based on these observations, the current thesis employs the term “representativeness” (in learning and testing designs) instead of the imprecise term of “realistic” (and related terms like “reality-based”). In the sport research domain, representative tasks allow the performer to search the environment for reliable information, integrate this information with existing knowledge, and complete an appropriate action (Broadbent et al., 2015). The representativeness of a given task consists of two critical components: functionality of the task and action fidelity (Broadbent et al., 2015).
2015; Pinder, Davids, Renshaw, & Araújo, 2011a). The former refers to whether the constraints a performer is exposed to and must act upon in the task are the same as in the performance environment. The latter requires that the performer is allowed to complete a response that is the same as in the performance environment. Central to these ideas is the relationship between perceptual-cognitive and motor processes as well as emotional responses associated with the task (Broadbent et al., 2015; Headrick, Renshaw, Davids, Pinder, & Araújo, 2015; Pinder, Davids, Renshaw, & Araújo, 2011b). This distinction reflects the described distinction between physical, cognitive, and emotional demands on officers in use of force incidents, which are described in Section 2.3.2. It has to be clearly noted, that the three components cannot be always clearly separated with regards to their causes and effects. However, with regards to designing and refining testing and learning environments a distinction of these components is useful.

Use of force environmental constraints that the performer must act upon (functionality) can be categorized in: (a) physical, (b) perceptual-cognitive, and (c) affective components. The physical design refers to components that mainly influence the intensity of attacks and attacker behaviour, which the defender has to cope with (functionality). This is connected to the intensity of executed motor skills of the defender (action fidelity). Perceptual-cognitive components impact the difficulty of decisions, which skill to perform and how to perform it (functionality). Therefore, such constraints mainly put load on the perception, decision-making, and problem solving abilities of the performer (action fidelity). Finally, affective components influence the emotional state, under which the defender has to perform (functionality). This allows the performer to experience the emotions associated with the task and how this impacted their thoughts and actions. Performers are able to learn (learning environment) or test (testing environment) their coping skills with these emotional demands (action fidelity). The matrix in table 1 shows aspects of functionality and action fidelity, related to the physical, perceptual-cognitive and affective design components.
Even though the functionality of the task is related to the action fidelity of the performer, it can be worth disconnecting them for learning and safety reasons. For example, in order to allow the performer to learn how to recognise cues that reveal an immediate attack (i.e. pre-attack indicators), the attacker may be allowed to attack very fast with a low level of contact (functionality – physical design). At the same time, the defender may be allowed to defend very fast with no level of contact (action fidelity – physical design). While high levels in every category cannot be achieved simultaneously without compromising health and safety issues (Wollert et al., 2011), the matrix allows to adjust single categories for optimal training effects and thus enables trainers to precisely design representative learning and testing environments.

Table 2: Functionality and Action Fidelity in Self-Defence Situations in Police Use of Force Simulations

<table>
<thead>
<tr>
<th>Functionality</th>
<th>action fidelity</th>
</tr>
</thead>
<tbody>
<tr>
<td>physical</td>
<td>physical</td>
</tr>
<tr>
<td>• speed / level of force of the attack (Staller &amp; Bertram, 2015)</td>
<td>• speed of the defence (Staller &amp; Bertram, 2015)</td>
</tr>
<tr>
<td>• spatial structure of the attack (Staller &amp; Bertram, 2015)</td>
<td>• spatial structure of the defence (Staller &amp; Bertram, 2015)</td>
</tr>
<tr>
<td>• contact-level of the attack (Staller, 2014)</td>
<td>• contact-level of the defence (R. Pfeiffer, 2014)</td>
</tr>
<tr>
<td>perceptual-cognitive</td>
<td>perceptual-cognitive</td>
</tr>
<tr>
<td>• valid cues (Staller &amp; Abraham, 2016)</td>
<td>• information processing (Staller &amp; Zaiser, 2015b)</td>
</tr>
<tr>
<td>• surprises (Jensen &amp; Wrisberg, 2014)</td>
<td>• problem-solving (Staller &amp; Zaiser, 2015b)</td>
</tr>
<tr>
<td>affective</td>
<td>affective</td>
</tr>
<tr>
<td>• anxiety / pressure (Nieuwenhuys, Caljouw, Leijser, Schmeits, &amp; Oudejans, 2009; Renden et al., 2014)</td>
<td>• pain-avoidance (Nieuwenhuys et al., 2011)</td>
</tr>
<tr>
<td>• emotion-laden (Headrick et al., 2015)</td>
<td></td>
</tr>
</tbody>
</table>

2.7.4 Health and Safety in Testing and Learning Environments

The designer of the learning or testing environment has to ensure the safety of participants as well as safety of training partners or role players. Since performance mistakes are going to happen, the instructor has to make sure that mistakes do not occur or, if they occur, that they have no serious consequences (e.g. injuries, death). This can be achieved by: (a) a reduction
of intensity, (b) a reduction of task complexity, and/or (c) environmental changes. Changes in intensity refer to measures that focus on making self-defence and combat techniques less dangerous in testing or training settings. Possible options include the reduction of permissible contact (as defender or as attacker), the exclusion of target areas or the reduction in speed and applied force. The reduction of task complexity aims at lowering the load of perceptual-cognitive processes of the performer. By reducing surprises, ambiguity, and available options, the probability of mistakes in the decision-making component in self-defence performance decreases, leaving the performer more attentional resources for the associated motor processes. Finally, environmental changes refer to measures by the task designer, which reduce the risk of injury by altering the physical structure of the training or testing environment. This can be achieved, for example, by using different forms of safety gear, using replica weapons that are less dangerous than the original weapons, modifying the training area by providing mats, and/or removing sharp or dangerous devices.

Since the design of any activity in police use of force training has to take into account the individual (Staller & Zaiser, 2015a; 2015b), the described safety options have to be tailored to the participant. For example, a role player attacks a participant with gloves and reduced force in his punches (environmental change; reduction in intensity), whereas a more skilled participant is attacked with full force and lighter gloves (lesser level of environmental change; no reduction in intensity). Because of the different skill level of the defenders, the risk of mistakes stays the same. The more skilled the instructor, the better will be his estimation about the probability of mistakes and injuries.

2.7.5 The Trade-Off Model of Police Use of Force Simulation Design

The analysis of representativeness and health/safety in the context of designing police use of force simulations leads to the conclusion that these two concepts are competitive in nature. For example, the use of the weapon systems that are employed in real world incidents cannot be equally used in
simulations in the training context due to health and safety issues. On the one hand simulations should be designed to accelerate expertise through active learner engagement, repetitive practice, the ability of vary difficulty and complexity levels, as well as performance measurement and intra-experience feedback (Issenberg, McGaghie, Petrusa, Gordon, & Scalese, 2005; Okuda et al., 2009; Staller & Abraham, 2016), on the other hand the safety of participants has to be ensured by omitting features that bear the risk of injuring participants (K. R. Murray, 2004; Wollert et al., 2011). The more health and safety features are implemented in a certain learning or testing environment, the more the level of overall representativeness will decline and vice versa.

The Trade-Off Model of Police Use of Force Simulation Design (see Figure 16) illustrates this relationship between representativeness and health and safety together with the skill level of the participants and conveys its implications for the design of effective self-defence learning and testing environments.

The different components of representativeness and the different components of health and safety in use of force learning and testing environments enable the designer to make informed and precise decisions about the “trade-off” between the two competing concepts. Since a 100% level of overall representativeness cannot be achieved (this would be the criterion environment, in which it is ethically impermissible to perform), the instructor may design a task, in which a higher level of representativeness can be achieved in one component, while representativeness would be reduced in another component, in order to ensure health and safety of the participants. For example, if the attacker attacks with a real knife, which reflects a high level of representativeness regarding the affective constraints under which the individual performs, the designer may consider reducing speed in the task, which reduces the intensity of the attack, in order to ensure health and safety.
2.7.6 Weapon Systems in Use of Force Training Settings

In police use of force learning and testing environments a variety of weapons systems are used in order to account for the described trade-off between representativeness of the task and health and safety in the design of simulations. Figure 17 provides an overview of common taxonomies and weapon systems that are used in use of force training. The current categorization is based on: (a) the projectile or emission out of the barrel from the weapon system, and (b) the target of that emission. Regarding the emission out of the barrel, weapons can be differentiated between four categories: First, weapons can emit potentially lethal projectiles like CA, potentially lethal training ammunition or unhardened structure target ammunition (K. R. Murray, 2004). Second, NLT weapons use projectiles that
exit the barrel with a reduced velocity that is great enough to enable shots over a certain distance but is small enough to not penetrate the skin of users. Common NLT systems include Simunition FX weapons, Airsoft, and Paintball. The third category consists of laser-based systems like the SAAB Small Arms Transmitter, Lasertag systems, Stress Vest, and the Beamhit Interactive Dry Fire System. These systems involve a sender unit (weapon) and a receiver unit that is mounted on the potential target (e.g. vest, head gear). Since no projectiles are emitted, there is no experience of pain, if a participant is hit by a laser-based weapon system (although the Stress Vest can be used with electric shock to induce pain). Finally, there are inert weapons that are non-functional by design, meaning they are not able to emit any projectiles during the training session. Examples include ASP RedGuns, rubber guns, but also real weapons that are temporarily rendered non-functional (e.g. by blocking the barrel, unloading the weapon).

The target, that is the potential receiver of the projectile or emission of the weapon, can either be non-human or human. Non-human targets can be stationary like traditional paper targets on shooting ranges or non-stationary with the ability to move, making target practice more difficult. Human targets in training settings are fellow trainees or trainers that can behave in a scripted or non-scripted manner. Scripted target behaviour is given, when the role player acts in a predefined way in order to generate situations, that the trainee has to solve. Non-scripted behaviour refers to actions that are geared towards winning against the training partner in the boundaries of pre-defined rules. Paintball and Lasertag matches are prototypical activities involving non-scripted behaviours of participants.

The two described broad categories of weapon systems (emission/projectile and target of emission/projectile) regularly define the boundaries under which the weapon system can be used with regards to representativeness and participants’ health and safety. For example, it is ethically impermissible to use live-fire against a fellow training partner that acts in a non-scripted way (Staller, 2015; Staller & Zaiser, 2016). On the other side the use of inert weapons in scenarios that involve the discharge of a firearm against the
police officer may not elicit the same psychophysiological effects as the use of NLT weapons in this context (Nieuwenhuys & Oudejans, 2011a; 2011b).
Figure 17: Weapon systems and taxonomies in police use of force training
2.7.7 Psychophysiological Effects of Weapon Systems in the Learning and Testing Environment

Researchers in the domain of police use of force training have started to investigate if the use of different weapon systems and training activities are representative of the real world situation (Andersen et al., 2016; Getty, 2014; Kratzig, 2013; Kratzig & Hudy, 2011; Lewinski, 2006; Nieuwenhuys et al., 2009; 2011; 2012; R. Pfeiffer, 2014; Renden et al., 2014; Staller & Abraham, 2016). For example, with regards to possible training activities, Andersen (2016) found that ‘highly realistic scenario based training’ was significantly correlated to the stress responses of active duty police officers as measured by maximum HR and the difference between resting HR and maximum HR. In another study (Lewinski, 2006) measured HR during a demanding scenario training culminating in a hostage situation. Even though he claimed that “realistic scenarios do produce extreme stress arousal that is at least in the range of what a real-life situation would provoke” (p. 3), data of real world incidents are still missing, as far as I am aware of. Both studies emphasize the use of scenario training as a training activity, since physiological arousal, operationalized through HR, seems to be similar to active duty operations. Yet, both studies fail to describe what ‘realistic’ in that context means and what features in terms of action fidelity and functionality of the task add to the representativeness of the training design.

Other studies have demonstrated the benefits of more distinct features of environmental design for learning and testing (Nieuwenhuys et al., 2012; Renden et al., 2014). For example, the use of affective learning design (e.g. training the threat of pain) adds to the representativeness of a task, by incorporating key aspects that are present in the criterion environment. The possibility of experiencing pain leads to emotional load resulting in behavioural changes (Nieuwenhuys et al., 2012; Renden et al., 2014). This in turn makes the task more representative, due to the increase in task functionality.

Regarding the use of different training systems, the use of laser-based weapon systems have been subject to recent investigations (Getty, 2014;
Kratzig, 2013). In the context of Navy handgun training, Getty (2014) compared the use of laser-based simulators to dry fire practices. Results showed, that the use of laser-based training weapons improved performance significantly compared to dry fire practice as indicated by qualification scores. Equally, Kratzig (2013) investigated potential benefits of laser-based simulation training compared to live-fire training, aiming at answering the question, “if simulation training could replace live-fire training” (p. 5). Over 24-weeks the experimental group worked stationary with dry-fire laser-based handguns (with a built-in recoil mechanism) on a digital 25 m shooting range, whereas the control group received identical training with live-fire weapons on a traditional shooting range. The experimental group only shot live rounds during three evaluations sessions. Retention tests one year after final qualification exam, results indicated that the experimental group increased their results on qualification scores, whereas performance of the control group decreased. A second experimental group also was introduced to live-fire shooting one training session before the first evaluation session. Results showed, that performance did not differ on the benchmark test, compared to the experimental group that used live-fire for the first time on that benchmark test. These results indicate that: (a) police officers can acquire handgun skills in the absence of live-fire, (b) the skills are transferable to a real world setting, and (c) that handgun skills learned through the use of laser-based simulation weapons are better maintained than the same skills learned through live-fire practice.

However, besides the striking benefits of laser-based simulation training, it has to be acknowledged, that both studies (Getty, 2014; Kratzig, 2013) only focused on stationary practice that is typically involved with marksmanship training (Behneman et al., 2012; Carlton, Orr, Stierli, & Carbone, 2013; Charles & Copay, 2001; Copay & Charles, 2001; Dulin, 1999; Swain, Ringleb, Naik, & Butowicz, 2011; Tharion, Shukitt-Hale, & Lieberman, 2003; C. R. White, Carson, & Wilbourn, 1991). In contrast, real world incidents are typically highly complex, dynamic situations, that must be dealt with under time-constraints and under the risk of injury or death (Alpert & Dunham, 2004; Andrew et al., 2013; Artwohl & Christensen, 1997; Ayoob, 1984; Blum,
Therefore, in order to account for representative task design in the learning or testing environment, key features have to be implemented with regards to functionality and action fidelity. The experience of the psychophysiological state under which the police officer has to operate is crucial if the learner has to develop skills for these environmental characteristics. Yet, if the use of different weapon systems alters the state under which the police officer operates, the representativeness of task would decrease making testing results less valid compared to a more representative environment. No studies exist so far investigating the effects of the use of training weapons and NLTA compared to live-fire weapons and CA on the psychophysiological state of police officers. This issue will be addressed in Chapter 5.

2.8 Summary of Chapter 2

Human defensive behaviour is associated with a range of psychophysiological changes that can be considered normal in the light of the prevalent constraints in situations of threat. Chapter 2 provided a review of the literature and theories associated with human defensive behaviour in the context of policing. As such the attention of the reader was drawn to human defensive behaviour with its complexity and its challenges. Furthermore, background about the police use of force was given. Training and explaining models were reviewed and the use of force framework in Germany – the country were the experiments took place – explained. Following this section, the literature about stress in the context of policing was reviewed with the aim to provide the reader with information about how acute situational demands can disrupt performance. Key aspects of decision making with an emphasis on NDM was explained afterwards. Since cognitive performance relies heavily on WM, this topic presented the theme of subsection five. Furthermore, the concept of self-control was reviewed with an emphasis on ego depletion and aggression, which is subject to further
investigations later on. Finally, the reader was provided with general information about the use and design of representative learning and testing environments, in order to facilitate understanding of the employed design of the majority of the studies in this thesis. The next chapter provides a general overview of these methods.
Chapter 3: General Methods
3 General Methods

3.1 Methods Employed in the Thesis

The main body of the current work consisted of several studies that were conducted at the Hessian Police Academy in Wiesbaden and Mühlheim/Germany (Chapter 4, 5, 6 and 7) with German police officers from different units. Additionally, data was collected in martial arts training gyms in Frankfurt and Niedernhausen (Chapter 5) and online (Chapter 6). In order to maintain external validity, the samples of the current studies consisted of professionals working in the domain under investigation making the findings applicable in the context of policing (Sue, 1999).

The current thesis employed a positivist/scientific approach meeting the requirements of quantitative research (Robson, 2011; Yardely & Bishop, 2008): (a) a clear testable hypothesis, which were derived from the relevant literature (e.g. working memory, threat attentional bias, risk taking, ego depletion), (b) strong internal validity (e.g. standardised and replicable) and (c) the use of appropriate statistical analysis (e.g. assumptions of data).

A general overview of the methods employed will be provided in this chapter and more precise details will be provided in the study chapters (Chapter 4, 5, 6 and 7). The ‘field setting’ allowed the data collection to remain in context, maintaining the ecological validity of the research. Principles from the NDM research domain, like testing in a fast paced dynamic environment, have been employed in the current project. Furthermore, standardised physiological measures that have been used successfully in previous police use of force research (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b; Lewinski, 2008; Nieuwenhuys & Oudejans, 2011b; Oudejans, 2008; Oudejans & Pijpers, 2009) have been built upon.

Standardised measures of cognition are rarely employed in a field setting (Hutchings, 2005). Recent research managed to successfully examine WM function during simulated armed confrontations (Condon, 2015; A. P. J. Roberts, 2012). The current thesis builds upon this work, using this reality based approach. The use of representative training scenarios maintains the
theoretical integrity of NDM in this time sensitive, highly dynamic environment, which stimulates the psychophysiological responses that are expected in real world encounters (Lewinski, 2006).

3.2 Difficulties Conducting Research within the Police Environment

Data collection within the police environment is a difficult endeavour due to the: (a) institutional, (b) organisational, and (c) procedural issues that the researcher has to deal with.

Institutionally, there is the problem that the law enforcement community historically does not believe that individuals outside the policing profession would assist them in performing their duties (Terrill et al., 2003). The lack of trust in academic endeavours may stem from the little or no useful return from the data collected compared to the resources (e.g. time, manpower) spent to support the research. Because of this, many agencies, commanders and trainers were reluctant to participate in the current project. Several German police agencies rejected research requests, except for the Hessian State Police, where I have been professionally affiliated. I was granted full access and able to work with the constraints outlined in the next two paragraphs.

Organisational issues involve: (a) problems regarding the restricted and limited access to training locations, (b) operational demands of participants, (c) the physical safety protocols, (d) sensitive data protection issues, and (e) the different levels of necessary authorisation in order to proceed with the experiments. Since I am a police officer myself, problems regarding the access to the training sites existed for my supervisor and colleagues that helped during data collection. These issues could be resolved by the professional relationship that has been built with the trainers of the Hessian Police Academy during the last decade. Due to the high demand for the training facilities in the State of Hesse, the number of testing dates had to be arranged before the series of studies started. In combination with the operational demands of participating officers, participant numbers varied
between the studies and could not be controlled. For example, in the CA vs. NLTA study (Chapter 4), several police officers were called to an operation before the data collection was finished. Therefore, the number of police officers tested was determined by the access to the training facilities and the operational demands of the participants and not by the more traditional effect-size based calculation of participant numbers.

The presence of “non-policing personnel” posed a particular problem around physical safety during testing, which were addressed by additional personnel that took care of the helpers and provided them with the equipment needed for their safety. The sensitivity of data, especially when operators of special units took part in the experiments, also proved to be difficult. The need for anonymity was addressed by assigning numbers to the participants and only referring to them by that number. Thus, the non-policing personnel were not able to identify the participating officers. The data itself was anonymised and stored on encrypted hard disks. After each experiment (after the data were analysed and interpreted) a presentation about the results and its implications for policing practice was given to the senior management at the Hessian Police Academy in order obtain clearance for the studies to continue.

Procedural issues arise when studies involve the use of live-fire weapons, NLTA, and/or scenarios with physical contact between the actor and the police officer. This is due to the fact that the firearms training facilities at the Hessian State Police have strict SOPs as soon as training involves the emission of projectiles (CA or NLTA). The same applies to the scenario training with physical contact. In both cases SOPs aim to maintain the health and safety of participating officers and trainers during training sessions. The SOPs include different levels of personal protective equipment (PPE) and safety gear depending on the kind of training or scenarios, respectively. Also significant safety protocols have to be met in order to have researchers present during testing.

The SOPs were responsible for two major issues concerning data collection in the field setting: Firstly, the use of PPE and safety gear in combination with
heavy force-on-force contact between participants and actors caused malfunctions in the recording of HR data via a HR belt during the scenarios. The initial use of HR belts in combination with wrist watches that stored the data for subsequent analysis lead to major data losses due to the lack of continuous data transmission from belt to watch while participating in a highly dynamic scenario. This issue was partly solved by the use of memory belts, which record the data directly on the belt. Data loss only occurred when the force-on-force contact was so heavy that the HR belt got out of place. Secondly, after highly dynamic scenarios or shooting exercises, SOPs demand that a safe environment is provided (e.g. weapons unloaded). This led to delays in the collection of data subsequent to the completion of a scenario as weapons were unloaded and so on. This is important to note, since research indicates that cognitive and physiological adaptations resulting from increased demand may start to reduce once the stimuli causing the demand has been removed (Lewinski, 2008). This issue was addressed by collecting the data directly in the shooting range or the training facility where the scenario took place. The drawback was that additional safety equipment had to be worn by the researchers, and the next participant could not be tested until the data collection (e.g. cognitive tests) had finished for their predecessor. This in turn, reduced the number of participants that were able to participate on one day, which posed a dilemma since the number of testing days have been pre-agreed with the Hessian Police Academy before the studies started.

3.3 Active Duty Equipment and Training Tools

Police Officers in the state of Hesse have variety of use of force tools at their disposal. In order to better understand what a Hessian police officer is required to carry, a brief outline of the use of force tools is given here. Furthermore, the training tools that allow for representative training design are outlined here.
3.3.1  Active Duty Equipment

3.3.1.1  Active Duty Belt

Active duty Hessian police officers are issued a duty belt with use of force tools and equipment such as flashlights and disposable gloves. Additionally, officers can add items to their duty belt (e.g. second flashlight) as long as they are in line with the guidelines issued by the state of Hesse. The standard issue of the active duty belt may vary from unit to unit and from officer to officer, according to their assignments and personal feeling of comfort. For example, both aspects can influence the decision to choose from different firearm holsters available (e.g. low vs. high carry).

A patrol officer’s duty belt typically holds a service weapon with an additional magazine (see 3.3.1.2), a telescopic baton, Oleoresin Capsicum (OC) spray, and handcuffs.

3.3.1.2  Service Issued Firearms

In accordance with the underlying principle of federalism, the 16 states of Germany are responsible for managing their police services. The two federal agencies (Bundeskriminalamt and Bundespolizei) receive its standards of care and regulations from the federal government.

Hessian police officers are issued a 9mm handgun. The model depends on the unit that the officer is assigned to. While patrol officers carry a Heckler & Koch P30, special units and personal protection officers are issued a Glock 21. Other models are also issued, depending on the assignment and the unit (e.g. undercover officers).

While on duty the issued handgun has to be loaded and carried in the appropriate holster. The spare magazine has to be carried with the officer. Personal modifications of service weapons in any manner are forbidden.

Live fire training and requalification are typically done with the officers’ issued weapon. Scenario training is done with replicas, which are identical to the officers’ duty weapon. Two types of replicas exist, which are used for different purposes. The replicas are coloured blue when they are specially...
prepared to fit coloured soap cartridges (i.e. Simunition FX NLTA). The red replicas are not able to emit any projectiles, yet mimicking exactly the functionality of the service issued firearm.

3.3.1.3 Body Armour

Hessian police officers are issued body armour that they are obliged to wear during deployments. Research indicates that the use of body armour affects cognitive functioning in healthy students (A. P. J. Roberts & Cole, 2013). It is not clear, if these effects are consistent with police officers, since they are used to wearing body armour on a daily basis. It is possible that officers would react differently because of physical conditioning (R. Norris, Carroll, & Cochrane, 1990). Yet, to account for any possible effects, participants had to wear their body armour in all scenarios in the current research project.

3.4 Training Equipment

The goal of training sessions in the police use of force context can be described as the learning of skills that transfer from the training environment (e.g. Police Academy) to the criterion environment (e.g. real world police operations or incidents). This transfer of learning is fostered by representative learning design (Broadbent et al., 2015; Pinder, Davids, Renshaw, & Araújo, 2011b). In the unique context of police use of force training, the major demand in any training or testing design is the trade-off between health and safety of participants (including role players) and the representativeness of the given task. The use of appropriate training equipment is a crucial factor in meeting this demand. Therefore, the training equipment used in the studies are briefly outlined here.

3.4.1 Non-Lethal Training Weapons

NLT weapons are designed for training that incorporates live target engagement (K. R. Murray, 2004). It is important to differentiate between live target engagement and scenario training, since not all scenario training requires the use of NLT equipment and weapons.
NLT weapons in the Hessian Police are identical to the live fire model, except that they are prepared to fit coloured soap cartridges (Simunition, FX NLTA). The use of this NLTA enables a high degree of representative learning and testing design, if the task or the scenario is geared towards the precise use of firearms (where hitting or missing a live target is essential in the assessment). For the use of NLT weapons additional PPE is required according to the standards of the Hessian Police. These include eye protection and additional safety equipment, depending on the physical proximity between officer and role player. This may include helmets, vests, shin and elbow guards, gloves or a full protection suit, which depends on the intended contact level. In the state of Hesse the leading instructor of the training is responsible for the decision about which equipment is mandatory for the training. There are high safety standards in place for the use of NLT weapons at the Hessian Police Academy. The SOPs for conducting training with NLT weapons has to be followed strictly. NLT weapons at the Hessian Police Academy are coloured blue.

Due to several accidents with firearms using CA on the shooting range, there is an ongoing discussion around whether NLT weapons provide a safe alternative in the learning process of basic weapons skills at the shooting range. Recent studies suggest, that CA is not needed in the learning process (Getty, 2014; Kratzig, 2013; Kratzig & Hudy, 2011). Chapter 4 aims at answering the question if the use of CA differs from the use of NLTA with regards to the psychophysiological demand put on the police officer during a training exercise.

3.4.2 Inert Weapons

Inert weapons are non-functional by design and cannot emit any projectiles (K. R. Murray, 2004). They are useful training devices, when the focus is on skill development that does not need the firing of any projectiles at a role player. Such training includes weapon retention training, choreographing team movement or scenario training where shots are expected to be fired from very close range.
The use of inert weapons requires a lower level of safety standards than NLT weapons. The lead instructor of the training has to take care that impact blows with the weapon is appropriately accounted for by the use of proper safety equipment (e.g. helmets, vests, shin and elbow guards, gloves).

Inert weapons at the Hessian Police Academy are red and are manufactured by the same company (Heckler & Koch, Glock) as the live weapons. The inert weapons used at the Hessian Police Academy allow for weapon manipulation (including changing the magazine), except for loading it with any projectiles.

3.4.3 Personal Protective Equipment in Police Use of Force Training

The PPE used in police use of force training depends on the role of the individual (learner or role player) and the scenario design. As soon as the scenario provides the possibility of physical contact, the role player has to wear a full protection suit according to the guidelines of the Hessian Police. The rationale for this is that the learning police officer will probably make mistakes and/or use force where it was not intended. The Hessian Police Academy uses the RedMan XP Instructor Suit (RedMan, RedMan Training Gear) for force-on-force training with inert weapons. As soon as NLT weapons are involved the RedMan WDS Instructor Suit (RedMan, RedMan Training Gear) is used, which allows for receiving full impact blows and being shot at with NLTA. The learning police officers are equipped with PPE according to the instructions of the lead instructor. This may include shin and elbow guards, groin protection, gloves, body armour and helmets. If the training involves NLT weapons, eye protection is mandatory.

In all studies the role player wore a RedMan Instructor Suit (XP suit in Chapter 7; WDS suit in Chapter 4 and 5). Participating police officers were issued eye protection in Chapter 4 and elbow and shin guards and a helmet in Chapter 5. Officers routinely wore their body armour. The decision to use tactical gloves was up to the participant (like on a routine deployment).
3.5 Ethical Considerations

Prior to any data collection and in accordance with the University of Liverpool’s ethical procedure guidelines, ethical approval was attained (see Appendix 1).

This included a full risk assessment of the planned research in the firearms and reality based scenario setting. Potential risks to researchers, role players and police officers were considered. All reasonable steps were taken to ensure health and safety of the involved personal, such as exactly following the SOPs and wearing appropriate PPE (e.g. high visibility vests, ear protection, eye protection).

Approval for attending training days and conducting research within the Hessian State Police was given by the head of the Hessian Police Academy. Access was granted based on the agreement of certain criteria. Emphasis was put on the sensitivity of the field of police use of force and the importance of police officer anonymity. Since everything discussed in the current thesis is available in the public domain, the anonymity of participating officers was paramount. In order to make the identification of participants impossible the demographic information collected was kept to a minimum. Sometimes this means, that even requesting minimal information, such as years of experience, was difficult, since identification of participants would be possible in small units. Collected data was stored on encrypted devices and identifying information was redacted from printed material. Only the supervisors and the author have access to the collected data.

Participants were provided with a detailed information and consent form prior to testing (see Appendix). In the “threat attentional bias study” (Chapter 6), which was conducted partly online, an information page occurred and consent has to be provided by the participant before data collection started via the software used (see Appendix). Because testing was conducted in conjunction with regular training sessions, participants were informed that their performance during testing; (a) was not indicative of their skills as a police officer, (b) would have no bearing on the outcome of their organizational training, and (c) would not been reported to any superiors or
colleagues. Participants were informed that they could withdraw their agreed participation from an experiment at any time.

### 3.6 Participants

Four quantitative studies with several sub-studies were completed with a total of 398 participations (see Table 3) of sworn police officers, martial artists, and civilians with no background in either simulated or real confrontations. All participants gave their informed consent prior to participating in the study. No participant was allowed to participate in more than one sub-study of one study. Fifty participants of Study 3 (Chapter 6) were based in the UK. Except for these, participants were German citizens living in Germany.

<table>
<thead>
<tr>
<th>Study</th>
<th>Chapter</th>
<th>Title</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>CA vs. NLTA</td>
<td>42</td>
</tr>
<tr>
<td>1.1</td>
<td></td>
<td>Phonemic Fluency</td>
<td>13</td>
</tr>
<tr>
<td>1.2</td>
<td></td>
<td>Digit Span Task</td>
<td>14</td>
</tr>
<tr>
<td>1.3</td>
<td></td>
<td>Corsi Block-Tapping Task</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Scenario vs Exercise</td>
<td>73</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>Threat Attentional Bias</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German Sample</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UK Sample</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>Ego Depletion</td>
<td>81</td>
</tr>
<tr>
<td>4.1</td>
<td></td>
<td>Video Task</td>
<td>44</td>
</tr>
<tr>
<td>4.2</td>
<td></td>
<td>Cold Pressor Task</td>
<td>37</td>
</tr>
</tbody>
</table>

### 3.7 General Design

All studies employed quantitative measures in between and within-subjects designs (see Table 4). The DV included various cognitive measures (Chapter 4, 5, 6 and 7), physiological measures (Chapter 4, 5 and 7) and quantitative data from the scenario performance itself (Chapter 7). The designs of the specific studies are outlined in more detail in the methods section of each chapter.
Table 4: Design, IVs and DVs of Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>IV</th>
<th>DV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Within</td>
<td>Time</td>
<td>Heart rate, heart rate variability</td>
</tr>
<tr>
<td>2</td>
<td>Between / within</td>
<td>Group / time</td>
<td>Heart rate, Blood pressure, temperature,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>phonemic fluency scores</td>
</tr>
<tr>
<td>3</td>
<td>Between</td>
<td>Group</td>
<td>Measures of different psychological tests</td>
</tr>
<tr>
<td>4</td>
<td>Between / within</td>
<td>Group / time</td>
<td>Heart rate variability, aggression time,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>measures of different psychological tests</td>
</tr>
</tbody>
</table>

3.8 Quantitative Data Collection

3.8.1 Scenario Training

In Hesse each police officer is obliged to participate in a scenario-based exercise at least once per year. The design of the training session including the practiced scenarios is the responsibility of the lead instructor of each session. Since training time is rare, the specific design of the scenarios had to be negotiated with the lead instructor. Before every study an agreement about the scenario was reached based on the different goals: on the one hand, letting the participants practice the specific goals the Hessian Police wanted to train on, and on the other hand a scenario design that allowed for optimal data collection based on the specific hypotheses.

In Studies 1 and 2 the scenario-based exercises were not subject to any measurements. By not focusing on the results of the scenario itself, trust was built up between the researchers and the Hessian Police Academy. This afforded the opportunity to run a scenario-based exercise where the results were based on the participants’ performance (Chapter 7). In Studies 4.1 and 4.2, the participants were provoked by the role player until they responded. The time from the start of the scenario until the police officer used offensive physical aggression was measured. A detailed description of the scenarios and exercises that had to be solved are given in the specific methods section of each chapter.

3.8.2 Psychological Tests Employed in the Thesis

The psychological tests employed in the thesis are displayed in Table 5.
### Table 5: Psychological tests employed in the thesis

<table>
<thead>
<tr>
<th>Test</th>
<th>Study</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonemic Fluency Task</td>
<td>1.1, 2, 4</td>
<td>4, 5, 7</td>
</tr>
<tr>
<td>Digit Span Task</td>
<td>1.2</td>
<td>4</td>
</tr>
<tr>
<td>Corsi Block-Tapping Task</td>
<td>1.3</td>
<td>4</td>
</tr>
<tr>
<td>Dot Probe Task</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Emotional Stroop Task</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Ballon Analogue Risk Task (BART)</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>State-Trait Anxiety Inventory (STAI)</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Brief Mood Introspection Scale (BMIS)</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>

The tests employed in the current thesis shall be reviewed in detail within the chapter that they are used, except for the Phonemic Fluency Task, which is used in Chapter 3, 4 and 6. The latter is reviewed in the following section.

#### 3.8.2.1 Phonemic Fluency Task

Tests of verbal fluency are commonly used for the evaluation of executive cognitive functioning by asking individuals to generate as many words as possible according to a specific rule (phonemic or semantic criteria) in a limited time frame (Benton, 1968; Newcombe, 1969). There are three distinct processes associated with the completion of the Verbal Fluency Task: First, when the task instructions are received, participants have to retrieve information from long-term memory stores relevant to the task (updating). Second, for the random generation of words beginning with a certain letter or within a category participants are likely to use both phonemic (e.g. similar sounding and rhyming words) and semantic (e.g. similar looking items) strategies. Once a particular cluster has been exhausted participants have to quickly generate another set (switching). Third, participants are instructed not to repeat the same word twice or repeat the same word altered where possible (inhibition). The three described processes (updating, switching, inhibition) are considered to be primary executive cognitive functions concerned with higher order control (Miyake et al., 2000; Miyake & Shah, 1999; Troyer, Moscovitch, & Winocur, 1997). The Verbal Fluency Task is a relatively general measure of executive function since it requires the
undertaking of three executive processes. Like other tests, measuring executive functioning, the verbal fluency task requires non-executive cognitive functions for completion. Compared to other tests, non-executive processing function can be separately scored within the Verbal Fluency Task (Troyer, 2000; Troyer et al., 1997; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998b).

Retrieving a cluster of words that are cued by similarity is a relatively automatic process. The base level neural activity will be higher for items similar in nature to those immediately recalled (J. R. Anderson, 1993; Lovett, Reder, & Lebiere, 1999). This is proposed to facilitate the production of similar items, as the neural activation that is required for retrieval is reduced. This process is likely to involve sub-systems of WM like the phonological loop, for the storage and maintenance of items until they are recited. Hence, the required executive cognitive processes for this part are minimal. Various different strategies may be used for the verbal fluency task, demonstrating why it is considered such a general test of WM. For example, strategies for spontaneous retrieval may involve sub-systems, outside of those proposed in Baddeley’s WM model (e.g. the olfactory system). The average cluster size reflects these non-executive retrieval processes in total word production.

Verbal Fluency Tasks allow for a variety of different measures reflecting partially dissociable processes. As described, tests of verbal fluency measure the ability to search lexical stores, to retrieve information from the semantic memory, and to switch cognitive set (Kavé, Kigel, & Kochva, 2008; Troyer, 2000). Troyer et al. (1997) argued that usually the number of correct words generated is measured, but since fluency is a multifactorial task, this score alone does not cover all of the important aspects of the participant’s performance. Therefore, they introduced two additional components: (a) clustering (defined as the production of words within semantic or phonemic categories), and (b) switching (defined as the ability to shift efficiently to another category). According to the findings of Troyer et al. (1997) clustering and switching contribute equally to semantic fluency, but switching made a greater contribution to phonemic fluency. The two additional components allow the measurement of partially dissociable processes. On the one hand,
strategic search, response initiation, monitoring, shifting, and flexibility within fluency tasks can be measured through the number of switches (Troyer et al., 1997) and the number of clusters (Kavé et al., 2008). On the other hand, the size of the generated cluster reflects the reliance on lexical stores (Kavé, Heled, Vakil, & Agranov, 2011). Previous research suggests that switching in phonemic fluency is mediated by frontal lobe and executive cognitive functioning whereas clustering is associated with non-executive temporal functioning (Troyer, 2000; Troyer et al., 1997).

In the current thesis WM capacity was measured with the Phonemic Fluency Task of the “Regensburger Wortflüssigkeitstest” (RWT (Aschenbrenner, Tucher, & Lange, 2000), which is the German complement of the Controlled Order Word Association Test (Benton & Hamsher, 1989). The test was administered to the participants before and after the scenario or exercise. Participants were asked to provide as many German words as possible within 120 seconds beginning with the letters B and M (pre-test), or P and K (post-test). They were told not to use any proper nouns (such as a name of a person or a place), or variations of the same root words (like “Kaufräusch” and “Kauften”; Troyer, 2000; Troyer et al., 1997). According to the guidelines of Troyer (1997), repetitions and mistakes were included in the scoring of clustering and switching. A cluster was counted when two consecutive words shared any of the following phonemic characteristics: (i) words beginning with the same first two letters (Bus – Bude), (ii) words that rhyme (kiffen – kläffen), (iii) words differing only by a vowel regardless of the actual spelling (See – sah), or (iv) words with two or more spellings (Seen – sehn). Three scores were calculated for each protocol. The total number of words (TW) comprised of all words, excluding repetitions and errors. The score was summed across the two letters. Mean cluster size (MCS) was calculated as follows: The number of words in a cluster was counted from the second word. That is, a single word was coded as 0, two words were coded as 1, three words were coded as 2, and so forth. Errors and repetitions were included. MCS across the two letters was computed for the two phonemic trials (BM and PK). The number of switches (NS) was the last score, to be calculated: The total number of transitions between clusters, including single
words, was counted for each letter. Switches produced by the participants were summed across the two letters (BM or PK). Errors and repetitions were included.

The results of the three scores have to be interpreted in the light of the connection between them: clustering as indicated by MCS, and switching (as indicated by NS) determine together TW (Troyer et al., 1997; 1998b). On the one hand, clustering in this context relies on temporal-lobe functioning (Troyer et al., 1998b), as indicated by previous research that has implicated the role of temporal lobes in semantic memory (Hodges, Patterson, Oxbury, & Funnel, 1992; Pietrini et al., 1988; Warrington & Shallice, 1984). Therefore, MCS reflects the non-executive retrieval processes that facilitate the total word production (A. P. J. Roberts, 2012). On the other hand, switching relies on frontal-lobe functioning, as indicated by studies providing evidence that frontal lobes are involved in set shifting (A. M. Owen et al., 1991; Vilkki & Holst, 1994) and that switching itself is sensitive to manipulations of attention and to focal frontal lobe lesions (Troyer et al., 1997; 1998b; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998a). Therefore, NS reflects the executive cognitive processes that are involved in the generation of words in the Phonemic Fluency Task.

As there is a lack of consensus within the literature on the effect of different letters on Phonemic Fluency Task performance, the letters used for pre- and posttesting were the same allowing direct comparison. Previous research has shown (A. P. J. Roberts, 2012), that using the same letters for pre- and posttesting allows for the differentiaten between practice effects and effects due to the manipulation. An increase indicates a potential practice effect, whereas the relationship between executive (NS) and non-executive measures (MCS) indicates the effects of the manipulation.

3.8.2.2 Digit Span Task

The Digit Span Task requires the participant to repeat forwardly (digit span forward; DF) or backwardly (digit span backward, DB) a list of random number sequence that have been read aloud at the rate on one number per second. The total digit span (TDS) is the sum of DF and DB. The Digit Span
Tasks examine the functioning of the phonological loop of WM, since the task requires the active maintenance of verbal information (Baddeley, 1992). The underlying processes of the DF and DB differ slightly (Cornoldi, Orsini, Cianci, Giofrè, & Pezzuti, 2013; Giofrè, Stoppa, Ferioli, Pezzuti, & Cornoldi, 2016; C. R. Reynolds, 1997). Phonological loop processing representing short term auditory memory (Baddeley & Hitch, 1974) is more involved in DF, which may reflect efficiency of attention rather than central executive functioning (C. R. Reynolds, 1997; Rosenthal et al., 2006). In contrast, DB is suggested to be a more sensitive test of executive function, since it requires monitoring of the number sequence representations stored in WM and active manipulation of that information in order to be able to repeat that sequence in reverse (Hedden & Yoon, 2006; Rosenthal et al., 2006).

Judging from differing correlations with intelligence (Cornoldi et al., 2013) and neuroanatomical data that indicates that brain regions involved in DF and DB differ to some degree (Rossi et al., 2013), it seems that the two tasks measure slightly different components of WM. In line with this, DF is considered a measure of the phonological loop, whereas DB is thought to measure the central executive component of WM (Alloway, Gathercole, Willis, & Adams, 2004). Hence, DF represents a purer measure of active phonological maintenance and the attention processes recruited for rehearsal (Conklin, Curtis, Katsanis, & Iacono, 2000). Furthermore, the maintenance of manipulated serial order in the DB can involve a right hemispheric recording. This allows for visual representation to facilitate performance in DB due to the increased executive demand (Larrabee & Kane, 1986; Rudel & Denckla, 1974).

Test sheets and instructions of DF and DB were taken from the German Version of the Wechsler Adult Intelligence Scale (Petermann, 2012; Wechsler, 2008). During the tests in the current thesis a sequence of numbers were read aloud to the participants by the experimenter using a monotonous, one second paced tone for the sequence of numbers. For the DF participants were asked to repeat back the numbers in the same order. Participants were told, that the numbers would increase in size through the trials. The DB was administered in exactly the same manner, except
participants were instructed to repeat the numbers in reverse order. During both tasks the number of correctly recited items were recorded. No time limits were placed on these tasks. The TDS was calculated out of the sum from the DT and DB score of each participant.

3.8.2.3 *Corsi Block-Tapping Task*

The Corsi Block-Tapping Task (Corsi, 1972; Milner, 1971) is used to assess visuo-spatial WM function and is widely used in experimental research (Berch, Krikorian, & Huha, 1998; Bo, Jennett, & Seidler, 2011; Brunetti, Del Gatto, & Delogu, 2014; M. H. Fischer, 2001; Kessels, van Zandvoort, Postma, Kappelle, & de Haan, 2000; M. M. Smyth, Pearson, & Pendleton, 1988). The task requires participants to remember the order in which a set of blocks is tapped and repeat that order tapping the relevant blocks. The score that is obtained through the task, is the number of correctly remembered sequences and the length of the longest sequence that was remembered correctly.

The Corsi Block-Tapping Task requires the visuo-spatial sketchpad for completion to recall the blocks which are tapped as well as their spatial location (Corsi, 1972; Milner, 1971; 1982; Petrides & Milner, 1982; T. P. Ross, Hanouskova, Giarla, Calhoun, & Tucker, 2007). The test is primarily employed for the assessment of non-verbal WM functioning (Spinnler, Sala, Bandera, & Baddeley, 2007; Vilkki & Holst, 1989). The executive demand of the task is minimal, since no manipulation of the information is required. The material is only stored and maintained.

The Corsi Block-Tapping Task is commonly viewed as the visuo-spatial equivalent of the verbal Digit Span Task detailed above (Kessels, van den Berg, Ruis, & Brands, 2008). The data regarding performance between the Corsi Block-Tapping Task and Digit Span Task is quite mixed suggesting they may not be as comparable as first thought, particularly the direct comparison of raw scores (Wilde, Strauss, & Tulsky, 2004). Visual tasks, even those requiring simple storage, are generally considered to be more complex than their verbal counter parts due to the integration of numerous information types (e.g. colour, texture, and spatial information). This, along
with the different strategies individuals may employ, can make direct comparisons of the task difficult (Kessels et al., 2008; Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000).

In the current study an electronic version of the Corsi Block-Tapping Task was used. The software and scripts used to run the tests were provided by Inquisit ("Inquisit 4.0.5.0.," 2014). The tests were administered on two Apple MacBook Pros. Screens were calibrated prior to the test. Participants were presented nine squares at fixed, pseudorandom positions on the screen. The task required participants to remember the order in which the blocks flashed and repeat that order by clicking the relevant blocks with the mouse. As the trials progress the number of blocks in one sequence increases. The test was terminated if the participants failed to reproduce two sequences of equal length. Two different scores are computed for each participant: First, the block span equals the length of the longest sequence that was remembered correctly. Second, the total score is the product of the block span and the number correctly repeated sequences until the test was discontinued (Kessels et al., 2000). Since the total score accounts for the performance on both trials of equal length, it is more reliable than the block span score alone.

### 3.8.3 Measuring Physiological Arousal

Physiological arousal is primarily measured by changes in HR, heart rate variability (HRV), systolic and diastolic blood pressure, core temperature, breathing rate, pupil size, perspiration, muscle tension and blood or salivary cortisol levels (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b; Charmandari, Tsigos, & Chrousos, 2005; HeartMath, 1999; Honig & Lewinski, 2008; Sibley & Beilock, 2007; Strahler & Ziegert, 2015; Tsigos & Chrousos, 2002; Vrij et al., 1994; Vrijkotte, van Doornen, & de Geus, 2000).

In the context of examining physiological arousal in police officers, researchers have successfully used HR (J. Armstrong et al., 2014; Kayihan, Ersöz, Özkan, & Koz, 2013; Lewinski, 2008; Nieuwenhuys & Oudejans, 2010), HRV (Brisinda, Fioravanti, Sorbo, Venuti, & Fenici, 2015a; Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b; Strahler & Ziegert, 2015) temperature (Condon, 2015; A. P. J. Roberts, 2012; Vrij et al., 1994) and
cortisol levels (Akinola & Mendes, 2011; 2012; Strahler & Ziegert, 2015; Taverniers et al., 2010; 2011; Walvekar, Ambekar, & Devaranavadagi, 2015).

HR, HRV, temperature and blood pressure are the most commonly used measures of physiological arousal, since they provide the greatest cost/benefit balance between accurate readings, invasiveness and cost efficiency (Dishman et al., 2000; Hardy & Parfitt, 1991; Klaperski, Dawans, Heinrichs, & Fuchs, 2013; G. F. Lewis et al., 2015; Mullen & Hardy, 2000; Teisala et al., 2014).

Even though several researchers suggest that cortisol levels are the best indicators of physiological arousal (Dickerson & Kemeny, 2004; Taverniers et al., 2010; 2011), they have not been used in the current project, because of the following reasons: Firstly, cortisol measure requires the researcher to take blood or salivary swab samples from participants, which is difficult whilst operating in a naturalistic environment (fast paced, unhygienic, officers wearing PPE). Secondly, the taking of blood or salivary swabs requires extra training, which was not possible because of the remote location of this project. At the time of the project, no additional researchers were available who would have been able to conduct such measures. Thirdly, the collecting and analysing of cortisol measures are extremely expensive. Due to the part-time design of the project and the associated self-funding, this was not possible.

Even though the measure of cortisol levels would have provided a direct comparison between physiological arousal and any cognitive adaptation, a linear relationship can be inferred from a vast amount of research in this area (T. W. Buchanan, Tranel, & Adolphs, 2006; Mattarella-Micke, Mateo, Kozak, Foster, & Beilock, 2011; Mehler, Reimer, Coughlin, & Dusek, 2009; Oei, Everaerd, Elzinga, van Well, & Bermond, 2006; Porcelli et al., 2008; Reimer & Mehler, 2011; Schoofs, Pabst, Brand, & Wolf, 2013; Schoofs, Preuß, & Wolf, 2008; Schoofs, Wolf, & Smeets, 2009; Steptoe, Wardle, & Marmot, 2005). For the outlined reasons, HR, HRV, blood pressure and core temperature were used as measures of physiological arousal through the
course of this research. Table 6 shows how the physiological measures were employed in the different studies.

**Table 6: Physiological measures employed in the thesis**

<table>
<thead>
<tr>
<th>Test</th>
<th>Study</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRV</td>
<td>1, 4</td>
<td>4, 7</td>
</tr>
<tr>
<td>HR</td>
<td>1, 2, 4</td>
<td>4, 5, 7</td>
</tr>
<tr>
<td>Core temperature</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Blood pressure (systolic and diastolic)</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

### 3.8.3.1 Heart Rate and Heart Rate Variability

HR has been a quantitative measure of stress in policing contexts for more than 20 years (Siddle, 1995). Siddle (1995) as well as other authors (D. Grossman & Christensen, 2007) suggest that the optimal tactical HR “zone” should be between 115 and 145 beats per minute (bpm), since above that limit complex motor skills (i.e. the ability to track and shoot a moving target) begin to deteriorate. When HR raises above 175 bpm tactical efficiency may be lost, because of a progressive trend toward irrationality, fostered by perceptual and memory distortions (Klinger & Brunson, 2009), reduced functioning of the prefrontal cortex, and prevalence of amygdala activation under uncontrolled fight-or-flight response (Sharps, 2010). It has to be acknowledged that Siddles research influenced police use of force training in a positive way, since instructors started to figure out, that training under these constraints is necessary to prepare for dealing with the real-life operational demands (R. Fenici, Brisinda, & Sorbo, 2011; Olson, 1998). However, empirical evidence that the absolute value of HR achieved can be an index of individual stress and predictive of operational behaviour is questionable at the moment (R. Fenici et al., 2011). Other studies (R. Fenici, Ruggieri, Brisinda, & Fenici, 1999; Meyerhoff et al., 2004) found that the elevation of HR cannot be correlated with performance efficiency, due to the overlap of physical effort and psychological strain, which both contribute to HR increases (R. Fenici, Brisinda, & Sorbo, 2011). Consequently a potential
HR increase has to be interpreted in the context of the event dynamics (R. Fenici et al., 2011).

HRV analysis provides a more sophisticated method for monitoring the cardiac autonomic response to increased demand and has received increased attention in the scientific literature (Acharya, Joseph, Kannathal, Lim, & Suri, 2006; Elkins et al., 2009; Hansen, Johnsen, & Thayer, 2009; R. Lane et al., 2009; Lehrer et al., 2013; Lucini, Norbiato, Clerici, & Pagani, 2002; Luft, Takase, & Darby, 2009; Melillo, Bracale, & Pecchia, 2011; C. A. Morgan, Aikins, Steffian, Coric, & Southwick, 2007; Pico-Alfonso et al., 2007; Prinsloo et al., 2011; Sandercock & Brodie, 2006; Saus et al., 2006; Schubert et al., 2009; J. F. Thayer & Lane, 2009; J. F. Thayer, Hansen, Saus, & Johnsen, 2009; A. G. Thompson, Swain, Branch, Spina, & Grieco, 2015; Tsuji et al., 1996).

HRV is the direct quantitative result of regulatory mechanisms with the autonomic nervous system (ANS), reflecting internal changes from reacting to acute or chronic conditions and stimuli (Acharya et al., 2006; R. Lane et al., 2009; Luft et al., 2009; Sandercock & Brodie, 2006; Tsuji et al., 1996; A. D. Walker, Muth, Switzer, & Rosopa, 2013).

More than 45 years ago, research (Jose & Collison, 1970) has found that when both cardiac vagal (the primary parasympathetic nerve) and sympathetic inputs are blocked pharmacologically, the intrinsic HR, is higher than the normal resting HR. These findings support the idea that the heart is under tonic inhibitory control by parasympathetic influences. Hence, the resting cardiac autonomic balance favours energy conservation through the means of parasympathetic dominance over sympathetic influences. Additionally, the intervals between heart beats (interbeat interval, IBI) is characterized by beat-to-beat variability over a wide range, which also implicates vagal dominance since the sympathetic influence on the heart is too slow to produce beat to beat changes (J. F. Thayer, Ahs, Fredrikson, Sollers, & Wager, 2012).
ANS adaptation to the demand that occurs in police tactical operations is a complex interplay of cognitive, emotional, behavioural changes and autonomic reactions, which are difficult to quantify in a laboratory setting (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b). Previous research suggests, that absolute values of HR cannot be assumed as a univocal index to quantify individual demand, nor to predict tactical efficiency under pressure (D. Grossman & Christensen, 2007; Meyerhoff et al., 2004). The recently more applied approach of HRV analysis is based on the scientific evidence that the sympathovagal balance can be assessed quantitatively, and that the effects of different emotions can be separated with HRV analysis (Kreibig, Wilhelm, Roth, & Gross, 2007; McCraty, Atkinson, Tiller, Rein, & Watkins, 1995; Montano et al., 2009; Pagani & Lucini, 2008; Sloan et al., 1994). Yet, it is still difficult to apply a mechanistic interpretation of HRV measures, which is reflected by the long lasting debate about the concept of sympathovagal balance (Eckberg, 1997; Montano et al., 2009; J. F. Thayer et al., 2012).

In the context of police scenario training the recent work (Brisinda, Fioravanti, Sorbo, Venuti, & Fenici, 2015a; Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b) showed that combinations of different computed HRV parameters allow for a differentiation between rest and demand conditions and between physical and mental load.

Regarding calculation of all measures of HRV, the basic data is the sequence of time intervals between heart beats. This IBI time series is used to measure variability in the timing of the heart beat. The heart is dually innervated by the ANS such that relative increases in sympathetic activity are associated with HR increases and relative increases in parasympathetic activity are associated with HR decreases. Hence, relative sympathetic increases cause the IBI to become shorter and relative parasympathetic increases cause the IBI to become longer. The parasympathetic influences are pervasive over the frequency range of the HR power spectrum. The sympathetic influences ‘roll-off’ at about 0.15 Hz (Saul, 1990). Thus, high frequency HRV represents primarily parasympathetic influences, whereas lower frequencies (below about 0.15 Hz) have a mixture of sympathetic and
parasympathetic autonomic influences. The differential effects of the ANS on the sinoatrial node and therefore the timing of the heart beats, are caused by the differential effects of the neurotransmitters for the sympathetic and parasympathetic nervous systems. The sympathetic effects are slow (seconds), whereas the parasympathetic effects are fast (milliseconds). Therefore, rapid changes in IBIs can only be produced by parasympathetic influences.

HRV is operationalised through the performance of time-domain, frequency-domain, and nonlinear methods (Tarvainen, Niskanen, Lipponen, Ranta-aho, & Karjalainen, 2014; Task Force of The European Society of Cardiology and The North American Society of Pacing and Heart rate Variability Standards of Measurement, Physiological Interpretation, and Clinical Use, 1996). Time-domain parameters (see Table 7) are statistical calculations, which are directly applied to a series of successive RR-intervals (i.e. the interval between two consecutive R waves), such as the root mean square of successive RR differences (RMSSD).

**Table 7: Time Domain HRV Parameters**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MeanRR</td>
<td>ms</td>
<td>Mean of RR intervals</td>
</tr>
<tr>
<td>SDNN</td>
<td>ms</td>
<td>Standard deviation of RR intervals</td>
</tr>
<tr>
<td>MeanHR</td>
<td>1/min</td>
<td>Mean HR</td>
</tr>
<tr>
<td>RMSSD</td>
<td>ms</td>
<td>Square root of the mean of the sum of the squares of differences between adjacent RR intervals</td>
</tr>
<tr>
<td>NN50</td>
<td>counts</td>
<td>Numbers of pairs of adjacent RR intervals differing by more than 50 msec in the entire recording</td>
</tr>
<tr>
<td>pNN50</td>
<td>%</td>
<td>NN50 count divided by the total number of all RR intervals</td>
</tr>
</tbody>
</table>

Frequency domain methods calculate a spectrum estimate for the RR interval series. In the current project the spectrum estimate is calculated using Welch’s periodogram. Within Welch’s periodogram, the RR series is divided into overlapping segments, each segment is windowed to decrease the leakage effect, and the spectrum estimate is obtained by averaging the fast Fourier transform spectra of these windowed segments (Tarvainen et al., 2014).
The spectrum estimates are divided into very low frequency (VLF), low frequency (LF), and high frequency (HF) bands. In the case of short-term HRV recordings, the used limits in normal human subjects are 0-0.04 Hz for VLF, 0.04-0.15 for LF and 0.15-0.4 for HF (Tarvainen et al., 2014). The HF component reflects cardiac vagal tone, whereas the LF component reflects a mixture of vagal and sympathetic influences. RMSSD and the HF component of the power spectrum are closely related, all reflecting vagal cardiac influence (J. F. Thayer et al., 2012). The calculated HRV parameters of frequency domain methods are displayed in Table 8.

Table 8: Frequency Domain HRV Parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LFpow</td>
<td>ms²</td>
<td>Total power of low frequency bands (0.04-0.15 Hz)</td>
</tr>
<tr>
<td>HFpow</td>
<td>ms²</td>
<td>Total power of high frequency bands (0.15-0.4 Hz)</td>
</tr>
<tr>
<td>LFpownu</td>
<td>%</td>
<td>Low frequency power in normalized units (0.04-0.15 Hz)</td>
</tr>
<tr>
<td>HFpownu</td>
<td>%</td>
<td>High frequency power in normalized units (0.15-0.4 Hz)</td>
</tr>
<tr>
<td>LFHF</td>
<td></td>
<td>Ratio between LF and HF band powers</td>
</tr>
</tbody>
</table>

Nonlinear methods are increasingly used in clinical studies. Yet, the physiological interpretation of the results is still difficult (Carrasco M J Gaitán R González, 2009; Zbilut, Thomasson, & Webber, 2002).

For all three methods (time domain, frequency domain, and nonlinear) the standard duration of the explored time window for short-term analysis is 300 sec (Task Force of The European Society of Cardiology and The North American Society of Pacing and Heart rate Variability Standards of Measurement, Physiological Interpretation, and Clinical Use, 1996). In the context of fast-paced, dynamic police training alterations in psychophysiological demand occur much faster. Therefore, it would be beneficial to analyse much shorter time segments. Recent work showed that time and frequency domain HRV parameters were not significantly affected by shortening the length of the explored time segments from 300 to 30 seconds (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b).
In the current thesis the time domain method of RMSSD was used for several reasons: First, RMSSD enables the comparison with other work, since this method is widely used in studies addressing stress and arousal effects (Haller et al., 2014; Hansen, Johnsen, & Thayer, 2003; Ivarsson, Anderson, Akerstedt, & Lindblad, 2009; Krypotos, Jahfari, van Ast, Kindt, & Forstmann, 2011; Luque-Casado, Zabala, Morales, Mateo-March, & Sanabria, 2013; Porto & Junqueira, 2009; Ruiz-Padial, Sollers, Vila, & Thayer, 2003; Scarpa, Haden, & Tanaka, 2010; Strahler & Ziegert, 2015). This is in line with the work of Brisinda et al. (2015b), who suggest, that time-varying HRV is an “efficient method to quickly evidence transient changes of autonomic modulations of the HR in police officers undergoing realistic tactical training” (p. 83). Second, vagal tone as the physiological mechanism that is linked with self-control (Porges, 2007; J. F. Thayer & Lane, 2009), is reflected by RMSSD (Laborde & Mosley, 2016).

Since HRV encompasses the potential calculation of all of the described parameter, the term of ‘vagal tone’ or ‘parasympathetic activity’ will be used, in order to refer to the measure of RMSSD (Laborde & Mosley, 2016). HRV will be used, when referring to various calculations in the time-domain, frequency-domain or nonlinear methods. HRV was measured in Study 1 (Chapter 4) and Study 4 (Chapter 5) using Suunto memory belts.

Special precautions were taken to maintain similar conditions across all experiments in these studies, since factors, such as circadian rhythm, body position, activity level prior to recording, medication, verbalization, and breathing condition may influence HRV (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b). These included, performing the training session at a similar time of day, counterbalancing participants (within-subjects design, Study 1) and having the resting basal recordings after an adaptation time of at least 10 minutes.

Due to the lack of HRV compatible equipment in Study 2 (see Chapter 4), HR was measured using an Omron HEM-650 wrist blood pressure device before and directly after the treatment.
In Study 1 Suunto HR monitors T6c, T6d and Suunto memory belts were used. All models allow for continuous recording of the HR with the option to record HRV as well (see 3.8.3.2). The Suunto T6c and T6d measure heartbeats and the beat-to-beat interval with a HR belt, which sends the data directly to the associated wrist watch, where the data is displayed and recorded. In accordance with manufacturers instructions, HR belts were placed on participants' chest and additionally fixated with tape. The Suunto memory belts record and store the data directly on the belt, which caused minimal data loss, since no transmission errors occurred. The only observed data loss occurred because of very dynamic fighting movements, where the memory belt lost its initial position.

After the recording started, participants were asked to sit quietly for 10 minutes in order to obtain a resting HR. After the treatment (scenario and cognitive testing) participants sat down for another 10 minutes before the HR belt was removed.

In Study 1.1 the connected wrist watch was placed on the wrist of participants’ non-weapon hand. Due to dynamic fighting movements whilst punching pads and fighting with an aggressor the connection between monitor and HR belt was lost in several runs, leading to a loss in HR data. In addition, many watches were damaged. This issue was addressed by attaching the monitor to the shoulder of the participants in Study 1.2. Yet, in some cases the connection between monitor and belts was lost during the experiment presumably due to the body armour participants were wearing. Hence these cases were excluded from the subsequent analysis of physiological data. In Study 1.3. Suunto memory belts were used in order to minimize data loss caused by transmission errors. Taken together, the Suunto memory belts provided the best option to record HR and HRV in highly dynamic situations. However, the memory belts were not available before Study 1.3.

In Study 2 HR was measured in conjunction with systolic and diastolic blood pressure with the use of an Omron HEM-650 wrist blood pressure device. Participants were asked to sit down before the measurement. According to
the manufacturer’s instruction the device was placed on the left hand of participants, who were then asked to cross their arm across their chest so that their left hand rested on their right shoulder. The cuff of the monitor would then start to inflate automatically and provide the results on the display screen. The decision to use the Omron HEM-650 was based on two considerations: Firstly, the Omron HEM-650 provides acceptable accuracy and reliability for cardiovascular functioning (K. K. Hung, Cocks, Rainer, & Graham, 2015). Secondly, the monitor could easily be placed on the wrist, which meant that only gloves and no further PPE had to be removed to start the measure. The removal of further PPE, which is required by more traditional monitors, would have caused more time delay. This would have been problematic, since I aimed for the smallest time delay as possible, since research indicates that the effect of physiological arousal start to diminish immediately once the demand causing the arousal (e.g. the scenario) diminishes (J. B. Carter, Banister, & Blaber, 2003; Lewinski, 2008; Sibley & Beilock, 2007). HR was measured before (pre) and directly after the scenario (post). In Study 4 Suunto memory belts were used to record HR and HRV.

Following the recording with Suunto devices (Study 1 and 4), data was transferred to Suunto Moveslink and exported in an Excel file, which was imported into Kubios HRV Analysis Software 2.1 (Tarvainen et al., 2014).

3.8.3.2 Heart Rate Variability Analysis

After importing the data into Kubios HRV Analysis Software 2.1 (Tarvainen et al., 2014), the R-R intervals were further analysed. The recordings were pre-processed to exclude artefacts (Malik, Cripps, Farrell, & Camm, 1989). Description of the artefact correction used in the specific study is provided in the respective chapter. Removed RR intervals were replaced by conventional spline interpolation so that the length of the data did not change (i.e., resulting in the same number of beats). The smoothness prior method with a Lambda value of 500 was used to remove disturbing low frequency baseline trend components (Tarvainen, Ranta-aho, & Karjalainen, 2002).

The method of analysis of the HRV data used in this study was through linear mathematical processes (i.e., time domain method). This method is based on
the mathematical calculation of the variations in time occurring between beats.

The following parameters were used to analyse the HRV within the time domain: the mean R-R interval (RRi), standard deviation of R-R interval (SDNN) and the root-mean-square difference of successive normal R-R intervals (RMSSD). The denotations and definitions for the HRV parameters in this study follow the guidelines given by the task force of the European society of cardiology and the North American society of pacing and electrophysiology (Task Force of The European Society of Cardiology and The North American Society of Pacing and Heart rate Variability Standards of Measurement, Physiological Interpretation, and Clinical Use, 1996).

The quantitative HRV parameters were calculated from multiple RR segments of 60 seconds instead of 300 seconds. Even though 300 seconds is the suggested duration of the explored time-window (Task Force of The European Society of Cardiology and The North American Society of Pacing and Heart rate Variability Standards of Measurement, Physiological Interpretation, and Clinical Use, 1996), this is not recommended for fast-paced and highly dynamic interactions in police scenario trainings, where the interaction between participants and role players often lasts less than one minute. Additionally, recent results suggest that the quantitative estimation of time-domain HRV parameters are not significantly affected by shortening the length of the explored time-segments from 300 to 30 seconds (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b; A.-L. Smith, Owen, & Reynolds, 2013). The Selection of time-windows in the studies, are described in detail in the relevant chapter for the respective study.

3.8.3.3 Systolic and Diastolic Blood Pressure

The Omron HEM-650 wrist blood pressure monitor was used in Study 2 to take blood pressure and HR readings. The rationale for choosing wrist monitor is provided in 3.8.3.1. The blood pressure device uses the oscillometric of blood pressure measurement. Via this method the blood movement trough the artery in the wrist is detected and converted into a
digital reading. Systolic and diastolic blood pressure was recorded before and after the treatment in Study 2.

3.8.3.4 Core Temperature

Core temperature was measured in Study 2 using a Braun ThermoScan Ear Thermometer. An infrared ear thermometer was chosen because of the non-invasive and quick measurements combined with its accuracy, that has been validated in several studies (Kocoglu, Goksu, Isik, Akturk, & Bayazit, 2002; Streets & Carr, 2002). Prior to subject enrolment each day, ear thermometers were verified in their respective calibration devices to assure that the thermometers were functioning properly and measuring within specifications. The thermometer was inserted into the outer cartilaginous part of the auditory canal, without touching the osseous part of the tympanic membrane. Measurements were taken before (pre) and after (post) the treatment.

3.9 Statistical Analysis

Statistical analysis was conducted using IBM SPSS Statistics version 24.0. Depending on the number and characteristics of independent variables (within subjects, between subjects, mixed design) and the number of dependent variables, different statistical analyses were performed. Assumption testing, analysis and correcting data were conducted according to standard procedural guidelines (A. Field, 2013).

3.9.1 Assumptions for Parametric Testing

The four basic assumptions for parametric tests of all data sets were investigated to examine if parametric analysis was appropriate. These included (a) additivity and linearity (i.e. interval or ratio data), (b) a normally distributed sampling distribution, (c) homogeneity of variance, and (d) independence. All quantitative measures in the presented studies are assumed to meet the assumptions of additivity and linearity. In case of repeated measures, the assumption of independence is likely to be violated. Therefore, an additional assumption, the assumption of sphericity, has been checked where repeated measures designs have been used (see 3.9.2).
Where available, normal distribution was investigated by ‘eye-ball ing’ histogram plots of the data against normal distribution and inspecting probability-probability and normal quantile-quantile plots (A. Field, 2013). Furthermore, the distribution of was data was investigated by examining the skewness of the data sets (K. Pearson, 1895) and the results of the Shapiro-Wilk test (A. Field, 2013). A non-significant result \( p > .05 \) indicates that the data is normally distributed (S. S. Shapiro & Wilk, 1965).

Homogeneity of variance was investigated using Levene’s test (Levene, 1960). A non-significant test statistic for Levene’s test \( p > .05 \) indicates, that the variances are roughly equal and that the assumption is tenable.

When the assumptions of parametric testing were not met, square root transformations can be used to stabilise the variance of a sample (A. Field, 2013). Subsequently, the assumptions of data sets were re-examined. If all parametric assumptions were met, the appropriate statistical analysis was performed. The legends in the results table in each chapter provide information about which data sets have been transformed. For the ease of interpretation, untransformed means are reported. If transformations did not normalise the data nonparametric analyses were conducted and interpreted with caution.

The comparison of means when group sizes are equal (particularly within repeated measures designs) controls well for type 1 error rates, even when data are not normally distributed, homogeneity of variance is not assumed and data sets are skewed (Donaldson, 1966; Keselman, Algina, & Kowalchuk, 2002).

### 3.9.2 ANOVA

Due to the possible violation of the assumption of independence, the assumption of sphericity has to be examined when conducting repeated-measures ANOVAs. This requires that the variance for each set of difference scores are equal and can be examined using Mauchly’s test (A. Field, 2013). A significant result \( p < .05 \) is an indicator that sphericity of the data sample has been violated. The majority of data sets within the presented studies met
the assumption of sphericity. However, continuing to analyse data that violates this assumption leads to an increased risk of making a type 1 error. In order to reduce such risk, data sets violating this principle were corrected using greenhouse-geisser which adjusts the degrees of freedom to produce a more accurate significance value (Greenhouse & Geisser, 1959).

If ANOVAs yielded a significant result, post hoc procedures were conducted to examine the relationship between different levels of the independent variable.

Pair wise comparisons were conducted to examine if changes between the different levels of the independent variable were significant. The performance of multiple t-tests can lead to inflated type 1 error rates, which is increased when the distribution of the data is not normal (e.g. transformed data) and when sphericity has been violated (e.g. use of greenhouse-geisser correction). In that case a Bonferroni correction was used, which adjusts the degrees of freedom to produce a more conservative significance \((p)\) value.

When parametric assumptions of the data sets were adequately met (e.g. near perfect distribution and sphericity), such a correction can lead to an increased risk of a type 2 error as a result of it being too conservative. In such instances Fishers least significant differences (LSD) was employed, which is equivalent to performing multiple t-tests on the data. Fishers LSD is considered to adequately control for both type 1 and type 2 errors for parametric data (Hayter, 1986).

3.9.3 MANOVA

If the statistical analysis required the use of MANOVA, the examinations of assumptions have been extended to the multivariate case. The assumption of multivariate normality (i.e. the residuals have multivariate normality) was checked by examining univariate normality of residuals for each dependant variable.

The assumption of homogeneity of covariance matrices proposes that the variances in each group are roughly equal for each dependent variable and
that the correlations between any two dependent variables is the same in all
groups (A. Field, 2013). The MANOVA test statistics are robust to violations
of this assumption when sample sizes are equal (Hakstian, Roed, & Lind,
1979). When group sizes differed, Box’s test was used to test this
assumption. A non-significant test statistic \((p > .05)\) indicates that covariance
matrices are roughly equal like assumed.

In case of significant results univariate ANOVAs were used to examine the
effects and interactions of independent variables on each dependent
variable.

3.9.4 Robust Methods

When assumptions for parametric data were not met and transformation
procedures did not normalise the data set, robust methods were used: these
included bootstrapping in the case of \(t\)-tests (A. Field, Miles, & Field, 2012)
and robust methods for ANOVA and MANOVAs following the guidelines of
(Wilcox, 2012). The exact procedures are described in the methods section
of the respective chapters.

3.9.5 Effect Sizes

Effect Sizes can be defined as “simply the amount of anything of interest”
(Cumming & Fiddler, 2009, p. 16). In that context it is a standardised
measure of the magnitude of the observed effect across the entire population
(A. Field, 2013; Levine & Hullett, 2002; J. T. E. Richardson, 2011) In the
current thesis “effect size” refers to the actual value calculated from certain
effect statistics (Nakagawa & Cuthill, 2007). According to the guidelines of
the American Psychological Association (2010) effect sizes should be
included in the results section in order for the reader to appreciate the
magnitude or importance of a study’s finding.

The reporting of effect sizes facilitates the interpretation of significant results
(American Psychological Association, 2010; Levine & Hullett, 2002) and
helps to reduce the risk of statistically non-significant results due to
inadequate sample size (S. Sun, Pan, & Wang, 2010). Furthermore, it
enables researchers to make informed decisions about the practical significance of the study and the importance of an effect (A. Field, 2013).

There is an ongoing debate about the reporting of statistical significance (null hypothesis significance testing, NHST) and the reporting of effect sizes. For example, it has been argued (Fan, 2001), that statistical significance and effect sizes “complement each other but they do not substitute for each other” (p. 275). On the other side, the use of “new statistics” recommend omitting NHST and only reporting effect sizes (Cumming, 2014; Cumming & Fidler, 2009).

Many measures of effect sizes have been proposed with strengths and weaknesses for particular purposes (Coe, 2002; Ferguson, 2009; A. Field, 2013). The particular estimate of effect size calculation employed depends on what is being examined.

In studies comprising of three conditions or more, eta squared ($\eta^2$) is a commonly used effect size measure, because it estimates the variance explained by the model of the sample population and does not correct for the variance of the general population. Hence, $\eta^2$ can produce a (biased) inflated effect size (Levine & Hullett, 2002). Omega squared ($\omega^2$) provides an alternative effect size measure that corrects for this bias when applied to the general population. Since police officers in Germany are assumed to be specialised population, it seems more appropriate to use $\eta^2$ to gain insight of the effect size specially for this population. Partial $\eta^2$ describes the proportion of variation attributable to the factor being tested and it can also be inflated when compared to $\eta^2$.

In studies comparing two means Pearson’s correlation coefficient ($r$) is a commonly used measure of effect size. There are widely used suggestions about the magnitude of effect sizes (J. Cohen, 1988; 1992): $r = .10$ is considered a “small effect”, accounting for 1% of the total variance, $r = .30$ represents a “medium effect” explaining 9% of the total variance, and $r = .50$ is considered a “large effect”, that accounts for 25% of the variance. When there is a discrepancy between group sizes this can produce a biased effect
size making the use of another effect size measure (e.g. Cohen’s $d$) more appropriate (McGrath & Meyer, 2006). Like with $r$, J. Cohen (1988, 1992) suggested what constitutes a large or a small effect: $d = 0.2$ (small), 0.5 (medium) and 0.8 (large). In the presented studies, $r$ was used, where groups were equal and a specific cohort was investigated.

The effect size statistic $\eta^2$ is comparable to $r^2$, when variance and distribution are adequately accounted for. Since not all data sets were normally distributed or had perfect variance (i.e. adjustments were made) the $\eta^2$ calculation was maintained as the effect size statistic for all ANOVA’s conducted ($\eta^2$ was not converted to $r$). A general conversion guide for $\eta^2/r^2$ states that .01 represents a small effect, .09 a medium effect and .25 a large effect (J. Cohen, 1988).

However, the presented benchmarks of effect sizes encourage the interpretation of findings without further considerations, which has to be avoided in the context of interpreting results (Baguley, 2004; A. Field, 2013). Therefore, results will be interpreted in the context of the effects and the particular research area, i.e. the police environment. To provide uniformity when reporting the data, formatting and standard procedures of the American Psychological Association (2010) were used.
Chapter 4: Conventional vs. Non-Lethal Training Ammunition
4 Conventional vs. Non-Lethal Training Ammunition

4.1 Introduction

Research indicates that use of force training may not directly result in improved performance in the field (Andersen, 2015; Andersen et al., 2015; Jager et al., 2013; Morrison & Vila, 1998; Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b). Therefore, practitioners and scholars in the use of force domain equally emphasize the need for training designs that foster the transferability of skills from the learning to the criterion environment (Andersen et al., 2016; J. Armstrong et al., 2014; Hoff, 2012; Jager et al., 2013; K. R. Murray, 2004; Oudejans, 2008; Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b; Staller & Bertram, 2015; Staller & Zaiser, 2015b; Wollert et al., 2011). From a skill acquisition viewpoint, this need for representative task design involves the functionality of the task and action fidelity (Broadbent et al., 2015; Pinder, Davids, Renshaw, & Araújo, 2011b).

In use of force training settings, functionality of the task refers to whether the constraints the trainee police officer is exposed to, and must act upon, in the task are the same as in the performance environment. Action fidelity requires that the police officer is allowed to complete a response that is the same as in a real world incident. As previously discussed, the relationship between perceptual-cognitive and motor processes as well as emotional responses associated with the task are central for the idea of representative task design (Broadbent et al., 2015; Headrick et al., 2015; Pinder, Davids, Renshaw, & Araújo, 2011a; Pinder, Renshaw, Headrick, & Davids, 2014).

In learning and testing environments that incorporate the use of human targets (role players), live-fire weapons using CA are regularly replaced by NLT weapons using NLTA for the sake of health and safety. Hence, the aim of Study 1 was to examine WM functioning and physiological arousal during the use of NLTA compared to CA. Study 1 aims at providing an evidence-base for the design of representative learning and testing environments by investigating the effects of the use of NLT weapons compared to the use of live-fire weapons on psychophysiological functioning in a demanding fighting and shooting exercise.
For that purpose, the current study assesses HR and vagal tone as measures of psychophysiological load while police officers engage in the exercise with either NLT or live-fire weapons. Furthermore, in order to account for any potential differences in cognitive load, tests of WM capacity, assessing executive and non-executive functioning, phonemic and visuo-spatial processing were used.

4.2 Hypotheses

In line with previous research, showing that in scenario training that does not involve live-fire shooting, physiological load (as measured via HR) does not differ from the load experienced on active duty operations (Andersen et al., 2016), the following hypotheses were generated:

- Physiological load will not differ between the NLTA and the CA condition (Hypothesis 1).
- Cognitive load does not differ between the two conditions (Hypothesis 2).

Regarding the effects of the training exercise on psychophysiological demand, previous research in the domain of simulated armed confrontations (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012) indicates, that cognitive and physiological load is linked to the task at hand: when the operational task at hand involved complex decision-making process executive functioning and visuospatial processing increased at the expense of phonological processing. On the other hand, when the operational task comprised of simple, well-learned behaviours, non-executive functioning increased. Based on these results, and given that the exercises in the current study comprised of simple, non-complex movements and decision-making processes, the following hypotheses were formulated:

- Physiological load will increase from pre- to post-testing, due to the psychophysiological demand of the fighting and shooting exercise. (Hypothesis 3)
• Executive cognitive functioning will not increase from pre- to post-test, since the fighting and shooting exercise involves no complex decision-making processes. (Hypothesis 4)

• Phonological processing will decrease due to the exercise as exhibited by a decrease in Digit Span Task scores (Hypothesis 5).

• Visual-spatial processing will be enhanced, since an adaptive response to threat would encompass the allocation of resources to visuo-spatial processing capacity (Hypothesis 6)

4.3 Methods

4.3.1 Participants

A total of 42 participants took part in Study 1. All participants were police officers of the Hessian state and were assigned to Special Operation Units (SOU). Thirteen SOU officers (age: $M = 30.23$, $SD = 6.44$), with experience as police officers of $M = 8.86$, $(SD = 7.55)$ and experience as SOU officer of $M = 4.63$ $(SD = 5.53)$ participated in Study 1.1. The participants of Study 1.2 consisted of 14 SOU officers with a mean age of $M = 30.50$ years $(SD = 3.84)$, a mean experience as police officer of $M = 8.50$ years $(SD = 2.95)$ and a mean experience as a SOU officer of $M = 4.14$ years $(SD = 3.20)$. Fifteen SOU officers participated in Study 1.3 (age: $M = 28.13$, $SD = 3.96$; experience as police officer: $M = 7.47$, $SD = 4.27$; SOU experience: $M = 4.07$, $SD = 4.01$). Due to operational demands 3 officers had to leave, leading to incomplete data sets for these participants. Therefore, the final sample, that was subject to further analysis in Study 1.3 consisted of 12 police officers with a mean age of $M = 27.42$ years $(SD = 3.65)$, mean experience as a police officer of $M = 6.67$ years $(SD = 3.89)$ and experience as a SOU officer of $M = 3.5$ years $(SD = 6.44)$. To protect the anonymity of the police officers no other identifying demographics were recorded.

Since participants were required to perform highly dynamic activities during Study 1, the HR monitor belts used did not manage to record data for every participant. After the exclusion of incomplete and missing data, cardiovascular data of 25 participants were subjected to further analysis. This
sample consisted of 25 police officers with a mean age of $M = 29.00$ years ($SD = 5.52$), experience as a police officer of $M = 7.94$ years ($SD = 7.94$) and SOU experience of $M = 4.12$ ($SD = 4.48$) years.

4.3.2 Materials and Apparatus

4.3.2.1 Cognitive Measures

The sub-studies of Study 1 differed in regards to the measures of cognitive functioning used. In Study 1.1 WM capacity was measured with the Phonemic Fluency Task of the “Regensburger Wortflüssigkeitstest” (RWT (Aschenbrenner et al., 2000). Study 1.2 employed the Digit Span Task taken from the Wechsler Adult Intelligence Scale v.4 (Petermann, 2012; Wechsler, 2008). In Study 1.3, participants were given a computerized version of the Corsi Block-Tapping Task (Corsi, 1972). Detailed descriptions of the cognitive measure are given in the general methods chapter (see Chapter 3).

4.3.2.2 Physiological Measures

HR and HRV was recorded using Suunto HR monitors T6c and T6d (Study 1.1. and Study 1.2.) and Suunto memory belts (Study 1.3.) as described in the General methods chapter (see Chapter 3).

4.3.2.3 Weapons and Safety Equipment

Participants used their duty weapon (Heckler & Koch P30) with CA throughout Study 1. When the use of NLTA was required, participants were equipped with the dedicated NLT weapons of the same model. Simunition FX Cartridges were used as the NLTA. The attackers in the exercise were played by experienced use of force trainers, who wore RedMan protection suits. Participants were required to wear their issued PPE.

4.3.3 Procedure

Each Study was carried out the same way, only differing in the cognitive task (Phonemic Fluency Task, Digit Span Task, Corsi Block-Tapping Task) participants had to complete.
Participants of each study were briefed about the purpose and the design of the study in the morning. Each participant completed two counterbalanced runs of the exercise, one with CA (using live-fire weapons) and one with NLTA (using NLT weapons). In each run participants followed exactly the same protocol. Participants started with a 10 minutes resting phase in a sitting position in order to record a baseline of HRV. Subsequently, they completed the cognitive task (pre-test) and prepared for the fighting exercise. The exercise consisted of three phases designed by an experienced police use of force trainer of the hessian police academy simulating a confrontation culminating in a firearm discharge (see Figure 18). In order to motivate participants for maximum performance, they were told, that results of the exercise would be recorded. However, results of participants’ exercise performance were not recorded.

In phase 1 participants ran as often as possible between two punching pads performing a series of attacks consisting of a minimum of three punches, blows or kicks on each pad for a total of 45 seconds. Directly after that, participants were required to pull a wrestling puppet (60 kg) to safety over a distance of 13 metres. While performing this task, participants were attacked by an attacker in a RedMan Suit. They had to fight off the attacker and continuing pulling the puppet to safety. When participants reached the finish line, they had to turn around and pull the puppet to safety again, while being attacked by another attacker. They had to perform this task as often as possible in 120 seconds. Immediately after the time is over, participants had to pick up eye and ear protection and run to the shooting position, where they started shooting at targets at a distance of 6 meters. Participants were told to hit each of the 6 red targets with 2 bullets without missing a single shot and ignore the green targets. During the exercise officers had to perform a quick reload, since they were only given 2 magazines with 6 rounds in each one.

After the last shot, participants had to holster the weapon and immediately sit down at a prepared desk near the shooting position in order to do the post-test of the cognitive task of the study they participated in.
4.3.4 **Analysis of HRV data**

HRV data was processed as described in the general methods chapter using Kubios HRV Analysis Software 2.1. (Tarvainen et al., 2014). Since measured HRs were very high in study 1, artefact correction level was set to very strong. The following criteria were chosen to optimize the selection of the time-windows for calculation of HRV parameters: (a) at rest (time point: baseline), the highest possible “stationarity” of the RR signal (defined as the absence of arrhythmias and of any kind of artefacts in visual analysis of corresponding Suunto memory belt recordings); (b) during the exercise (time point: exercise), the measurements were centred on the maximum peak of HR; (c) after the scenario one time segment 2 minutes after the ending of the exercise, when cognitive testing took place (time point: cognitive task), was explored and (d) one time window 8 minutes after the ending of the exercise was chosen, in order to obtain a recovery measure (time point: recovery).
4.3.5   Statistical Analysis

A Shapiro-Wilk’s test (Razali & Wah, 2011; S. S. Shapiro & Wilk, 1965) and a visual inspection of their histograms, normal Quantile-Quantile plots and box plots was performed on each data set (Doane & Seward, 2011). Results showed that the phonemic fluency scores (Study 1.1) were not normally distributed. In order to normalise the data, a square root transformation was performed on the data set. Re-examination of transformed values revealed that the data was normally distributed, allowing for parametric testing. Exploratory results for Digit Span Task scores showed that the assumption of normal distributed data was tenable. Results for the Corsi Block-Tapping Task scores indicated, that block span scores were not normally distributed. Transformation of the data set did not normalise the data. Therefore, a robust method for MANOVA conducted on the ranked data (Munzel & Brunner, 2000), was chosen to analyse the main effect of condition on the data set. Due to the lack of any tests that deal with factorial repeated-measures designs (A. Field et al., 2012), the effect of time on blockspan and total score was assessed using the Wilcoxon signed rank-test (non-parametric data) and paired samples t-tests (parametric data) on time within each condition separately.

Furthermore, results of the exploratory data inspection showed, that mean HR data was normally distributed, allowing for the computation of a 2 x 2 factorial ANOVA (Time x Condition). However, Mauchly’s test indicated that the assumption of sphericity had been violated for the main effect of time, \( \chi^2(5) = 25.40, p < .001 \), and the Time x Condition interaction effect, \( \chi^2(5) = 34.20, p < .001 \). Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (\( \varepsilon = .66 \) for the main effect of time and \( \varepsilon = .54 \) for the interaction effect between time and condition).

RMSSD as a measure of HRV and parasympathic activity was not normally distributed. Transformation procedures did not normalise the RMSSD data set. Since according no robust factorial ANOVAs for repeated-measures designs exist so far (A. Field et al., 2012), separate robust ANOVAs for each condition using the rmanova() function (based on 20% trimmed means)
implemented in R (Wilcox, 2012) were computed. In order to investigate potential differences in cardiovascular measures between the two conditions on each time point, robust t-test using bootstrapping was performed on mean HR and RMSSD data.

4.4 Results

4.4.1 Phonemic Fluency (Study 1.1)

Means, standard deviations and 95% confidence intervals around the mean of the three Phonemic Fluency Task scores over all combinations of condition and time are displayed in Table 9.

A 2 x 2 MANOVA with repeated measures on 3 DVs (TW, MCS, NS) was performed. Pillai’s trace yielded one statistically significant main effect for time, $V = 0.70$, $F(3, 10) = 7.62$, $p = .006$, $\eta_p^2 = .696$. There was no significant main effect for condition, $V = 0.18$, $F(3,10) = 0.73$, $p = .557$, $\eta_p^2 = .180$, and no significant interaction effect for Condition x Time, $V = 0.32$, $F(3, 10) = 1.57$, $p = .256$, $\eta_p^2 = .321$.

The significant main effect of time was followed up with separate univariate mixed ANOVAs for each DV. There was a significant main effect of time on TW, $F(1, 12) = 24.41$, $p < .001$, $r = .82$, showing that TW increased from pre- to post-testing. There was also a significant main effect of time on MCS, $F(1, 12) = 8.97$, $p = .011$, $r = .65$, indicating that MCS was higher after the exercise than before. There was no significant main effect of time on NS, indicating that switching scores were similar before and after the exercise, $F(1, 12) = 0.45$, $p = .513$, $r = .19$.

Figure 19 displays the significant increase of TW and MCS for the whole sample (and the two groups)
Figure 19: Means and 95% Confidence Intervals of Phonemic Fluency Scores of Study 1.1. ** indicates significance at $p < 0.1$ from pre- to posttesting.
4.4.2  Digits Span (Study 1.2)

Means, standard deviations and 95 % confidence intervals around the mean of DF, DB and TDS are displayed in Table 10.

A 2 x 2 MANOVA with repeated measures was performed. The IV comprised of time (pre versus post) and condition (CA versus NLTA); the DVs were the three Digit Span Task scores (DF, DB, TDS). Pillai’s trace yielded one statistically significant main effect for time, $V = 0.69$, $F(2, 12) = 13.05$, $p = .001$, \(\text{partial } \eta^2 = .685\). There was no significant main effect for condition, $V = 0.12$, $F(2, 12) = 0.84$, $p = .457$, \(\text{partial } \eta^2 = .122\), and no significant interaction effect for Condition x Time, $V = 0.05$, $F(2, 12) = 0.34$, $p = .719$, \(\text{partial } \eta^2 = .054\).

Separate univariate 2 x 2 ANOVAs were conducted, in order to follow up the significant main effect for time on Digit Span Task scores. There was a significant main effect of time on DF, $F(1, 13) = 23.40$, $p < .001$, $r = .80$, showing that DF decreased from pre- to post-testing. There was no significant main effect of time on DB, $F(1, 13) = 3.22$, $p = .096$, $r = .45$. This indicates that DB was similar from pre- to post-testing. Finally, the results showed a significant main effect of time on TDS, $F(1, 13) = 12.85$, $p = .003$, $r = .71$, revealing that TDS decreased from before to after the exercise.

Figure 20 displays the summary data for Digit Span Task scores over all conditions over time, with significant changes of DF (increase) and TDS (decrease) over time.
Figure 20: Means and 95% confidence intervals of Digits Span Task scores of Study 1.1. ** indicates significance at $p < .01$; *** indicates significance at $p < .001$ from pre- to posttesting.
4.4.3  Corsi Block-Tapping Task (Study 1.3)

Means, standard deviations and 95 % confidence intervals around the mean of block span and total score are displayed in Table 11.

Since no robust MANOVAs with two independent variables exist so far (A. Field et al., 2012), the difference from pre- to post-testing was calculated for each DV (block span, total score), reducing the independent variables. The remaining IV comprised of condition and the differences of time of block span scores as DVs. Subsequently a MANOVA as conducted on the ranked data using Munzel and Brunner’s (Munzel & Brunner, 2000) method, implemented in R using the `mulrank()` function (Wilcox, 2012). There was no significant main effect of condition on the difference of scores between pre- and post-testing, $F = 1.81, p = .178$.

Results of the Wilcoxon signed-rank test revealed, that for the CA condition, block span scores did not differ significantly from before ($Mdn = 6.50$) to after the exercise ($Mdn = 7.00$), $T = 30, p = .084, r = .24$. In the NLT condition police officers also showed no significant difference from pre- ($Mdn = 6.50$) to post-test ($Mdn = 6.50$), $T = 5, p = .480, r = -.14$. Results from the paired-samples $t$-test revealed, that in the CA condition, on average, total score did not change significantly from pre- ($M = 63.17, SE = 5.54$) to post-test ($M = 71.67, SE = 4.69$), $M_{diff} = -8.50$, 95% CI [-21.35, 4.37], $t(11) = -1.46, p = 0.173, r = .40$. In the NLT condition, total scores did on average not change significantly from pre- ($M = 61.92, SE = 3.21$) to post-test ($M = 59.75, SE = 4.33$), $M_{diff} = 2.17$, 95% CI [-5.19, 9.52], $t(11) = 0.64, p = 0.530, r = .19$. Figure 21 displays the non-significant changes of block span and total score between groups.
4.4.4 Heart Rate Variability

A 2 x 4 factorial ANOVA on repeated measures (Condition x Time) on the mean HR was performed. There was no significant main effect of condition on the mean HR, $F(1, 24) = 1.86, p = .185, \eta^2_p = .072$. There was also no significant interaction effect between the condition and the time point of mean HR measure, $F(1.61, 38.54) = 1.43, p = .251, \eta^2_p = .056$. A significant main effect was discovered for time, $F(1.97, 47.16) = 634.68, p < .001, \eta^2_p = .964$. 

Figure 21: Means and 95% confidence intervals of block span and total score of Study 1.1.
Figure 22: Mean HR and 95% confidence intervals across all time points for the CA and NLTA condition (T1: baseline; T2: exercise; T3: cognitive task; T4: recovery). *** indicates significance at $p < .001$ between time points.

Bonferroni post hoc tests revealed significant differences between all time points (all $ps < .001$), indicating that mean HR scores changed between the different time points, with the baseline measure being the smallest behind recovery and cognitive task. Mean HR peaked at the timepoint of the exercise (see Figure 22). Summary data is provided in Table 12.

In order to investigate the effects of time on RMSSD, robust ANOVAs using the `rmanova()` function (based on 20% trimmed means) implemented in R (Wilcox, 2012) were computed for each condition separately.

For the CA condition, results revealed a significant main effect of time on RMSSD, $F(2.28, 31.96) = 158.49, p < .001$. Pairwise comparisons based on 20% trimmed means showed, that there were significant differences between the following time points: Baseline compared to exercise, $psihat = 12.96$...
(9.12, 16.81), \( p < .001 \), and cognitive task, \( 
\text{psihat} = 19.32 \ (16.09, \ 22.56) \), \( p < .001 \), and recovery, \( 
\text{psihat} = 17.72 \ (12.65, \ 22.79) \), \( p < .001 \), and exercise compared to cognitive task, \( 
\text{psihat} = 5.74 \ (4.37, \ 7.01) \), \( p < .001 \), and recovery, \( 
\text{psihat} = 3.65 \ (0.72, \ 6.57) \), \( p < .001 \). There was no significant effect between the time point of cognitive task and recovery, \( 
\text{psihat} = -1.47 \ (-3.67, \ 0.72) \), \( p > .05 \).

For the NLTA condition, results also yielded a significant main effect for time on RMSSD measures, \( F(2.03, \ 28.40) = 49.97 \), \( p < .001 \). Pairwise comparisons based on trimmed means (20%) showed significant differences between baseline and exercise, \( 
\text{psihat} = 8.65 \ (4.56, \ 12.73) \), \( p < .001 \), baseline and cognitive task, \( 
\text{psihat} = 13.30 \ (10.86, \ 15.74) \), \( p < .001 \), baseline and recovery, \( 
\text{psihat} = 11.50 \ (6.53, \ 16.47) \), \( p < .001 \), exercise and cognitive task, \( 
\text{psihat} = 4.01 \ (1.66, \ 6.37) \), \( p < .001 \), and exercise and recovery, \( 
\text{psihat} = 2.70 \ (0.24, \ 5.16) \), \( p < .001 \). There was no significant effect between the time point of cognitive task and recovery, \( 
\text{psihat} = -0.98 \ (-2.63, \ 0.67) \), \( p > .05 \).

Figure 23 shows the summary data for RMSSD scores at the four time points for CA and the NLTA condition.
Figure 23: RMSSD and 95% confidence intervals across all time points for the CA and NLTA condition (T1: baseline; T2: exercise; T3: cognitive task; T4: recovery). *** indicates significance at $p < .001$ between time points.

Potential differences of cardiovascular data between CA and NLTA condition were analysed using separate paired-samples t-tests for each DV at the four different time points (baseline, exercise, cognitive task, recovery). Results yielded no significant differences between CA and NLTA conditions at any time point. Summary data for cardiovascular measures are shown in Table 12.
Table 9: Means, Standard Deviations and Confidence Intervals of Phonemic Fluency Scores of Study 1.1

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td>Pre</td>
<td>13</td>
<td>40.92</td>
<td>11.44</td>
<td>34.01</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>13</td>
<td>44.77</td>
<td>9.47</td>
<td>39.07</td>
</tr>
<tr>
<td></td>
<td>Diff</td>
<td>13</td>
<td>3.85</td>
<td>4.41</td>
<td>1.18</td>
</tr>
<tr>
<td>NLTA condition</td>
<td>Pre</td>
<td>13</td>
<td>41.31</td>
<td>11.44</td>
<td>34.39</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>13</td>
<td>48.62</td>
<td>13.6</td>
<td>40.62</td>
</tr>
<tr>
<td></td>
<td>Diff</td>
<td>13</td>
<td>7.31</td>
<td>7.12</td>
<td>3.00</td>
</tr>
<tr>
<td><strong>MCS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td>Pre</td>
<td>13</td>
<td>0.32</td>
<td>0.14</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>13</td>
<td>0.66</td>
<td>0.55</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Diff</td>
<td>13</td>
<td>0.34</td>
<td>0.49</td>
<td>0.04</td>
</tr>
<tr>
<td>NLTA condition</td>
<td>Pre</td>
<td>13</td>
<td>0.38</td>
<td>0.19</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>13</td>
<td>0.49</td>
<td>0.31</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>Diff</td>
<td>13</td>
<td>0.11</td>
<td>0.25</td>
<td>-0.04</td>
</tr>
<tr>
<td><strong>NS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td>Pre</td>
<td>13</td>
<td>29.77</td>
<td>8.76</td>
<td>24.48</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>13</td>
<td>28.15</td>
<td>9.20</td>
<td>22.59</td>
</tr>
<tr>
<td></td>
<td>Diff</td>
<td>13</td>
<td>-1.62</td>
<td>7.44</td>
<td>-6.11</td>
</tr>
<tr>
<td>NLTA condition</td>
<td>Pre</td>
<td>13</td>
<td>29.85</td>
<td>6.12</td>
<td>26.15</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>13</td>
<td>34.15</td>
<td>10.86</td>
<td>27.59</td>
</tr>
<tr>
<td></td>
<td>Diff</td>
<td>13</td>
<td>4.31</td>
<td>7.09</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Table 10: Means, Standard Deviations and Confidence Intervals of Digit Span Scores of Study 1.2

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>M</td>
<td>SD</td>
<td>LL</td>
<td>UL</td>
<td></td>
</tr>
<tr>
<td>DF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td>Pre</td>
<td>14</td>
<td>10.43</td>
<td>1.65</td>
<td>9.48</td>
<td>11.38</td>
</tr>
<tr>
<td>Post</td>
<td>14</td>
<td>9.79</td>
<td>1.72</td>
<td>8.79</td>
<td>10.78</td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>14</td>
<td>-0.64</td>
<td>1.08</td>
<td>-1.27</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>NLTA condition</td>
<td>Pre</td>
<td>14</td>
<td>11.07</td>
<td>1.94</td>
<td>9.95</td>
<td>12.19</td>
</tr>
<tr>
<td>Post</td>
<td>14</td>
<td>10.00</td>
<td>1.41</td>
<td>9.18</td>
<td>10.82</td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>14</td>
<td>-1.08</td>
<td>1.21</td>
<td>-1.07</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>DB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td>Pre</td>
<td>14</td>
<td>8.36</td>
<td>2.68</td>
<td>6.81</td>
<td>9.90</td>
</tr>
<tr>
<td>Post</td>
<td>14</td>
<td>7.57</td>
<td>2.59</td>
<td>6.07</td>
<td>9.07</td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>14</td>
<td>-0.79</td>
<td>2.83</td>
<td>-2.42</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>NLTA condition</td>
<td>Pre</td>
<td>14</td>
<td>8.86</td>
<td>1.79</td>
<td>7.82</td>
<td>9.89</td>
</tr>
<tr>
<td>Post</td>
<td>14</td>
<td>8.07</td>
<td>1.49</td>
<td>7.21</td>
<td>8.93</td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>14</td>
<td>-0.79</td>
<td>1.63</td>
<td>-1.72</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>TDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td>Pre</td>
<td>14</td>
<td>18.79</td>
<td>4.06</td>
<td>16.44</td>
<td>21.13</td>
</tr>
<tr>
<td>Post</td>
<td>14</td>
<td>17.36</td>
<td>3.65</td>
<td>15.25</td>
<td>19.46</td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>14</td>
<td>-1.43</td>
<td>2.82</td>
<td>-3.06</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>NLTA condition</td>
<td>Pre</td>
<td>14</td>
<td>19.93</td>
<td>2.89</td>
<td>18.26</td>
<td>21.60</td>
</tr>
<tr>
<td>Post</td>
<td>14</td>
<td>18.07</td>
<td>2.06</td>
<td>16.88</td>
<td>19.26</td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>14</td>
<td>-1.86</td>
<td>2.21</td>
<td>-3.14</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>
Table 11: Means, Standard Deviations and Confidence Intervals of Corsi Block-Tapping Task Scores of Study 1.3

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
<th>Mdn</th>
<th>LQ</th>
<th>UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Block span</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>12</td>
<td>6.33</td>
<td>1.07</td>
<td>5.66</td>
<td>7.02</td>
<td>6.50</td>
<td>6.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Post</td>
<td>12</td>
<td>7.00</td>
<td>0.85</td>
<td>6.46</td>
<td>7.54</td>
<td>7.00</td>
<td>6.25</td>
<td>7.00</td>
</tr>
<tr>
<td>Diff</td>
<td>12</td>
<td>0.67</td>
<td>1.23</td>
<td>-0.12</td>
<td>1.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLTA condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>12</td>
<td>6.42</td>
<td>0.67</td>
<td>5.99</td>
<td>6.84</td>
<td>6.50</td>
<td>6.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Post</td>
<td>12</td>
<td>6.25</td>
<td>0.87</td>
<td>5.70</td>
<td>6.80</td>
<td>6.50</td>
<td>5.25</td>
<td>7.00</td>
</tr>
<tr>
<td>Diff</td>
<td>12</td>
<td>-0.17</td>
<td>0.83</td>
<td>-0.70</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>12</td>
<td>63.17</td>
<td>19.20</td>
<td>50.97</td>
<td>75.36</td>
<td>61.50</td>
<td>51.00</td>
<td>82.25</td>
</tr>
<tr>
<td>Post</td>
<td>12</td>
<td>71.67</td>
<td>16.23</td>
<td>61.35</td>
<td>81.98</td>
<td>69.00</td>
<td>56.25</td>
<td>82.25</td>
</tr>
<tr>
<td>Diff</td>
<td>12</td>
<td>8.50</td>
<td>20.22</td>
<td>-4.35</td>
<td>21.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLTA condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>12</td>
<td>61.92</td>
<td>11.12</td>
<td>54.85</td>
<td>68.98</td>
<td>65.00</td>
<td>54.00</td>
<td>70.00</td>
</tr>
<tr>
<td>Post</td>
<td>12</td>
<td>59.75</td>
<td>15.00</td>
<td>50.22</td>
<td>69.28</td>
<td>65.00</td>
<td>43.50</td>
<td>70.00</td>
</tr>
<tr>
<td>Diff</td>
<td>12</td>
<td>-2.17</td>
<td>11.57</td>
<td>-9.52</td>
<td>5.19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 12: Comparison between CA and NLTA Condition of Cardiovascular Variables of Study 1 (n=25)

<table>
<thead>
<tr>
<th>Variable</th>
<th>LF condition</th>
<th>NLT condition</th>
<th>Difference</th>
<th>95% CI</th>
<th>95% CI</th>
<th>BCa 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>LL</td>
<td>UL</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Mean HR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>77.89</td>
<td>11.22</td>
<td>73.26</td>
<td>82.52</td>
<td>75.21</td>
<td>13.47</td>
</tr>
<tr>
<td>Exercise</td>
<td>200.22</td>
<td>19.49</td>
<td>192.17</td>
<td>208.26</td>
<td>194.49</td>
<td>17.68</td>
</tr>
<tr>
<td>Cognitive task</td>
<td>127.47</td>
<td>13.38</td>
<td>121.95</td>
<td>132.99</td>
<td>127.21</td>
<td>12.60</td>
</tr>
<tr>
<td>Recovery</td>
<td>106.92</td>
<td>11.19</td>
<td>102.31</td>
<td>111.54</td>
<td>109.16</td>
<td>11.49</td>
</tr>
<tr>
<td>RMSSD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>21.58</td>
<td>6.61</td>
<td>18.86</td>
<td>24.31</td>
<td>20.84</td>
<td>7.06</td>
</tr>
<tr>
<td>Exercise</td>
<td>10.62</td>
<td>6.17</td>
<td>8.07</td>
<td>13.16</td>
<td>10.46</td>
<td>3.87</td>
</tr>
<tr>
<td>Cognitive task</td>
<td>5.06</td>
<td>2.74</td>
<td>3.93</td>
<td>6.18</td>
<td>4.89</td>
<td>2.44</td>
</tr>
<tr>
<td>Recovery</td>
<td>8.11</td>
<td>7.23</td>
<td>5.13</td>
<td>11.09</td>
<td>7.39</td>
<td>5.73</td>
</tr>
</tbody>
</table>
4.5 Discussion

In Study 1 the effects of the use of NLTA compared to CA on cardiovascular and cognitive functioning was investigated. The results showed a number of relevant findings: First, no effect between the use of NLTA and CA on mean HR, RMSSD (overall results of Study 1), Phonemic Fluency Task scores (Study 1.1), Digit Span Task scores (Study 1.2) and Corsi Block-Tapping Task scores (Study 1.3) were discovered. Second, mean HR increased from baseline to the peak at the time point of the exercise, with a subsequent drop at the cognitive task and recovery measure (but without returning to baseline). This indicates that the psychophysiological demands of the fighting and shooting exercise lead to physiological load. Third, RMSSD measures dropped from the highest score at baseline over the time point at the exercise to the lowest score at cognitive testing. Measures did not change significantly from the time point at cognitive testing to the time point at recovery measure. This indicates that vagal tone declined from baseline to the exercise and from the exercise to the cognitive testing, reflecting an increasing withdrawal of parasympathetic activity. Fourth, Study 1.1 showed that TW generated and MCS increased from pre- to post-testing, whereas NS stayed the same, indicating that non-executive functioning was affected by the exercise whereas non-executive functioning was not affected absolutely (but relatively). Fifth, in Study 1.2 DF and TDS decreased from before to after the exercise. Digit span backwards did not change. This suggests a decrease of phonological processing capacity due to the shooting and fighting exercise. Sixth, in Study 1.3 neither block span, nor total score changed from pre- to post-testing, indicating that visuospatial processing capacity was not affected by the exercise.

4.5.1 Differences Between the Use of CA and NLTA in Training and Testing Environments

The first result fails to reject the hypothesis that physiological arousal and cognitive functioning did not differ between the two ammunition conditions. However, it is noteworthy, that this does not equal that the hypothesis of no difference (null hypothesis) is accepted. The results rather provide support
for the hypothesis, that there is no difference between the two conditions. Therefore, future studies should aim at providing further support for this hypothesis. With regards to the comparison of other NLT weapons and life-fire weapons, research on the use of laser-based handgun training systems found that the use of live-fire weapons is not necessary for the acquisition of marksmanship skills (Getty, 2014; Kratzig, 2013). The current results extend these findings, in that NLTA can be used in order to elicit the same psychophysiological and cognitive effects than the use of CA does. Since unintentional firearm-related injuries and deaths during training are a serious risk (Barber & Hemenway, 2011; Fowler, Dahlberg, Haileyesus, & Annest, 2015; Heim, 2009; Heim, Schmidbleicher, & Niebergall, 2006a; 2006b; Karger, Wissmann, Gerlach, & Brinkmann, 1996; K. R. Murray, 2004; Örnehult & Eriksson, 1987; Stansfield & Rushforth, 2009; Therbo & Osten, 2009), fostering safer training activities is extremely valuable. The use of NLTA instead of CA in skill acquisition training can enhance the health and safety of participants without compromising representativeness, therefore fostering optimal environments for skill acquisition and/or testing (Staller, 2015; Staller & Zaiser, 2016).

4.5.2 Psychophysiological Demand Under Physical Pressure

The hypothesis that physiological load would increase from pre- to post-testing (Hypothesis 3) was accepted because of the second and third result. Mean HR at the peak of physiological load were near 200 bpm, with some police officers passing this threshold. Such high HR recordings have not yet been reported in the context of policing. For example, Andersen et al. (2016) reported an average maximum HR of around 148 bpm for real-life scenarios and 146 bpm for real-life incidents. J. Armstrong et al. (2014) reported average HRs between 112 and 121 bpm at the time point of physical contact with the role player within four different scenarios. In another study investigating the psychobiological stress response to simulated school shooting, the authors found average HRs of 141 bpm during the scenario (Strahler & Ziegert, 2015). Other studies investigating HR in scenarios that require the discharge of weapon showed average HRs between 140 and 150
bpm (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b; Haller et al., 2014).

Yet, the current exercise was a combination of fighting, carrying the dummy and shooting. For real life incidents, where a police officer had to fight for his life combining unarmed with armed combat, there are no data in the public domain. Data gathered in real life incidents so far (Andersen et al., 2016) do not involve the process of overcoming an individual that is determined to fight back for as long as he can. Data from the martial arts domain, where cardiovascular functioning can be tested in the field (e.g. competition, sparring) in an ethically acceptable way, maximum HR regularly reaches between 195 and 205 bpm (Bouhlel et al., 2006; Haddad, Chaouachi, Wong, Castagna, & Chamari, 2011; Markovic, Vucetic, & Cardinale, 2008; L. Santos et al., 2011). In two studies (Bouhlel et al., 2006; Markovic et al., 2008) average HR dropped to between 110 and 125 bpm after three minutes of rest. In the current study, police officers showed mean HRs between 121 and 132 bpm two minutes and between 102 and 113 bpm eight minutes after finishing the exercise. The results of the current study combined with measures from the martial arts domain indicate that HR of about 200 bpm will be reached when physical fighting occurs. Thus, in terms of representativeness of scenarios, use of force trainers should design very physically demanding exercises that elicit the described HRs in order to allow the learners to develop their skills under these constraints. However, it has to be noted that even the cause of the HR increase may play a crucial role. High HR due to fighting behaviour may have different effects than high HR due to physical exercise (without fighting). If fighting led to adaptive behaviour of the organism in order to cope with the demand, a cognitive shift of executive processes might be possible (Boulton, 2014; A. P. J. Roberts, 2012). This issue is subject to investigation in the next chapter.

Despite enabling the organism to fight for survival, high HR is also associated with some drawbacks in the context of policing. High arousal as shown by high HR has been shown to reduce the cognitive complexity of perceptions and limit the capacity to notice cues that are in the periphery of central tasks (Chajut & Algom, 2003; Easterbrook, 1959; Paulhus & Lim,
This phenomenon, which is based on the Yerkes-Dodson Law (Yerkes & Dodson, 1908), has been conceptualized as “attentional myopia” (T. Mann & Ward, 2004; 2007) and is relevant for the control of aggressive behaviour (A. Ward et al., 2008). Haller et al. (2014) recently showed that high HRs do not necessary result in excessive use of force. Their results of police officers engaging in two different scenarios indicated that the combination of high HR with low vagal tone as measured by RMSSD was associated with inappropriate aggressive behaviour. They suggest that vagal tone is a valuable external symptom of internal motivational states and decision making processes. The individual characteristics of the responsiveness to demand could be a predictor of aggressive behaviour (Haller et al., 2014).

In the current study, RMSSD as a HRV marker of parasympathetic activity, decreased markedly from baseline to the exercise, and indicating a suppression of vagal activity confirming previous reports in policing contexts (Haller et al., 2014; Strahler & Ziegert, 2015). Recently, it has been suggested that HRV is the peripheral, vagus-mediated reflection of central decision-making processes, in particular the reflection of prefrontal-subcortical inhibitory functions (G. Park & Thayer, 2014). The suppression of HRV (via the downregulation of the cardiac vagal tone) signals impairment of central decision-making processes, while an increase in HRV indicates efforts of self-regulation. Together with high HRs, that are associated with bias towards hostile attributions and aggression (Lorber, 2004; Patrick, 2008; S. C. Williams, Lochman, Phillips, & Barry, 2003), this may result in inappropriate aggression in the context of policing (Haller et al., 2014).

It is noteworthy that parasympathetic activity declined more after the ending of the physically very demanding exercise, when cognitive testing took place. Research has shown that an increase in cognitive task difficulty often results in a decrease in HRV measures, that reflect parasympathetic activity (HF-HRV and RMSSD). This has been shown for tasks that involve executive functioning (Byrd, Reuther, McNamara, DeLucca, & Berg, 2015; Duschek, Muckenthaler, Werner, & Reyes del Paso, 2009; Fairclough & Roberts, 2011; Hansen et al., 2003; Makersie & Calderon-Moultrie, 2016; Mukherjee,
With regards to non-executive functioning, Hansen et al. (2003) showed that individuals with low vagal tone performed worse on cognitive tasks measuring executive functioning whereas no performance decrements were observed on non-executive components of the tasks. This may provide an explanation for the observed results: It may be possible that the fighting and shooting scenario placed lower (executive) cognitive load on participants than the subsequent WM tasks.

**4.5.3 Cognitive Functioning**

The current findings that MCS increased during the exercise suggests that non-executive functioning increased from pre- to post-test. The increase in TW can be explained by the observed increase of MCS. NS did not change due to the exercise (wheras MCS and TW did), can be interpreted as a relative decrease of executive functioning.

One potential reason for the difference in phonemic fluency performance between pre- and post-test, may be the demands that the fighting and shooting exercise placed upon the officers. All officers, as members of SOUs, were extremely well trained, which was also indicated by the cardiovascular measures. Therefore, the three components of the fighting and shooting task (Phase 1 - 3) may have required less attentional control leading to more autonomous completion of the task (Magill & Anderson, 2014). Once a task becomes autonomous its completion will require less executive control leaving more WM capacity free for the completion of other tasks (Petersen et al., 1998; Shallice & Burredes, 1993). It is reasonable to assume that participants were able to employ simple heuristics during the exercise, since the task itself was not complex. The use of heuristics such as pattern matching the unfolding attack of the aggressor by using previous experience and training may be effectively employed for the production of an optimal solution (Kahneman & Klein, 2009; G. A. Klein, 2009; G. A. Klein, Orasanu, Calderwood, & Zsambok, 1993; Lipshitz et al., 2001; K. G. Ross et al., 2004). If information processing was devoted to following simple heuristics, this would explain the increase in non-executive functioning. In the
context of human defensive behaviour this could be accounted for as an adaptive response of the organism in order to cope with the demand at hand (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012). A. P. J. Roberts (2012), investigating phonemic fluency performance of authorized firearms officers before and after different scenarios showed that in a scenario where officers had to rely on well-learned and practiced tactical behaviour, a relative increase in non-executive functioning occurred, as shown by an increase in TW generated while simultaneously NS decreased. It has to be noted that the current findings are in line with this, even though the current results yielded a relative decrease in executive functioning and an absolute increase in non-executive functioning.

The absolute increase of non-executive functioning in the current study compared to the relative increase of A. P. J. Robert’s work (2012), could be explained by the different physiological load placed upon participants in the different studies. Average HRs in the tactical scenarios were between 88 and 94 bpm directly after the completion of the scenario, while in the current study the fighting and shooting exercise elicited average HRs between 195 and 205 at the peak. It is reasonable to assume, that participants with these HR levels, had considerable levels of blood lactate, as shown for different martial arts immediately after fighting bouts (Branco et al., 2013; Joel et al., 2014). Increases in blood lactate levels is associated to a worsening in attentional capabilities involving the prefrontal cortex (Coco et al., 2009; Valentina Perciavalle et al., 2013; Valentina Perciavalle, Di Corrado, Scuto, Perciavalle, & Coco, 2014; Valentina Perciavalle, Maci, Perciavalle, Massimino, & Coco, 2015). Thus, if executive functioning is hard to maintain under this level of physiological load, an increase in non-executive processes may facilitate performance of the trained tasks at hand (fighting and shooting) and therefore seems beneficial.

However, practice effects may also account for these results. Practice effects in the Phonemic Fluency Task have been observed for TW (E. Strauss, Sherman, & Spreen, 2006). Regarding switching (NS) and clustering (MCS) practice effects have been lower (E. Strauss et al., 2006), suggesting that search processes can vary with each test administration. It is reasonable to
assume, that a combination of practice effects and changes in WM processing lead to the obtained results, since TW (practice effects) and the relationship between the measures of executive and non-executive functioning changed (manipulation). In this context, the finding that no absolute increase in NS was found could be interpreted as a relative decrease in NS.

Regarding phonological and visuospatial processing the results showed, that the fighting and shooting exercise had an effect of Digit Span Task scores (DF and TDS decreased, while DB did not change) but no effect on Corsi Block-Tapping Task scores.

In demanding situations that bear the potential for cognitive overload, numerous sources of interference compete for brain resources, resulting in decrements to the attentional resources that are available for perception and decision making (Baddeley, 2000; McElree, 2001; Sliwinski, Smyth, Hofer, & Stawski, 2006). In order to cope with the increased demand, it is necessary for certain cognitive system to reduce output. The phonological loop as one of the most recently evolved cognitive systems, as it is involved in the development of spoken language (M. H. Christiansen & Kirby, 2003a). As such it is likely that a reduction in capacity is experienced first, since during highly demanding situations the most recently evolved higher cortical systems are likely to reduce output first (Baddeley, 2007; Fuster, 2002; Huttenlocher & Dabholkar, 1997; Jerison, 1973; G. Roth & Dicke, 2005). This would explain the decrease in performance on phonological tasks, as exhibited by the decrease in DF and TDS. This is consistent with previous findings of Digit Span Tasks in demanding scenarios (A. P. J. Roberts, 2012).

However, the finding that DB, which involves more higher order processing and visuo-spatial manipulation compared to DF (Alloway et al., 2004), did not equally decrease, may stem from the output of: (a) the central executive, which still operated at same level as findings from the Phonemic Fluency Task have shown, and (b) the visuo-spatial sketchpad, that equally operated without decrements during the exercise.
The visual system is the primary sensory processing system for threat detection in humans, since it is the most automatically processed sensory modality (M. H. Christiansen & Kirby, 2003a; Howard & Templeton, 1966; Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999; Klinge, Röder, & Büchel, 2010; Liddell et al., 2005; Parasuraman & Galster, 2013; Pick, Warren, & Hay, 1969). As such it remains largely intact even under high levels of physiological load, with hyper-vigilance and the completion of simple search tasks almost always improved (Charmandari et al., 2005; Lambourne et al., 2010; Lambourne & Tomporowski, 2010; Lewinski, 2008; A. P. J. Roberts, 2012; Tomporowski, 2003). This is reflected by the current findings of no change in visual processing performance in the Corsi Block-Tapping Task from pre- to post-testing.

A. P. J. Roberts (2012) found an increase in block span and total score with authorized firearms officers completing a tactical scenario, indicating that visuo-spatial processing increased due to the task. With average HRs between 98 and 112 bpm measured directly after the completion of the scenario, the current study was more physiologically demanding. This may provide a possible explanation for the differences in the findings, in the sense that the psychophysiological load (at around 200 bpm) does not enhance visual processing. There are two aspects, that support this interpretation: First, research confirming enhanced visual sensory processing under psychological load had lower levels of demand placed upon participants as indicated by lower HR (Lambourne et al., 2010; Lewinski, 2008; A. P. J. Roberts, 2012). Second, ‘perceptual narrowing’ (tunnel vision) has been reported for real world shooting incidents of police officers (Engel & Smith, 2009; Klinger & Brunson, 2009). Since no data of psychophysiological arousal for real world armed confrontations exists in the public domain, it can be hypothesised that the ‘perceptual narrowing’ occurred under very high physiological load similar to the HRs measured in the current study. Taken together, this would indicate that visuospatial processing increases as an adaptive response to the task at hand due to psychophysiological arousal, dropping off later on, when arousal reached the limit of individuals. Thus, the initial increase could be equalized by a subsequent decrease in visual processing.
processing capacity, resulting in no changes in the recorded visual processing task. This relationship would then best be described an inverted-U relationship, as it has been described for the workload of exercise and the activation level of the central nervous system (Easterbrook, 1959; Hüttermann & Memmert, 2014; Sternberg, 1969).

Additionally, recent research in the sports domain has shown that expert athletes maintain performance on an attentional task during increasing workload up to high levels of exercise intensity (measured via HR), while non-athletes show decrements in performance later on after an initial increase (Hüttermann & Memmert, 2014). The authors concluded that the enhanced cognitive capacity was mediated through the physical fitness level of participants (Etnier, Salazar, Landers, & Petruzzello, 1997; Tomporowski & Ellis, 1986). In the current study, only police officers from SOU participated. As such they regularly demonstrate high levels of physical fitness (Angeltveit, Paulsen, Solberg, & Raastad, 2016; M. R. Davis et al., 2016; Halson & Nichols, 2015; A. P. Hunt, Orr, & Billing, 2013; Kelly, 2016; Pryor, Colburn, Crill, Hostler, & Suyama, 2012; Solberg et al., 2015; Stanfill, 2011). Thus, it is reasonable to assume, that potential decrements in cognitive performance will start at a much higher level of workload (at higher levels of physical intensity) compared to individuals with lower levels of physical fitness (Hüttermann & Memmert, 2014).

4.5.4 Weaknesses and Limitations

There were several limitations with these studies. First, the highly dynamic exercise led to several losses in cardiovascular data. The reasons are that: (a) HR belts with paired wrist devices were used in Study 1.1. and 1.2. for recording the data, (b) ballistic vests posed difficulties for the continuous connection between sender (belt) and receiver unit (wrist monitor), and (c) highly dynamic fighting motions were required during the exercise, leading to disruptions in the cardiovascular recording. The missing data did not allow for sound statistical testing of each sub-study due to the small remaining sample sizes of cardiovascular data. Therefore, mean HR and HRV was only analysed over the whole study. Even though the exercise stayed the same, it
has to be taken into account that cardiovascular data measured at the time point of cognitive testing included participants performing three different cognitive tests (Study 1.1 – 1.3).

Second, in Study 1 police officers were using CA and NLTA only on static non-human targets and officers knew in advance that they would not be shot at. Research has demonstrated that participation in a representative training scenario delivers close to the same levels of physiological arousal as would be expected from an actual incident (Andersen et al., 2016; Lewinski, 2006; 2008). Participants were exposed to cognitive, emotional and physical demands during the exercise that are present in actual encounters involving high levels of threat or lethal force (Jager et al., 2013; Jensen & Simpson, 2014; Jensen & Wrisberg, 2014). Although the training has been demonstrated to be demanding, the emotional demand related to real danger is not present, it therefore has to be inferred that the demand placed on officers operationally is even greater.

4.5.5 Summary of Chapter 4

In conclusion Study 1 provides evidence, that the use of NLTA produces the same levels of psychophysiological demand as CA. In terms of the following experiments and in terms of learning and testing environments in general, the use of NLTA provides a safe yet representative alternative to CA when studying the use of issue weapons.

Furthermore, the current study suggests, that SOU officers operating at the high anaerobic levels are able to maintain executive functioning and visuospatial processing capacity at the expense of phonological processing. Reliance on retrieval processes as non-executive functioning is enhanced. Altogether the changes in cognitive functioning seem to be an adaptive response to the task at hand.
Chapter 5: Scenario vs. Exercise
5 Scenario vs Exercise

5.1 Introduction

Executive cognitive functioning abilities are required in situations that involve planning, goal-directed behaviour, decision-making, error correction, trouble shooting, attentional control, self-control or inhibition of habitual responses (Baddeley, 1996; Giancola, Mezzich, & Tarter, 1998; Hawkins & Trobst, 2000; Hofmann, Schmeichel, & Baddeley, 2012; A. Luria, 1980; A. B. Morgan & Lilienfeld, 2000; Norman & Shallice, 1986). In the context of policing, optimal performance in situations like armed confrontations require a number of complex processes that are linked to the described situations. For example, police officers have to decide if approaching the suspect is the better option than to stay behind cover. Also, when there are two competing goals like saving an injured police officer under the risk of injury to oneself or to stay in a relatively safe zone while risking the death of the injured officer.

Furthermore, with regards to excessive use of force executive cognitive functioning plays an important role in policing since deficits are hypothesised to be an aetiological factor in aggression and violent behaviour through its influence on inhibitory control (Giancola, 1995; 2000; 2004; Giancola & Chermack, 1998; Hawkins & Trobst, 2000; Hoaken, Allaby, & Earle, 2007; Micaï, Kavussanu, & Ring, 2015; Moffitt, Lynam, & Silva, 1994; A. B. Morgan & Lilienfeld, 2000; Paschall & Fishbein, 2002; Pennington & Bennetto, 2008; Raaijmakers et al., 2008; Raine, 1993; Séguin, Pihl, Harden, Tremblay, & et al, 1995; van Rijn & Swaab, 2015). For example, limited WM capacity in threatening environments has been significantly related to an increased risk of police officers’ “aggressive shooting behaviour” toward unarmed bystanders (Kleider & Parrott, 2009). Additionally, reduced WM processing capacity seems to be associated with an increased overall error risk and a higher likelihood of reaction failure against armed opponents by police officers (Kleider, Parrott, & King, 2010). With reports about excessive use of force being a continuous issue in society, scholars and practitioners alike aim to reduce this phenomenon (Alpert & Smith, 1994; Atherley & Hickman, 2014; Gül, Hekim, & Terkeşli, 2013; Hutson et al., 2009; Kargin, 2016;
In this context, police officers have to display control over their aggressive behaviour in order to act within the legal framework and not use excessive force, while simultaneously demonstrating enough aggression to complete the task at hand, like restraining the suspect or defending oneself.

Taken together, performance in situations that require the officer to make sound tactical decisions and apply a reasonable amount of force should be linked to executive functioning. Previous studies in this area suggest a link between the demands that are placed by use of force situations on police officers and executive functioning (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012). Yet, taking into account the results in the previous Chapter, it is currently unclear what kind of psychophysiological demand (threat or physical activity) produces the effects on WM function. Since use of force incidents regularly incorporate higher levels of physiological demand, both aspects (coping with threat/demand, physiological arousal) have to be taken into consideration.

Therefore, this chapter aims to investigate what kind of psychophysiological demand produces the observed effects on executive functioning by comparing the effects of different activities (physical activity versus scenario training).

5.2 Executive Functioning in Police Use of Force Events

An increasing body of research suggests that stress affects WM function in human performance (Lupien, Gillin, & Hauger, 1999; Lupien, Maheu, Tu, Fiocco, & Schramek, 2007; Oei et al., 2006; Qin, Hermans, van Marle, Luo, & Fernández, 2009; Schoofs, 2009; Schoofs et al., 2008; 2009; 2013). Some studies found no influence of stress on WM (R. Grossman et al., 2006; Kuhlmann, Kirschbaum, & Wolf, 2005; Porcelli et al., 2008), whereas other findings indicate a WM performance increase (Al'Absi, Hugdahl, & Lovatto, 2002; R. S. Lewis, Nikolova, Chang, & Weekes, 2008) or decrease (Elzinga & Roelofs, 2005; Lupien et al., 1999; Oei et al., 2006; Schoofs et al., 2008; 2009). Schoofs et al. (2009) suggested that the results vary due to varying...
samples (e.g. age, sex, and health status) and the application of different WM tasks and stressors. Concerning different WM tasks, they assumed that challenging WM tasks (including processing of stored information; e.g. n-back Task, Sternberg paradigm, Operation Span Task, Phonemic Fluency Task) require more cognitive resources and are therefore more prone to the impact of stress than simple WM tasks. This indicates that stress impairs performance in demanding WM tasks that require maintenance and executive processing of information (Oei et al., 2006; Schoofs et al., 2008; 2009).

Since executive functioning is linked to performance in police use of force situations, several studies have investigated the effects of various simulated use of force situations and training activities on WM functioning, especially executive processing (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012; Taverniers et al., 2011).

5.2.1.1 Exploratory Work of A. P. J. Roberts (2012) on Use of Force Situations and Executive Functioning

5.2.1.1.1 Effects on Verbal Fluency Tasks

A. P. J. Roberts (2012) used a variety of different WM tasks in order to assess the effect of different situational judgment tests of authorized firearms officers on cognitive functioning. He employed both phonemic and the semantic versions of Verbal Fluency Tasks (Benton, 1968; Newcombe, 1969) before and after a scenario that required the use of three different tactics. The first scenario comprised of a short car chase with a subsequent apprehension of the suspect that lasted for approximately 10 minutes (scenario A). A slow building search with a dynamic entry in order to find a suspect, call him out, and apprehend him, was the goal of the second scenario lasting approximately 45 minutes (scenario B). The third scenario required participants to dynamically search and confront an armed aggressor shooting at civilians (scenario C). This scenario took about 45 minutes.

Results showed that executive functioning decreased in scenario A and B, while an increase during scenario C was observed. The author concluded
this difference occurred because of the tactics used by the police officers: Scenario A and B required the use of well-learned tactics, whereas a novel tactic was used in scenario C.

5.2.1.1.2 Effects on Stroop Task

He furthermore employed the Stroop Task (Stroop, 1935) as a common measure for assessing the executive component of inhibition (Archibald & Kerns, 1999; Miyake et al., 2000) in order to investigate the effects of a live-fire exercise on cognition (time of live-fire exercise: approximately 20 minutes). The Stroop Task requires the participants read aloud the colour the word is printed in and neglect the colour meaning of the word. Thus it examines the ability to inhibit an automatic or predisposed response to certain stimuli in favour of a deliberate and unusual response (Shiffrin & Schneider, 1977). Results of incongruent trials showed a decrease in completion time of the task, while errors increased from before to after the exercise. However, these results were only measures comparing baseline testing in the morning to post-testing after the final assessment of the scenario. Participants received a morning session of tactical training and then had lunch, before participating in the final assessment scenario. The last part did not yield any effects on Stroop performance.

5.2.1.1.3 Effects on Digit Span Task, Letter Number Sequencing Task, Symbol Search Task and Trail Making Test

In another study A. P. J. Roberts (2012) also used the Digit Span Task, the Letter Number Sequencing Task, the Symbol Search Task and the Trail Making Task in order to investigate the effects of an active shooter scenario on phonological loop processing.

The Letter Number Sequencing Task requires individuals to repeat back stimuli strings comprised of a mix of letters and numbers, which requires executive processing for completion (Larrabee & Kane, 1986; Rudel & Denckla, 1974). The Symbol Search Task requires participants to examine visually presented patterns, involving visuo-spatial but little executive processing (Thomason et al., 2008). The Trail Making Test (C. Armstrong,
Allen, Donohue, & Mayfield, 2008) is administered in two parts, where participants are required to connect a series of circles (a) with numbers in an ascending order (part A) or (b) by alternating between a numerical and alphabetical order (part B). Part A is considered a visual attention task which involves conceptual and tracking abilities, whereas part B more accurately measures executive functioning due to its complexity and greater reliance on WM to simultaneously remember what numbers and letters have passed in order to successfully decide which circle is next to be connected in the sequence (J. C. Davis, Marra, Najafzadeh, & Liu-Ambrose, 2010; Hedden & Yoon, 2006).

The results showed that Digits Span Task performance decreased from before to after the scenario. Furthermore, time taken to complete part B of the Trail Making Test increased, the number of digits recalled in the Letter Number Sequencing Task and completion time for part A of the Trail Making Test decreased. Completion time for the Symbol Search Task increased from before to after the scenario.

5.2.1.1.4 Effects of Corsi Block-Tapping Task and Self Ordering Pointing Task

A. P. J. Roberts (2012) used the Corsi Block-Tapping Task (as employed in Study 1.3 in Chapter 4) and the Self Ordering Pointing Task (Petrides & Milner, 1982; T. P. Ross et al., 2007) to assess the effects of an active shooter scenario primarily on visuo-spatial processing. The latter requires participants to point to one item out of a set and then to point at a different item at a different spatial location until the opportunity to point to every item has passed. The completion of the Self Ordering Pointing Task requires visuo-spatial sketchpad processing in combination with executive cognitive processing, which is suggested to be greater than during the Corsi Block-Tapping Task (Petrides & Milner, 1982). This is mainly due to the need to internally initiate and execute a sequence of responses, while maintaining record of these responses. The findings showed, that block span and total span of the Corsi Block-Tapping Task increased from before to after the active shooter scenario. In contrast, the sum of error scores of the Self
Ordering Pointing task decreased significantly. A. P. J. Roberts (2012) concluded that “visuo-spatial functioning of authorized firearms officers (AFO) increased both when the task involves little executive demand and when the task involves high executive demand” (p. 166).

5.2.1.2 Subsequent Work of Boulton (2014) on Use of Force Situations and Different Components of WM and the Influence of Experience

Boulton (2014) building on A. P. J. Robert’s work (2012) investigated the effects of simulated use of force incidents on the different aspects of executive functioning: (a) updating, (b) manipulation and (c) inhibition. She furthermore explored the effects of experiences on these components of executive functioning.

In order to assess the ability of the the updating component of executive functioning before and after a tactical scenario assessment of authorized firearms officers in the UK, Boulton (2014) employed the Adaptive Digit Ordering Task (J. A. Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991), the Digit Span Task with its two subsidiary tasks: DF and DB, the Trail Making Test and the Stroop Task.

The Adaptive Digit Ordering Task requires the participant to repeat a sequence of random number, that are read aloud by the experimenter, in ascending numerical order. The task is considered to be a more accurate measure of the central executive function of WM than the Digit Span Task, since participants have to manipulate information to complete the task successfully (Sair, Welsh-Bohmer, Wagner, & Steffens, 2006). Finally the Stroop Task as a common measure for assessing the executive component of inhibition (Archibald & Kerns, 1999; Miyake et al., 2000) was employed.

The findings of Boulton (2014) showed that performance on the Digit Span Task (DF and DB) and the Adaptive Digit Ordering Task (both employed as a marker of the updating function of the central executive) significantly decreased during a scenario that involved the pursuit of a car with the subsequent apprehension of the suspects. On average these scenarios lasted approximately 5 minutes. In a second scenario where authorized
firearms officers were required to search a building in order to locate, secure, and subsequently call out any suspect to be present inside and to affect their arrest (average scenario time: 45 minutes), no changes in the Trail Making Test, as a measure of the shifting component of executive functioning were found. The third scenario involved the tactical approach, entering, and the dynamic search of a building in which a suspect is believed to have already injured members of the public (average scenario time: 30 minutes). Performance on the Stroop Task, as a marker of the inhibition function of the central executive, showed that time measures of the Stroop Task were significantly decreased and error rates were significantly increased after the scenario compared to before the scenario. This may indicate a speed-accuracy trade-off (McMorris, Sproule, Turner, & Hale, 2011).

In all, these findings demonstrated a decrease in Digit Span Tasks scores in scenario 1, a faster completion of part A of the Trial Making test, but a slower completion of part B in scenario 2 and a faster, but more error prone Stroop completion in scenario 3. According to (Boulton, 2014) these findings indicate that the updating (scenario 1) and inhibition (scenario 3) processes of executive functioning (J. C. Davis et al., 2010; Hedden & Yoon, 2006; Miyake et al., 2000) are significantly reduced in authorized firearms officers after the scenarios compared to before. Regarding the influence of experience, Boulton (2014) concludes that “rather than a difference in the functioning of certain neuropsychological processes, expertise in this context may reflect interactions related to the control of behavioural adaptation and cognitive flexibility” (p. 186). Yet, it has to be acknowledged that the notion of expertise in this context yield different interpretations. While the study employed a longitudinal design during the period of a firearms qualification course, this may not be enough to elicit expertise in any developed skills. The identification of experts and the subsequent capturing of expert performance in order to gain insights about expertise (Causer, Barach, & Williams, 2014; Ericsson, 2003; 2013; A. M. Williams & Ericsson, 2005), seems to be a difficult endeavour in other domains as well (Abraham, Collins, & Martindale, 2006; Gobet, 2016; Nash, Martindale, Collins, & Martindale, 2012): While there are domains, where expertise can by seen by experts clearly
outperforming non-experts (e.g. in sports, chess), there are fields, like expertise in armed conflicts, where reliable measures of expertise are difficult to find.

5.2.1.3 **Work of Condon (2015) on Physical Combat and Executive Functioning**

Condon (2015) investigated effects of two different simulated physical combat exercises on executive cognitive functioning of police recruits using the Dual 2-Back Continuous Performance Task (M. J. Kane et al., 2007; M. J. Kane & Engle, 2002). It requires participants to monitor a sequence of items (audio and visual cues) and responding when presented item matches the one that appeared 2 items ago. Participants have to continuously monitor the information being presented and update its representation with each presentation of new stimuli. It is suggested that Dual 2-Back Task performance is a key indicator of executive functioning such as attentional control, inhibition, updating, and task shifting (Conway et al., 2005; Jaeggi et al., 2008; M. J. Kane et al., 2007; M. J. Kane & Engle, 2002). The exercises in the study of Condon (2015) differed with regards to the task at hand and the additional load that was placed on recruits: the first exercise required participants to exhibit appropriate combat skills by reacting to cues presented by a use of force instructor, acting as a suspect, who continuously increases and decreases aggressive behaviour. Participants had to complete three rounds of one minute in a confined area. The second exercise required police officers to demonstrate tactical behaviour in a scenario with multiple environmental variables such as loud music, multiple actors, low lighting, and a belligerent, uncooperative suspect displaying violent behaviour. Due to inconsistencies in data reporting (scores change between tables, figures and text), findings have to be interpreted in a conservative way. For the scenario Condon (2015) reported a decrease in one component of executive processing in combinations with phonological loop processing. Executive tasks needing the visuo-spatial sketchpad for completion were not affected by the scenario.
5.2.1.4 Work of Taverniers et al. (2011) on the Effects of representative Testing Environments in Police Use of Force Training

Taverniers and colleagues (2011) investigated the effects of two different kinds of scenario activities on executive functioning, as measured by the Digit Span Task with DB. Two randomly assigned groups completed a scenario training, where participating police officers were required to search a house for an armed offender with a history of violence. The experimental group was equipped with NLT weapons and told they will maybe be shot at. In contrast, the control group was briefed that they will only be assessed on their tactical behaviour, while being equipped with inert weapons. Furthermore, they would not be shot at. The results showed that DB of the control group was significantly higher than the experimental group. The differences of the groups could be explained by the level of representative design of the scenario with regards to the risk of being shot at. However, the study only measured post-treatment executive functioning.

5.2.1.5 Summary of Findings and Rationale for Study 2

The outlined findings provide an initial overview over different use of force training activities on the effects of executive functioning (see Table 13). The displayed studies differ with regards to (a) the employed test of executive functioning, (b) the operational task (scenario / exercise) that had to be completed by the participating officers, (c) the duration of the operational task and (d) the skill level of police officers (participants). Furthermore (e) the time, when executive functioning was assessed varied within the studies. In some studies (A. P. J. Roberts, 2012) baseline measures were taken in the morning, before the training started. In these studies participants had lunch before returning for the final assessment, where executive functioning was measured before and after the operational task. Therefore, changes of executive functioning can be seen from baseline to pre-testing (effect of training and lunch), from baseline to post-testing (effects of training, lunch and assessment) and from pre- to posttesting (effects of assessment). In all other studies employing a pre-/post-design, executive functioning was only measured before and after the operational task.
Preliminary interpretations of the results suggest, that executive functioning processing is affected by use of force training and simulated scenarios, and also differs with regards to the task at hand. Furthermore, performance on psychological tests that additionally rely on phonological loop processing seem to decrease if the task at hand (scenario/exercise) relies heavily on visuo-spatial processing. Likewise, performance on tests requiring visuo-spatial processing increases in such situations. Some authors (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012) suggest that executive processing is facilitated due to the task at hand, representing an adaptive response to cope with the demand upon, or threat being posed by the role players to, participants.

However, some alternative possible explanations have to be considered. For example, studies from the sports domain suggest that physical activation may have an impact on executive functioning (Ashryn & Alison, 2015; Y.-K. Chang, Tsai, Hung, So, & Chen, 2011; Cheng, Donoghue, Gilfoyle, & Eppich, 2012; A. Diamond, 2015; Etnier & Chang, 2009; Hsieh, Chang, Hung, & Fang, 2016; Hüttermann & Memmert, 2014; Soga, Shishido, & Nagatomi, 2015; Tsukamoto et al., 2016; Weng, Pierce, Darling, & Voss, 2015). Since: (a) police officers in the described studies (see Table 13) always showed higher increased cardiovascular activity, (b) studies were conducted without control groups (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012), and (c) without the employment of a pre-post design (Taverniers et al., 2011), it may be possible that the effects on executive functioning are mediated by the psychophysiological load of participants.

5.2.2 Aim of Study 2

The current study aims at investigating whether changes in executive functioning performance (as measured by the Phonemic Fluency Task) occur as a result of tactical training or physical activity. Hence, the current study employs a between-subjects design of two groups, both having experience in physical combat and scenario training: a group of police officers performing in a demanding scenario training, where tactical decisions have to be made in order to successfully complete the scenario, compared to a control group
of martial artists, who were required to physically exercise without any additional cognitive demand for the time-period.
<table>
<thead>
<tr>
<th>Author</th>
<th>Test employed</th>
<th>Scenario / exercise</th>
<th>Duration</th>
<th>Participants</th>
<th>Post HR</th>
<th>Measure</th>
<th>Effect of EF</th>
<th>Additional component of WM</th>
<th>Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taverniers et al. (2011)</td>
<td>Digit Span Task</td>
<td>Building search; apprehension of armed suspect</td>
<td>n.a.</td>
<td>19 officers (experimental) vs. 17 officers (control)</td>
<td>n.a.</td>
<td>DB</td>
<td></td>
<td></td>
<td>Experimental: &quot;might be shot at&quot;, use of NLT weapons; control: &quot;will not be shot at&quot;; use of NLT weapons</td>
</tr>
<tr>
<td>Roberts (2012)</td>
<td>Phonemic Fluency Task</td>
<td>Active shooter incident; dynamic building search</td>
<td>45 min</td>
<td>25 AFOs</td>
<td></td>
<td>$M = 107.00, SD = 15.36$</td>
<td>TW</td>
<td></td>
<td>Participants received tactical training in the morning; after lunch final assessment took place</td>
</tr>
<tr>
<td>Roberts (2012)</td>
<td>Semantic Fluency Task</td>
<td>Dynamic car stop; apprehension</td>
<td>10 min</td>
<td>28 AFOs</td>
<td></td>
<td>$M = 93.21, SD = 22.17$</td>
<td>TW</td>
<td></td>
<td>No information about the training day</td>
</tr>
<tr>
<td>Roberts (2012)</td>
<td>Semantic Fluency Task</td>
<td>Slow building search; apprehension of suspect</td>
<td>45 min</td>
<td>25 AFOs</td>
<td></td>
<td>$M = 88.64, SD = 15.30$</td>
<td>TW</td>
<td></td>
<td>No information about the training day</td>
</tr>
<tr>
<td>Roberts (2012)</td>
<td>Stroop Task</td>
<td>LF exercise</td>
<td>20 min</td>
<td>20 AFOs</td>
<td></td>
<td>$M = 108.55, SD = 24.84$</td>
<td></td>
<td></td>
<td>No information about the training day</td>
</tr>
<tr>
<td>Roberts (2012)</td>
<td>Digit Span Task</td>
<td>Active shooter incident; dynamic building search</td>
<td>45 min</td>
<td>29 AFOs</td>
<td></td>
<td>$M = 118.38, SD = 16.44$</td>
<td>DF</td>
<td></td>
<td>No information about the training day</td>
</tr>
<tr>
<td>Roberts (2012)</td>
<td>Symbol Search Task</td>
<td>Completion time</td>
<td></td>
<td></td>
<td></td>
<td>Completion time</td>
<td></td>
<td></td>
<td>VSS</td>
</tr>
<tr>
<td>Roberts (2012)</td>
<td>Trail Making Test</td>
<td>Active shooter incident; dynamic</td>
<td>45 min</td>
<td>28 AFOs</td>
<td></td>
<td>$M = 118.71, SD = 16.61$</td>
<td></td>
<td></td>
<td>PL</td>
</tr>
<tr>
<td>Task</td>
<td>Description</td>
<td>Duration</td>
<td>Participants (AFOs)</td>
<td>Mean (M)</td>
<td>Standard Deviation (SD)</td>
<td>Number of recalled digits</td>
<td>Block span △</td>
<td>Total span △</td>
<td>VSS</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------</td>
<td>---------------------</td>
<td>----------</td>
<td>-------------------------</td>
<td>--------------------------</td>
<td>----------</td>
<td>----------</td>
<td>-----</td>
</tr>
<tr>
<td><strong>Roberts (2012)</strong></td>
<td>Corsi Block-Tapping Task</td>
<td>45 min</td>
<td>29</td>
<td>M = 111.92, SD = 18.37</td>
<td>△</td>
<td>△</td>
<td>PL</td>
<td>VSS</td>
<td>VSS</td>
</tr>
<tr>
<td><strong>Roberts (2012)</strong></td>
<td>Self Ordering Pointing Task</td>
<td>45 min</td>
<td>28</td>
<td>M = 98.32, SD = 25.29</td>
<td>△</td>
<td>△</td>
<td>VSS</td>
<td></td>
<td>VSS</td>
</tr>
<tr>
<td><strong>Boulton (2014)</strong></td>
<td>Digit Span Task</td>
<td>5 min</td>
<td>10</td>
<td>M = 92.50, SD = 22.09</td>
<td>DF</td>
<td>△</td>
<td>VSS</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boulton (2014)</strong></td>
<td>Digit Span Task</td>
<td>5 min</td>
<td>10</td>
<td>M = 91.40, SD = 9.34</td>
<td>DF</td>
<td>△</td>
<td>PL</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boulton (2014)</strong></td>
<td>Trail Making Test</td>
<td>45 min</td>
<td>12</td>
<td>M = 101.08, SD = 24.26</td>
<td>△</td>
<td>△</td>
<td>VSS</td>
<td>PL</td>
<td></td>
</tr>
<tr>
<td><strong>Boulton (2014)</strong></td>
<td>Trail Making Test</td>
<td>45 min</td>
<td>12</td>
<td>M = 89.29, SD = 17.81</td>
<td>△</td>
<td>-</td>
<td>VSS</td>
<td>PL</td>
<td></td>
</tr>
<tr>
<td><strong>Boulton (2014)</strong></td>
<td>Stroop Task</td>
<td>30 min</td>
<td>17</td>
<td>M = 95.53, SD = 18.10</td>
<td>△</td>
<td>△</td>
<td>speed accuracy trade-off</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boulton (2014)</strong></td>
<td>Stroop Task</td>
<td>30 min</td>
<td>17</td>
<td>M = 90.06, SD = 15.41</td>
<td>△</td>
<td>△</td>
<td>speed accuracy trade-off</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boulton (2014)</strong></td>
<td>Stroop Task</td>
<td>30 min</td>
<td>17</td>
<td>M = 90.06, SD = 15.41</td>
<td>△</td>
<td>△</td>
<td>speed accuracy trade-off</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Task</td>
<td>Description</td>
<td>Participants</td>
<td>M</td>
<td>SD</td>
<td>Visual correct</td>
<td>Visual wrong</td>
<td>Audio correct</td>
<td>Audio wrong</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>-------------</td>
<td>--------------</td>
<td>----</td>
<td>----</td>
<td>----------------</td>
<td>--------------</td>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Condon (2015)</td>
<td>Dual 2-Back Continuous Performance Task</td>
<td>Displaying combat skills by reacting to cues presented by a use of force instructor</td>
<td>19 Recruits</td>
<td>124.48</td>
<td>22.70</td>
<td>-</td>
<td>-</td>
<td>▼</td>
<td>-</td>
</tr>
<tr>
<td>Condon (2015)</td>
<td>Dual 2-Back Continuous Performance Task</td>
<td>Scenario, where participants had to display tactical behavior against a resisting, violent suspect</td>
<td>n.a. 14 Recruits</td>
<td>97.74</td>
<td>33.39</td>
<td>-</td>
<td>-</td>
<td>▼</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. A = Assessment, AFO = Authorized firearms officers, EF = Executive functioning, L = Lunch (between T and A), PL = Phonological loop, T = Training, VSS = Visuo-spatial sketchpad, △ = increase, ▼ = decrease, - = no difference.
5.3 Hypotheses

In line with previous research, linking effects on executive functioning to the task at hand, it is hypothesised that participants will show different changes on executive functioning in relation to the required task. To this end, the following experimental hypotheses were generated:

1. The demands of the scenario and physical exercise will result in participants' experiencing physiological arousal (Hypothesis 1).

2. The cognitive demands of the tactical scenario training will result in higher measures of executive functioning than the non-cognitive demands of the physical exercise. (Hypothesis 2)

5.4 Methods

The study employed a mixed-design with two independent variables: group as a between-subjects factor (scenario group vs. exercise group) and time as the repeated-measures factor (pre- and post-test measures). The dependent variables were the various measures of phonemic fluency and measures of HR, blood pressure, and core temperature.

5.4.1 Participants

A total of 73 participants took part in the study. The scenario group (scenario) consisted of 41 police officers from the state of Hesse. The exercise group (exercise) consisted of 32 practitioners of a German krav maga club, with 4 participants working as police officers. To protect the anonymity of police officers, no other identifying demographics were recorded.

One case of missing data was identified in regard to Phonemic Fluency Task scores. The case was excluded from further analysis resulting in a total sample size of 72 for the analysis of Phonemic Fluency Task scores ($n_{scenario} = 41$, $n_{exercise} = 31$).

Concerning measures of physiological arousal, 7 cases of missing data were identified and excluded from further analysis. Initial evaluation of equality of covariance matrices using Box’s test revealed a violation of this assumption.
Since robustness of MANOVA cannot be assumed when sample sizes are unequal (A. Field, 2013), two additional cases of the data set were randomly excluded resulting in a final sample size of 64 ($n_{\text{scenario}} = 32, n_{\text{exercise}} = 32$).

5.4.2 Materials and Apparatus

5.4.2.1 Phonemic Fluency as a Measure of Executive Functioning of WM

In the current study executive functioning of WM was measured with the Phonemic Fluency Task of the RWT (Aschenbrenner et al., 2000), which is the German complement of the Controlled Order Word Association Test (Benton & Hamsher, 1989). Further details of this test are provided in the General methods chapter (see Chapter 3).

5.4.2.2 Physiological Arousal

The degree to which the scenario influenced the physiological arousal of the participants was measured using HR, blood pressure and core temperature. These parameters were measured before and after with several electronic devices. HR and blood pressure were measured via a blood pressure wrist unit (HEM-650). Core temperature was measured using an infrared thermometer (Braun ThermoScan Ear Thermometer).

5.4.2.3 Safety Equipment

During the scenario, participants were equipped with safety gear (PPE, elbow- and knee pads, Simunition helmet) and a 9mm handgun, identical to their duty weapon (Heckler & Koch P30) and specially prepared to fit coloured soap cartridges (Simunition, FX Marking Ammunition). The exercise group did not wear any safety equipment.

5.4.3 Procedure

The general procedure for the current study was identical for the experimental and the control group, except for the treatment and the location of the experiment.
5.4.3.1 Treatment for Scenario and Exercise Group

Testing for the scenario group took place at the training facility of the Hessian Police Academy in Wiesbaden, Germany, whereas the control group was tested at the gym of a local krav maga club in Frankfurt, Germany.

The treatments differed as follows: In the scenario condition, police officers were paired up, equipped and briefed about the upcoming scenario. Their orders were to rescue a shot officer in the first floor of an apartment. In order to approach the officer, participants had to climb up a ladder and enter the apartment through a window. Inside the apartment, the participants had to enter two rooms before they could see the bleeding officer lying on the stairs. When the participants approached the officer, the perpetrator started to shoot at the officer from the ground level. The officers had to eliminate the threat in order to administer first aid to their fellow officer. The scenario was finished when the participants shot the perpetrator. Directly after the final shootout, the participants were led to a separate room in the immediate vicinity, where post-testing measures took place. In the exercise condition, participants were tested individually. They were required to perform a running-on-the-spot exercise, mimicking the duration and intensity of exercise that police officers may be exposed to during a critical incident (A. P. J. Roberts & Cole, 2013). The exercise lasted 5 minutes, consisting of a 1-minute jog, 1-minute sprint bringing their knees up in front of them, 30-seconds jog, followed by a further minute of sprinting bringing their knees up in front of them, and then another 30-second jog followed by a final 1-minute sprint bringing their knees up in front of them. Immediately after finishing the final 1-minute sprint, participants were asked to sit down for post-testing measures.

5.4.3.2 General Procedure

All participants were tested individually in separate rooms on physiological parameters. HR and blood pressure were recorded in a standing position on the left arm. Simultaneously, core temperature was measured non-invasively via the ear canal. Directly afterwards the phonemic fluency test with the letters M and B was administered. Then, participants engaged in the treatment. Immediately after the scenario or the exercise, respectively, they
were tested again individually in separate rooms on the physiological parameters, the same way as before the scenario, followed by the Phonemic Fluency Task (letters: P and K). On average, the time between the ending of the scenario and the testing was 30 to 60 seconds, depending on the location of the shootout, and how fast the safety officer of the scenario training could ensure safety of the Simunition weapons. After participants had finished the second testing, they were debriefed in a separate room to share their experiences.

5.4.4 Analytic Strategy

Differences on phonemic fluency scores between groups were analysed by using a mixed MANOVA (Group x Time) with repeated measures on the second factor. The dependent variables (DV) were TW, MCS and NS. The different groups formed the between-subjects independent variable (IV): the scenario and the exercise group. The within-subjects IV treated multivariately was the two sessions of the Phonemic Fluency Task: before (pre) and after (post) the scenario. Since assumptions for multivariate normality and homogeneity of covariance matrices were not met, a log transformation was applied on the data. Results of evaluation of assumptions of mixed MANOVA were satisfactory after transformation of the data. For the ease of interpretation, untransformed means are reported. Follow up mixed ANOVAs were used to examine the effects and interactions of status and time on each DV. A Bonferroni correction ($\alpha = .017$) was applied to significant results to control for type I statistical errors. Bonferroni corrected post hoc comparisons were used to determine the source of significant effects.

To check if the arousal increased during the scenarios, differences in physiological parameters (HR, systolic and diastolic blood pressure, core temperature) from before to after the scenario were tested using a mixed MANOVA (status x time) with repeated measures on the second factor. The DVs were HR, systolic blood pressure, diastolic blood pressure, and core temperature. Results of evaluation of assumptions of mixed MANOVA were satisfactory. Follow up mixed ANOVAs were used to examine the effects and interactions of status and time on each physiological component. A
Bonferroni correction ($\alpha = .013$) was applied to significant results to control for type I statistical errors.

5.5 Results

5.5.1 Sample Characteristics

A Shapiro-Wilk's test (Razali & Wah, 2011; S. S. Shapiro & Wilk, 1965) and a visual inspection of their histograms, normal Quantile-Quantile plots and box plots was performed on each data set (Doane & Seward, 2011). Results showed that the phonemic fluency scores were not normally distributed. In order to allow for parametric analysis, a square root transformation was performed on the data. The data set was then re-examined. Results revealed that assumptions for parametric testing were met allowing for using a mixed MANOVA on the data (transformed data not shown).

Data of physiological arousal (systolic and diastolic blood pressure, HR, core temperature) were normally distributed.

5.5.2 Phonemic Fluency Task

Means, standard deviations, and 95% confidence intervals around the mean of the three phonemic fluency scores (TW, MCS, NS) over all combinations of condition and time are displayed in Table 14.

A 2 (group: scenario versus exercise) x 2 (time: pre versus post) mixed MANOVA on 3 DVs (TW, MCS, NS) was performed. Pillai’s trace yielded one statistically significant within subjects effect for time, $V = 0.28$, $F(3, 68) = 9.00$, $p < .001$, partial $\eta^2 = .284$. There was not a statistically significant between subjects effects for group, $V = 0.09$, $F(3, 68) = 2.14$, $p = .104$, partial $\eta^2 = .086$, and for Group x Time interaction, $V = 0.02$, $F(3, 68) = 0.55$, $p = .65$, partial $\eta^2 = .024$.

In order to follow up the significant within-subjects effect of time, three follow up mixed ANOVAs, one for each DV, were calculated. There was a significant main effect of time on TW, $F(1, 70) = 22.49$, $p < .001$, $r = .49$, indicating that TW score increased significantly from pre- to post-testing. A significant main effect of time on NS was observed, $F(1, 70) = 21.80$, $p <$
.001, $r = .49$, indicating that NS scores increased significantly from pre- to post-testing. There was no significant main effect on time on MCS, indicating that MCS scores were similar before and after the treatment, $F(1, 70) = 0.61, p = .437, r = .09$.

**Table 14: Descriptive Statistics of Phonemic Fluency Scores**

<table>
<thead>
<tr>
<th></th>
<th>Scenario</th>
<th>Pre</th>
<th>41</th>
<th>38.15</th>
<th>10.37</th>
<th>34.87</th>
<th>41.42</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>41</td>
<td>43.30</td>
<td>10.53</td>
<td>39.98</td>
<td>46.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>41</td>
<td>5.15</td>
<td>8.26</td>
<td>2.55</td>
<td>7.76</td>
</tr>
<tr>
<td></td>
<td>Exercise</td>
<td>Pre</td>
<td>31</td>
<td>33.45</td>
<td>10.25</td>
<td>29.69</td>
<td>37.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>31</td>
<td>36.61</td>
<td>10.27</td>
<td>32.85</td>
<td>40.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>31</td>
<td>3.16</td>
<td>6.11</td>
<td>0.92</td>
<td>5.40</td>
</tr>
<tr>
<td></td>
<td>Scenario</td>
<td>Pre</td>
<td>41</td>
<td>0.37</td>
<td>0.19</td>
<td>0.32</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>41</td>
<td>0.35</td>
<td>0.17</td>
<td>0.30</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>41</td>
<td>-0.02</td>
<td>0.19</td>
<td>-0.08</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Exercise</td>
<td>Pre</td>
<td>31</td>
<td>0.36</td>
<td>0.21</td>
<td>0.28</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>31</td>
<td>0.34</td>
<td>0.04</td>
<td>0.25</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>31</td>
<td>0.01</td>
<td>0.25</td>
<td>-0.09</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Scenario</td>
<td>Pre</td>
<td>41</td>
<td>28.90</td>
<td>8.53</td>
<td>26.21</td>
<td>31.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>41</td>
<td>33.79</td>
<td>9.65</td>
<td>30.74</td>
<td>36.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>41</td>
<td>4.89</td>
<td>6.53</td>
<td>2.83</td>
<td>6.95</td>
</tr>
<tr>
<td></td>
<td>Exercise</td>
<td>Pre</td>
<td>31</td>
<td>26.23</td>
<td>9.07</td>
<td>22.90</td>
<td>29.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>31</td>
<td>28.87</td>
<td>8.78</td>
<td>25.65</td>
<td>32.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>31</td>
<td>2.65</td>
<td>7.08</td>
<td>0.05</td>
<td>5.24</td>
</tr>
</tbody>
</table>

Figure 24 displays the significant increase of TW and NS scores for the whole sample (and the two groups).
Figure 24: Means and 95% confidence intervals of Phonemic Fluency Task scores of Study 2. *** indicates significance at $p < .001$ from pre-to posttesting.
5.5.3 Physiological Arousal

Means and standard deviations of systolic and diastolic blood pressure, HR, and core temperature over all combinations of group and time are displayed in Table 15.

Table 15: Descriptive Statistics of Physiological Arousal Measures

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LL</td>
</tr>
<tr>
<td><strong>Systolic blood pressure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>127.44</td>
<td>9.16</td>
<td>124.14</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>138.38</td>
<td>11.85</td>
<td>134.10</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>10.94</td>
<td>12.72</td>
<td>6.35</td>
</tr>
<tr>
<td>Exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>136.91</td>
<td>15.91</td>
<td>131.91</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>179.97</td>
<td>25.36</td>
<td>170.83</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>43.06</td>
<td>18.73</td>
<td>36.31</td>
</tr>
<tr>
<td><strong>Diastolic blood pressure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>81.91</td>
<td>9.90</td>
<td>78.34</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>88.66</td>
<td>10.94</td>
<td>84.71</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>6.75</td>
<td>12.96</td>
<td>2.08</td>
</tr>
<tr>
<td>Exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>84.47</td>
<td>9.70</td>
<td>80.97</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>98.16</td>
<td>14.00</td>
<td>93.11</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>13.69</td>
<td>10.44</td>
<td>9.92</td>
</tr>
<tr>
<td><strong>HR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>82.03</td>
<td>13.05</td>
<td>77.32</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>107.13</td>
<td>17.55</td>
<td>100.83</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>25.09</td>
<td>14.59</td>
<td>19.83</td>
</tr>
<tr>
<td>Exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>73.94</td>
<td>8.51</td>
<td>70.87</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>131.38</td>
<td>18.24</td>
<td>124.80</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>57.44</td>
<td>18.43</td>
<td>50.79</td>
</tr>
<tr>
<td><strong>Core temperature</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>36.66</td>
<td>0.53</td>
<td>36.47</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>36.44</td>
<td>0.44</td>
<td>36.28</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>-0.22</td>
<td>0.43</td>
<td>-0.38</td>
</tr>
<tr>
<td>Exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>32</td>
<td>36.89</td>
<td>0.37</td>
<td>36.76</td>
</tr>
<tr>
<td>Post</td>
<td>32</td>
<td>36.62</td>
<td>0.42</td>
<td>36.47</td>
</tr>
<tr>
<td>Diff</td>
<td>32</td>
<td>-0.27</td>
<td>0.24</td>
<td>-0.36</td>
</tr>
</tbody>
</table>

A 2 (group: scenario versus exercise) x 2 (time: pre versus post) mixed MANOVA on 4 DVs (systolic blood pressure, diastolic blood pressure, HR, and core temperature) was performed.
Pillai’s trace yielded one statistically significant within subjects effect for time, $V = 0.91$, $F(4, 59) = 141.22$, $p < 0.001$, partial $\eta^2 = .905$, for group, $V = 0.50$, $F(4, 59) = 14.76$, $p < 0.001$, partial $\eta^2 = .500$, and for Group x Time interaction, $V = 0.639$, $F(4, 59) = 26.07$, $p < 0.001$, partial $\eta^2 = .639$.

In order to follow up the significant effects of time, condition and time x condition interaction, separate univariate mixed ANOVAs on each DV were performed.

### 5.5.3.1 Systolic Blood Pressure

There was a significant main effect of time on systolic blood pressure, indicating an increase of systolic blood pressure from pre- to post-testing, $F(1, 62) = 182.02$, $p < .001$, $r = .86$. The results yielded a significant main effect of group on systolic blood pressure, which indicates that the scenario group had lower systolic blood pressure than the exercise group, $F(1, 62) = 48.30$, $p < .001$, $r = .66$. There was a significant interaction effect between time and group on systolic blood pressure indicating that the change of systolic blood pressure differed according to the group, $F(1, 62) = 64.42$, $p < .001$, $r = .59$.

![Figure 25: Means and 95% confidence intervals of systolic blood pressure of Study 2.](image)
5.5.3.2 Diastolic Blood Pressure

There was also a significant main effect of time on diastolic blood pressure, which indicates that diastolic blood pressure was significantly higher after the treatment than before, $F(1, 62) = 48.24, p < .001, r = .66$. There was also a significant main effect of group on diastolic blood pressure, indicating that lower diastolic blood pressure was measured in the scenario group than in the exercise group, $F(1, 62) = 6.30, p = .015, r = .30$. There was a significant Group x Time interaction effect on diastolic blood pressure, $F(1, 62) = 5.56, p = .022, r = .29$. This suggests, that the change of diastolic blood pressure from pre- to post-testing differs according to the group (scenario vs. exercise).

![Figure 26: Means and 95% confidence intervals of diastolic blood pressure of Study 2.](image)

5.5.3.3 Heart Rate

The main effect of time on HR was significant, $F(1, 62) = 394.48, p < .001, r = .93$. This indicates that HR increased from pre- to post-testing. There was also a significant main effect of group on HR, indicating that groups differ concerning the measured HR, $F(1, 62) = 6.92, p = .011, r = .32$. Results also yielded a significant Group x Time interaction on HR, indicating that the increase in HR differs between the groups, $F(1, 62) = 60.59, p < .001, r = .70$. 

186
5.5.3.4 Core Temperature

There was a significant main effect of time on core temperature, indicating that this variable also decreased during the treatment, $F(1, 62) = 32.67, p < .001, r = .59$. There was no significant main effect of group on core temperature, indicating that this variable does not differ between the groups, $F(1, 62) = 4.00, p = .050, r = .25$. There was no significant interaction effect of group and time on core temperature, $F(1, 62) = 0.30, p = .589, r = .07$. This indicates, that the change in core temperature does not differ between the scenario and the exercise group.

Figure 27: Means and 95% confidence intervals of mean HR of Study 2

Figure 28: Means and 95% confidence intervals of core temperature of Study 2. *** indicates significance at $p < .001$ from pre- to posttesting.
5.6 Discussion

Study 2 aimed at investigating the effects of a cognitively demanding training scenario on executive functioning compared to physical exercise. It was hypothesised that measures of physiological arousal would increase from pre- to post-testing in both groups (Hypothesis 1) and that the increase in executive functioning would be higher in the scenario group, due to the cognitive demands of the associated tactical decision-making in the scenario (Hypotheses 2). The results confirmed Hypothesis 1, since measures of physiological arousal increased significantly in both groups. Hypothesis 2 was disconfirmed in the current study: even though executive functioning was enhanced after the treatment in both groups, there was no observed group effect on TW and NS, indicating that the two types of treatment did not differ in their effect on executive functioning. Additionally, the obtained group x time interaction effect on mean HR, systolic and diastolic blood pressure indicated that the 5 minutes running-on-the-spot exercise was more physiologically arousing than the scenario.

5.6.1 Physiological Arousal Under Physical and Tactical Demand

As a centrally mediated measure, HR is thought to be a sensitive indicator of cognitive and/or emotional demand (Myrtek, 2004). In both groups HR increased significantly due to the treatment indicating that both the scenario and exercise are physiologically arousing experiences. During the scenario, participants were required to move around the training environment, which represents physical load (Hillman et al., 2002; 2004; Hillman, Snook, & Jerome, 2003; Tomporowski, 2003). Furthermore, they had to make appropriate tactical decisions based on recollections of previously learnt skills in combination with creating novel solutions to new problems encountered in that situation, while inhibiting inappropriate responses and maintaining cognitive focus, which represents cognitive load (Critchley, Melmed, Featherstone, Mathias, & Dolan, 2001b; Hillman et al., 2002; 2004; Tomporowski, 2003). Finally, termed as emotional load, they had to control their levels of anxiety, fear, and aggression (Charmandari et al., 2005; Chrousos, 1992a; Critchley et al., 2001b; A. P. J. Roberts, 2012; E. M. Roth
et al., 2010). HR increases in the scenario group could represent any of these physical, cognitive, and affective factors individually or collectively. Since it is not possible to separate the influence of these factors with the current data, the identification of causal sequences in these circumstances is difficult.

Concerning temperature, several studies revealed that psychological demands affect core temperature, leading to a rise in this parameter after watching films (Kleitmann, 1945), immediately before boxing contests (Renbourn, 1960) or examinations (Briese, 1995; Marazziti, Di Muro, & Castrogiovanni, 1992). Results from A. P. J. Roberts (2012) showed that psycho-physiological demands during tactical training lead to a significant increase in temperature. Furthermore, an increased heat load resulting from inadequate heat dissipation in a warm environment or an overproduction of body heat, can cause a rise in temperature (T. Oka, Oka, & Hori, 2001). The results of the current study are inconsistent with these findings, showing that core temperature drops significantly during the scenario. A possible explanation is provided by an anticipatory effect of the scenario. It could be possible that the participants experienced an increased level of anxiety by thinking about the upcoming scenario. This "anticipatory anxiety stress" can lead to a rise in temperature (T. Oka et al., 2001), flattening after the completion of the scenario. Nevertheless, the decrease of core temperature seems inconsistent with the results of heart rate and systolic blood pressure. Future studies would benefit from additionally measuring core temperature at the beginning of the day of the experiment. Doing so would possibly provide an explanation for the decrease of core temperature during the scenario (e.g. anticipatory anxiety stress) could be provided more easily.

Likewise, the changes in blood pressure measures are consistent with previous studies in simulated use of force situations (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012). However, inconsistencies as observed in in studies by Boulton (2014) and A. P. J. Roberts (2012) may reflect the larger variability and peripheral responsiveness of blood pressure measures (G. S. Anderson et al., 2002; Hjortskov et al., 2004).
The observed Group x Time interaction effect in HR and blood pressure measures indicates that the exercise induced higher levels of physiological arousal than the scenario. The implicit assumption of the scenario (and the exercise) was that this treatment induced some kind of load on the participants leading to a change in cognitive functioning. The risk of being shot in a simulated scenario is known to increase the emotional demand placed upon officers (Meyerhoff, 2004; Nieuwenhuys et al., 2011; Nieuwenhuys & Oudejans, 2010; 2011a; Oudejans, 2008; Taverniers et al., 2011; Vickers & Lewinski, 2012). The parameters of physiological arousal of the participants showed a clear rise in HR, blood pressure and core temperature, indicating that the scenario placed a certain load on the police officers. Nevertheless, it could be possible that the increase of these physiological parameters was only due to the physical demands of the scenario (e.g. climbing up a ladder to access the first floor of the building, adopting unusual shooting stances because of the stairs, and wearing PPE).

5.6.2 Executive Functioning Under Physical and Tactical Demand

Concerning the influence of switching and clustering on phonemic fluency performance, Troyer and colleagues (1997) found switching was more highly correlated with phonemic fluency than clustering. The results of the current study are consistent with this. It is suggested that the higher score in the TW of the post-test is due to the higher score in NS, since MCS did not change significantly. Since patients with frontal lobe lesions are more impaired in switching than in clustering regarding the Phonemic Fluency Task (Troyer et al., 1997; 1998b), the results indicate that the treatments (scenario and exercise) enhance executive functioning. The increase in switching score indicates an increase in executive processes, since it is related, first, to the ability to disengage from previous strategy, and second, to the ability to initiate a search for a new strategy (Troyer et al., 1997). Taking into account the underlying process of clustering, the results showed that retrieval processes were not impaired due to the scenario and exercise. This possibly indicates that the ability to retrieve stored information, such as SOPs, was not affected by the scenario. This has to be qualified by mentioning that, if all phonemic fluency scores rose because of a possible practice effect, the
static MCS score indicates a relative and selective decrease in retrieval processes.

The results showed that HR increased to a level of a mean of 107.13 bpm in the scenario and to a mean of 131.38 bpm in the exercise group. This is consistent with other research findings that these levels of physiological arousal may not produce negative effects on cognition (Levitt & Gutin, 1971; Salmela & Ndoye, 1986; Tomporowski, 2003; Tomporowski & Ellis, 1986). However, other studies found that executive functioning decreased due to tasks resulting in increases in HR (Boulton, 2014; A. P. J. Roberts & Cole, 2013). For example, A. P. J. Roberts and Cole (2013) investigating the effects of wearing body armour and exercise on executive functioning with students, employed 2 exercise conditions (light jog vs. the running-on-the-spot exercise of the current study) and the Phonemic Fluency Task. Their results showed, that both the light jog (mean HRs of 118.65 bpm) and the running-on-the-sport group (mean HRs of 168.85 bpm) showed an increase in MCS and a decrease in NS, while TW generated did not change. This indicates a cognitive shift toward non-executive processing at the cost of executive processing (A. P. J. Roberts & Cole, 2013). The results of A. P. J. Roberts and Cole (2013) seem to be inconsistent with the current findings, since the latter showed an increase in the executive functioning in the exercise group. For now, a possible explanation for the inconsistencies could lie in the differing fitness levels of participants in the studies, with the martial artists and the police officers in the current study possibly being physically fitter than the students in the study of A. P. J. Roberts and Cole (2013). Hüttermann and Memmert (2014) provided evidence that athletes with higher fitness levels were able to maintain cognitive performance on a measure of attentional behaviour at higher exercises intensities than non-athletes. The physical fitness level is suggested to play a major role in maintaining cognitive skills under high exercise intensities (Y.-K. Chang, Labban, Gapin, & Etnier, 2012) and with enhancements in cognitive functioning in general (Etnier et al., 1997; Pesce, 2009; Tomporowski & Ellis, 1986). Several research works (Brisswalter et al., 2002; Hüttermann & Memmert, 2014; McMorris & Graydon, 2000) have shown a positive effect of exercise on
cognitive behaviour up to a certain point, where cognitive performance decreases, illustrated by an inverted-U function (Easterbrook, 1959). The point to which cognitive performance increases before dropping off is thought to be close to the anaerobic (lactate) threshold (Chmura, Nazar, & Kaciuba-Uścilko, 1994), indicating that exercise intensity acts as moderator in this exercise–cognition relationship (Hüttermann & Memmert, 2014). Cognitive performance improvements due to moderate physical exercise are generated by an improvement in the vertebral blood circulation and an alteration in the level of neurotransmitters in CNS—acetylcholine, dopamine, norepinephrine, epinephrine, adrenocorticotropic hormone, and vasopressin (Kashihara, Maruyama, Murota, & Nakahara, 2009; Radosevich et al., 1989; Rikli & Edwards, 1991; Spirduso, 1980). A facilitated regional cerebral blood flow positively affects performance on cognitive tasks. The blood lactate concentration increases rapidly as soon as the anaerobic threshold is reached while the production of hormones and catecholamines is activated (Chmura, Krysztofiak, Ziemba, Nazar, & Kaciuba-Uścilko, 1997; Chwalbińska-Moneta et al., 1998) resulting in a cognitive performance decrease. Therefore, it may be possible that participants of the current study may have had a higher level of physical fitness than participants of A. P. J. Roberts and Cole (2013). This hypothesis is supported by the mean HR immediately after exercise, which were on average 30 bpm lower in the current study. Thus, it is likely that participants of A. P. J. Roberts and Cole (2013) already performed at the anaerobic threshold, resulting in cognitive shift from executive to non-executive functioning without general performance increases, while participants of the current study, assumingly working under the anaerobic threshold, elicited performance improvements in the Phonemic Fluency task with an increase in the executive component.

However, inconsistencies in findings regarding changes in cognitive functioning in general after demanding tasks or exercises are regularly obtained. There are a number of studies that suggest that physical exercise intensity does not have any influence on cognitive performance (for a review see Y.-K. Chang et al., 2012; Lambourne & Tomporowski, 2010). A possible explanation for the failure to unequivocally demonstrate the inverted-U effect
are the cognitive and physical task characteristics (type, intensity, and duration) as well as cognitive performance tasks (Y.-K. Chang et al., 2012; Lambourne & Tomporowski, 2010; Pesce, 2009) that are used in the different studies. Likewise, depending on the cognitive test used and the task that had to be performed by participants, results in the use of force domain differ as well (see Table 13). For example, results of the Tactical Decision Making Research Group (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012) addressing the effects of police use of force scenarios and exercises on the executive functioning component of WM, indicated that the demands placed on the officers led to a cognitive shift. In situations of high levels of demand, the executive cognitive processing of AFOs increased whilst non-executive processing decreased. In simulated use of force situations involving a reduction of demand, a decrease of executive cognitive processing was observed whilst non-executive processing decreased. In one particular study A. P. J. Roberts (2012) employed the Verbal Fluency Task before and after the tactical training of the officers. The results showed that “AFOs are able to maintain above average executive cognitive function, even during periods of intense cognitive and physical load” (p. 112). Additionally, an increase in non-executive cognitive functioning occurred, when well established standard operational procedures were used during the tactical training. At the same time executive cognitive processing decreased. The author suggests that these cognitive shifts that can be interpreted as “adaptive defensive responses, which result in behavioural adaptions to cope with and remove the threat/demand” (p. 223).

In this light, the current study provides evidence that the cognitive demands of the task at hand (scenario with tactical decisions vs non-cognitive demanding exercise) did not differ with regards on the effect on executive functioning performance. It seems reasonable to assume that the effect on cognitive functioning was due to the physiological load induced by the scenario and exercise. The current results would also indicate that changes in WM performance due to increased physiological demand would primarily be in the executive component. In situations where retrieval processes are warranted, such as the reliance on standard operational procedures, non-
executive processes may be enhanced. Further studies are clearly needed in order to obtain a full picture about changes in cognitive processing, and especially executive functioning, due to various use of force situations.

When comparing participants’ TW scores to a large normative sample of ranked scores (Aschenbrenner et al., 2000) the average score was below normative average at pre-testing. At post-testing TW was average, inferring that performance changes reflect a cognitive shift rather than a deterioration of cognitive functioning. These results differ from the results of A. P. J. Roberts (2012), where authorised firearms officers’ performance was above average at all times. At post-testing TW scores were ranked between the 95th and 98th quartile, “suggesting that the officers were operating near the limits of WM capacity” (Roberts, 2012, p. 107). Roberts further argues that the above average performance may be due to regular training and operational activity, since the daily routine of authorized firearms officers involves planning, problem solving and dealing with new situations (Greenwood-Ericksen et al., 2004). Participants of the current study routinely do not carry out such high cognitive demanding activities compared to authorised firearms officers. The training and operational work of participating police officers mainly focuses on community policing or designing and supervising training programs or keeping themselves physically fit.

5.6.3 Limitations

Some limitations of the current study have to be acknowledged.

The sample consisted of participants with experience in simulated physical confrontations. However, participants could not be randomly assigned to one group or another. The scenario group comprised police officers, whereas the exercise group consisted of martial artists. It was not possible to task police officers with exercise instead of the scenario since data collection took place during a regular training period and police officers were required to participate in scenario training (and not in physical exercise). This difference in the group characteristics may be an influential factor. For example, it may be possible that the police officers had more experience in real conflict
situations than the martial artists. Further studies should reduce this potential confound by randomly assigning participants of the same population (police officers) to different groups.

A specific problem in the performance assessment during simulated use of force situations involving the discharge of a weapon is that Simunition rounds and protective equipment were used in order to prevent injuries. Even though it has been demonstrated the physiological arousal in use of force scenarios is close to the same level in real life incidents (Andersen et al., 2016; Lewinski, 2006; 2008) the emotional demand of real threat is not present. It can be hypothesised that the demand placed on police officers in a real life incident is even greater. Nevertheless, an emotional response to threat has been generated in the setting of the scenario by the fear of negative evaluation. The tactical performance of the participants was assessed and evaluated by a trainer team. Research (Wagstaff et al., 2008) has demonstrated that fear of appraisal interferes with the completion of executive processes.

Another limitation of the current study is that “stress” hormones, such as cortisol, were not measured. This would be desirable to better compare the acute “stress” level of participants to previous studies assessing WM functioning in the context of naturalistic stressors in the law enforcement and military communities (Tavernier et al., 2011, 2010; Morgan et al., 2006; Morgan et al., 2001, 2000a, 2000b).

Also due to the lack of equipment, HRV was not measured. These measures could provide an alternative to further measure the influences of different demands (psychological vs. physiological) on HR (Brisinda, Venuti, Cataldi, Efremov, Intorno, & Fenici, 2015b).

Finally, one limitation of the employed experimental design is the interval between the treatment and the WM testing. Participants were requested to proceed as fast as possible to the testing room directly after the end of the scenario. For safety reasons participants could only leave the scenario area after checking their guns and handing them over to the safety officers. Due to
the relatively long time interval between the RBT and WM measurement, it
seems possible that arousal decreased, influencing the effects on the
cognitive performance measurement of subjects.

5.6.4 Summary

Study 2 showed that an increase in executive functioning and a relatively
decrease of non-executive functioning occurred due to participations in a
scenario training and in a running-on-the-spot exercise. The changes in
executive functioning did not differ between the two groups. In light of
previous research on WM performance after acute exercise, this suggests
that the effects on WM performance occurred due to physiological arousal
induced by the two different treatments. As such, this shift in cognitive
functioning may reflect an adaptive response of the human body to the
demands at hand.
Chapter 6: Threat Attentional Bias and Risk Taking
6 Threat Attentional Bias and Risk Taking

6.1 Introduction

The previous chapter investigated the effects of scenario training compared to physical exercise in police officers and martial artists. Acknowledging the limitation that the two involved groups could have had different experiences with (simulated) violent confrontations, possible effects of attentional processes caused by these experiences are subject to investigation in the current chapter.

One critical factor in making a situationally adequate decision is generating and maintaining SA, which refers to the cognitive processes that are involved in perceiving and comprehending the meaning of a given situation (Endsley, 1995; Endsley & Robertson, 2000; G. A. Klein, 2000; Saus et al., 2006). A critical component of SA is the fundamental perception of the elements of a particular environment (Saus et al., 2006). In this context, the allocation of attention to threat cues is a fundamental requirement of any police officer as identifying real threats can result in the prevention of harm (College of Policing, 2013; Füllgrabe, 2014).

The perception of biologically-relevant stimuli, particularly life-threatening ones are essential for human survival. The monitoring of threat allows for efficient and accurate detection of, and behavioural responses to, potential threats. It involves a continuous balancing of various cognitive resources and response patterns (Adolphs, 2013; M. Davis & Whalen, 2001; Liddell et al., 2005; Naim et al., 2015; Pessoa & Adolphs, 2010; Zald, 2003) in which the amygdala is central (Gur et al., 2002; Hariri, Bookheimer, & Mazziotta, 2000; Morris et al., 1996; M. L. Phillips et al., 1998; L. M. Williams et al., 2001). Healthy adaptation requires the individual to allocate attention to genuine threats while simultaneously ignoring similar but non-threatening stimuli (Naim et al., 2015). This delicate balance can be set off due to traumatic events and give rise to threat avoidance and threat-related hypervigilance, which are often referred to as ‘clinical symptoms’ (Das et al., 2005; Ehlers & Clark, 2000). For example, cognitive biases like the threat-related attentional
bias have been identified as one of the most consistently demonstrated cognitive correlates of anxiety disorders (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Naim et al., 2015; Yiend, 2010). However, cognitive biases have been demonstrated to exist in populations that are exposed to life-threatening dangers on a regular basis (Bar-Haim et al., 2010; R. M. Todd et al., 2015). Therefore, the valence of cognitive bias per se is neutral. Depending on the environment and the tasks individuals have to fulfil on a regular basis, the development of cognitive biases can be functional, such as normal defense, or dysfunctional, like psychological disorders (D. J. Stein & Nesse, 2011). For example, compared to other professions, the proper allocation of attention to such threats on a regular basis is unique to the work of police (and military) personal, since effective and proportionate responding of officers requires the rapid and efficient detection of lethal threats in police-citizen encounters. As such, it represents an adaptive (functional) use of ‘genetically hard wired’ structures involved in human defensive behaviour coupled with experiences of the environment individuals operate in. On the other hand, if people inhabit environments that are mostly free of lethal threats the tendency to interpret ambiguous stimuli as potentially life-threatening no longer serves the adaptive use that it once did. Cognitive resources that have been previously used for the constant assessment of threat have been freed up allowing for the allocation into other aspects of cognition (Cisler & Koster, 2010; M. W. Eysenck, Derakshan, Santos, & Calvo, 2007; M. W. Eysenck, Payne, & Derakshan, 2005), again reflecting an adaptive development. As a consequence, constant feelings of uneasiness and danger in environments that are mostly free of lethal threats, are nowadays considered to be a pathological trait (S. J. Bishop, 2008a).

Similarly, risk-taking behaviours that place an individual at risk for deleterious health or safety outcomes have received particular attention in the literature (Banducci, Felton, Dahne, Ninnemann, & Lejuez, 2015; Boronova-lova, Gwadz, Kahler, Aklin, & Lejuez, 2008; Hanson, Thayer, & Tapert, 2014; Lejuez, Simmons, Aklin, Daughters, & Dvir, 2004; Zuckerman, Ball, & Black, 1990). Risk-taking behaviours per se are not positive nor negative in the...
sense they are neither functional nor dysfunctional. Rather it encompasses a broad range of behaviours that fall along both positive and negative dimensions (Byrnes, Miller, & Schafer, 1999; Foersterling, 1980; Leigh, 1999; Lejuez et al., 2002). In the context of policing, taking risks is a common task for police officers, especially when decisions must be made in a split-second time frame and the result can have serious impact on the safety of the officer and those around.

In this context, Study 3 aims to assess whether experience in occupations involving life-threatening hazards on a routine basis can sequentially lead to the development of cognitive biases. Specifically, the aim is to assess whether exposure to real threat that is typical for police officers lead to the development of an increase in risk taking behaviour and threat-related attentional biases. The study compares police officers to martial artists as they differ from normal populations due to experiences of being in hostile situations where threat-detection and risk assessment may serve some benefit. However, martial artists are not exposed to real threat on a regular basis per se. This comparison allows for the assessment of the effects of real versus representative danger. A third group consisted of police officers with martial arts experience, allowing for the investigation of real threat experience combined with training experiences as martial artists.

6.1.1 Threat-Related Attentional Bias

The allocation of attentional resources towards threat is the first stage in the Distant-Dependent-Defence-Hierarchy (Hendrie, Weiss, & Eilam, 1996; Rodgers & Randall, 1987). In this stage of “attentive immobility” (freezing) the organism is passive, but attentive and engages in selective scanning (Marks, 1987; Volchan et al., 2011). With being motionless and vigilant the chances of not being noticed by the predator/threat increases. However, the organism is automatically mobilized and primed to respond, but not yet active (P. J. Lang et al., 2000).

There are three observable components in the investigation of threat-related attentional biases: (a) an attentional bias towards threat (i.e. threat stimuli are detected faster than non-threat stimuli), (b) an attentional bias away from
threat (i.e. attention is allocated towards opposite the location of threat), or (c) difficulties in the disengagement from threatening stimuli (Cisler & Koster, 2010).

An attentional bias towards threat refers to the tendency to allocate more attention towards threatening stimuli relative to neutral stimuli (Bar-Haim et al., 2007; Cisler & Koster, 2010; H. J. Eysenck, 1992a; MacLeod, Mathews, & Tata, 1986; Mogg & Bradley, 1998). A vast amount of research demonstrates that anxious individuals show an attentional bias towards threatening sources of information, while this effect is less consistent or not observed non-anxious individuals (Bar-Haim et al., 2007; 2010; Mogg & Bradley, 1998; Shechner et al., 2012; J. M. G. Williams, Mathews, & MacLeod, 1996). Since anxiety serves the function of threat anticipation (J. M. G. Williams, Watts, MacLeod, & Mathews, 1997), attentional biases have mostly been observed in relation to differing level of pathology (Bar-Haim et al., 2007; Beck & Clark, 1997; Blanchette & Richards, 2010; Cisler & Koster, 2010; H. J. Eysenck, 1992a). According to these findings, highly anxious individuals are oversensitive in the detection of threat – a phenomenon referred to as hypervigilance (Bar-Haim et al., 2007; Koster, Crombez, Verschuere, & De Houwer, 2004; Naim et al., 2015; Sheppes, Luria, Fukuda, & Gross, 2013).

Even though most studies find an attentional bias towards threat in anxiety disorders, results from laboratory-based research show that acute stress can also lead anxious individuals to shift their attention away from threats (Amir et al., 1996; M. Garner, Mogg, & Bradley, 2006; Helfinsteihn, White, Bar-Haim, & Fox, 2008; Mansell, Clark, Ehlers, & Chen, 1999; Mathews & Sebastian, 2008). Similar results have been shown in combat veterans with post-traumatic stress disorder (Constans, McCloskey, Vasterling, Brailey, & Mathews, 2004; Sipos, Bar-Haim, Abend, Adler, & Bliese, 2014; Wald, Lubin, et al., 2011a; Wald, Shechner, et al., 2011b), and also civilians that are regularly exposed to life-threatening danger (Bar-Haim et al., 2010).

A third observational characteristic of attentional bias - besides facilitating attention and attentional avoidance – is the difficulty in disengagement, which
refers to that it is harder to disengage attention from a threat stimulus relative to a neutral stimulus (Cisler & Olatunji, 2010; Mogg, Holmes, Garner, & Bradley, 2008; Salemink, van den Hout, & Kindt, 2007).

Several underlying mechanisms have been found to mediate the production of observable characteristics of threat-related attentional bias. For example, the amygdala, as a neural mechanism, has been found to be a mediator of automatic vigilance for threat (A. K. Anderson & Phelps, 2001; M. Davis & Whalen, 2001; Öhman, 2005). Regarding difficulties in the disengagement of threat, attentional control, as the cognitive ability to allocate attention (Derryberry & Reed, 2002; M. W. Eysenck et al., 2007; Posner & Rothbart, 2000), has been shown to be a possible mechanism (M. W. Eysenck et al., 2007). The mediating mechanism the observable biases operate under specific stages of information processing, which (Cisler & Koster, 2010) divide into automatic and strategic stages (McNally, 1995; Moors & De Houwer, 2006; Shiffrin & Schneider, 1977). The automatic processing of stimuli refers to processing that is capacity free and occurs without control or conscious awareness. On the contrary, strategic processing is an intentional, controllable, capacity-limited process that depends on awareness (Cisler & Koster, 2010).

Reviewing evidence regarding the three layers of attentional bias (observable bias, mediating mechanisms, information processing) (Cisler & Koster, 2010) concluded that interactions between the mediating mechanisms and attentional components seem to be relatively well supported, while the interaction between the stage of information processing, mediating mechanism, and attentional components remains unclear. Research investigating the relationship between facilitated attention and disengagement showed that on the one hand difficulties in disengagement can be observed without facilitated attention (Amir, Elias, Klumpp, & Przeworski, 2003; E. Fox, Russo, & Dutton, 2002; E. Fox, Russo, Bowles, & Dutton, 2001; Rinck, Becker, Kellermann, & Roth, 2003; Rinck, Reinecke, Ellwart, Heuer, & Becker, 2005; Yiend & Mathews, 2001), and on the other hand facilitated attention almost regularly occurs with difficulty in disengagement (Byrne & Eysenck, 1995; Gilboa-Schechtman, Foa, & Amir,
Since evidence supports the claim that facilitated attention is linked to automatic stages of processing (S. J. Bishop, 2008b; S. Bishop, Duncan, Brett, & Lawrence, 2004), whereas difficulties in disengagement are linked with higher order control mechanism, the conclusion that when both occur together facilitated attention precedes difficulty in disengagement seems reasonable (Cisler & Koster, 2010). A temporal structure of attentional biases seems to be supported by the vigilance-avoidance hypothesis (Mogg, Bradley, Miles, & Dixon, 2004), that proposes that anxious individuals first demonstrate facilitated attention to threat before avoiding of attention to the threat stimulus is shown.

### 6.1.1.1 Attentional Bias and Fear and Anxiety

Several cognitive theories have been put forward to explain the relationship between attentional bias and fear or anxiety (Bar-Haim et al., 2007; Beck & Clark, 1997; M. W. Eysenck, 1992b; 1997; Mogg & Bradley, 1998; J. M. G. Williams et al., 1997; J. M. G. Williams, Williams, Watts, MacLeod, & Mathews, 1988). For example, Bar-Haim et al. (2007) proposed an integrative model of cognitive mechanisms that underlie threat processing, comprising four systems (see Figure 29).

![Cognitive mechanisms underlying threat processing](image)

**Figure 29:** Cognitive mechanisms underlying threat processing according to Bar-Haim et al. (2007)
According to the model, the threat value of any incoming stimuli is assessed in the bottom-up preattentive threat evaluation system that constantly assesses the surrounding stimuli and diverts attention towards those deemed to be a threat. If the threat is high, the resource allocation system is activated, which leads to interruption of ongoing activities, attentional orientation towards the location of the threat stimulus, and physiological alertness. The outcome of this resource allocation system serves as an input for the guided threat evaluation system. Here the stimulus is compared with memory and prior learning experiences. Furthermore, context and possible coping mechanisms are taken into consideration within the guided threat evaluation system. The outcome of this system is a conscious evaluation of the threat value of the current stimulus. If the consciously evaluated threat level is high, a goal engagement system will interrupt the pursuit of current goals and the primary goal of the individual will orient towards the threat. The model explains attentional biases as the result of a hypersensitive resource allocation system that allocates attentional resources even to stimuli that are evaluated as only mildly threatening by the preattentive threat evaluation system.

Regarding the relationship between attentional bias and fear and anxiety the model of Bar-Heim et al. (2007) adopts a strict linear interpretation of that relation. In a recent review of the evidence for the causal impact of attention bias on fear and attention (Van Bockstaele et al., 2014), the authors concluded that such a strict unidirectional cause-effect model in unlikely to hold. According to Van Bockstaele et al. (2014), the “relations between attentional bias and fear and anxiety is best described as a bidirectional, maintaining, or mutually reinforcing relation” (p. 682).

Yet, within the context of policing, where attentional biases may serve a functional aspect, a hypersensivity towards threat may be the result of prior learning (Öhman, 1996; Öhman & Wiens, 2004). According to Öhman (1996) a feature-detection model of attention to threat serves as an evolutionarily adaptive process that sensitises the expectancy system of the individual to
specific stimuli based on emotional memories. Hence, if a person regularly encounters threats in his/her typical environment, attentional biases would be expected to develop as a function of experiential learning. Since police officers are regularly exposed to situations that have the potential for harmful or life-threatening events compared to civilians, threat-related attentional biases may have developed. Study 3 tries to assess whether individuals that are routinely exposed to potentially life-threatening scenarios on a routine basis would display similar attentional biases to those found in anxious populations in the absence of significant anxiety, which is typically not found in police officers (B. J. Evans, Coman, & Stanley, 1992).

### 6.1.1.2 Temporal Structure of Attentional Bias

Studies examining the time course of attentional biases show a typical vigilance-avoidance pattern, where earlier processing is connected to orientation, whereas motivational processes are responsible for maintenance of attention (Zvielli, Bernstein, & Koster, 2014). It has been proposed that separate neural subsystems underlie these different cognitive operations of attentional shifting and maintenance (M. Field, Mogg, Zetteler, & Bradley, 2004b; LaBerge, 1995). Initial shifting is a subconscious effort and only measurable through paradigms that present stimuli at very brief exposures (M. Field et al., 2004b).

For instance, an attentional bias for alcohol-related words in abstinent alcoholics was found, when the stimuli were presented for 100 msec (Stormark, Field, Hugdahl, & Horowitz, 1997). This suggests that attentional biases can be demonstrated for very briefly presented stimuli. On the other hand, abstinent alcohol dependent participants showed after treatment a bias to avoid alcohol-related cues as shown by faster reaction times to targets invalidly cued by alcohol-related words at a 500 msec interval presentation. This seems to represent a bias in maintained attention that is more likely to be revealed when stimuli are presented for longer durations (M. Field et al., 2004b).

Most studies investigating threat-related attentional biases involve clearly visible stimuli, that are presented for at least 500 msec, in order to allow for
consciously processing (Bar-Haim et al., 2007). Results of recent studies using longer exposure times found attentional bias towards and/or away from threat stimuli (Bar-Haim et al., 2010; Naim et al., 2015; Zvielli et al., 2014), which is in line with the vigilance-avoidance pattern. Yet, the finding of an attentional bias in response to a stimulus that is presented of a longer duration (supraliminal exposure) does not allow for the distinction between the contributions of a preconscious bias to the threatening stimulus and a bias that requires awareness of that stimulus. The exposure of stimuli for very brief durations (subliminal exposure) prevents the stimuli from reaching awareness (Di Lollo, Enns, & Rensick, 2000), allowing for the investigation of early automatic rather than later conscious processing of threat-related stimuli. Finding a bias using subliminal exposure of the stimuli can be accounted for only by an early preconscious bias (Bar-Haim et al., 2007).

6.1.1.3 Measure of Attentional Bias

There are several experimental paradigms that are used for the investigation of threat-related attentional bias (Van Bockstaele et al., 2014). Among the most prominent ones are the Dot Probe Task and the Emotional Stroop, which have both been employed in a varieties of studies.

6.1.1.3.1 The Dot Probe Task

The Dot Probe Task (MacLeod et al., 1986) requires participants to determine the location of a small probe which is presented in a location previously occupied by one of two words. In this task, one threatening and one neutral stimulus are presented on different sides of a screen and then one of the stimuli is followed by a probe at its location. Selective attention towards a motivationally-relevant stimulus will result in faster reaction times (RT) when a probe appears in the spatial location of the target stimulus and RT will be slower when a probe appears in the spatial location opposite of the target stimulus. This reflects an attentional bias towards the threat and vice-versa, an attentional bias away from the threat is reflected when selective attention is allocated away from the motivationally-relevant stimulus (Zvielli et al., 2014).
The reliability of the Dot Probe Task has recently become subject for criticism: Several studies have shown poor test-retest reliability (Schmukle, 2005; 2002; Staugaard, 2009) and poor internal reliability (R. M. Cooper et al., 2011; Kappenman, MacNamara, & Proudfit, 2015; Schmukle, 2005; 2002; Waechter, Nelson, Wright, Hyatt, & Oakman, 2013). However, contradictory findings of high internal reliability were reported by Bar-Haim et al. (2010).

6.1.1.3.2 The Emotional Stroop Task

The Emotional Stroop Task is a measure of supraliminal attention since exposure times of the stimulus enable conscious perception of word meaning (P. Smith & Waterman, 2005). This allows for effective examination of cognitive-emotional processes of attentional biases (Bar-Haim et al., 2007). The Emotional Stroop is a modification of the classic colour-naming Stroop interference paradigm (Stroop, 1935). The Stroop effect refers to the differences in naming the colour of a word when the stimulus is congruent (e.g. the word yellow is printed in yellow) compared to an incongruent stimulus (e.g. the word yellow is printed in blue). The presence of the Stroop effect reflects the failure of the participants to exclusively focus on the target dimension of colour. Mathews and MacLeod (1985) developed a modified version of the Stroop task, called the Emotional Stroop. In this task, the word valence compared to the semantic congruence with the printed word colour, as in the classic Stroop task, are manipulated (Mathews & MacLeod, 1985). Hence, the response latency to name the printed colour of a word is compared, when the meaning of the word is threat-related (e.g., “dead”) compared to a neutral meaning (e.g., “house”). Participant’s demonstrate threat-related biases when colour naming takes longer with a threat stimulus relative to a neutral stimulus (Bar-Haim et al., 2007; MacLeod, 1991; J. M. G. Williams et al., 1996).

The Emotional Stroop Task has been shown to be highly reliably with a split-half reliability of $r = .76$ (Dresler, Mériaux, Heekeren, & Meer, 2008). In contrast, several researchers reported low test-rest reliability for the difference between colour-naming latencies of neutral and emotional words.

6.2 Risk Taking

Risk-taking behaviours can be defined as behaviours that involve some potential for danger or harm while simultaneously providing an opportunity to obtain a reward of some kind (Leigh, 1999). In the context of policing, the management of risk is an essential task that has to be dealt with on a routine basis (Hoyle, 1998; R. J. Kane, 1999; 2000; Perez Trujillo & Ross, 2008). This process of making a choice from a set of options, with the consequences of that choice being crucial, refers to decision-making per se (Bar-Eli, Plessner, & Raab, 2011). In such situations, where every alternative regularly has their drawbacks (Yates & Tschirhart, 2006) and the environment is uncertain, time pressured, and dynamic (Endsley et al., 2007; Lipshitz et al., 2001; Parent & Verdun Jones, 1998; Staller & Zaiser, 2015b), experienced decision-makers are thought to rely on intuitive decision-making based on recognition of situational cues (K. G. Ross et al., 2004). Therefore, police officers have to take risks since the environment is uncertain and balancing the potential rewards of their behaviour against potential negative outcomes. In situations of threat to the police officer or against others, the acting officer is faced with making quick and complex decisions. Potential courses of action are compared to contextual constraints and the familiarity of the unfolding situation (G. A. Klein, 2008; G. A. Klein et al., 1993). Thus, experience plays a crucial role in this context. In line with this, it is suggested that the process by which experts make decisions is based on experience and an increased ability to assess risk (Wickens, 1992). Furthermore, through the use of heuristics experts are able to quickly select the best tool or strategy for a given task in an uncertain world (Gigerenzer & Gaissmaier, 2011; P. M. Todd, Gigerenzer, ABC Research Group, 2012). These strategies ignore “part of the information, with the goal of making decisions more quickly, frugally, and/or accurately than more complex methods” (Gigerenzer & Gaissmaier, 2011, p. 454). With this in mind, a propensity for risk-taking behaviour in police officers may be affected by past experiences where risk-taking has been used effectively. This may lead to an
increased accessibility to heuristics related to risk-taking behaviour (P. Fischer, Guter, & Frey, 2008).

One potential tool for gauging risk taking behaviour within a laboratory setting is the Balloon Analogue Risk Task (BART), a computerised assessment of risk taking propensity (Hopko et al., 2006; Lejuez et al., 2002; Lejuez, Aklin, Zvolensky, & Pedulla, 2003b; Lejuez et al., 2007; T. L. White, Lejuez, & de Wit, 2008). It is one of the most widely used and tested sequential risk-taking tasks and has been shown to be a sensitive measure to real world risk-taking in different populations (Aklin, Lejuez, Zvolensky, Kahler, & Gwadz, 2005; Banducci et al., 2015; Boronvalova et al., 2008; J. A. Campbell, Salvartgis, & Crowe, 2013; Crowley, Raymond, Mikulich-Gilbertson, Thompson, & Lejuez, 2006; Hanson et al., 2014; Lauriola, Panno, Levin, & Lejuez, 2014; Lawyer, 2013; Lejuez et al., 2002; 2004; 2007; 2003a; Lejuez, Aklin, Boronvalova, & Moolchan, 2005). The BART requires participants to inflate a series of balloons. The more an individual balloon is inflated the more money a participant will earn, with the caveat that all money earned for a given balloon will be lost if the balloon is inflated too much and explodes. Hence, the BART is able to determine the likelihood of behaving in a risky manner by measuring the average number on pumps that balloons are inflated with in the context of earning monetary rewards (Banducci et al., 2015).

Since performance on the BART is associated with substance and alcohol use, smoking, psychopathy, and risky sex among adults (J. A. Campbell et al., 2013; Gonzalez et al., 2012; Hanson et al., 2014; Hopko et al., 2006; M. K. Hunt, 2005; Lejuez et al., 2002; 2003a; Schuster, Crane, Mermelstein, & Gonzalez, 2012), it can be assumed that it captures adults’ likelihood of behaving in a risky manner across a variety of contexts (Banducci et al., 2015). Therefore, performance on the BART can serve as a proxy of the participants’ risk taking propensity across multiple contexts.

6.3 Hypotheses

Linking previous research with theoretical considerations the following hypotheses were generated:
• Police officers will demonstrate higher levels of attentional biases, both subliminally and supraliminally shown by significant levels of attentional vigilance and difficulties to disengage from threatening stimuli (Hypothesis 1).

• Previous studies measuring state and trait anxiety among law enforcement personnel found that state and trait anxiety did not differ compared to other adults (Newman & LeeAnne Rucker-Reed, 2004; B. L. Patterson, 1992; M. Pendleton, Stotland, Spiers, & Kirsch, 1989; Storch & Panzarella, 1996; Territo & Vetter, 1981). Hence, state and trait anxiety of police officers will show no differences compared to the other groups (Hypothesis 2).

• Police officers will show differences in risk-taking behaviour compared to the control group (Hypothesis 3).

6.4 Methods

6.4.1 Participants

A total of 205 participants from Germany (n = 155) and the UK (n = 50) took part in Study 3. Police officers were recruited through an opportunistic sampling method. Martial artists and the control group were recruited via social media (facebook). Participants were asked about experience as police officers and martial arts and were subsequently assigned to different groups for the study. These consisted of (a) police officers (n = 74) with no experience in martial arts, except for police training, (b) martial artists (n = 50), that had regular training experience in martial arts, and were not part of a police force, (c) police officers with additional martial arts experience (n = 33), and (d) a control group (n = 48) with no experience in the policing sector and no experience in any martial arts or combat sports. Demographical data of participants is shown in Table 16. Two cases of missing data occurred in the Dot Probe Task, leaving the final sample at 203 participants in this task. Likewise, two cases of missing data were identified in the BART, resulting in 203 participants in the BART.
### Table 16: Demographical Data of Participants of Study 3

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Sex</th>
<th>Age</th>
<th>Experience police</th>
<th>Experience martial arts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>male</td>
<td>female</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Germany</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police officer</td>
<td>155</td>
<td>male</td>
<td>female</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>61</td>
<td>13</td>
<td>30.11</td>
<td>8.16</td>
</tr>
<tr>
<td>Martial artist</td>
<td></td>
<td>46</td>
<td>3</td>
<td>31.67</td>
<td>11.30</td>
</tr>
<tr>
<td>Police officer &amp; martial artist</td>
<td>28</td>
<td>27</td>
<td>1</td>
<td>29.75</td>
<td>5.93</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td>20</td>
<td>12</td>
<td>34.05</td>
<td>10.13</td>
</tr>
<tr>
<td><strong>UK</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police officer</td>
<td>50</td>
<td>13</td>
<td>0</td>
<td>35.77</td>
<td>7.73</td>
</tr>
<tr>
<td>Martial artist</td>
<td></td>
<td>4</td>
<td>3</td>
<td>34.75</td>
<td>19.81</td>
</tr>
<tr>
<td>Police officer &amp; martial artist</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>36.20</td>
<td>5.26</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td>28</td>
<td>21</td>
<td>31.46</td>
<td>13.80</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police officer</td>
<td>205</td>
<td>male</td>
<td>female</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>74</td>
<td>13</td>
<td>31.11</td>
<td>8.32</td>
</tr>
<tr>
<td>Martial artist</td>
<td></td>
<td>50</td>
<td>6</td>
<td>31.92</td>
<td>11.92</td>
</tr>
<tr>
<td>Police officer &amp; martial artist</td>
<td>33</td>
<td>32</td>
<td>1</td>
<td>30.73</td>
<td>6.22</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td>48</td>
<td>33</td>
<td>32.54</td>
<td>12.35</td>
</tr>
</tbody>
</table>
6.4.2 Materials

6.4.2.1 State-Trait Anxiety Inventory

State and trait anxiety were measured using the STAI (Spielberger, Gorsuch, & Lushene, 1970) in the German Version (Laux, Glanzmann, Schaffner, & Spielberger, 1981), which measures stable tendencies of anxiety and feelings of anxiety at the time of the testing. Trait anxiety refers to the relatively stable levels of anxiety in an individual, with high levels leading to the tendency to perceive situations at stressful. State anxiety is the experience of unpleasant feelings towards a particular situation and as a result fluctuates invariably as a result of context.

Trait anxiety, as a function of STAI scores, has been employed as a marker for attentional bias in several studies (Zvielli et al., 2014) and differences in trait anxiety are linked to threat-related attentional bias as proposed by theories of attentional bias (Mathews & Mackintosh, 1998; Mogg & Bradley, 1998). A vast amount of published studies that examine anxiety and threat-related attentional bias have focused on adults with high scores on trait anxiety (Bar-Haim et al., 2007), so using the STAI in the current study facilitates comparison to past work.

The STAI consists of two parts (see Appendix), each containing 20 items, which require participants to rate how they feel on a 4-point Likert scale in terms of “right now” (state) and “in general” (trait). The scale has been shown to be high in convergence and discriminant validity (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) and demonstrates a high level of internal consistency (Grös, Antony, Simms, & McCabe, 2007).

6.4.2.2 Dot Probe Task

The Dot Probe Task was used as a measure of subliminal attentional bias. The task was programmed according to MacLeod, Soong, Rutherford and Campbell (2007) using Inquisit (“Inquisit 4.0.5.0.,” 2014) allowing for the measurement of response times with millisecond accuracy (De Clercq, Crombez, Buysse, & Roeyers, 2003). The task was configured do deliver 96 trials of stimulus word pairs, each containing one threatening and one neutral
word that were matched on length and frequency (Kučera & Francis, 1967; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; MacLeod, Soong, Rutherford, & Campbell, 2007). Words were translated from English to German and then back-translated by a separate party using structured guidelines (Brislin, 1970; Geisinger, 1994). Words used with the UK and the German samples are shown in the Appendix.

The mechanism responsible for the initial shift of gaze is a rapid procedure that can only be assessed when stimuli are presented at brief exposure times (M. Field, Mogg, & Bradley, 2004a). Since it is difficult to disentangle the effect of initial orientation and cognitive-emotional processes, two exposure times of the stimulus word pair (50 msec and 200 msec) were used, which are proposed to represent thresholds of subliminal attention.

Participants were presented with a fixation cross (500 msec), followed by the simultaneously presented stimuli (word pair) for the duration of the set exposure time (50 msec or 200 msec). Across these trials threat word position (upper vs. lower screen) position and probe type (“<” vs. “>”) were balanced such that after every eight participants, each word pair had been presented once under each unique experimental condition. The order of trial presentation was randomized for each participant. Participants were instructed to indicate whether a “<” or “>” probe was presented by pressing either the left arrow or the right arrow key, respectively. Furthermore, responses had to be as quickly but as accurate as possible. Participants were told to place their fingers on the two response keys to allow for quick responses.

After each trial, the fixation cross of the next trial was presented on the screen. On incongruent trials (IT), the probe appeared in the location of the neutral stimulus, where on congruent trials (CT) the probe appeared in the location of the threat stimulus (Zvielli et al., 2014). The dependant measure was the response latency for IT and CT, which was timed from the appearance of each probe until detection of the associated response. Bias scores were computed by subtracting mean response times of CT from mean response times of IT (Bar-Haim et al., 2007; Zvielli et al., 2014). This bias
score was computed per participant for the 50 msec and the 200 msec condition.

A positive bias score indicates attentional vigilance towards threat as response times are shorter on CT, whereas negative scores indicate an attentional avoidance of threat as attention must be focused back to the vicinity of the threat, increasing the response times (MacLeod et al., 2007).

Reaction time outliers were dealt with using an upper cut-off of two standard deviations above the mean (Mansell et al., 1999; Mogg, Bradley, & Williams, 1995) and a lower cut-off of 250 msec (Ratcliff, 1993).

6.4.2.3 Emotional Stroop Task

The Emotional Stroop Task as described by P. Smith and Waterman, 2005) was used as a measure of supraliminal attention. The task consisted of 125 experimental trials: 25 aggression themed words, 25 positive emotion words, 25 negative emotion words, 25 neutral words and 25 colour words (see Appendix). The colour words were always incongruent with the colour presented to the participants.

Since participants were English or German, the Emotional Stroop Task was presented in the language of the participants. The words used were translated from English to German and then back-translated by a separate party using structured guidelines (Brislin, 1970; Geisinger, 1994).

Participants were given 10 practice trials. In each trial, a fixation cross appeared for 500 msec in the centre of the screen. Participants were asked to ignore the word and press a key assigned to a specific colour (red, green, blue or yellow). Participants were instructed to keep their fingers rested on the key to allow for faster RTs. Measures produced by the task show RT for a correct response. RT outliers were dealt with using an upper cut-off of two standard deviations above the mean (Mansell et al., 1999; Mogg et al., 1995) and a lower cut-off of 250 msec (Ratcliff, 1993). Mean RT on a particular word class are computed afterwards. Bias scores are calculated by subtracting the mean RT to neutral word presentations from each of the
mean RT to stimulus words (aggression, positive, negative, and colour), producing four bias scores for each participant (P. Smith & Waterman, 2003). Higher semantic interference on emotional or colour trials is reflected by higher RTs in these conditions compared to the neutral condition. Hence, higher bias scores indicate a higher amount of interference on colour naming caused by the processing of the semantic meaning of the stimulus word.

6.4.2.4 Balloon Analogue Risk Task

The BART requires the participants to pump up a simulated balloon on the screen by the click of a mouse. The computer screen shows a small simulated balloon, that gets bigger with every click of the mouse. Participants are asked to gain as much money as possible on each trial and cashing out by clicking “Collect $$$” on the screen. Furthermore, a permanent money-earned display labelled “Total Winnings”, and display listing the money that could be earned, when the participants decides to be cashed out (“Potential earnings”), and the trial number (“Ballon number”) are presented on the screen (see Figure 30).

![Figure 30: On-Screen Presentation of the Balloon Analogue Risk Task](image)

With each pump, 5 cents were accumulated in a temporary reserve (“Potential earnings”). When a balloon was pumped past its individual explosion point, a sound-effect combined with the explosion of the balloon was presented. In that case, participants lost all money in the temporary
bank and a new balloon appeared on the screen. Therefore, participants had to decide when to stop inflating the balloon.

After 30 trials the adjusted average pump count (AAPC) was calculated as the behavioural statistic of the task. The AAPC is the mean number of pumps on trials that do not end in an explosion (Pleskac, Wallsten, Wang, & Lejuez, 2008).

6.4.3 Procedure

Participants accessed a website (hosted on www.millisecond.com), which outlined information regarding the experiment and issued a consent form. If accepted, the Inquisit web software (“Inquisit 4.0.5.0.,” 2014) was downloaded and screen resolution was automatically set to 800 x 600. Before testing began, participants were informed that the test would take approximately 40 minutes and should be completed in a quiet setting. The order of tasks was: Dot Probe (50 msec), Emotional Stroop, BART and Dot Probe (200 msec) and was constant across all participants. In order to reduce possible priming effects due to potential conscious processing of the words in the Dot Probe Task at 200 msec, it was decided against counterbalancing the order and to start with an exposure time of 50 msec, which does not allow for conscious processing of the meaning at all, followed by an exposure time of 200 msec at the end. After the tasks were finished, participants were directed to a website (hosted on www.soscisurey.de), which recorded demographic information and any experience of being part of a police force or martial art group. The STAI was then completed and participants were thanked for their time and presented with debrief information.

6.5 Results

6.5.1 Sample Characteristics

Normality of the data sets were checked using Shapiro-Wilk’s test (Razali & Wah, 2011; S. S. Shapiro & Wilk, 1965) and a visual inspection of histograms, normal Q-Q plots and box plots (Doane & Seward, 2011).
Results showed that the STAI scores were sufficiently normally distributed, allowing for parametric analysis. Scores from the Dot Probe and the Emotional Stroop Task were not normally distributed. After transformation procedures, the data sets were still not normally distributed. Hence, robust statistical analysis based on ranked data was performed on these data sets. Finally, results showed, that the AAPC score of the Bart Task was normally distributed.

6.5.2 State-Trait Anxiety Inventory

Participant scores for both state and trait anxiety scores of the STAI are displayed in Table 17.

Table 17: Means, Standard Deviations and 95% Confidence Intervals around the Mean for State and Trait Anxiety Scores of Study 3

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State anxiety</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>41.09</td>
<td>4.86</td>
<td>39.97</td>
<td>42.22</td>
</tr>
<tr>
<td>Martial arts</td>
<td>50</td>
<td>43.76</td>
<td>0.64</td>
<td>42.48</td>
<td>45.04</td>
</tr>
<tr>
<td>Police &amp; martial arts</td>
<td>33</td>
<td>41.21</td>
<td>4.99</td>
<td>39.44</td>
<td>42.98</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>44.85</td>
<td>4.82</td>
<td>43.45</td>
<td>46.26</td>
</tr>
<tr>
<td><strong>Trait anxiety</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>43.64</td>
<td>4.37</td>
<td>42.62</td>
<td>44.65</td>
</tr>
<tr>
<td>Martial arts</td>
<td>50</td>
<td>43.78</td>
<td>4.77</td>
<td>42.43</td>
<td>45.13</td>
</tr>
<tr>
<td>Police &amp; martial arts</td>
<td>33</td>
<td>42.42</td>
<td>3.79</td>
<td>41.08</td>
<td>43.77</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>47.52</td>
<td>4.71</td>
<td>46.15</td>
<td>48.89</td>
</tr>
</tbody>
</table>

Results of separate one-way ANOVAs on STAI scores, revealed a significant effect of group on state anxiety, $F(3, 204) = 7.88, p < .001, \eta^2 = .105$, and on trait anxiety, $F(3, 204) = 11.00, p < .001, \eta^2 = .141$.

Post hoc comparisons using Bonferroni corrected $t$-tests showed that the ‘police and martial arts’ group had significantly lower scores on state anxiety than the control group, $p = .006$. The police group reported significantly less state anxiety than the control group, $p < .001$, and the martial arts group, $p = .016$. For trait anxiety, the control group had significantly higher scores than the police and martial arts group, $p < .001$, than the police group, $p < .001$, than the martial arts group, $p < .001$. 
and than the martial arts group, \( p < .001 \). Means and 95% confidence intervals around for STAI score across all groups are depicted in Figure 31.

Due to variations between groups, state and trait anxiety scores will not be treated as covariates in subsequent analysis (A. Field, 2013; Lord, 1967; 1969; G. A. Miller & Chapman, 2001; Wildt & Ahtola, 1978).

![Figure 31: Means and 95% confidence intervals of state and trait anxiety scores across all groups of Study 3 (P: police; MA: martial arts; P&MA: police & martial arts; C: control group). * indicates significance at \( p < 0.05 \); ** at \( p < .01 \); *** at \( p < .001 \), between groups.](image)

### 6.5.3 Dot Probe Task

Summary data, including means, standard deviations and 95% confidence intervals, for Dot Probe Tasks across all groups are shown in Table 18 and in Figure 32 - Figure 35.

RT data from the Dot Probe Task were subjected to a two-way mixed ANOVA on 20% trimmed means using the `bwtrim`-function, implemented in R (A. Field et al., 2012; Wilcox, 2012). Group was the between-subjects factor,
whereas the location of the probe (CT, IT) was accounted for as a repeated within-subjects factor. Threat bias scores were calculated as mean RT for targets in neutral word locations (IT) minus mean RT for targets in threat word locations (CT) to simplify presentation.

At 50 msec exposure time, there was no significant main effect of group on RT, $Q = 0.45$, $p = .716$, no significant main effect of location of the probe, $Q = 0.33$, $p = .569$, and no significant interaction effect (Group x Target location of the probe) on RT, $Q = 1.86$, $p = .141$.

At 200 msec exposure times, reaction time performance yielded no significant main effect of group $Q = 0.99$, $p = .400$, no significant main effect of the target location of the probe $Q = 0.37$, $p = .546$, and no significant interaction effect between group and the target location of the probe, $Q = 0.93$, $p = .430$.

![Figure 32: Means and 95% confidence intervals of 50 msec Dot Probe Task scores across all groups of Study 3.](image)
**Figure 33:** Means and 95% confidence intervals of 200 msec Dot Probe Task scores across all groups of Study 3.

**Figure 34:** Means and 95% confidence intervals of mean bias score of 50 msec Dot Probe Task of Study 3.
Figure 35: Means and 95% confidence intervals of mean bias score of 200 msec Dot Probe Task of Study 3.
<table>
<thead>
<tr>
<th></th>
<th>Mean RT(IT)</th>
<th></th>
<th>Mean RT(CT)</th>
<th></th>
<th>Bias score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>LL</td>
<td>UL</td>
<td>Mdn</td>
</tr>
<tr>
<td>Dot probe 50 msec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police (n = 73)</td>
<td>523.72</td>
<td>104.98</td>
<td>499.23</td>
<td>548.22</td>
<td>498.82</td>
</tr>
<tr>
<td>Martial Arts (n = 50)</td>
<td>532.11</td>
<td>159.49</td>
<td>486.78</td>
<td>577.44</td>
<td>504.61</td>
</tr>
<tr>
<td>Police &amp; Martial Arts (n = 32)</td>
<td>517.26</td>
<td>106.15</td>
<td>478.99</td>
<td>555.53</td>
<td>484.76</td>
</tr>
<tr>
<td>Control (n = 48)</td>
<td>526.99</td>
<td>147.94</td>
<td>484.04</td>
<td>569.95</td>
<td>478.79</td>
</tr>
<tr>
<td>Dot probe 200 msec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police (n = 73)</td>
<td>469.85</td>
<td>65.25</td>
<td>454.63</td>
<td>485.08</td>
<td>452.76</td>
</tr>
<tr>
<td>Martial Arts (n = 50)</td>
<td>479.52</td>
<td>103.46</td>
<td>450.11</td>
<td>508.92</td>
<td>463.19</td>
</tr>
<tr>
<td>Police &amp; Martial Arts (n = 32)</td>
<td>471.84</td>
<td>83.53</td>
<td>441.72</td>
<td>501.95</td>
<td>448.06</td>
</tr>
<tr>
<td>Control (n = 48)</td>
<td>497.59</td>
<td>97.58</td>
<td>469.26</td>
<td>525.93</td>
<td>466.54</td>
</tr>
</tbody>
</table>
6.5.4 Emotional Stroop Task

Participants scores of the Emotional Stroop Task across the different groups are shown in Table 19.

Reaction times of the Emotional Stroop Task were entered into a 4 x 2 mixed ANOVA on 20% trimmed means with repeated measures on the second factor. Results were computed using the *bwtrim*-function, implemented in R (A. Field et al., 2012; Wilcox, 2012). Group was the between-subjects factor, comprising of 4 levels. The second factor was the repeated measure, comprising of the RT of neutral words versus RT of the emotionally themed word under investigation. Four separate ANOVAs, one for each word class (aggression-themed, colour, negative, positive) were computed. Bias scores were calculated as mean RT for emotionally labels words minus mean RT for neutrally labelled words to simplify presentation.

For aggression themed words, RT were not significantly affected by group, $Q = 0.41, p = .746$, by the difference between the emotional value of the words, $Q = 0.02, p = .885$, or by the interaction between group status and emotional value of the words, $Q = 0.95, p = .421$. For colour words, results showed no significant main effect for group status, $Q = 0.85, p = .469$, no significant main effect for the emotional value of the words, $Q = 1.87, p = .174$, and no significant Group x Emotional Value interaction effect, $Q = 0.07, p = .975$ on RT. For words with negative value, results yielded no significant main effect for group status, $Q = 0.70, p = .556$, no significant main effect for the emotional value of the words, $Q = 0.00, p = .966$, and no significant interaction effect (Group x Emotional Value). $Q = 0.58, p = .629$ on RT. For positive words, RT were not significantly affected by group status, $Q = 0.62, p = .597$, by the emotional value of the words, $Q = 2.73, p = .100$, or by the interaction between group status and the emotional value of negative words compared to neutral ones, $Q = 0.61, p = .61$
### Table 19: Means, Standard Deviations and 95% Confidence Intervals of Emotional Stroop Scores of Study 3

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
<th>Mdn</th>
<th>LQ</th>
<th>UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RT aggression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>731.95</td>
<td>157.65</td>
<td>695.43</td>
<td>695.43</td>
<td>703.71</td>
<td>617.82</td>
<td>809.00</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>719.34</td>
<td>142.32</td>
<td>678.89</td>
<td>678.89</td>
<td>694.67</td>
<td>607.99</td>
<td>839.60</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>690.80</td>
<td>129.68</td>
<td>644.82</td>
<td>644.82</td>
<td>686.46</td>
<td>584.24</td>
<td>810.92</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>717.93</td>
<td>161.43</td>
<td>671.05</td>
<td>671.05</td>
<td>678.91</td>
<td>614.40</td>
<td>779.65</td>
</tr>
<tr>
<td><strong>RT color</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>761.73</td>
<td>186.20</td>
<td>718.59</td>
<td>718.59</td>
<td>721.15</td>
<td>628.93</td>
<td>847.79</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>716.26</td>
<td>150.89</td>
<td>673.38</td>
<td>673.38</td>
<td>691.10</td>
<td>593.30</td>
<td>823.45</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>696.22</td>
<td>125.77</td>
<td>651.62</td>
<td>651.62</td>
<td>688.00</td>
<td>603.19</td>
<td>772.00</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>742.72</td>
<td>202.22</td>
<td>684.00</td>
<td>684.00</td>
<td>705.05</td>
<td>637.60</td>
<td>807.20</td>
</tr>
<tr>
<td><strong>RT negative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>733.67</td>
<td>159.84</td>
<td>696.64</td>
<td>696.64</td>
<td>711.65</td>
<td>604.49</td>
<td>822.05</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>702.26</td>
<td>130.88</td>
<td>665.06</td>
<td>665.06</td>
<td>699.70</td>
<td>593.32</td>
<td>772.15</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>679.89</td>
<td>115.44</td>
<td>638.96</td>
<td>638.96</td>
<td>686.05</td>
<td>580.33</td>
<td>764.12</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>743.74</td>
<td>217.02</td>
<td>680.72</td>
<td>680.72</td>
<td>694.20</td>
<td>623.80</td>
<td>840.47</td>
</tr>
<tr>
<td><strong>RT neutral</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>736.31</td>
<td>157.39</td>
<td>699.85</td>
<td>699.85</td>
<td>719.19</td>
<td>628.02</td>
<td>806.45</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>711.17</td>
<td>137.60</td>
<td>672.06</td>
<td>672.06</td>
<td>695.71</td>
<td>600.27</td>
<td>794.84</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>680.41</td>
<td>102.98</td>
<td>643.90</td>
<td>643.90</td>
<td>678.74</td>
<td>609.23</td>
<td>745.73</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>724.97</td>
<td>175.72</td>
<td>673.95</td>
<td>673.95</td>
<td>675.50</td>
<td>616.55</td>
<td>802.17</td>
</tr>
<tr>
<td><strong>RT positive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>711.45</td>
<td>133.40</td>
<td>680.54</td>
<td>680.54</td>
<td>691.97</td>
<td>610.68</td>
<td>786.23</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>711.17</td>
<td>137.60</td>
<td>672.06</td>
<td>672.06</td>
<td>695.71</td>
<td>600.27</td>
<td>794.84</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>676.38</td>
<td>118.26</td>
<td>634.45</td>
<td>634.45</td>
<td>664.46</td>
<td>595.35</td>
<td>761.42</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>723.39</td>
<td>209.77</td>
<td>662.47</td>
<td>662.47</td>
<td>668.98</td>
<td>594.85</td>
<td>769.21</td>
</tr>
<tr>
<td><strong>MBS aggression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>-4.36</td>
<td>82.93</td>
<td>-23.58</td>
<td>14.85</td>
<td>-4.45</td>
<td>-65.41</td>
<td>49.51</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>-2.17</td>
<td>77.71</td>
<td>-24.26</td>
<td>19.91</td>
<td>5.49</td>
<td>-40.86</td>
<td>43.07</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>10.39</td>
<td>76.67</td>
<td>-16.79</td>
<td>37.58</td>
<td>-7.08</td>
<td>-38.17</td>
<td>60.09</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>-7.04</td>
<td>71.69</td>
<td>-27.86</td>
<td>13.77</td>
<td>-0.63</td>
<td>-36.15</td>
<td>35.99</td>
</tr>
<tr>
<td><strong>MBS color</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>25.42</td>
<td>82.50</td>
<td>6.30</td>
<td>44.53</td>
<td>18.62</td>
<td>-31.05</td>
<td>67.08</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>-5.25</td>
<td>10.01</td>
<td>-36.23</td>
<td>25.73</td>
<td>7.23</td>
<td>55.40</td>
<td>48.37</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>15.81</td>
<td>79.94</td>
<td>-12.54</td>
<td>44.15</td>
<td>9.71</td>
<td>-30.33</td>
<td>62.95</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>17.75</td>
<td>78.76</td>
<td>-5.12</td>
<td>40.62</td>
<td>10.96</td>
<td>-38.34</td>
<td>73.07</td>
</tr>
<tr>
<td><strong>MBS negative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>-2.64</td>
<td>89.83</td>
<td>-23.45</td>
<td>18.17</td>
<td>-7.72</td>
<td>-52.63</td>
<td>42.22</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>-0.52</td>
<td>58.29</td>
<td>-21.19</td>
<td>20.15</td>
<td>7.33</td>
<td>-40.44</td>
<td>31.88</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>18.77</td>
<td>102.22</td>
<td>-10.92</td>
<td>48.45</td>
<td>15.98</td>
<td>-23.10</td>
<td>66.57</td>
</tr>
<tr>
<td><strong>MBS positive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>74</td>
<td>-24.86</td>
<td>87.73</td>
<td>-45.19</td>
<td>-4.54</td>
<td>-27.06</td>
<td>-52.57</td>
<td>20.65</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>-10.34</td>
<td>95.67</td>
<td>-37.53</td>
<td>16.85</td>
<td>-3.75</td>
<td>-54.26</td>
<td>52.03</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>-4.03</td>
<td>65.33</td>
<td>-27.19</td>
<td>19.14</td>
<td>-16.74</td>
<td>-35.69</td>
<td>35.85</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>-1.59</td>
<td>92.17</td>
<td>-28.35</td>
<td>25.18</td>
<td>-23.43</td>
<td>-41.23</td>
<td>47.92</td>
</tr>
</tbody>
</table>

RT: reaction time; MBS: mean bias score
Separate Kruskal-Wallis tests on the emotional Stroop scores revealed no significant effects of group on the RT of aggression-themed words, $H(3) = 1.10, p = .777$, colour words, $H(3) = 2.81, p = .422$, negative words, $H(3) = 2.50, p = .475$, neutral words, $H(3) = 2.17, p = .537$ and positive words, $H(3) = 1.39, p = .708$. Also, the group variable did not significantly affect the average bias scores for aggression-themed words, $H(3) = 0.69, p = .875$, colour words, $H(3) = 1.87, p = .599$, negative words, $H(3) = 3.94, p = .268$, and positive words, $H(3) = 3.06, p = .382$.

The difference in RT of all words classes across all groups is depicted in Figure 36, the differences in bias scores in Figure 37.

![Figure 36: Mean and 95% confidence intervals of RTs of word classes across all groups of Study 3.](image-url)
Figure 37: Mean and 95% confidence intervals of bias scores across all groups of Study 3.

### 6.5.5 Balloon Analogue Risk Task

Means, standard deviations and 95% confidence intervals of the AAPC are displayed in Table 20.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Police</td>
<td>72</td>
<td>32.26</td>
<td>14.69</td>
<td>28.81</td>
<td>35.72</td>
</tr>
<tr>
<td>Martial Arts</td>
<td>50</td>
<td>32.89</td>
<td>13.86</td>
<td>28.95</td>
<td>36.83</td>
</tr>
<tr>
<td>Police &amp; Martial Arts</td>
<td>33</td>
<td>29.17</td>
<td>14.13</td>
<td>24.16</td>
<td>34.19</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>30.64</td>
<td>11.51</td>
<td>27.30</td>
<td>33.98</td>
</tr>
<tr>
<td>Total</td>
<td>203</td>
<td>31.53</td>
<td>13.66</td>
<td>29.64</td>
<td>33.42</td>
</tr>
</tbody>
</table>

Table 20: Means, Standard Deviations and 95% Confidence Intervals of the Adjusted Average Pump Count across all Groups of Study 3
In order to investigate if groups differed in the AAPC, a one-way ANOVA was performed. Results yielded no significant difference between the groups, $F(3, 202) = 0.63, p = .600$. Means and 95% confidence intervals of the AAPC are depicted in Figure 38.

Figure 38: Means and 95% confidence intervals of Adjusted Average Pump Count of all groups of Study 3.
6.6 Discussion

The current study investigated the effects of job-related threat exposure on various aspects of cognition. Specifically, whether regular exposure to real threats which is typical of police officers or regular exposure to threats in training settings could lead to the development of threat-related attentional biases and an increased propensity for risk-taking. The study demonstrated two main results: First, neither police officers nor martial arts practitioners show a threat-related attentional bias or a propensity for risk taking, as indicated by the Dot Probe Task at 50 msec and 200 msec exposure times, the Emotional Stroop Task and the BART. Thus, Hypothesis 1 and 3 have to be rejected. Second, individuals that deal with threat on a regular basis, be it mainly in real (police officers) or in training settings (martial artists), showed less trait anxiety than the control group, fails to support Hypothesis 2.

6.6.1 Threat-Related Attentional Biases in Police Officers

There are several possible reasons for the null result regarding the hypothesised differences in threat-related attentional biases in police officers. It has to be acknowledged, that possible explanations of the results are due to a true lack of difference between the groups or due to noise in the data. For example, some participants took part in the experiment via the internet, providing many potential confounds that could not be controlled for (e.g. quiet environment, no distractions). Therefore, all further explanations of the findings have to be viewed in the light of the interpretation of a null result.

There are two additional interpretations of the failure to find a threat-related attentional bias in the current sample: (i) regular exposure to threat does not lead to any attentional bias, or (ii) regular exposure does lead to attentional bias but the current experiment did not capture that for some reason.

The findings contradict previous claims (Öhman, 1996; Öhman & Mineka, 2001; Öhman & Wiens, 2004), that biases develop as a result of survival relevance and that they are a conditioned response to decrease the likelihood of harm that is caused by these threats. In line with this, Bar-Haim
et al. (2010) presented data of the impact of war-related stress gathered during ongoing rocket attacks on RT in the Dot Probe Task. Their results clearly indicated that individuals who are exposed to severe life-threatening risks manifest attentional biases, specifically, biases away rather than towards the threat at exposure times of 1000 msec. Likewise R. M. Todd et al. (2015) showed (using an attentional blink task) that combat veterans (who served in Afghanistan) with and without post-traumatic stress disorder, both showed an attentional bias towards combat-related words. Again, this suggests that exposure to life-threatening situations facilitate the development of cognitive biases towards threat. In the described cases (inhabitants of areas that are regularly attacked by rockets and combat experienced soldiers) a threat-related attentional bias can be considered as functional.

In the current study, the sample under investigation consisted of police officers. Exposure to regular threat was assumed due to the experience in routine activity as a police officer. Even though studies have confirmed that police officers are at risk of being victims of violence (Bochenek & Staller, 2014; Ellrich & Baier, 2014; 2016; Ellrich, Baier, & Pfeiffer, 2010a), experience as a police officer does not necessarily account for experience with violent confrontations and the associated competencies (Schmalzl, 2008). Hence, officers in the current study had a mean of 9.28 years of work experience (7.35 years for the police officer and martial arts group), which does not necessarily reflect the amount of exposure to violent threats. Concerning martial artists, to my knowledge, no studies have looked at the prevalence of engagement in violent confrontations of this population so far. It is reasonable to assume, that individuals, that work in different occupations than law enforcement, military or security are exposed physical conflict situations on a lower level. In sum, a lack of regular exposure to life-threatening situations may serve as a possible explanation for the results.

Furthermore, the simulation of violent situations as conducted in police use of force training and martial artist classes has not lead to the development of threat-related attentional biases either. This could be explained by: (a) a lack of representativeness in the training settings, and/or (b) the time that
individuals have been exposed to such situations. Various researchers and practitioners in the martial arts and the police use of force domain have pointed out that training for real world encounters has to be more representative (Jager et al., 2013; K. R. Murray, 2004; W. A. Norris & Wollert, 2011; Renden, Nieuwenhuys, Savelbergh, & Oudejans, 2015b; Staller & Zaiser, 2016; J. Wagner, 2005; Wollert et al., 2011; D. Young & Bedard, 2009), implying that current practices are not representative enough to foster the acquisition of skills that are needed for violent encounters.

Furthermore, it may be possible, that individuals did not spend enough time in representative learning environments that allow for the functional development of threat-related attentional biases. This may be due to the low amount of practice activities within a given training setting or due to the low amount of occasions the individual participates. For example, a recent study investigating self-defence systems in Germany showed that martial arts schools advertising self-defence practises do not regularly incorporate representative training activities (Staller, Bertram, Althaus, Heil, & Klemmer, 2016). In the context of policing, the amount of practice, and as such exposure to simulated violent encounters, is limited by agencies’ policies. In the state of Hesse, regular police officers are subjected to use of force training for two days per year.

Another reason explaining the non-existence of any threat-related attentional bias in police officers may lie in the used words in the Dot Probe and Emotional Stroop Task. Threat related words in the Dot Probe Task were taken from MacLeod et al. (2007). The word set included words such as ‘panicky’, ‘confused’, ‘embarrassed’, as well as words as ‘assault’, ‘attack’ and ‘strangled’. The latter ones are clearly more combat-related than the former ones. The same is true for the Emotional Stroop Task, which was used as described by (P. Smith & Waterman, 2005): Aggression-themed words included words such as ‘temper’, ‘guilt’ or ‘annoyed’ as well as more combat-related words as ‘kick’, ‘slash’ or ‘kill’. In terms of a functional development of a threat-related attentional bias, it may be possible that the words used were too unspecific in the context of violent encounters. Since in both tasks combat related words were very rare compared to threat-themed
words in general, RTs of these words were not subject to further statistical analysis. In order to foster optimal decision making, the learner in any context has to be provided with the correct cues (Maran & Glavin, 2003; Staller & Abraham, 2016). If the development of threat-related attentional biases in police officers is considered as functional, then cues are only valid as long as they reflect ‘correct’ ones (i.e. cues that are present during a violent encounter). Therefore, a functional threat-related attentional bias may not be elicited because of the lack of representativeness in the psychological tests used. This line of argument is supported by the results of R. M Todd et al. (2015), who used combat-related words and subsequently found a threat-related attentional bias in combat-experienced personal. Results of this study seem to stem from the individual’s experience in potentially life-threatening environments.

However, besides being a result of experiences, human defensive responses are ‘hard-wired’, involving different brain structures for processing threats (Adolphs, 2013). Even though a small set of necessary and sufficient structures has not yet emerged, several brain regions, such as the amygdala, the periaqueductal gray, the hypothalamus and the prefrontal cortices, have been shown to be involved in responding to threats (Adolphs, 2013; Fanselow & Lester, 1988; Glascher et al., 2012; McNaughton & Corr, 2004; Ongur, 2000; Pessoa & Adolphs, 2010; Price, 2003; L. M. Williams et al., 2001; Zald, 2003). Through evolution several threat detection parameters have developed, such as the visual presentation of snakes or spiders (Öhman & Mineka, 2001). However, individuals are able to learn through experience to perceive other stimuli as threatening and/or dangerous (Adolphs, 2013; R. M. Todd et al., 2015). Lienard (2011) argues that this is mainly because of the slow life history of humans, allowing for the adaption of precautionary and defence systems to the individual’s age, status, and conditions in life (Abramowitz, Schwartz, Moore, & Luenzmann, 2003; Boyer & Lienard, 2006a; 2006b; Leckman et al., 2004; Lienard, 2011; Rachman & de Silva, 1978). Taken together, functional human defensive behaviour in violent encounters with regards to threat perception comprises two components: ‘hard-wired’ threat-processing structures with pre-set
parameters and threat-detecting parameters forged by specific experiences (with valid cues).

In sum, the current results indicate, that the groups did not show dysfunctional threat-related attentional bias, and thus any symptoms of a disorder of the defensive system (i.e. anxiety disorder). Yet, they did not show any functional threat-related attentional bias either, suggesting further research to examine the influence of: (a) training practices in police and martial arts training, (b) exposure to real threats of police officers, and (c) optimisation of the Dot Probe and Emotional Stroop Task in order to account for more specific threats for violent encounters.

6.6.2 Risk Taking in Police Officers

Risk taking behaviour did not significantly differ across groups, fails to support Hypothesis 3 that an exposure to routine threat leads to an increase in the overall propensity for risk taking. This finding has two implications: (i) exposure to situations of mild threat such as combat scenarios which martial artists routinely engage in (Staller et al., 2016) does not lead to an increase in risk taking, and (ii) exposure to real threat is also inadequate in producing an increase in risk-taking behaviour as indicated by the scores of police officers. Considering a substantial component of police training and martial arts training consists of applying reasonable force, the role of training may act as a mediating factor in situations of real and simulated threat and reduces risk-taking behaviour.

6.6.3 Differences in State and Trait Anxiety

Trait anxiety was significantly higher in the control group than in all the other tested groups disproving Hypothesis 2. A possible moderator of this effect could be the training participants take part in. Martial arts training has been shown to be capable of reducing anxiety in participants (Fuller, 1988; Trulson, 1986). Since police use of force training also contains ‘hand to hand’ combat techniques, it may be possible that the perceived improvement in self-protection skills may lead to reduced anxiety. However, several studies investigating state and trait anxiety of officers compared to adult normative
samples (Newman & LeeAnne Rucker-Reed, 2004; Storch & Panzarella, 1996) showed no differences between the groups. Hence, the current findings contradict these previous results.

Another possible explanation for the differences in trait anxiety between the control and the other groups may be due to a self-selection process due to the motivation to pursue a career in policing and/or to engage in martial arts. Since the engagement in (simulated) physical conflict is part of both domains, it may be possible that individuals with higher trait anxiety may be less motivated to work as a police officer or practice martial arts.

Furthermore, it is noteworthy, that previous research (Newman & LeeAnne Rucker-Reed, 2004) found a trait anxiety mean of 32.94 in US marshals compared to a mean of 35.55 in the normative sample of working male aged 19 to 39 (Spielberger et al., 1983), whereas the trait anxiety mean of police officers in the current study was 43.64 compared to the control group with a mean of 47.52. With these score the police officers are in the upper 20% of the German normative sample of working males aged 30 to 59. The scores of the control group in the current study reflects the upper 10% of the same normative sample. Taken together, the current groups display much higher trait anxiety than the normative sample, even if the police officer and martial artists demonstrate significant lower score than the current control group. A possible explanation for the difference between police officers (and martial artists) and the control group has been presented above. The difference between the current groups and the normative sample cannot be explained easily. Since all tests were administered the same way, and thus the STAI was filled out after threat-related attentional bias and risk-taking were measured, future studies should measure STAI at the beginning of the test battery in order to account for possible priming effects.

6.6.4 Limitations

There are several limitations that have to be acknowledged. First, since various locations were used for testing, a standardised environment could not be established. This has been discussed as being a potential confound, especially when relating to measures of attention and RT (Bimbaum, 2004).
Second, as discussed earlier, it may be possible, that the words used in the Dot Probe Task and the Emotional Stroop Task do not account for functional threat-related attentional bias in the context of violent confrontations. Future studies should consider using combat-related words (R. M. Todd et al., 2015) in order to capture possible effects. Third, the actual exposure to real life-threatening situations of the participants was not measured. Similarly, participants were only asked about their years of experience in martial arts training, which did not account for the kind of practices and how often participants engaged in them. Therefore, future research on this topic should include data about the exposure to real life threats and the amount and the representativeness of combat training (martial arts and police use of force training). Fourth, threat-related attentional measures in the current study measured cognitive biases towards and away from threat. However, variability within each session was not accounted for. Since attentional bias variability is suggested to be a useful marker of attentional impairment with regards to post-traumatic stress disorder (Iacoviello et al., 2014), future studies should incorporate threat-related attentional bias variability (Naim et al., 2015; Zvielli et al., 2014). Finally, since some of the participants completed the testing battery at home via the internet and given that the whole session took about 40 minutes, there is the possibility of distractions during testing. Although participants were asked to make sure that they will not get distracted, this cannot be controlled for (Birnbaum, 2004).

6.7 Summary of Chapter 6

Study 3 showed no differences between police officers, martial artists, and controls for threat-related attentional biases and the propensity of risk-taking, indicating no signs of a functional development of a cognitive bias. However, in conjunction with previous research on combat-related attentional bias, it may be possible that a possible functional attentional bias towards combat-related cues was not discovered. Furthermore, the control group showed higher trait anxiety than police officers and martial artists, which may account for the lowering effect of use of force and martial arts training on anxiety.
Chapter 7: Ego Depletion
7 Ego Depletion

7.1 Introduction

Humans display a remarkable capacity to inhibit aggressive urges that would be associated with negative outcomes in the long run. For example, despite the temptation to shout at his wife in an emotional discussion, a husband may choose to suppress emotionally-driven behaviour in order not to risk negative influences on the romantic partnership. Likewise, a football player may have the strong desire to retaliate a foul against him, but be able to resist these urges in order to continue playing instead of being red-carded due to aggressive behaviour towards another player.

In the context of policing, low levels of self-control are related to police misconduct in general (Donner & Jennings, 2014). Even in situations where use of force may be legitimate, the self-regulation of police officers by using non-aggressive behaviours may prove beneficial (Zaiser & Staller, 2015). Although police officers have the ability to achieve compliance through physical force, such tactics generally result in only momentary compliance and can induce disrespect and violence from the subject (Reisig et al., 2004; L. W. Sherman, 1993; Tyler & Huo, 2002; Wolfe, 2011). To date, there is a lack of research addressing possible relationship between self-control failure and the use of force. Results of Klukkert et al. (2009) investigating the motivation of German police officers to use force, provided the first results that behavioural conflict situations in daily police work can be emotionally driven as compared to behaviour that is determined by legal and institutional guidelines, reflecting rational behaviour. The researchers concluded that “the more the officers are drawn into the whirlpool of conflicts between maintaining authority, on the one hand, and the fear of escalation on the other hand, the more their actions will be determined by emotions, and the higher the probability that legal guidelines will be ignored and that a violent response will serve as a mechanism to resolve the conflict” (p. 199).

To date only a few studies investigated the effects of a loss of self-control on aggressive behaviour. The current chapter aims at investigating the loss of
self-control in police officers through ego depletion and whether that results in aggressive behaviour.

7.1.1 Aim of the Study

The primary aim of Study 4 was to extend the findings of previous research on self-control and aggression that has demonstrated that ego depletion affected aggressive behaviour in terms of indirect or covert forms of aggression. The current study investigated if these findings transfer to the policing domain, where aggression is displayed directly and in physical forms. The secondary aim was to investigate depletion manipulations that have been found to work with different populations, would also work with police officers. Third, the taxing of self-control resources may temporarily impair aspects of executive cognitive functioning is involved in behavioural regulation, such as inhibitory control (Miyake et al., 2000). Therefore, the current study also employs measures of executive functioning in order to assess the effect of inhibitory control on executive functioning performance. The fourth aim of the study was to investigate the effects of exercising self-control in a naturalistic task on HRV, since this variable has been suggested as a biological correlate of self-control (Baumeister, Vohs, & Tice, 2007; Zahn et al., 2015).

Study 4 consisted of two sub-studies. In Study 4.1, the ego depletion manipulation consisted of watching a disgusting video whilst suppressing any emotional reaction to it (Video Task). The Cold Pressor Task was used in Study 4.2 as the manipulation. Participants then would engage in a use of force scenario with a provocative role player.

7.1.2 Hypotheses

To this end, the following hypotheses were put forward for Study 4.1 and Study 4.2:

1. The depleting effect of the experimental manipulation (Video Task in Study 4.1, Cold Pressor Task in Study 4.2) would produce changes in the manipulation check measures (Hypothesis 1).
2. Ego depleted participants would display offensive aggression earlier in the course of the scenario than non-depleted participants (Hypothesis 2).

3. Ego depletion would impair all measures of executive function (Hypothesis 3).

4. Ego depletion would affect vagal tone as a biological marker of self-control (Hypothesis 4).

7.2 Methods

7.2.1 Participants

A total of 44 police officers (male: \( n = 39 \), female: \( n = 5 \)) of the Hessian state participated in Study 4.1. Participants were randomly assigned to either the experimental or control group, each consisting of 22 police officers. Officers of the experimental group had a mean age of \( M = 28.18 \) years \( (SD = 4.17) \) and a mean of experience as a police officer of \( M = 6.51 \) years \( (SD = 3.87) \). Three officers were female. The control group was on average \( M = 26.97 \) (\( SD = 3.73 \)) years old and worked as police officers for \( M = 5.79 \) (\( SD = 3.12 \)) years. Two female officers were assigned to the control group. In Study 4.1 two cases of missing of cardiovascular data were identified. Therefore, analysis of RMSSD and mean HR data was conducted with a reduced sample size of 42 police officers with 22 individuals in the experimental group (age: \( M = 28.18, SD = 4.17 \), experience: \( M = 6.51, SD = 3.87 \)) and 20 individuals in the control group (age: \( M = 26.36, SD = 5.23 \), experience: \( M = 5.23, SD = 1.97 \))

In study 4.2 a total of 37 police officers (\( n = 30 \), female: \( n = 7 \)) took part. Participants were randomly assigned to two different groups: ego depletion \( (n = 19) \) and control group \( (n = 18) \). Police officers of the experimental group were on average \( M = 28.06 \) \( (SD = 3.87) \) years old and worked as police officers for on average of \( M = 5.98 \) (\( SD = 2.57 \)) years. Four of them were female. The control group consisted of police officers with a mean age of \( M = 27.00 \) (\( SD = 3.23 \)) years and average work experience of \( M = 5.54 \) (\( SD = 1.76 \)) years. Three officers of the control group were female. Three sets of
missing cardiovascular data were identified and excluded from analysis. Hence, for the analysis of RMSSD and mean HR data the sample size was reduced to 34 police officers with 18 participants in the experimental group (age: $M = 27.94$, $SD = 3.95$; experience: $M = 5.95$, $SD = 2.64$) and 16 participants in the control group (age: $M = 27.27$, $SD = 3.38$; experience: $M = 5.95$, $SD = 2.64$).

7.2.2 Materials

7.2.2.1 Brief Mood Introspection Scale

The Brief Mood Introspection Scale (BMIS) is a self-report measure of mood and arousal (see Appendix). The BMIS (Mayer & Gaschke, 1988) consists of 16 adjectives which are responded to on 4 point Likert scales ranging from ‘definitely do not feel’ (1) to ‘definitely feel’ (4). Four underlying mood factors are derived: Pleasant-Unpleasant, Arousal-Calm, Positive-Tired, and Negative-Relaxed. Factor scores are derived by adding or subtracting scores from relevant items. For example, the Pleasant-Unpleasant factor is computed by subtracting values for unpleasant adjectives (e.g. Grouchy, Sad) from scores derived from pleasant adjectives (e.g. Content, Happy). Cronbach’s alpha for the subscales of the BMIS range between 0.76 and 0.83 (i.e. acceptable to good).

7.2.2.2 Cardiovascular Measures

Cardiovascular function was measured throughout the studies. Vagal tone, as the physiological parameter that is linked to self-control, was measured via the HRV parameter of RMSSD (Laborde & Mosley, 2016; Porges, 2007; J. F. Thayer et al., 2009). Time intervals of 60 seconds were chosen for HRV analysis due to the dynamics of the scenario (see general methods chapter for more detail). Furthermore, mean HR at the time points of measurements was calculated as a measure of psychophysiological arousal.

7.2.2.3 Video Stimuli for Ego Depletion Manipulation (Video Task) in Study 4.1

An 18 minute section from the film Audition (Miike, 1999) was selected to use for the ego depletion manipulation. The entire 18-minute section contained
numerous flashbacks which were edited so the final clip was 5 minutes long. The clip contained graphic depictions of torture (for example, a character has his feet amputated with wire). This clip was used as disgust is more easily and ethically manipulated than emotions such as happiness or sadness (Gross & Levenson, 1993) and has been shown to be an effective stimuli for the experimental condition (P. Christiansen, Cole, & Field, 2012).

7.2.2.4 Cold Pressor Task for Ego Depletion Manipulation in Study 4.2
An ice-cooled water bath (24 litre) with a temperature set at 1°C was used to induce acute pain. The time of cold water immersion was measured using a stop watch in seconds.

7.2.2.5 Employed Scenario in Study 4.1 and 4.2
The scenario, participants were required to complete in study 4.1 and 4.2., was developed by experiences use of force trainers of the Hessian State Police. Participants were required to fulfil their assignment as police officers on a routine basis.

Participants were assigned to a fixed spot in the training environment and briefed about their assignment for the upcoming scenario training. The assignment involved a cordon on a specific area behind the officer because of an unspecified police operation that is taking place there. As soon as the scenario started a role player presented himself at the scene. The role player followed a pre-defined script, involving several levels of behaviour for 30 seconds each. The general intention of the role player was to pass through the cordoned area. Every 30 seconds the level increased in terms of provocation and threatening behaviour towards the participants in a scripted manner. After each 30 seconds the role player immediately went on to the next level and showed the pre-defined behaviour. Characteristics of each level are displayed in Table 21 and Figure 39. Each scenario culminated in a situation where the participant had to control the role player physically, since he wanted to pass through the cordoned area. Scenarios were filmed using a GoPro Hero4 camera (GoPro Inc.). The dependent variable was the time from the onset of the scenario, when the role player appeared for the first
time until the participants showed offensive aggression for the first time.
Offensive aggression in the context of the scenario was defined as any visible behaviour to physically control the role player (e.g. take down) or to obtain dominance over him (e.g. closing the distance to the role player by leaving the pre-defined spot, where the area should be cordoned). The dependent variable was timed based on the video footage of participants’ performance.
<table>
<thead>
<tr>
<th>Level</th>
<th>Time (seconds)</th>
<th>Distance (meters)</th>
<th>General behaviour of role player</th>
<th>Verbal communication by role player</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 – 30</td>
<td>8</td>
<td>Does not talk; provocative look in the direction of police officer; does not react to any</td>
<td>No verbal communication</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>communication initiate by the police officer</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>30 – 60</td>
<td>8</td>
<td>Wants explanation for the police measures; does not relate to the police officer; has the</td>
<td>“What's the point here?”; “Explain yourself!”; “I only want to go home”;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>impression that measure is only to annoy him</td>
<td>“It's always the same with the police”; “You're feeling strong with your</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>weapon, don’t you?”; “You think you can do anything because of your</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uniform?”</td>
</tr>
<tr>
<td>3</td>
<td>60 – 90</td>
<td>8</td>
<td>Uses personal provocations; does not react if he is sent off by the police officer</td>
<td>“You only want to annoy me!”; “You cops are kind of stupid”; “You think</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I am stupid, don’t you?”; “Seems like they employ everybody as an police</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>officer!”; “Standing around – that’s what you were made for!”; “You’re</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>provokating me!”</td>
</tr>
<tr>
<td>4</td>
<td>90 – 120</td>
<td>5</td>
<td>Shortens the distance; uses personal provocation</td>
<td>See Level 3;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Additionally: “You’re a looser – you cannot do anything!”</td>
</tr>
<tr>
<td>5</td>
<td>120 – 150</td>
<td>3</td>
<td>Shortens the distance; uses personal provocation</td>
<td>See Level 3 and Level 4;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Additionally: “You are all stupid fucks!”</td>
</tr>
<tr>
<td>6</td>
<td>150 – 180</td>
<td>3</td>
<td>Maintains the distance; shouts at the police officer</td>
<td>“Fuck off!”; “Take your stuff and piss off!”; “Take your buddies and hit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>the road!”; “I help getting your stuff!” If police officer says “Calm</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>down!”, role player shouts: “I am calm, you shout!”</td>
</tr>
<tr>
<td>7</td>
<td>180 – 210</td>
<td>0-1</td>
<td>Pushes police officer at the arm; firm voice (but not shouting)</td>
<td>“Get off!”; “I help you getting away from here!”; “Take your buddies and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hit the road!”</td>
</tr>
<tr>
<td>8</td>
<td>210 – 240</td>
<td>0-1</td>
<td>Continues pushing police officer; shouts loudly</td>
<td>“Take you stuff and fuck off!”; “Get lost!”; “Take you buddies and piss</td>
</tr>
<tr>
<td>9</td>
<td>240 – 270</td>
<td>0-1</td>
<td>Pushes officer on the chest; continues shouting</td>
<td>See Level 8</td>
</tr>
</tbody>
</table>
Figure 39: Levels of the scenario design.
7.2.3 Procedure

7.2.3.1 General Procedure of Study 4.1 and Study 4.2

General procedures for Study 4.1 and Study 4.2 were the same. Studies only differed in the task that was used to induce a state of ego depletion in the participants. Testing sessions took place at the Hessian State Police Academy between 8am and 6pm. Each participant only attended one study. Diagrams illustrating key features of both experimental sessions (Study 4.1 and 4.2) are shown in Figure 40 and Figure 41.

Participants were briefed about the study and provided their informed consent in the morning, before being equipped with a Suunto HR memory belt. They were asked to sit quietly for 10 minutes in order to obtain baseline measure of physiological data. They then were tasked with the RWT, which is the German Version of the Controlled Order Word Association Test (see general methods for more detail). At pretesting (before the scenario) participants were asked to provide as many words a possible within 120 seconds beginning with the letters B and M. Participants were randomly assigned to either the experimental or the control condition. The experimental group then completed a task that involves some kind of self-control. The control group received a similar task, but were not tasked to control their behaviour, thoughts or emotions. In Study 4.1 the Video Task was employed, in Study 4.2 participants had to complete the Cold Pressor Task. After the ego depletion and control treatment, respectively, participants had to complete a scenario involving the handling of a provocative citizen, played by an experienced use of force trainer. Immediately after the scenario, participants completed the post-testing of the RWT with the letters P and K. Upon completion of this task, they were assigned to a resting area, where they should relax in a sitting position for 10 minutes in order to obtain a recovery measure for physiological data. Participants then were debriefed and thanked for their participation.
Figure 40: Schematic overview of experimental procedure in Study 4.1

Figure 41: Schematic overview of experimental procedure in Study 4.2
7.2.3.2 Experimental vs. Control Condition in Study 4.1 (Video Task)

In Study 4.1 the experimental condition consisted of the Video Task. Both groups of participants were told that they were to watch a film clip. Participants in the experimental condition were informed that they should try not to respond to the clip in any way (no facial expressions or turning away), and that they should suppress any thoughts, feelings, or emotions that they may experience while watching the clip. The experimenter remained in the room throughout the task in order to observe participants’ emotional expressions and to remind them of the task instructions if necessary. When the clip was finished, participants were required to complete a manipulation check questionnaire. They were asked to rate perceived effort put into suppression, perceived difficulty of suppression, emotionality of the task, and feelings of being emotionally drained, and tiredness on 25 point Likert scales. Afterwards, participants completed a cognitive suppression task. Participants in the experimental group were told to write down any thoughts that came to mind over the next five minutes but not to think about anything that they had just seen in the clip. Participants of the control group were told to write down any thoughts that they had over the next five minutes, but without the film clip being mentioned by the experimenter. Once the cognitive suppression task was completed, participants completed a second manipulation check and the BMIS.

7.2.3.3 Experimental vs. Control Condition in Study 4.2 (Cold Pressor Task)

Pain was induced by the Cold Pressor Task (Birnie, Petter, Boerner, Noel, & Chambers, 2012; McParland, Knussen, & Murray, 2016; McParland, Knussen, Lawrie, & Brodie, 2013). Participants were seated next to the cold water bath asked to immerse their non-dominant hand above the wrist into the water. Different instructions were provided for the ego depletion and the control group. The experimental group was asked to keep the hand in the water as long as possible. The control group was told, that they should take the hand out of the water as soon as they experience pain. Participants were given a towel to dry their hands and handed over to the use of force trainer
for the briefing of the scenario. A 5-minute limit was set for hand immersion. Participants were not informed about the limit.

7.2.4 Analysis of HRV data

HRV data was processed as describe in the general methods chapter using Kubios HRV Analysis Software 2.1. (Tarvainen et al., 2014). Artefact correction level was set to medium. The following criteria were chosen to optimize the selection for the time-windows.

In Study 4.1 and Study 4.2 the following criteria were chosen to extract four samples which were subjected to subsequent analysis: The baseline sample showed the highest possible “stationary” of the RR signals (defined as the absence of arrhythmias and of any kind of artefacts in visual analysis of corresponding Suunto memory belt recordings). The treatment sample (Study 4.1: video task; Study 4.2: cold pressor task) was set in the centre of the time period of the video task or the cold pressor task, respectively. The scenario samples were centred on the maximum peak of HR during the scenario. A recovery sample was computed 8 minutes after completion of the scenario. In Study 4.2, two additional samples (cognitive task pre, cognitive task post), measuring HRV while performing the Phonemic Fluency Task, have been analysed. The samples were centred at the time period, where participants were engaged in the Phonemic Fluency Task. Therefore, Study 4.1 comprised four time points, whereas Study 4.2 comprised six time-points per participant.

7.2.5 Statistical Analysis

A Shapiro-Wilk’s test (Razali & Wah, 2011; S. S. Shapiro & Wilk, 1965) and a visual inspection of their histograms, normal Quantile-Quantile plots and box plots was performed on each data set of Study 4.1 and Study 4.2 (Doane & Seward, 2011).

Results for Study 4.1 showed that the phonemic fluency scores were not normally distributed. Transformation of the data set did not normalise the data. Therefore, a robust method for MANOVA conducted on the ranked data (Munzel & Brunner, 2000) was chosen to analyse the main effect of condition
on the data set. The effect of time on phonemic fluency scores was assessed using separate univariate two-way mixed ANOVAS based on M-estimators and bootstrap (for TW and MCS) and a parametric two-way mixed ANOVA (for NS).

Preliminary analysis of the manipulation check questions and the BMIS data indicated that the data set was not normally distributed. Bootstrapping was performed on the data set to allow for robust testing (independent t-tests) of the data.

The time until offensive aggression was shown in the scenario (aggression time) was also not normally distributed. Hence, bootstrapping was performed on this data before analysing the data using independent t-tests.

Results for the HRV measures showed, that RMSSD was normally distributed among the groups. Yet, the assumption for homogeneity of variances was not met. Hence, a robust 2 x 4 (Condition x Time) ANOVA with repeated measures on the second factor was performed using the \textit{bwtrim()} function implemented in R (Wilcox, 2012). Mean HR measures were not normally distributed. Transformation procedures did not normalize the data set. Therefore, the \textit{bwtrim()} function as a robust method for a 2 x 4 ANOVA with one repeated measure (time) was chosen to analyse potential effects of condition or time on mean HR. Significant results both for RMSSD and mean HR were followed up by separate robust one-way repeated measures ANOVAs on 20% trimmed means using the \textit{rmanova()} function. Post hoc were computed by the use of \textit{rmmcp()} function as suggested by A. Field et al. (2012).

For study 4.2 the results of exploratory analysis showed that the phonemic fluency scores were not normally distributed. Transformation of the data set did not normalise the data. Therefore, a robust method for MANOVA conducted on the ranked data (Munzel & Brunner, 2000) was chosen to analyse the main effect of condition on the data set. The effect of time on phonemic fluency scores was assessed using separate robust univariate
two-way mixed ANOVAS based on M-estimators and bootstrap (for MCS) and a parametric two-way mixed ANOVA (for TW and NS).

Analysis of the manipulation check questions and the BMIS data indicated that the data set was not normally distributed. Bootstrapping was performed on the data set to allow for robust testing (independent \( t \)-tests) of the data.

The time until offensive aggression was shown in the scenario (aggression time) was also not normally distributed. Hence, bootstrapping was performed on this data before analysing the data set using independent \( t \)-tests.

The results of the exploratory analysis of cardiovascular data showed, that both mean HR and RMSSD were not normally distributed. Transformation procedures did not normalize the data set. Hence, a robust method for a 2 x 6 (Condition x Time) ANOVA with repeated measures on the second factor was chosen to analyse the data set. For this purpose, the \texttt{bwtrim()} function implemented in R was used. Significant results were followed up by separate robust one-way repeated measures ANOVAs on 20\% trimmed means using the \texttt{rmanova()} function. Post hoc test were computed by the use of \texttt{rmmcp()} function (A. Field et al., 2012).

### 7.3 Results of Study 4.1 – Video Task

#### 7.3.1 Effects of Manipulation on Manipulation Check Questions and Mood

In order to assess the success of the ego depletion manipulation, the manipulation check and BMIS data were analysed using a series of independent \( t \)-tests. Participants reported no significant differences between ego depletion and control condition in terms of the BMIS subscale, in emotional and cognitive suppression, and the perceived emotionality of the film clip. This indicates that the employed manipulation (Video Task) did not lead to a state of ego depletion in the participants. Summary data for manipulation check and mood variables are shown in Table 24.
7.3.2 Effects of Manipulation on Offensive Aggression

On average, participants in the experimental condition showed offensive aggression towards the role player earlier ($M = 174.09 \text{ sec}, SE = 9.89$), than participants in the control condition ($M = 191.59 \text{ sec}, SE = 9.12$). This difference, $-17.50 \text{ sec}, \text{BCa 95\% CI} [-45.84, 9.47]$, was not significant, $t(42) = -1.30, p = .200$; however, it does represent a small to medium sized effect, $r = 0.20$. Figure 42 displays means and 95\% CI of this effect.

![Figure 42: Means and 95\% confidence intervals of the time (in seconds) when participants displayed offensive aggression in Study 4.1.](image)

7.3.3 Phonemic Fluency Task

Means, standard deviations, and confidence intervals of measures Phonemic Fluency Task scores are displayed in Table 22.

Since no robust MANOVAs with two independent variables exist so far (A. Field et al., 2012), the difference from pre- to post-testing was calculated for each DV (TW, MCS, and NS), reducing the independent variables. The remaining IV comprised of condition and the differences of time of phonemic
fluency scores as DVs. Subsequently a MANOVA as conducted on the ranked data using Munzel and Brunner’s (Munzel & Brunner, 2000) method, implemented in R using the `mulrank()` function (Wilcox, 2012). There was no significant main effect of condition on the difference of scores between pre- and post-testing, \( F = 2.32, p = .096 \).

**Table 22:** Means, Standard Deviations, and 95% Confidence Intervals around the Mean of Phonemic Fluency Task Scores (Study 4.1)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
<th>Mdn</th>
<th>LQ</th>
<th>UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>Pre 22</td>
<td>34.64</td>
<td>8.12</td>
<td>31.04</td>
<td>38.23</td>
<td>35.00</td>
<td>26.75</td>
</tr>
<tr>
<td></td>
<td>Post 22</td>
<td>40.77</td>
<td>11.70</td>
<td>35.58</td>
<td>5.96</td>
<td>42.50</td>
<td>28.75</td>
<td>52.00</td>
</tr>
<tr>
<td></td>
<td>Diff 22</td>
<td>6.14</td>
<td>6.51</td>
<td>3.25</td>
<td>9.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control 22</td>
<td>Pre 38.77</td>
<td>7.53</td>
<td>35.43</td>
<td>42.11</td>
<td>37.00</td>
<td>35.00</td>
<td>42.25</td>
</tr>
<tr>
<td></td>
<td>Post 22</td>
<td>41.64</td>
<td>11.46</td>
<td>36.56</td>
<td>46.72</td>
<td>40.50</td>
<td>30.75</td>
<td>40.50</td>
</tr>
<tr>
<td></td>
<td>Diff 22</td>
<td>2.86</td>
<td>8.09</td>
<td>-0.72</td>
<td>6.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>Pre 0.36</td>
<td>0.20</td>
<td>0.28</td>
<td>0.45</td>
<td>0.31</td>
<td>0.19</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Post 22</td>
<td>0.28</td>
<td>0.13</td>
<td>0.23</td>
<td>0.34</td>
<td>0.25</td>
<td>0.20</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Diff 22</td>
<td>-0.08</td>
<td>0.19</td>
<td>-0.16</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control 22</td>
<td>Pre 0.38</td>
<td>0.18</td>
<td>0.30</td>
<td>0.46</td>
<td>0.34</td>
<td>0.27</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Post 22</td>
<td>0.39</td>
<td>0.24</td>
<td>0.28</td>
<td>0.49</td>
<td>0.33</td>
<td>0.23</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Diff 22</td>
<td>0.01</td>
<td>0.26</td>
<td>-0.11</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>Pre 27.14</td>
<td>8.71</td>
<td>23.27</td>
<td>31.00</td>
<td>27.50</td>
<td>19.00</td>
<td>35.00</td>
</tr>
<tr>
<td></td>
<td>Post 22</td>
<td>32.55</td>
<td>9.97</td>
<td>28.13</td>
<td>36.96</td>
<td>32.50</td>
<td>26.00</td>
<td>41.25</td>
</tr>
<tr>
<td></td>
<td>Diff 22</td>
<td>5.41</td>
<td>3.92</td>
<td>3.67</td>
<td>7.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control 22</td>
<td>Pre 29.36</td>
<td>7.40</td>
<td>26.08</td>
<td>32.65</td>
<td>28.00</td>
<td>25.00</td>
<td>34.00</td>
</tr>
<tr>
<td></td>
<td>Post 22</td>
<td>31.27</td>
<td>9.67</td>
<td>26.99</td>
<td>35.56</td>
<td>30.00</td>
<td>23.50</td>
<td>37.00</td>
</tr>
<tr>
<td></td>
<td>Diff 22</td>
<td>1.91</td>
<td>7.47</td>
<td>-1.40</td>
<td>5.22</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Separate robust univariate two-way mixed ANOVAS based on M-estimators and bootstrap for TW and MCS were conducted, in order to look at main and interaction effects of time and condition. Analysis was performed in R using the `sppb()` function (Wilcox, 2012).

For TW, results yielded a significant main effect for time, \( psihat = -4.79, p = .002 \), revealing that TW increased form pre- to post-testing. There was no
significant main effect of condition, $\psi_{ihat} = -0.85$, $p = .758$, and no significant interaction effect of Time x Condition, $\psi_{ihat} = -2.71$, $p = .306$.

For MCS, results yielded no significant main effect for condition, $\psi_{ihat} = -1.07$, $p = .506$, no significant main effect for time, $\psi_{ihat} = 0.43$, $p = .429$, and no significant Time x Condition interaction effect, $\psi_{ihat} = 0.93$, $p = .648$.

For NS, results of the 2 x 2 (Time x Condition) mixed ANOVA revealed a significant main effect of time, $F(1, 42) = 16.55$, $p < .001$, partial $\eta^2 = .283$, showing that scores for NS were higher at post-testing than at pre-testing, $M_{diff} = 3.66$, 95% CI [1.84, 5.47], $r = 0.53$.

There were no significant effects for condition, $F(1, 42) = 0.04$, $p = .853$, partial $\eta^2 = .001$, and for Time x Condition interaction, $F(1, 42) = 3.79$, $p = .058$, partial $\eta^2 = .083$.

Summary data for Phonemic Fluency Task scores are depicted in Figure 43.
Figure 43: Means and 95% confidence intervals of Phonemic Fluency Task scores of Study 4.1. ** indicates significance at $p < 0.01$; *** at $p < 0.001$; from pre- to posttesting.
7.3.4 Cardiovascular Measures

A robust 2 x 4 (Condition x Time) ANOVA with repeated measures on the second factor were performed using the `bwtrim()` function implemented in R (Wilcox, 2012).

For mean HR, results yielded no significant main effect for condition, $Q = 3.14, p = .09$, and no significant Condition x Time interaction effect, $Q = 0.31, p < .817$. There was a significant main effect of time on mean HR, $Q = 234.62, p < .001$. In order to follow up the effects of time, separate robust one-way repeated measures ANOVAs on 20% trimmed means using the `rmanova()` function implemented in R were used on each condition.

For the experimental condition, results yielded a significant main effect of time on mean HR, $F(1.97, 25.66) = 172.19, p < .001$. Post hoc tests using the `rmmcp()` function (20% trimmed means) revealed a significant difference between baseline and scenario, $psihat = -77.40 (-90.79, -64.02), p < .05$, video and scenario, $psihat = -70.66 (-92.53, -48.79), p < .05$, scenario and recovery, $psihat = 71.60 (58.93, 84.27), p < .05$. There were no significant differences between baseline and video, $psihat = -4.75 (-11.54, 2.03), p < .05$, baseline and recovery, $psihat = -2.36 (-9.19, 4.47), p < .05$ and video and recovery, $psihat = 3.06 (-4.59, 10.70), p < .05$.

For the control condition, results yielded a significant main effect of time on mean HR, $F(2.14, 23.59) = 197.11, p < .001$. Post hoc tests using the `rmmcp()` function (20% trimmed means) revealed a significant difference between baseline and video, $psihat = -8.92 (-14.23, -3.60), p < .05$, baseline and scenario, $psihat = -78.67 (-96.71, -60.63), p < .054$, video and scenario, $psihat = -70.21 (-93.25, -47.19), p < .05$ and scenario and recovery, $psihat = 74.29 (58.85, 89.73), p < .05$. There were no significant differences between baseline and recovery, $psihat = -3.02 (-12.09, 6.06), p < .05$ and video and recovery, $psihat = 4.99 (-0.98, 10.96), p < .05$. Summary data for mean HR is depicted in Figure 44.
Figure 44: Mean HR at the four time points across experimental and control condition in Study 4.1 (T1: baseline; T2: Video Task; T3: scenario; T4: recovery). * indicates significance at p < .01 between time points.

For RMSSD, results revealed no significant main effect of condition, $Q = 0.78, p = 0.384$, and no interaction effect between condition and time, $Q = 0.43, p = .736$. A significant main effect of time on RMSSD was observed, $Q = 6.20, p = .004$, indicating that RMSSD scores changed significantly between measured time points. In order to follow up the effects of time, separate robust one-way repeated measures ANOVAs on 20% trimmed means using the `rmanova()` function implemented in R were used on each condition.

For the experimental condition, results yielded a significant main effect of time, $F(2.45, 31.89) = 3.96, p = .022$. Post hoc tests based on 20% trimmed means using the `rmmcp()` function revealed, that RMSSD differs significantly from baseline to scenario, $psihat = 19.59 (0.59, 38.59), p < .05$. There were
no significant differences between baseline and video, $\psi_{ihat} = -4.90$ (-16.52, 6.72), $p > .05$, baseline and recovery, $\psi_{ihat} = 5.25$ (-10.93, 21.44), $p > .05$, video and scenario, $\psi_{ihat} = 20.57$ (-7.67, 48.81), $p > .05$, video and recovery, $\psi_{ihat} = 6.36$ (-8.73, 21.45), $p > .05$, and scenario and recovery, $\psi_{ihat} = -11.82$ (-30.66, 7.01), $p > .05$.

For the control condition, results showed no significant main effect of time on RMSSD, $F(2.52, 27.68) = 2.57, p = .08$. Summary data for RMSSD scores are depicted in Figure 45.

Figure 45: RMSSD measures at the four time points across experimental and control condition in Study 4.1. in Study 4.1 (T1: baseline; T2: Video Task; T3: scenario; T4: recovery). * indicates significance at $p < .01$ between time points.
7.4 Results of Study 4.2 – Cold Pressor Task

7.4.1 Effects of Manipulation on Manipulation Check Questions and Mood

In order to assess the success of the ego depletion manipulation, the time of cold water immersion, manipulation check, and BMIS data were analysed using a series of t-tests on bootstrapped data. Participants in the experimental condition immersed their hand longer in cold water ($M = 160.72$, $SE = 30.68$) than participants of the control group ($M = 34.78$, $SE = 3.45$). This difference (125.94 seconds) was significant, $t(34) = -4.08$, $p = .002$, representing a large effect, $r = .57$.

Participants in the experimental condition reported more difficulties in keeping their hand in the cold water ($M = 16.44$, $SE = 1.41$), than participants in the control group ($M = 10.61$, $SE = 1.56$). The difference (5.83) was significant, $t(34) = 2.78$, $p = .012$, representing a medium to large sized effect, $r = .43$. Figure 46 depicts means and 95% CI of self-reports of behavioral and emotional supressing.

There was no significant effect of condition for the BMIS subscales, self-reports of emotionally drained and tiredness, and the perceived effort of keeping the hand in the cold water.

In sum these results indicate that the Cold Pressor Task induced a state of ego depletion in participants. Summary data for manipulation check and mood variables are shown in Table 25.
Figure 46: Means and 95% confidence intervals of self-reports of behavioral and emotional supressing in Study 4.2. * indicates significance at $p < .05$ between groups.

7.4.2 Effects of Ego Depletion on Offensive Aggression

On average, ego depleted police officers showed offensive aggression earlier ($M = 92.00, SE = 19.27$) than participants in the control condition ($M = 160.44, SE = 15.19$). This difference, $-68.44$, BCa 95% CI [-112.60, -22.33], was significant, $t(34) = -2.79$, $p = .014$ representing a medium effect, $r = 0.43$. The results are depicted in Figure 47.
Figure 47: Means and 95% confidence intervals of the time (in seconds) when participants displayed offensive aggression in Study 4.2. * indicates significance at $p < .01$ between groups.

### 7.4.3 Phonemic Fluency Task

Means, standard deviations and confidence intervals of Phonemic Fluency Task scores are displayed in Table 23.
Table 23: Means, Standard Deviations, and 95% Confidence Intervals Around the Mean of Phonemic Fluency Task Scores of Study 4.2

<table>
<thead>
<tr>
<th>TW</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
<th>Mdn</th>
<th>LQ</th>
<th>UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Pre 19</td>
<td>37.95</td>
<td>7.66</td>
<td>34.25</td>
<td>41.64</td>
<td>38.00</td>
<td>31.00</td>
<td>44.00</td>
<td></td>
</tr>
<tr>
<td>Post 19</td>
<td>39.00</td>
<td>10.20</td>
<td>34.08</td>
<td>43.92</td>
<td>38.00</td>
<td>38.00</td>
<td>46.00</td>
<td></td>
</tr>
<tr>
<td>Diff 19</td>
<td>1.05</td>
<td>6.28</td>
<td>-1.98</td>
<td>4.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Pre 18</td>
<td>37.56</td>
<td>9.46</td>
<td>32.85</td>
<td>42.25</td>
<td>39.00</td>
<td>30.75</td>
<td>42.25</td>
<td></td>
</tr>
<tr>
<td>Post 18</td>
<td>42.44</td>
<td>11.86</td>
<td>36.55</td>
<td>48.34</td>
<td>42.50</td>
<td>33.25</td>
<td>47.00</td>
<td></td>
</tr>
<tr>
<td>Diff 18</td>
<td>4.89</td>
<td>7.79</td>
<td>1.01</td>
<td>8.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MCS</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
<th>Mdn</th>
<th>LQ</th>
<th>UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Pre 19</td>
<td>0.51</td>
<td>0.43</td>
<td>0.30</td>
<td>0.71</td>
<td>0.43</td>
<td>0.19</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Post 19</td>
<td>0.60</td>
<td>0.76</td>
<td>0.24</td>
<td>0.97</td>
<td>0.35</td>
<td>0.26</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Diff 19</td>
<td>0.10</td>
<td>0.48</td>
<td>-0.13</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Pre 18</td>
<td>0.36</td>
<td>0.28</td>
<td>0.22</td>
<td>0.50</td>
<td>0.28</td>
<td>0.18</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Post 18</td>
<td>0.30</td>
<td>0.14</td>
<td>0.23</td>
<td>0.37</td>
<td>0.31</td>
<td>0.20</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Diff 18</td>
<td>-0.06</td>
<td>0.28</td>
<td>-0.19</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NS</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>LL</th>
<th>UL</th>
<th>Mdn</th>
<th>LQ</th>
<th>UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Pre 19</td>
<td>26.42</td>
<td>7.18</td>
<td>22.96</td>
<td>29.88</td>
<td>29.00</td>
<td>19.00</td>
<td>31.00</td>
<td></td>
</tr>
<tr>
<td>Post 19</td>
<td>27.05</td>
<td>7.86</td>
<td>23.26</td>
<td>30.84</td>
<td>27.00</td>
<td>22.00</td>
<td>32.00</td>
<td></td>
</tr>
<tr>
<td>Diff 19</td>
<td>0.63</td>
<td>5.42</td>
<td>-1.98</td>
<td>3.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Pre 18</td>
<td>28.89</td>
<td>8.73</td>
<td>24.55</td>
<td>33.23</td>
<td>29.00</td>
<td>21.75</td>
<td>34.25</td>
<td></td>
</tr>
<tr>
<td>Post 18</td>
<td>33.50</td>
<td>10.10</td>
<td>28.48</td>
<td>38.52</td>
<td>32.50</td>
<td>26.50</td>
<td>39.50</td>
<td></td>
</tr>
<tr>
<td>Diff 18</td>
<td>4.61</td>
<td>8.76</td>
<td>0.26</td>
<td>8.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since no robust MANOVAs with two independent variables exist so far (A. Field et al., 2012), the difference from pre- to post-testing was calculated for each DV (TW, MCS, and NS), reducing the independent variables. The remaining IV comprised of condition and the differences of time of phonemic fluency scores as DVs. Subsequently a MANOVA was conducted on the ranked data using Munzel and Brunner’s (Munzel & Brunner, 2000) method, implemented in R using the `mulrank()` function (Wilcox, 2012). There was no significant main effect of condition on the difference of scores between pre- and post-testing, $F = 1.07, p = .0.35$.

A separate robust univariate two-way mixed ANOVA based on M-estimators and bootstrap for MCS was conducted, in order to look for main and interaction effects of time and condition. Analysis was performed in R using the `sppb()` function (Wilcox, 2012).
For MCS, there was no significant main effect of condition, $\psi_{ihat} = 0.49, p = .766$, no significant main effect of time, $\psi_{ihat} = 0.70, p = .228$, and no significant interaction effect of time and condition, $\psi_{ihat} = -0.06, p = .978$.

For TW, results of a 2 x 2 (Time x Condition) mixed ANOVA yielded a significant main effect of time, $F(1, 35) = 6.55, p = .015$, partial $\eta^2 = .158$, showing that scores for NS were higher at post-testing than at pre-testing, $M_{diff} = 2.97, 95\% \text{ CI} [0.62, 5.33], r = 0.40$.

There were no significant effects for condition, $F(1, 35) = 0.252, p = .618$, partial $\eta^2 = .007$, and for Time x Condition interaction, $F(1, 35) = 2.73, p = .107$, partial $\eta^2 = .072$.

For NS, results of the 2 x 2 (Time x Condition) mixed ANOVA revealed a significant main effect of time, $F(1, 35) = 4.85, p = .034$, partial $\eta^2 = .122$, showing that NS were higher at post-testing than at pre-testing, $M_{diff} = 2.62, 95\% \text{ CI} [0.21, 5.08], r = 0.35$.

There were no significant effects for condition, $F(1, 35) = 3.10, p = .087$, partial $\eta^2 = .081$, and for Time x Condition interaction, $F(1, 35) = 2.80, p = .103$, partial $\eta^2 = .083$.

Figure 48 displays the means and 95\% CI of Phonemic Fluency Task scores of pre- and post-testing of the two groups.
Figure 48: Means and 95% confidence intervals of Phonemic Fluency Task scores of Study 4.2. * indicates significance at $p < .05$ from pre- to posttesting.
7.4.4 Cardiovascular Measures

Cardiovascular data was subjected to separate robust 2 x 6 (Condition x Time) ANOVAs with repeated measures on the second factor performed using the `bwtrim()` function implemented in R (Wilcox, 2012). For RMSSD results showed no significant main effect of condition, $Q = 2.46, p = .13$, and no significant interaction effect between condition and time, $Q = 0.53, p = .75$. However, there was a significant main effect of time on RMSSD, $Q = 3.87, p = .019$. This effect was followed up by separate robust one-way repeated measures ANOVAs on 20% trimmed means using the `rmanova()` function implemented in R. The follow up analysis for the experimental condition yielded no significant main effect of time on RMSSD, $F(2.78, 30.63) = 2.52, p = .081$. For the control condition the results showed also showed no significant main effect of time on RMSSD, $F(4.36, 39.24) =1.14, p = .35$. Figure 49 displays summary data for RMSSD in Study 4.2.

![Figure 49: Means and 95% confidence intervals of RMSSD for the experimental and control condition across all time points in Study 4.2. (T1: baseline; T2: Phonemic Fluency Task pretest; T3: Video Task; T4: scenario; T5 Phonemic Fluency Task posttest; T6: recovery).](image)
For mean HR, the results yielded no significant main effect for condition, $Q = 0.01$, $p = .923$, and for Condition x Time interaction, $Q = 0.72$, $p = .615$. There was a significant main effect of time, $Q = 69.94$, $p < .001$. In order to follow up the effects of time, separate robust one-way repeated measures ANOVAs on 20% trimmed means using the `rmanova()` function implemented in R were used on each condition.

For the experimental condition, results showed a significant effect of time on mean HR, $F(2.46, 27.04) = 90.75$, $p < .001$. Post hoc procedures (`rmmcp()` function on 20% trimmed means) revealed significant differences between T1 and T2, $psihat = -10.04$ (-14.91, -5.17), $p < .05$, T1 and T4, $psihat = -75.86$ (-97.05, -54.67), $p < .05$, T1 and T5, $psihat = -19.78$ (-32.29, -7.27), $p < .05$, T2 and T4, $psihat = -64.41$ (-86.13, -42.68), $p < .05$, T2 and T5, $psihat = -8.89$ (-16.31, -1.46), $p < .05$, T3 and T4, $psihat = -51.23$ (-78.57, -23.88), $p < .05$, T4 and T5, $psihat = 53.59$ (35.34, 71.84), $p < .05$, T4 and T6, $psihat = 69.52$ (56.58, 82.45), $p < .05$ and T5 and T6, $psihat = 16.14$ (9.38, 22.90), $p < .05$. There were no significant differences between T1 and T3, $psihat = -20.64$ (-42.66, 1.39), $p > .05$, T1 and T6, $psihat = -4.86$ (-12.34, 2.62), $p > .05$, T2 and T3, $psihat = -9.12$ (-30.27, 12.02), $p > .05$, T2 and T6, $psihat = 5.39$ (-0.22, 11.01), $p > .05$, T3 and T5, $psihat = 0.17$ (-19.45, 19.79), $p > .05$, T4 and T5, $psihat = 63.72$ (37.99, 89.44), $p < .05$, T4 and T6, $psihat = 16.38$ (-5.94, 38.70), $p > .05$.

There was a significant difference between T1 and T3 using the `rmmcp()` function on 15% trimmed means, T1 and T3, $psihat = -21.45$ (-41.11, -1.79), $p < .05$. This explains the graph which was plotted without trimmed means.

For the control condition, results yielded a significant main effect of time on mean HR, $F(1.68, 15.14) = 39.22$, $p < .001$. Post hoc tests using the `rmmcp()` function (20% trimmed means) revealed a significant difference T1 and T2, $psihat = -11.24$ (-21.63, -0.85), $p < .05$, T1 and T3, $psihat = -25.73$ (-42.25, -9.20), $p < .05$, T1 and T4, $psihat = -85.61$ (-130.58, -40.63), $p < .05$, T2 and T4, $psihat = -74.25$ (-110.56, -37.94), $p < .05$, T3 and T4, $psihat = -65.22$ (-116.45, -13.99), $p < .05$, T4 and T5, $psihat = 63.72$ (37.99, 89.44), $p < .05$,
T4 and T6, $\psi_{ihat} = 75.89 (41.21, 110.57)$, $p < .05$, and T5 and T6, $\psi_{ihat} = 12.19 (5.17, 19.20)$, $p < .05$. There were no significant differences between T1 and T5, $\psi_{ihat} = -19.86 (-42.24, 2.52)$, $p > .05$, T1 and T6, $\psi_{ihat} = -8.49 (-22.21, 5.23)$, $p > .05$, T2 and T3, $\psi_{ihat} = -15.26 (-34.06, 3.55)$, $p > .05$, T2 and T5, $\psi_{ihat} = -8.14 (-23.84, 7.57)$, $p > .05$, T2 and T6, $\psi_{ihat} = 2.75 (-7.77, 13.27)$, $p > .05$, T3 and T5, $\psi_{ihat} = 3.70 (-25.22, 32.62)$, $p > .05$, T3 and T6, $\psi_{ihat} = 15.40 (-4.31, 35.12)$, $p > .05$.

There was a significant difference between T1 and T5 using the `rmmcp()` function on 15% trimmed means, $\psi_{ihat} = -19.19 (-37.46, -0.92)$, $p < .05$. This explains the graph which was plotted without trimmed means.

Summary data for mean HR measures of Study 4.2 are depicted in Figure 50.

![Figure 50](https://example.com/figure50.png)

**Figure 50:** Means and 95% confidence intervals of mean HR for the experimental and control condition across all time points in Study 4.2. (T1: baseline; T2: Phonemic Fluency Task pretest; T3: Video Task; T4: scenario; T5 Phonemic Fluency Task posttest; T6: recovery). * indicates significance at $p < .05$ between time points.
Table 24: Group Comparisons of Manipulation Check and Mood Variables of Study 4.1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental (n = 22)</th>
<th>Control (n = 22)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>LL</td>
</tr>
<tr>
<td>Emotional suppression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effort suppressing emotions</td>
<td>9.59</td>
<td>5.67</td>
<td>7.08</td>
</tr>
<tr>
<td>Difficulty suppressing emotions</td>
<td>8.59</td>
<td>5.73</td>
<td>6.05</td>
</tr>
<tr>
<td>Emotionality of the clip</td>
<td>13.09</td>
<td>6.80</td>
<td>10.07</td>
</tr>
<tr>
<td>Emotionally drained</td>
<td>5.09</td>
<td>5.55</td>
<td>2.63</td>
</tr>
<tr>
<td>Tiredness</td>
<td>4.55</td>
<td>4.54</td>
<td>2.53</td>
</tr>
<tr>
<td>Cognitive suppression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effort suppressing thoughts</td>
<td>7.68</td>
<td>6.45</td>
<td>4.82</td>
</tr>
<tr>
<td>Difficulty suppressing thoughts</td>
<td>8.09</td>
<td>7.12</td>
<td>4.93</td>
</tr>
<tr>
<td>Emotionally drained</td>
<td>5.32</td>
<td>5.30</td>
<td>2.97</td>
</tr>
<tr>
<td>Tiredness</td>
<td>5.77</td>
<td>4.98</td>
<td>3.57</td>
</tr>
<tr>
<td>BMIS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pleasant - unpleasant</td>
<td>2.86</td>
<td>5.63</td>
<td>0.37</td>
</tr>
<tr>
<td>arousal - calm</td>
<td>15.09</td>
<td>4.82</td>
<td>12.95</td>
</tr>
<tr>
<td>positive - tired</td>
<td>6.00</td>
<td>3.48</td>
<td>4.46</td>
</tr>
<tr>
<td>negative - relaxed</td>
<td>6.64</td>
<td>3.59</td>
<td>5.04</td>
</tr>
</tbody>
</table>
**Table 25: Group Comparisons of Manipulation Check and Mood Variables of Study 4.2**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental (n = 18)</th>
<th>95% CI</th>
<th>Control (n = 18)</th>
<th>95% CI</th>
<th>Difference</th>
<th>BCa 95% CI</th>
<th>t(34)</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time in cold water (sec)</td>
<td>M</td>
<td>SD</td>
<td>LL</td>
<td>UL</td>
<td>Mdn</td>
<td>LQ</td>
<td>UQ</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>160.72</td>
<td>130.15</td>
<td>96.00</td>
<td>225.4</td>
<td>132.0</td>
<td>40.00</td>
<td>300.0</td>
<td>34.78</td>
<td>14.65</td>
</tr>
<tr>
<td>Behavioral suppression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effort suppressing behavior</td>
<td>11.28</td>
<td>6.36</td>
<td>8.12</td>
<td>14.44</td>
<td>12.00</td>
<td>7.00</td>
<td>15.00</td>
<td>8.78</td>
<td>5.55</td>
</tr>
<tr>
<td>Difficulty suppressing behavior</td>
<td>16.44</td>
<td>5.99</td>
<td>13.46</td>
<td>19.42</td>
<td>17.00</td>
<td>11.00</td>
<td>21.00</td>
<td>10.61</td>
<td>6.60</td>
</tr>
<tr>
<td>Emotional suppression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotionally drained</td>
<td>3.00</td>
<td>3.05</td>
<td>1.48</td>
<td>4.52</td>
<td>2.00</td>
<td>1.00</td>
<td>3.00</td>
<td>3.61</td>
<td>4.25</td>
</tr>
<tr>
<td>Tiredness</td>
<td>3.11</td>
<td>2.87</td>
<td>1.69</td>
<td>4.54</td>
<td>2.00</td>
<td>1.00</td>
<td>4.00</td>
<td>3.89</td>
<td>5.43</td>
</tr>
<tr>
<td>BMIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pleasant - unpleasant</td>
<td>4.72</td>
<td>7.30</td>
<td>1.09</td>
<td>8.35</td>
<td>4.00</td>
<td>-2.00</td>
<td>12.00</td>
<td>4.78</td>
<td>3.62</td>
</tr>
<tr>
<td>arousal - calm</td>
<td>15.00</td>
<td>3.18</td>
<td>13.42</td>
<td>16.58</td>
<td>16.00</td>
<td>13.00</td>
<td>17.00</td>
<td>15.00</td>
<td>4.33</td>
</tr>
<tr>
<td>positive - tired</td>
<td>7.06</td>
<td>3.89</td>
<td>5.12</td>
<td>8.99</td>
<td>6.00</td>
<td>4.00</td>
<td>11.00</td>
<td>7.06</td>
<td>2.94</td>
</tr>
<tr>
<td>negative - relaxed</td>
<td>5.83</td>
<td>2.66</td>
<td>4.51</td>
<td>7.16</td>
<td>5.00</td>
<td>4.00</td>
<td>8.00</td>
<td>5.67</td>
<td>2.66</td>
</tr>
<tr>
<td>Variable</td>
<td>Experimental (n = 22)</td>
<td>Control (n = 20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>LL</td>
<td>UL</td>
<td>Mdn</td>
<td>LQ</td>
<td>UQ</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Mean HR</td>
<td></td>
<td></td>
<td>95 % CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>95 % CI</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>70.58</td>
<td>11.91</td>
<td>65.30</td>
<td>75.87</td>
<td>69.16</td>
<td>63.87</td>
<td>74.08</td>
<td>73.79</td>
<td>15.01</td>
</tr>
<tr>
<td>T2</td>
<td>78.22</td>
<td>16.01</td>
<td>71.13</td>
<td>85.32</td>
<td>76.99</td>
<td>65.51</td>
<td>85.95</td>
<td>82.13</td>
<td>15.28</td>
</tr>
<tr>
<td>T3</td>
<td>146.84</td>
<td>19.11</td>
<td>138.36</td>
<td>155.31</td>
<td>148.51</td>
<td>130.55</td>
<td>160.02</td>
<td>152.44</td>
<td>23.32</td>
</tr>
<tr>
<td>T4</td>
<td>73.90</td>
<td>14.56</td>
<td>67.45</td>
<td>80.36</td>
<td>71.62</td>
<td>63.32</td>
<td>85.00</td>
<td>76.92</td>
<td>13.37</td>
</tr>
<tr>
<td>RMSSD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>49.78</td>
<td>24.23</td>
<td>39.04</td>
<td>60.52</td>
<td>49.07</td>
<td>35.48</td>
<td>67.84</td>
<td>45.41</td>
<td>26.98</td>
</tr>
<tr>
<td>T2</td>
<td>56.31</td>
<td>36.47</td>
<td>40.14</td>
<td>72.48</td>
<td>50.45</td>
<td>26.38</td>
<td>83.15</td>
<td>44.06</td>
<td>24.63</td>
</tr>
<tr>
<td>T3</td>
<td>29.63</td>
<td>13.22</td>
<td>23.76</td>
<td>35.49</td>
<td>31.25</td>
<td>20.07</td>
<td>37.04</td>
<td>29.16</td>
<td>7.65</td>
</tr>
<tr>
<td>T4</td>
<td>44.73</td>
<td>25.75</td>
<td>33.31</td>
<td>56.14</td>
<td>36.14</td>
<td>22.13</td>
<td>64.96</td>
<td>44.33</td>
<td>28.74</td>
</tr>
</tbody>
</table>
Table 27:  Group Comparisons of Cardiovascular Data of Study 4.2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental (n = 18)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Control (n = 16)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>LL</td>
<td>UL</td>
<td>Mdn</td>
<td>LQ</td>
<td>UQ</td>
<td>M</td>
<td>SD</td>
<td>LL</td>
</tr>
<tr>
<td>Mean HR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>72.92</td>
<td>7.48</td>
<td>69.21</td>
<td>76.64</td>
<td>72.03</td>
<td>69.08</td>
<td>77.22</td>
<td>68.80</td>
<td>8.64</td>
<td>64.99</td>
</tr>
<tr>
<td>T2</td>
<td>83.52</td>
<td>8.80</td>
<td>79.14</td>
<td>87.90</td>
<td>82.86</td>
<td>76.17</td>
<td>92.35</td>
<td>80.21</td>
<td>10.44</td>
<td>74.65</td>
</tr>
<tr>
<td>T3</td>
<td>95.24</td>
<td>16.78</td>
<td>86.89</td>
<td>103.58</td>
<td>90.28</td>
<td>82.71</td>
<td>111.43</td>
<td>94.82</td>
<td>19.16</td>
<td>84.61</td>
</tr>
<tr>
<td>T4</td>
<td>151.59</td>
<td>28.50</td>
<td>137.42</td>
<td>165.76</td>
<td>150.74</td>
<td>136.81</td>
<td>159.71</td>
<td>159.27</td>
<td>39.27</td>
<td>138.35</td>
</tr>
<tr>
<td>T5</td>
<td>92.62</td>
<td>12.08</td>
<td>86.61</td>
<td>98.63</td>
<td>89.56</td>
<td>82.93</td>
<td>102.93</td>
<td>89.17</td>
<td>16.45</td>
<td>80.40</td>
</tr>
<tr>
<td>T6</td>
<td>77.31</td>
<td>9.08</td>
<td>72.80</td>
<td>81.83</td>
<td>74.78</td>
<td>72.14</td>
<td>80.37</td>
<td>77.78</td>
<td>13.12</td>
<td>70.79</td>
</tr>
<tr>
<td>RMSSD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>39.87</td>
<td>20.73</td>
<td>29.56</td>
<td>50.18</td>
<td>35.58</td>
<td>26.30</td>
<td>43.06</td>
<td>38.67</td>
<td>15.15</td>
<td>30.60</td>
</tr>
<tr>
<td>T2</td>
<td>31.64</td>
<td>14.24</td>
<td>24.56</td>
<td>38.72</td>
<td>27.16</td>
<td>21.27</td>
<td>38.65</td>
<td>34.46</td>
<td>13.39</td>
<td>27.33</td>
</tr>
<tr>
<td>T3</td>
<td>22.38</td>
<td>10.18</td>
<td>17.31</td>
<td>27.44</td>
<td>22.13</td>
<td>15.18</td>
<td>27.79</td>
<td>36.05</td>
<td>25.83</td>
<td>22.28</td>
</tr>
<tr>
<td>T4</td>
<td>26.45</td>
<td>16.98</td>
<td>18.01</td>
<td>34.89</td>
<td>25.56</td>
<td>14.18</td>
<td>34.53</td>
<td>36.20</td>
<td>11.24</td>
<td>30.21</td>
</tr>
<tr>
<td>T5</td>
<td>27.82</td>
<td>22.12</td>
<td>16.82</td>
<td>38.82</td>
<td>21.87</td>
<td>13.96</td>
<td>29.48</td>
<td>28.18</td>
<td>16.29</td>
<td>17.50</td>
</tr>
<tr>
<td>T6</td>
<td>34.29</td>
<td>21.80</td>
<td>23.45</td>
<td>45.14</td>
<td>27.31</td>
<td>15.57</td>
<td>54.63</td>
<td>41.41</td>
<td>27.59</td>
<td>26.71</td>
</tr>
</tbody>
</table>
7.5 Discussion

In the current study it was predicted that the experimental manipulations (Video Task in Study 4.1, Cold Pressor Task in Study 4.2) would lead to a state of ego depletion in participants displayed via changes in the manipulation check measures (Hypothesis 1). Only the results of Study 4.2, where the Cold Pressor Task was employed confirmed this hypothesis. Therefore, the experimental manipulation in Study 4.1. failed to produce a state of ego depletion, whereas in Study 4.2. participants were ego depleted after the manipulations. It was hypothesized, that ego depleted participants would display offensive aggression towards the provoking role player earlier compared to the control (Hypothesis 2). The results of Study 4.2. (manipulation confirmed this hypothesis. Also the results of Study 4.1. (manipulation failed) provide support for this hypothesis, since no differences with regards to the onset of offensive aggression were found between the experimental and the control condition. Regarding the effects on executive functioning, it was predicted that ego depletion would impair executive functioning (Hypothesis 3). Contrary to the hypothesis, executive functioning increased in both studies, as indicated by an increase in TW and NS with a decrease in MCS. Finally, it was hypothesised that vagal tone as a measure of self-control would be affected by ego depletion (Hypothesis 4). The current results disconfirmed this hypothesis. No significant group effect was observed in both studies. However, in Study 4.1 vagal tone changed across time, with the lowest scores during the scenario, whereas in Study 4.2 a significant main effect of time on RMSSD indicated that the demands placed upon participants led to a temporary decrease of vagal tone.

7.5.1 Ego Depletion and Offensive Aggression

As predicted, successful ego depletion manipulation (Study 4.2., Cold Pressor Task) resulted in an earlier display of offensive aggression. In line with this effect was not present when ego depletion manipulation did not work (Study 4.1., Video Task).

The failing of the ego depletion manipulation in Study 4.1., where the results of the manipulation check showed that effort and difficulty to suppress
emotions and thoughts in the Video Task manipulation did not differ between the ego depletion and the control group is inconsistent with previous results employing the same ego depletion manipulation. Findings of P. Christiansen et al. (2012) showed, that emotional and cognitive suppression following a video clip, which is likely to illicit feelings of disgust, was rated more effortful and more difficult by the experimental group than by the control group. However, participants involved individuals that were recruited via word of mouth and intranet from the University of Liverpool, suggesting that mainly students took part in the experiment. In the current study the participants were police officers. Since police officers experience situations in their daily activities that can involve negative feelings, like disgust, it is possible that they are used to such visual stimuli that are known to illicit disgust in non-police individuals. Another possibility is that the police population in general consists of individuals that are less prone to illicit feelings of disgust after such a video clip. It is likely that only individuals that are able to cope with such stimuli on a regular basis become police officers.

The main effect of ego depletion on offensive aggression supports current theories of self-control. With regards to the strength model of self-control the results support this model in that relative to a control (non-depleting) manipulation, offensive aggression was displayed earlier by police officers when confronted with provoking behaviour by a role player. Therefore, this finding adds to the growing body of evidence which suggests that ego depletion affects aggressive behaviour (Barlett et al., 2016; DeWall et al., 2007; Stucke & Baumeister, 2006; Vohs et al., 2011). However, the current study is the first one that provides evidence that ego depletion decreases the time until physical forms of aggression are displayed after provoking behaviour. Since there is evidence that self-control performance varies across contexts (J. R. Cohen & Lieberman, 2010; Cortes, Kammrath, Scholer, & Peetz, 2014), the current results are substantial regarding aggression and the police use of force domain.

The results are also in line with the newly proposed Dual Component Theory of Inhibition Regulation (J. J. Reynolds & McCrea, 2016). The theory proposes, that there are numerous domain-specific modules working to carry
out specific behaviours. Some of them are “impulsive” modules that motivate immediate behaviour. According to the theory these modules are regulated by a computational inhibition module, which comprises a monitor and a threshold component. In police-citizen encounters, the impulsive module to aggress against a provoking citizen may conflict with more long term goals such as the rise of later costs in the police-citizen relationship (Reisig et al., 2004; L. W. Sherman, 1993; Tyler & Huo, 2002; Wolfe, 2011). The monitor module detects that an impulsive module needs to be inhibited. The “cost” or output of the impulsive module is then measured, reflecting the temptation to aggress experienced by the police officer. A signal in the form of an inhibition effort is produced by the monitor component and then sent to the threshold component. This signal reflects the sensation of effort that the individuals feels when inhibiting. The threshold component then processes the signal of inhibition effort to determine whether the threshold has been met. According to the model, this threshold represents the individual’s tolerance for applying inhibitory effort. If the threshold has not been met, the signal is sent from the threshold component to continue to inhibit the “impulsive” module. With regards to the current study, non-depleted participants were able to maintain inhibitory effort in the light of provoking behaviour. If the magnitude of inhibitory effort reached the tolerance for applying inhibition, then the inhibition module stops inhibiting, and proponent behaviour (the “impulsive” module) is carried out. Therefore, the ego depletion effect of the current study reflects a reduction in the threshold for inhibition of the impulsive module to aggress.

In animals and humans alike, aggression is also elicited by pain (J. Archer, 1989; Berkowitz, Cochran, & Embree, 1981; R. J. Blanchard, Kleinschmidt, Fukunaga-Stinson, & Blanchard, 1980). Participants in the manipulation condition in Study 4.2 were exposed longer to pain than participants in the control condition, since they suppressed the impulse to pull the hand out of the cold water. Therefore, it may be possible that the experienced pain has been a mediating factor in the earlier onset of offensive aggression in the subsequent scenario. As such, the Cold Pressor Task employed in Study 4.2. might not be ideal with regards to clearly separating the effects of
depleted self-regulatory resources and pain on participants' behavior. However, police officers have to apply self-control in the face of aversive stimuli such as pain. For example, excessive use of force may follow the legitimate use of force after restraining an actively resisting suspect, leading to minor injuries and/or the experience of pain on the side of the officer. In such situations, the officer has to regulate his/her behavior. As such, self-regulatory tasks involving pain may prove fruitful with regard to real-world application in the context of policing. Furthermore, the manipulation and control group did not differ with regard to BMIS subscales, self-reports of emotionally drained and tiredness, and the perceived effort of keeping the hand in the cold water and measures of physiological parameters (HR and HRV). The experimental group reported more difficulties in keeping their hand in the cold water than the control group. This indicates that self-control (withstanding the impulse to pull the hand out of the water) puts more load on police officers than the actual experience of pain. However, future studies employing the Cold Pressor Task in police settings should incorporate manipulation check measures targeting the intensity of the experiences of pain.

The observed behavior of participants is in line with current knowledge about the biology of aggression. With reference to antecedents, offensive aggression occurs in the context of resource, including territorial and dominance disputes (D. C. Blanchard & Blanchard, 2005). Challenges to the individual's authority may directly call into question the current dominance hierarchy and may involve insults to the subject (Averill, 1982). As such the role player in the current experiment may be regarded as a resource challenger to the police officer, that tries to protect resources (area of operation of other police officers in the scenario; authority). The function of offensive aggression in this context is closely tied to the outcomes of possible behaviors, with the best outcome being, that the resource / dominance challenge is terminated without any offensive attack. Therefore, preceding offensive attack with offensive threats may prove functional. Such threats have not been recorded systematically in the current study, however they were observed regularly by the author.
Another aspect, that has to be taken into consideration, when interpreting the observed behaviours of participants is, that offensive aggression is very sensitive to the fighting capabilities of the opponent (D. C. Blanchard & Blanchard, 2005). Fear, reflecting that damage by the target of the offensive attack on the attacker can prove disastrous to the latter, inhibits offensive attack. This can be seen as a highly adaptive response on the part of the attacker. As such, the decision to aggress by the police officer may be influenced by the fighting capabilities of both parties. In the current study, individual characteristics that maybe influence the evaluation of one’s fighting capabilities (size, muscle mass, fighting experience) were not recorded. Future studies should account for the relationship of fighting capabilities between the aggressor and the target of the aggression by focusing on specialised units with similar fighting abilities.

Concerning theories of aggression, the current results also provide further evidence for the newly developed $I^3$ theory of aggression. This meta theory proposes that aggression will only be displayed when inhibition resources are small compared to the sum of instigating factors, like provocations, and impelling factors, such as trait aggressiveness (Denson et al., 2012; DeWall et al., 2011; Finkel, 2014; Finkel et al., 2012; Slotter et al., 2012). The current results have shown in the domain of police use of force that a state of ego depletion diminishes resources to inhibit the impulse to aggress offensively when provoked. Interestingly, executive functioning as a predicted resource of inhibition did not decrease throughout the study (see next paragraph for possible explanations).

From a practical point of view, there are several implications for the findings that are based on this explanation of the ego depletion effect. Since ego depletion seems to negatively affect the actual behavioural response in a police-citizen encounter, it would be beneficial to focus on different strategies in order to circumvent these effects. First, the individual tolerance level can be bolstered by training. For example, it has been shown, that students who practiced self-control for 2 weeks by using their non-dominant hand for everyday tasks showed reduced anger when subsequently provoked by a fictitious fellow student (Denson et al., 2012). Second, strategies aiming at
reducing post-depletion aggression, like mindfulness practices (Yusainy & Lawrence, 2015), could be learned by police officers in order to reduce behavioural aggression following a depleting task. Third, reducing the inhibitory effort and reducing its accumulation via adapting to common self-regulatory tasks that require self-control (Dang, Dewitte, Mao, Xiao, & Shi, 2013) may also prove fruitful. For police officers, this would comprise of identifying self-regulatory tasks in the daily routine and designing programs that aim at adapting to these demands. Fourth, procedural guidelines and tactics should account for the prevention of the accumulation of inhibitory effort, by for example changing the role of the communicating officer in a police-citizen interaction according to previous amounts of self-regulatory effort. For example, if a police officer had to suppress pain during a shift, his partner may initiate the communication in the subsequent encounter with a citizen. However, in the light of context specificity, future studies should focus on validating these counter-measures to the ego depletion effect for the police use of force domain.

7.5.2 Ego Depletion and Executive Functioning

For both studies the results showed an increase in executive functioning as indicated by a rise of TW and NS between pre- and post-testing. MCS as an indicator of non-executive functioning decreased from pre- to post-testing. Contrary to the hypotheses there was no significant effect between groups on these measures. These results contradict predictions that impairments in self-regulation could be explained via state reductions in executive functioning (Hagger, Wood, Stiff, & Chatzisarantis, 2010; Hofmann et al., 2012). However, the results are in line with findings of P. Christiansen et al. (2012), who focused on ad-lib alcohol consumption, that phonemic fluency measures as an indicator of executive functioning, did not differ between depleted and non-depleted individuals. The researchers put forward the experimental design as an explanation for their findings, since the effects of ad-lib alcohol consumption was tested in one session, whereas effects of ego depletion on executive functioning were tested in a second session. Additionally, manipulation check data revealed that the ego depletion manipulation was less effective in the second session. The current
experimental design controlled for possible practice effects that could weaken the effectiveness of the manipulation, by assessing the variables in a single session. Therefore, the experimental design could not account for the results.

A possible explanation is the increase of physiological arousal as indicated by the rise in mean HR above 145 bpm during the scenario. This would be consistent with the results of Study 2 (see Chapter 5), where performance in the Phonemic Fluency Task also increased after both scenario training and running-on-the-spot exercise. It may be possible that executive functioning was impaired by ego depletion but the psychophysiological demands of the scenario meant both groups were affected which masked the effect. In order to control for possible effects of psychophysiological load, future experiments could be stopped, when the role player is approached (before the takedown). This would enable a more precise interpretation about the effects of ego depletion on executive functioning as measured by the Phonemic Fluency Task (Hofmann et al., 2012).

7.5.3 Ego Depletion and Heart Rate Variability

Vagal tone and parasympathetic activity as measured by RMSSD, did not differ between the experimental and control group in Study 4.1 and 4.2. However, vagal tone changed over time in Study 4.1, with the lowest scores at the time point of the scenario training. In Study 4.2 there was a significant main effect of time on RMSSD in the overall analysis that diminished once separate analyses had to be performed due to non-parametric data. This indicated that the demands placed upon participants led to a temporary decrease of vagal tone. Even though non-significant, the visual inspection of HRV reactivity over the different time points suggest that the decrease in vagal tone during the Cold Pressor Task and the scenario was more pronounced in the experimental group than in the control group.

Taken together vagal tone and parasympathetic activity was the lowest in Study 4.1 at the time point of the scenario and in Study 4.2 at the manipulation time point. Given that the manipulation did not work in the first study, it seems reasonable to assume that vagal tone was the lowest at the
time point where self-control was exercised the most, which was the scenario in Study 4.1 and the manipulation in Study 4.2.

A recent meta-analysis (Zahn et al., 2015) focusing on resting HRV and self-control found that higher trait HRV is associated with better self-control performance in laboratory settings. However, due to publication bias the authors suggest interpreting the results very cautiously, suggesting no significant association between HRV and self-control. Since vagal reactivity plays a critical role in adaptation and evolution (Beauchaine, 2001; Porges, 2007), resting vagal tone alone does not serve as an optimal indicator of self-control performance (Laborde & Mosley, 2016).

With regards to vagal tone reactivity a recent study (Geisler, Kleinfeldt, & Kubiak, 2016) focused on the interaction between vagal tone and self-control during food exposure of restrained eaters in an ego depleted and a control condition. The authors found a significant increase from baseline to food exposure in vagal tone and neither a significant main effect of condition nor an interaction effect. The current results contradict these findings, since a decrease (even though not significant) was observed in the experimental condition during the manipulation. However, vagal tone in the experimental condition of Study 4.2 increased afterwards. Therefore, it could be possible that vagal tone during the ego depletion manipulation was lower than the measure at the time point of food exposure. Yet, in the current study vagal tone at the scenario timepoint was not higher compared to the baseline.

In line with current recommendations (Laborde & Mosley, 2016) future studies should focus on vagal tone reactivity when focusing on self-control in police use of force events. Furthermore, different time intervals for the different tasks (e.g. manipulation and self-control performance) should be investigated to draw a more complete picture of the link between HRV and self-control in these situations.

7.5.4 Limitations

There are some limitations regarding Study 4 that have to be acknowledged.
First, as mentioned in the methods section, the datasets vary slightly regarding the dependent variables, since complete cardiovascular data were not available for all participants. Therefore, interpretations about the relationship between different variables (e.g. HRV and offensive aggression) have to be done cautiously. However, since missing cardiovascular data affected less than 10% of the data sets (i.e. 4.5% in Study 4.1, and 8.1% in Study 4.2), it was decided to conduct the statistical analyses of offensive aggression and executive functioning with the full data set.

Second, the naturalistic setting of the study made it hard to disentangle possible effects of physiological demand (see Chapter 5) and self-control on executive functioning. Since the offensive aggression against the role player, as the most physiologically demanding part of the exercise, immediately followed ego depletion, the effects on measures of executive function have to be interpreted with caution.

Third, it has to be acknowledged that the current study did not focus on impelling factors of aggression, like trait aggressiveness, to determine possible interaction effects as predicted by the I^3 theory of aggression (Finkel, 2014) Future studies could benefit from focusing on these interaction effects in police use of force settings.

7.5.5 Conclusion

The two studies in this chapter investigated the effects of ego depletion on offensive aggression of police officers. The results indicated that ego depleted police officers engaged in offensive aggression earlier when provoked by a role player in a scenario. Furthermore, the studies showed that the display of a disgusting video combined with the task to supress emotions did not induce a state of ego depletion. Future studies should instead rely on the Cold Pressor Task to induce ego depletion. Additionally, the results of the current chapter showed that: (a) executive functioning increased from pre- to post-testing without significant differences between the manipulation and the control group and (b) that vagal tone (although not statistically significant) decreased more in the ego depletion compared to the control group during the initial self-control task.
Taken together the results provide support for current theories of self-control and aggression. From a practical point of view, mastering self-control and reducing the effects of ego depletion seems fruitful avenues of further investigation in order to reduce aggressive behaviour in police officers.
Chapter 8: General Discussion
8 General Discussion

The identification of factors that underlie tactical decision making of police officers and how these processes can be influenced is crucial for two main reasons: First, the police use of force can be improved for both the public and the police. For example, the risk of injury to, and death of, law enforcement officers can be reduced in both police-citizen encounters and training settings (Boulton, 2014). Second, the public’s attitude and behaviours towards the legitimacy of the police and the justice system can positively influenced (Friedrich, 1980; J. H. Garner, Maxwell, & Heraux, 2002). Even though its crucial role in effective policing has been acknowledged (Boulton, 2014; Boulton & Cole, 2016; Condon, 2015; Johnsen, Espevik, Saus, Sanden, & Olsen, 2015; A. P. J. Roberts, 2012; Vrij, 1995; Vrij et al., 1995; Vrij & Dingemans, 1996; M. D. White, 2001; 2003), very little research has been conducted with regards to the tactical decision making of German police officers (Flin et al., 2007). The current thesis addresses this gap by examining the impact of various factors on tactical decision making of mainly German police officers.

The findings indicate that: (a) physiological arousal plays a key role in eliciting shifts in cognitive processing, (b) the psychophysiological effects of using different ammunition (i.e. CA vs. NLTA) in training settings does not differ despite the inherent risks in using CA suggesting that representative training environments are more important than risk, (c) experience as a police officer does not lead to threat-related attentional biases or risk taking, and (d) ego depletion leads to earlier offensive aggression in police-citizen encounters. The results have been discussed in detail in the respective chapters. Hence, the following sections chapter discusses key findings and integrates them into the existing literature, how these findings can be applied by the Police, and the implications of such applications.

8.1 Key Findings

The current body of research revealed several key findings:
1. In training settings NLTA produces the same levels of psychophysiological demands as CA. This provides strong evidence that the use of NLTA provides a safe yet representative alternative to CA in training, but also in research settings. Contrary to the belief of many police trainers, it is the representativeness of the training that is key to the psychophysiological reaction of the trainees. This should help to reduce the number of training-related fatalities around the world.

2. Cognitive functioning, especially executive functioning, changed under psychophysiological demand. In Study 1 SOU officers operating at high anaerobic levels (mean HR between 185 bpm and 205 bpm) in situations that require simple decisions and reliance on learned SOPs were able to maintain executive functioning and visuospatial processing capacity at the expense of phonological processing. Reliance on retrieval processes as non-executive functioning was enhanced. In Study 2 both scenario training involving complex tactical decision making and non-cognitively demanding physical exercise (mean HR between 105 and 135 bpm) equally resulted in WM changes: executive functioning increased, while non-executive functioning decreased relatively. In the light of previous research on WM performance after acute exercise, this suggests that the effects on WM performances occurred due to the physiological arousal induced by the two different treatments. In Study 4 executive functioning increased from pre- to post-testing (mean HR between 145 and 155 bpm) without significant differences between the ego depleted and control groups. Taken together, the results indicate that the observed shifts in cognitive functioning in this thesis and previous research (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012) may reflect an adaptive response of the human body to the demands physical exercise. The higher rates in Study 1 suggest that defensive behaviour may induce greater psychophysiological demand than exercise alone. However, it is extremely difficult to disentangle psychophysiological load from the physical movements from the
cognitive load of the decisions being made. Clearly further work is needed to answer this question.

3. Neither police officers nor martial artists showed any threat-related attentional biases or a propensity for risk-taking. This indicated no signs of a functional development of a cognitive bias. However, in conjunction with previous research on combat-related attentional bias, it may be possible that a possible functional attentional bias towards combat-related cues was not discovered due to the methodology used.

4. Compared to a control group, police officers and martial artists show less trait anxiety, indicating (a) that exposure to, and coping with, real and simulated threat situations may result in lower trait anxiety or (b) individuals with lower trait anxiety are more likely to pursue careers in the police and/or engage in martial arts.

5. Ego depleted police officers engaged in offensive aggression earlier when provoked by a role player in a scenario. This indicates that depleted self-control resources influence policing behaviour and tactical decisions.

8.1.1 Cognitive Shifts Under Threat

The results of Studies 1, 2, and 4 elicited changes to the executive functioning of police officers. It was found that executive functioning was enhanced when physiological arousal reached a medium level (mean HR between 105 and 155 bpm) due to the task to be performed (Study 2 and Study 4). When the physiological arousal level reached high levels (mean HR between 185 and 205), executive functioning stayed the same (Study 1). The most logical way to interpret these findings is framework of the Yerkes-Dodson law (D. Diamond, 2005; D. M. Diamond, Campbell, Park, Halonen, & Zoladz, 2007; Yerkes & Dodson, 1908). Yerkes and Dodson, investigating visual discrimination tasks under different intensities of electrical stimulation, observed a linear relationship between learning a task and the strength of electrical stimulation but a curvilinear relationship between learning a task of
moderate difficulty and simulation strength. Their results have been replicated (K. J. Anderson, 1994; Dickman, 2002; Telegdy & Cohen, 1971) and widely applied to performance in many contexts (L. C. Chang et al., 2013; Giddings, Thomas, & Little, 2013; Jeong & Biocca, 2012; C. Stinson & Bowman, 2014). The Yerkes-Dodson law states that moderate arousal can enhance performance (Broadhurst, 1957; Ni, 1934; Salehi, Cordero, & Sandi, 2010), whereas high levels of arousal impair performance (D. M. Diamond, Park, Heman, & Rose, 1999; Kim & Diamond, 2002).

Due to the moderate physiological arousal participants experienced in Study 2 and Study 4 the observed effects could be explained by the Yerkes-Dodson law. The high levels observed in Study 1 and the observed effects on cognition would also fit this curvilinear relationship. Even though executive functioning did not change from pre- to post-test, it might be possible that the post-test results reflect the decrease of executive functioning after it initially increased. Contradictory to the findings in Study 2, A. P. J. Roberts and Cole (2013) observed a decrease of executive functioning in university students after employing the same exercise manipulation. As discussed in Chapter 5, the physical fitness level of participants could account for this effect (Y.-K. Chang et al., 2012; Hüttermann & Memmert, 2014) underpinning the relationship between physiological arousal and executive functioning as described the Yerkes-Dodson law (see Figure 51). In line with this hypothesised relationship, studies have shown: (a) a positive effect of exercise on cognitive performance (Brisswalter et al., 2002; Hüttermann & Memmert, 2014; McMorris et al., 2011), and (b) a point to which cognitive performance increases before dropping off close to the anaerobic threshold (Chmura et al., 1994).

However, it has to be noted that there are a number of studies that suggest physical exercise intensity does not have any influence on cognitive performance (for a review, see Y.-K. Chang et al., 2012; Lambourne & Tomporowski, 2010). A possible explanation for the failure to unequivocally demonstrate the inverted-U effect are the cognitive and physical task characteristics (type, intensity, and duration) as well as cognitive
performance measures (Y.-K. Chang et al., 2012; Lambourne & Tomporowski, 2010; Pesce, 2009).

Besides the shape of the function of the relationship between physiological arousal and physical fitness as a moderator of this relationship, account has to be taken of the other cognitive shifts that take place in simulated (un)armed confrontations. The results in Studies 1, 2, and 4, equally demonstrate a cognitive shift towards the WM systems that are most functional in dealing with the threat that is placed in the specific situation upon the officer. In Study 1 under very high physical workload, where simple decisions had to be made by participants, a cognitive shift towards non-executive processes occurred, while executive functioning capacity remained the same. Furthermore, visuospatial processing was enhanced at the expense of phonological processing. The visuospatial processing system evolved much earlier than the phonological processing system (Fuster, 2001; Petersen & Posner, 2012; Posner & Petersen, 1990) and is assumed to be the primary sensory modality employed by humans when under increased threat or demand (Fuster, 2001; Klinge et al., 2010; Liddell et al., 2005; Parasuraman & Galster, 2013). Given that participants had to react to visual

Figure 51: The hypothesised relationship between executive functioning and physiological arousal
stimuli in the experiment (i.e. seeing when the roleplayer attacks and target identification while shooting), this cognitive shift can be interpreted as an adaptive response to the task at hand. In Study 2 complex tactical decision making under ‘moderate’ physical load was required of the officers. The observed increase in executive functioning reflects a functional and adaptive response to the perceived threat. In Study 4, complex decision making, involving the inhibition of offensive aggression under ‘moderate’ physical load was required from the officers. The increase in executive functioning under these environmental circumstances can again be termed functional, and thus adaptive.

The observed cognitive shifts in the current thesis are in line with other research involving novel or threatening encounters (Boulton, 2014; Condon, 2015; Kleider et al., 2010; Kleider & Parrott, 2009; Nieuwenhuys & Oudejans, 2011a; A. P. J. Roberts, 2012). Defensive behaviours in animals comprise several adaptive responses that are functional to the specific context, serving to reduce the perceived threat (D. C. Blanchard et al., 2001; 2011). Likewise, participants in the current studies displayed automatic shifts in cognitive functioning that could be interpreted as contextually appropriate and functional defensive responses. The proposed hypothesis that increases in executive functioning are mediated by physiological load (being the result of psychophysiological demand) also are in line with the argument of the observed changes being an adaptive response to threat.

When the organism has to cope with threat, arousal is the first step in activating the set of defensive responses, displaying vigilant behaviour (Kozlowska et al., 2015). An increase in executive functioning capacity at this stage would be functional and may be mediated by physiological arousal, as it has been shown in Study 2. The subsequent freezing and ‘fight or flight’ stages are different with regards to the needs of executive functioning capacity and physiological arousal. Because freezing (defined as being the ‘fight or flight’ response being put on hold (Kozlowska et al., 2015) has to prepare the organism for a possible upcoming ‘fight or flight’ response, physiological changes prepare the organism for maximal physical load (fighting or running). Simultaneously the individual has to focus attention on
the threat and think about possible evasion strategies depending on the contextual factors. As such, this process requires complex decision making that is assumed to be positively influenced by increased executive functioning. The ‘fight or flight’ stage requires maximum physical capacity. The required energy is provided through anaerobic metabolism thus producing high amounts of lactate, which will result in the termination of maximal physical load at some point, depending on the fitness level of the individual (Janković et al., 2015; Pekünlü, Ozkaya, & Yapıcıoglu, 2016). At this stage, the execution of simple, learned behaviours with maximal effort appear more functional than starting a creative decision-making process from the scratch. This is supported by the experiences of people who had to engage in lethal defensive behaviour (Jensen & Simpson, 2014; Jensen & Wrisberg, 2014) and models that claim that at high arousal states complex combat-related behaviour deteriorates (D. Grossman & Christensen, 2007; Siddle, 1995). As such, a cognitive shift towards non-executive functioning at the expense of executive functioning can be termed adaptive in the ‘fight or flight’ stage. Figure 52 depicts the hypothesised relationship between physiological arousal, executive functioning, and defensive responses.

![Figure 52](image)

**Figure 52:** A model linking the relationship of physiological arousal and executive functioning capacity to defensive responses
Even though the current thesis broadened the understanding of cognitive shifts under threat, there are several aspects of the hypothesised model that need further investigation and could be an avenue for further research in this area.

First, future studies should systematically take account of the fitness level of participants. As studies have employed students (A. P. J. Roberts & Cole, 2013), police officers from different countries and units (Study 1, 2, and 4; Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012), and marital artists (Study 2) a wide variety of fitness level can be assumed. If fitness level acts as a moderator of the shift of the functional relationship between physiological load and executive functioning, results of systematic investigations taking account of the participants’ fitness levels should reflect that.

Second, the status of offensive aggression should be further elaborated within the hypothesised model and other models of human defensive behaviour in general. Police officers have to be vigilant towards threat cues and display ‘classic’ defensive behaviour when attacked. However, there are situations when officers have to display offensive aggression first, such as in raids, pursuing an active shooter, or taking somebody down because of threatening cues (Study 4). These situations are much more difficult to place within a hierarchical model of human defensive behaviour. Since offensive aggression can be displayed with many different levels of force, even when under threat (thus rendering it a defensive behaviour as well), the fight response in general encompasses more behavioural responses than the fight for survival. Therefore, further investigations of cognitive, physiological and affective changes under threat should take account of these differences and systematically manipulate them.

8.1.2 Representative Learning Environments for Police Use of Force Training

The purpose of police use of force training is to develop transferable skills that help officers prevent and deal with (un)armed confrontations in the field (Staller & Zaiser, 2015b). The transferability of skill is dependent on the
representativeness of the learning environment, with its two components: (a) functionality, and (b) action fidelity (see Chapter 2 for details). The results of the current thesis help to broaden the understanding on how to design representative learning and testing environments in the context of police use of force training.

Study 1 provided evidence that the use of NLTA and CA in training settings does not differ with regards to psychophysiological load, when employed in a training setting. With regards to training settings that employ a human target, health and safety considerations do not allow for the use of CA. Study 1 suggests that the constraints under which the officer must operate (with regards to the internal state of the police officer) do not differ between the two systems. As long as weapon manipulation procedures stay the same (action fidelity), NLT weapons operate at the same level of representativeness as live-fire weapons. It is noteworthy that these results support the use of NLT weapons on human (i.e. scenario training) and non-human targets (i.e. marksmanship training), reflecting previous results (Getty, 2014; Kratzig, 2013) that question the amount of CA use in police use of force training.

The functionality component of representative learning and testing design requires the participant to perform under the constraints of the criterion context. Study 1, 2, and 4, supporting previous results of cognitive shifts taking place in (simulated) use of force incidents (Boulton, 2014; Condon, 2015; Kleider et al., 2010; Kleider & Parrott, 2009; Nieuwenhuys & Oudejans, 2011a; A. P. J. Roberts, 2012) indicate that: (a) different tasks (i.e. fight for survival using firearms or fighting skills or achieving compliance) elicit different levels of psychophysiological load, (b) physiological load may be a moderator of executive functioning, and (c) cognitive shifts supporting systems that are required to cope with the threat (reduction of phonological processing to enhance resources for visuospatial processing) take place. In this context, representative learning environments could manipulate the levels of psychophysiological load under which any given policing task has to be performed. Furthermore, physiological load could be artificially induced (i.e. exercise in order to get the HR to a specific level) before completing a
policing task in a training setting. Finally, it is important to design learning environments that are as representative as possible in order to facilitate the cognitive shifts that take place. Police officers have to learn to cope with demands under these alterations in cognitive processing, thus making skill transfer more efficient as suggested in recent research in the use of force domain (Jager et al., 2013; Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b). Likewise, as shown in Study 4, self-control demands that are placed upon the officer before engaging in a police-citizen encounter influence tactical decision making and thus represent an internal state (constraint) under which the officer must operate. In terms of representativeness this reflects the functionality of the task, whereas the inhibitory effort that the police officer has to continuously display refers to the action fidelity of the task. The systematic use of the ego depletion state in training settings can thus enhance the level of representativeness in learning and testing designs.

With regards to valid cue design, the current work also offers additional insight. Valid cues are a crucial component in learning environments fostering the development of tactical decision making (Staller & Abraham, 2016). Study 3 supports this argument by showing that training and exposure to (simulated) violent confrontations does not lead to the development of traditional threat-related attentional biases measured through existing laboratory measures (Naim et al., 2015). In order to fully investigate the original hypothesis for Study 3 new laboratory measures, similar to those used for combat-related attentional bias, will need to be developed and more appropriate experimental designs used.

8.1.3 Threat-Related Attentional Bias

Study 3 provided evidence that neither police officers nor martial artists developed threat-related attentional bias as measured via the Dot Probe (MacLeod et al., 1986) and the Emotional Stroop Task (P. Smith & Waterman, 2005). However, the experience of police officers does not automatically involve exposure to threat (Schmalzl, 2008). Likewise, experience in martial arts does not necessarily equate to a similar amount of
exposure to simulated threat situations, like representatively designed scenario trainings (Staller et al., 2016). Future studies should address this issue by accounting for the actual amount of exposure to (simulated) violent confrontations.

Furthermore, as discussed in the section above, there are reasons to believe that the nonspecific (in terms of violent confrontations) tests of threat-related attentional bias that have been used are not capable of detecting any functional threat- or combat-related attentional bias. In line with this argument, R. M. Todd et al. (2015) observed functional threat-related attentional bias in individuals exposed to real threats (combat experience in Afghanistan) measured with combat-related words in the test (instead of nonspecific threat- and aggression-themed words). Hence, future research should employ the use of combat-related cues in order to capture functional cognitive biases. However, since human defensive behaviour comprises ‘hard wired’ structures in combination with learned behaviour (Adolphs, 2013; R. M. Todd et al., 2015), the non-existence of any threat-related attentional bias in possibly less-exposed police officers and martial artists is still consistent with the original hypothesis.

8.1.4 Ego Depletion and Self-Regulation

The results of Study 4 add to the body of knowledge about the effects of ego depletion on aggression (Barlett et al., 2016; DeWall et al., 2007; Stucke & Baumeister, 2006; Vohs et al., 2011). By showing that ego depleted police officers engaged earlier in offensive aggression than non-depleted officers the results provided evidence that demands on self-control affect aggressive behaviour in a subsequent police-citizen interaction. This is the first demonstration that depleted self-control resources transfer to observable physical aggression. Since there is evidence that self-control performance varies across contexts (J. R. Cohen & Lieberman, 2010; Cortes et al., 2014) the obtained results are important for both the aggression and the police use of force domain.

In line with self-control theories (Baumeister et al., 1998; J. J. Reynolds & McCrea, 2016) and theories explaining aggressive behaviour (C. Anderson &
Bushman, 2002; Barlett et al., 2016; Ferguson & Dyck, 2012; Finkel, 2014; Finkel et al., 2012; Stucke & Baumeister, 2006), the inhibition of impulsive behaviour (i.e. aggression towards a provocative citizen) plays a crucial role in maintaining self-regulated behaviour.

It is suggested that executive functioning plays a crucial role in self-regulation (Hofmann et al., 2012). However, there is considerable debate about the relationship between these two constructs (Hofmann et al., 2012). In Study 4, executive functioning increased and did not differ between the ego-depleted and the non-depleted group, suggesting that the effect of ego depletion was not due to a change in executive function. Even though in Study 1, the actual decisions of SOU officers were not recorded, the lack of inhibitory regulation was observed. For example, some of the officers shot at the wrong targets, even though they have been told otherwise. Compared to Study 4, participants in Study 1 showed no increase in executive functioning and operated near their physical limits (as shown by mean HRs). Even though this cannot count as scientific evidence, this anecdotal report provides an idea how to further investigate self-regulation in police use of force settings. By measuring the actual outcome of the decisions made in the experimental setting in Study, more information about the results of inhibitory efforts could be obtained. Thus, future research could focus on the role of executive functioning in maintaining self-regulation in police-citizen encounters.

8.2 Limitations

8.2.1 Human Defensive Behaviour

Research in the domain of human defensive behaviour is a challenging topic as it is unethical to place humans in life threatening situations. The current thesis employed the simulated (un)armed confrontation paradigm (Staller & Cole, 2016) in order to investigate human defensive behaviour in an ethical manner. However, the simulation of confrontations in scenario-based experiments assumes that the observed psychophysiological load occurs under conditions that are similar to those found in real police-citizen encounters. However evidence-based data are extremely limited (Andersen et al., 2016), especially physiological data from real life-threatening combat
situations are missing. It can be assumed that the adaptations that occur during real violent conflict situations would be much greater as a reflection of greater emotional demand (e.g. fear for one’s life). Future studies may record physiological data of real world confrontations by continuously recording physiological data of police officers during their normal activities (e.g. with a memory belt) over a longer period, aiming at capturing physiological data when real conflict situations unfold.

8.2.2 Physiological Measures

Even though it is reasonable to assume that the reported studies produced psychophysiological load measures of cortisol would add weight to such assumptions as it is a widely used biological marker (Charmandari et al., 2005; Chrousos, 1992a; Tsigos & Chrousos, 2002). However, studies have suggested that HR and temperature are positively correlated with increases in cortisol implying the added expense, participant resistance, and ethical problems argue against measuring cortisol (Steptoe et al., 2005).

In the current studies physical, cognitive, and emotional demand led to similar patterns of physiological arousal. Hence, the effects of the specific demands are difficult to separate (Dolan, 2002; Frijda, 1987; LaBar, Gitelman, Parrish, & Mesulam, 1999; Paus, Koski, Caramanos, & Westbury, 1998), so it is difficult to determine which demand led to the rise in physiological arousal in the current studies. However, in real use of force incidents these demands are deeply intertwined representing constraints under which the officer has to operate (Driskell & Johnston, 1998; K. R. Murray, 2004; Wollert et al., 2011) and so it is questionable whether they can ever be separated. However, future studies should consider methods for separating these components and thus examining the effects of the specific load that is placed upon an individual, while acknowledging drawbacks in terms of ecological validity.

8.2.3 Cognitive Measures

Due to the multifaceted nature of cognitive tasks, there is a lack of precise understanding of the constructs that they measure (Jacoby, 1991; A. B.
Morgan & Lilienfeld, 2000). Furthermore, there is the possibility of practice effects, when tests were applied in pre- and post-est sessions (E. Strauss et al., 2006). Some tests in the current project (Study 3) were administered online at home and while under supervision of the experimenter. Standardised procedures, for example making sure that the participant does not get distracted could not be assured for those who got tested at home (Birnbaum, 2004).

Another key concern of the current work is the domain specificity of tests. Test protocols were used that were validated for use on the general public, where deviant behaviour is seen as pathological and not adaptive. Hence, recent findings suggest that adaptive responses are domain-specific, like for threat-related attentional bias (R. M. Todd et al., 2015) and risk taking propensity (J. E. C. Lee & Blais, 2014). Further research is required using domain specific measures in order to capture adaptive effects. Similarly, ego depletion manipulations have to take into the account the special populations, such as police officers. For example, the use of disgusting video stimuli has been proven to be an effective manipulation for students (P. Christiansen et al., 2012), but it did not work with police officers (see Study 4.2).

Finally, a potential limitation of the current research is that except for Study 4 measures of tactical decision making performance were not taken. It was pre-agreed with police gatekeepers that the actual decisions made would not be recorded. This was important in order to gain trust and build the relationship with the Hessian Police Academy. However, as trust was established, actual decisions (Study 4) could be recorded and reported. Future research will benefit from studying the relationship between individual tactical decision making ability and cognitive performance.

8.2.4 Police Data Challenges

Collecting data within police use of force training presented some significant challenges, such as limited testing time. Due to pre-agreed data collection sessions during officers’ training days, data could only be collected on these days. This limited the number of participants, especially when some
participants had spontaneously to fulfil operational obligations. In order to test as many participants as possible on these days, the organisation of the experiments was tight, leaving no personnel for observing participants at the resting phases. Therefore, it could be that police officers were not as quiet and sitting still as they were supposed to be.

8.3 Theoretical Implications

8.3.1 Implications for Theories of Human Defensive Behaviour and Executive Functioning

It could be inferred from the current work that there is validation for several defensive behaviour and central executive theories suggesting resource allocation limits and flexibility of cognitive functioning and behaviour. The hypothesised model of the relationship between physiological arousal and executive functioning in situations of threat (see Section 8.1.1) supports findings from the literature that suggest that once resource allocation has reached capacity limits automatic processes (non-executive functioning) becomes the favoured modality as they do not create significant demand (Garavan, 1998; McElree, 2001; Sliwinski et al., 2006). It seems that executive processing is enhanced as long as the limited WM capacity system can re-allocate resources at the expense of others (i.e. phonological processing) and as long as resources are available. As physiological demand increases (in combat situations), individuals expend a disproportionate amount of attentional capacity attending to them (A. P. J. Roberts, 2012), switching between task relevant and task irrelevant cognitions (Driskell & Johnston, 1998), thus limiting executive processing capacity for the task at hand.

8.3.2 Implication for Theories of Self-Regulation and Ego Depletion

The current work adds to the growing body of research that suggests that ego depletion affects aggressive behaviour (Barlett et al., 2016; DeWall et al., 2007; Stucke & Baumeister, 2006; Vohs et al., 2011). It is the first study providing evidence that depleted self-control resources actually result in physical forms of aggression. As such, the simulated (un)armed
confrontation paradigm provides the opportunity to operationalise physical aggression after depletion manipulation.

The results provide validation for theories of self-control management (Baumeister & Heatherton, 1996; J. J. Reynolds & McCrea, 2016; Vohs et al., 2011). After exercising self-control in a first task, inhibition of impulsive responses (offensive aggression towards a provocative citizen) becomes more difficult, as shown by earlier offensive aggression in Study 4.2. These observations can be explained as depleted self-control resources that lead to impulsive behaviour (Baumeister, 2014; Muraven & Baumeister, 2000) or as problems of maintaining inhibitory effort in the face of impulsive modules that seek to carry out proponent behaviour (J. J. Reynolds & McCrea, 2016).

The data of police officers in conjunction with previous research on students (P. Christiansen et al., 2012) also indicates, that ego depletion manipulations are domain specific. While students had to display inhibitory effort to suppress emotions while watching a disgusting movie clip, police officers did not. It can be assumed that regular experiences with situations, eliciting disgust, had provided the officers with coping mechanism for such situations. Future studies should consider such effects when designing experiments.

8.3.3 Implications for Theories of Aggression

Concerning theories of aggression, the current work also provide further evidence for the newly developed \( I^3 \) theory of aggression (Finkel, 2007; 2014; Finkel et al., 2012). This meta theory proposes that a combination of: (a) instigating factors, (b) impelling factors, and (c) a lack of inhibiting factors, give rise to the display of aggression (Denson et al., 2012; DeWall et al., 2011; Finkel, 2014; Finkel et al., 2012; Slotter et al., 2012).

In Study 4, the personally provoking behaviour of the role player (instigating factor) in conjunction with diminished resources to inhibit the impulse to aggress offensively (inhibiting factors) lead to the display of offensive aggression. However, it is not clear how impelling factors (i.e trait aggressiveness) will affect behaviour in this naturalistic context. Furthermore, there is a lack of evidence base for enhancing inhibiting factors to aggress in
the context of police-citizen interaction. Future avenues of enquiry could focus on these issues in order to further validate the theoretical and practical application of the model.

8.3.4 Applications to Methodology

The current work has theoretical application in facilitating a better conceptual understanding of WM, executive functioning, threat-related attentional bias, ego depletion, methodology in cognitive psychology, and the ethological study of defensive behaviours in humans. In line with previous work on simulated (un)armed confrontations (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012), this thesis demonstrates that the collection of data in a naturalistic setting is possible and that future research does not need to be limited to laboratory settings with little real world application.

Furthermore, the integration of different disciplinary knowledge bases has been proven to be a fruitful endeavour and strengthens the argument for interdisciplinary collaboration for investigating real world problems. Combined with awareness of appropriate use of methodologies, future interdisciplinary research could provide work with greater theoretical scope and practical application.

8.3.5 Future Research

Overall, the current thesis provides a body of results that future research can build on. The proposed model of the hypothesised relationship between physiological arousal and executive functioning in the context of defensive behaviour provides a substantial number of routes for potential future study. For example, research could begin to fully examine individual fitness levels and its effects on executive functioning under threatening conditions. As it was found that executive functioning was affected by physiological load, further studies could systematically manipulate arousal within the same subjects in the context of defensive behaviour in order to obtain a full picture of the relationship.

In the context of representative learning design and the use of NLT weapons, simunition weapons have found to lead to the same psychophysiological
arousal levels as the use of CA. Future investigations could aim at replicating the results with other NLT and laser-based systems such as airsoft or the Saab ATES System. Furthermore, it would be valuable to examine if there are differences in skill development depending on the use of the systems. Systematic interventions with retention and transfer tests would answer the question of whether learning has occurred (see Kratzig, 2013).

The results with regards to threat-related attentional bias also provide various routes for further research: for instance, the non-existence of threat-related bias could be replicated with the use of angry faces instead of words. Another option lies in the use of combat-related words (R. M. Todd et al., 2015), in order to account for domain specific bias. Another line of inquiry could be to systematically control for various amounts and intensities of exposure to life-threatening events, in order to obtain a full picture of when and under what circumstances threat-related attentional bias occurs.

Finally, with regards to ego depletion and offensive aggression in police use of force incidents there are three main lines of further research that are practically relevant: (a) studies could examine different ego depletion manipulations, aiming at replicating the effect with a more naturalistic demand than the Cold Pressor Task from the policing perspective, (b) a line of research could cover potential training procedures that could reduce the effect of ego depletion in police-citizen encounters (Denson, Capper, Oaten, Friese, & Schofield, 2011) showed that self-control training could help overcome aggressive impulses in individuals high on trait aggressiveness. It would be valuable to replicate these findings in with police officers in the current naturalistic setting), and (c) the influence of impelling factors such as trait aggressiveness on the enactment of offensive aggression against a provocative citizen could be examined.

8.4 Practical Implications

The current thesis provides scientifically derived findings that could be applicable to police policy and police training.
8.4.1 Applications to Police Policy

The presented body of results demonstrate the complexities of tactical decision making in police-citizen encounters. An increased understanding of these processes may be useful in understanding performance in use of force encounters. Post-incident investigations, as well as the support by superiors and colleagues, should take account of these processes. Likewise, knowledge about the complexities of such situations may improve public trust in the accountability of police decision making.

Specifically, the maintenance of visual processing at the expense of auditory processing which has also been shown in previous research projects (Boulton, 2014; Condon, 2015; A. P. J. Roberts, 2012), indicate that a reduction of verbal inputs (from fellow officers and from higher command) might be beneficial for overall performance. Limiting these communications to an absolute minimum will prevent cognitive overload in such situations.

Furthermore, understanding the effects of exerting self-control on subsequent deployments to police-citizen encounters could help facilitate strategic planning before the encounter, in the sense that non-depleted officers primarily engage in the communication.

With regards to firearms training policy, the current research together with previous work on firearms training (Getty, 2014; Kratzig, 2013) provides evidence that in order to prepare for potential violent police-citizen encounters the use of CA can be massively reduced. By using NLTA firearms training accidents can be reduced and more representative training tasks can be implemented (see next section).

Finally, the results of the current work can be applied to police officer selection policy by identifying psychological traits and abilities that facilitate optimal functioning in situations under extreme demands. For example, identifying thresholds of the ability to maintain inhibitory effort in impulse-laden policing situations and applying the Cold Pressor Task as a practical ego depletion manipulation measure, selection tests could be improved. After further investigating the influence of trait aggressiveness on self-control in
police-citizen encounters this could be a variable that could be looked at in the selection process. Similarly, other variables, that have been identified to be an indicator of high levels of cognitive functioning (Boulton, 2014; Boulton & Cole, 2016; Condon, 2015; A. P. J. Roberts, 2012) could be used to identify potentially successful SOU officers and reduce drop outs from specialised courses.

8.4.2 Applications for Police Use of Force Training

Police officers have to sometimes operate in highly threatening and cognitively demanding environments (Greenwood-Ericksen et al., 2004). The results of the current thesis suggest that police use of force training is a crucial aspect in order to prepare the officer to optimally deal with potentially violent situations. Based on the obtained results, it is suggested that: (a) to further optimize the use of representative learning and testing environments, (b) to use training methods that facilitate self-control, and (c) to ensure that police officers develop and maintain a high level of physical fitness.

8.4.2.1 Representative Learning and Testing Environments

This work emphasises the use of representative learning environments in order to facilitate skill transfer to real world incidents. Previous work has also advocated for the use of representative learning design without employing this term (Boulton, 2014; Colin et al., 2013; Condon, 2015; Helsen & Starkes, 1999; Jager et al., 2013; K. R. Murray, 2004; Neal et al., 2006; Nieuwenhuys et al., 2009; Nieuwenhuys & Oudejans, 2010; Oudejans, 2008; Oudejans & Pijpers, 2009; 2010; Renden, Nieuwenhuys, Savelsbergh, & Oudejans, 2015b; A. P. J. Roberts, 2012; Saunders et al., 1996). This work provided evidence that NLTA induces the same psychophysiological arousal level as CA. Thus firearms drills on the shooting range can be made much more dynamic and less linear, since safety considerations allow for far more variations in practice drills than CA. Furthermore, the use of NLTA allows for shooting at human targets, enabling the learner to be presented with valid cues compared to non-human targets. With regards to perceptual-cognitive skill development the presentation of valid cues is crucial for skill development (Abernethy, Schorer, Jackson, & Hagemann, 2012; D. Y. Mann,
Williams, Ward, & Janelle, 2007; Roca, Ford, McRobert, & Williams, 2013; Staller & Abraham, 2016; Staller & Zaiser, 2015b; P. Ward & Williams, 2003; A. M. Williams & Ericsson, 2005). Therefore, it is noteworthy that classic shooting range exercises do not allow for the development of skills that are warranted for highly dynamic, fast changing situations.

The results also emphasise, that use of force trainers have to pay close attention that the cues that are presented to the learner (in drills or scenario training by the training partner or the trainer) are reflecting cues in a real world situation. With a learning environment (the training partner or role player), that presents these valid cues and the opportunity of the trainee to learn these cues through practice, pattern recognition abilities can build up (Fadde & Klein, 2010; Kahneman & Klein, 2009), facilitating the development of tactical decision making capabilities in violent situations.

With regards to the functionality of practice activities, the trainer also has to design learning environments that reflect the constraints under which the officer has to operate. Thus, if cognitive shifts are taking place under specific demands, it is important for the trainee to operate under these demands with the respective cognitive shifts taking place. Furthermore, if different situational demands lead to different levels of physiological arousal, the officer has to train under differing physiological arousal conditions. For example, this can be achieved by inducing physiological arousal artificially thought the use of aerobic or anaerobic exercises.

8.4.2.2 Self-Control Training

The current work showed that depleted self-control resources result in difficulties in maintaining inhibitory control of the impulse to aggress. Hence, police use of force training would benefit from training methods aimed at increasing the self-control capacity. By tailoring training protocols from self-control training (Denson et al., 2011; E. Miles et al., 2016) to police use of force training settings, impulsive aggressive behaviour of police officers could be reduced. Furthermore, and in line with the argument of training under ecologically valid constraints, training under ego depleted conditions may
provide the individual with the experience to cope with this internal state while displaying optimal behaviour.

8.4.2.3 Physical Fitness of Police Officers

The fitness level of the individual seems to be a crucial factor in maintaining executive functioning at high levels of physiological load (Hüttermann & Memmert, 2014). On the other hand, the fitter the individual the less physiological load will be induced by the physical workload (D. S. Holmes & Roth, 1985; McEwen, 2000; R. Norris et al., 1990; C. H. Turner & Pavalko, 1998). Hence, the maintenance of reserve capabilities is a beneficial factor for reducing the effects of demand (physical, cognitive, and emotional) on cognitive functioning (A. P. J. Roberts, 2012). In this scenario physiological arousal is reduced, resulting in less reduced impact on executive functioning than on unfit police officers (Baddeley, 2007; Driskell & Salas, 1991; 2006; Keinan, 1987; Sliwinski et al., 2006; Verhaeghen & Basak, 2005). Taken together, physically fitter officers are able to maintain executive functioning under: (a) higher physical workloads, and (b) higher physiological load than unfit individuals. The developing and maintaining physical fitness in police officers is mandatory in order to ensure effective operation in demanding situations.

8.5 Conclusion

The current work provided evidence, that (a) tactical decision making is influenced by physiological load and ego depletion, and (b) in training settings the use of NLTA and CA do not differ with regards to psychophysiological demand. The goal of the current work was not to provide an exhaustive description of the cognitive processes taking place in police use of force incidents, but to elicit key aspects that contribute to tactical decision making and to training settings within this context. Overall, the current thesis employed a quantitative experimental approach to examine the influence of various factors on police tactical decision making. The key findings from the current thesis reflect the contribution from a variety of research areas in order to provide a rich narrative about cognition in violent encounters and how it is affected. In sum, the current body of research has
led to a broader understanding of key aspects in the field of tactical decision making and provides several avenues for further research and the optimisation of police use of force training.
References


special forces police officers. *SAGE Open, 5*(2), 1–8.  
http://doi.org/10.1177/2158244015590446


http://doi.org/10.1016/j.adolescence.2003.10.009

http://doi.org/10.1111/1467-9280.00061


http://doi.org/10.1108/13639510010355611

http://doi.org/10.1108/13639510210429437

http://doi.org/10.1108/13639510110382232


Executive Forum, 3, 41, 49.
http://doi.org/10.3389/conf.fpsyg.2015.66.00027
http://doi.org/10.1016/S0079-7421(08)60422-3
. New York: Springer.
http://doi.org/10.1016/0167-8760(94)90042-6
http://doi.org/10.1080/14640746808400159
http://doi.org/10.1126/science.1736359


Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the


http://doi.org/10.1002/ab.21648

the research meet the needs? In L. J. Alison (Ed.), The forensic
psychologist’s casebook: Psychological profiling and criminal
investigation (pp. 47–66). Devon: Willan Publishing.


force during simulated incidents. Legal and Criminological Psychology,
5(1), 107–121.

Barton, J., Vrij, A., & Bull, R. (2002). Questions, preconceptions and
reactions: Police use of lethal force in Britain. International Journal of

damage and selective impairment of auditory verbal short-term memory.
http://doi.org/10.1016/0028-3932(82)90101-4

high-reliability occupations: Live-fire exercises for civilian firefighters.
http://doi.org/10.1177/0018720811418224

stress, and performance: A model of decision making under acute stress.
In E. Salas & G. A. Klein (Eds.), Linking expertise and naturalistic
decision making (pp. 139–158). Mahwah, NJ: Lawrence Erlbaum

Neuropsychologia, 65, 313–319.
http://doi.org/10.1016/j.neuropsychologia.2014.08.012

http://doi.org/10.1207/s15327965pli0701_1

depletion: Is the active self a limited resource? Journal of Personality and


Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and


Bochenek, A., & Staller, M. S. (2014). Gewalt gegen Polizeibeamte – Zum Bedarf eines Kompetenzmodells zur Abwehr von gewalttätigen Angriffen...


Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C.,


verbal working memory, 7(5-6), 549–561.  
http://doi.org/10.1080/096582193877742
http://doi.org/10.1016/S0010-9452(75)80026-8


Ellrich, K., & Baier, D. (2014). *Gewalt gegen niedersächsische Beamteinnen*
und Beamte aus dem Einsatz und Streifendienst: Zum Einfluss von
personen-, arbeits- und situationsbezogenen Merkmalen auf das
Gewaltopferrisiko [Violence against North-Rhine-Westphalian police
officers: The influence of personal, work-related and situational
characteristics on the risk of victimization] (No. 123). Hannover:
Kriminologisches Forschungsinstitut Niedersachsen e.V.

Ellrich, K., & Baier, D. (2016). The influence of personality on violent
victimization—A study on police officers. Psychology, Crime & Law, 1–

Ellrich, K., Baier, D., & Pfeiffer, C. (2010a). Gewalt gegen Polizeibeamte:
Ausgewählte Befunde zu den Tätern der Gewalt [Violence against police
officers: Selected results regarding the perpetrators]. Hannover:
Kriminologisches Forschungsinstitut Niedersachsen e.V.

Forschungsberichts des Projekts „Gewalt gegen Polizeibeamte“ zu den
Tätern der Gewalt [Main results of the research project ‘Violence against
police officers’ regarding perpetrators]. Hannover: Kriminologisches
Forschungsinstitut Niedersachsen e.V.

Befunde zu Einsatzbeamten, Situationsmerkmalen und Folgen
von Gewaltübergriffen [Violence against police officers: Results regarding
police officers, situational parameters and consequences of violence].
Hannover: Kriminologisches Forschungsinstitut Niedersachsen e.V.

Ellrich, K., Pfeiffer, C., & Baier, D. (2010c). Gewalt gegen Polizeibeamte:
Gewalt gegen Polizeibeamte: Begleittext zu "7 Thesen zur Gewalt gegen
Polizeibeamte. Erste Ergebnisse einer Polizeibefragung in den
Bundesländern Berlin, Brandenburg, Bremen, Niedersachsen,
Mecklenburg-Vorpommern, Rheinland- Pfalz, Sachsen-Anhalt, Saarland,
Schleswig-Holstein und Thüringen" [Violence against police officers:
Supplement to "7 theses regarding violence against police officers"].
Hannover: Kriminologisches Forschungsinstitut Niedersachsen e.V.

dominant and subordinate mice. Hormones and Behavior, 10(2), 156–


Human Kinetics.


http://doi.org/10.1037/a0015433

http://doi.org/10.1006/brcg.2000.1221


http://doi.org/10.1177/154230500706100103


http://doi.org/10.1093/brain/121.7.1239


http://doi.org/10.1108/13639510710753289


http://doi.org/10.1080/02699938708408043


http://doi.org/10.1111/j.1467-9280.2009.02372.x


http://doi.org/10.1080/1354678042000046


http://doi.org/10.1177/002242788001700107


Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of
http://doi.org/10.1007/BF02289823

http://doi.org/10.1177/154193120404801634

http://doi.org/10.1177/10986110250365


http://doi.org/10.1017/jmo.2015.31


Inventory (STAI). *Psychological Assessment, 19*(4), 369–381.  
http://doi.org/10.1037/1040-3590.19.4.369


http://doi.org/10.1016/j.jflm.2014.12.007


Haller, J., Raczkewy-Deak, G., Gyimesine, K. P., Szakmary, A., Farkas, I., & Vehq, J. (2014). Cardiac autonomic functions and the emergence of
violence in a highly realistic model of social conflict in humans. *Frontiers in Behavioral Neuroscience, 8*(278), 364.
http://doi.org/10.3389/fnbeh.2014.00364

http://doi.org/10.1123/IJSPP.2015-0017


http://doi.org/10.1080/14640749108401001


http://doi.org/10.1080/10615800802272251

http://doi.org/10.1177/0269881114550352


http://doi.org/10.1097/00001756-200001170-00009


352


Inquisit 4.0.5.0. (2014). Inquisit 4.0.5.0. Seattle, WA: Millisecond Software LLC.


http://doi.org/10.19044/esj.2016.v12n1p488

http://doi.org/10.2114/jpa2.28.155

http://doi.org/10.1016/S0896-6273(00)80734-5

http://doi.org/10.1016/S0149-7634(01)00042-2

http://doi.org/10.1007/BF02855682


http://doi.org/10.1080/13803390701416197


England: Harvard University Press.
http://doi.org/10.4159/harvard.9780674183940

http://doi.org/10.3389/fpsyg.2016.00653


http://doi.org/10.1249/MSS.0b013e3181cbee11

http://doi.org/10.1016/j.brainres.2007.08.057


http://doi.org/10.1002/bdm.1784


Lejuez, C. W., Aklin, W., Daughters, S., Zvolensky, M., Kahler, C., & Gwadz,


Löw, A., Weymar, M., & Hamm, A. O. (2015). When threat is near, get out of here: Dynamics of defensive behavior during freezing and active


Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A.


http://doi.org/10.1016/j.neubiorev.2010.08.011


or don't shoot? Why police officers are more inclined to shoot when they are anxious. *Emotion.* http://doi.org/10.1037/a0025699


Pennington, B. F., & Bennetto, L. (2008). Main effects or transactions in the neuropsychology of conduct disorder? Commentary on “The neuropsychology of conduct disorder.” *Development and*


Pfeiffer, C., Ellrich, K., & Baier, D. (2010). *7 Thesen zur Gewalt gegen Polizeibeamte* [*7 theses regarding violence against police officers*]. Hannover: Kriminologische Forschungsinstitut Niedersachsen e.V.


Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F.,


Neuropsychological Society, 18(05), 827–833.
http://doi.org/10.1017/S1355617712000586
Intelligence, 29(2), 169–186.
http://doi.org/10.1037//0021-843X.104.4.614
http://doi.org/10.1007/s11606-008-0892-6
http://doi.org/10.1080/00335557043000203
http://doi.org/10.1093/biomet/52.3-4.591
Sharps, M. J. (2010). Processing under pressure: Stress, memory and


Smith, P. K. (2007). Why has aggression been thought of as maladaptive? In P. Hawley, T. Little, & P. Rodkin (Eds.), *Aggression and adaptation: The bright side to bad behavior* (pp. 65–83). Mahwah, NJ. http://doi.org/10.1037/e552522012-131


reisende-im-zug-mehrere-schwerverletzte-a-1103596.html


Reviews, 35(4), 1075–1079. 
http://doi.org/10.1016/j.neubiorev.2010.11.012

http://doi.org/10.1073/pnas.0409174102

http://doi.org/10.1017/CBO9781316257081.004

http://doi.org/10.1016/0001-6918(69)90055-9


http://doi.org/10.1109/TVCG.2014.23


404


Sun, I. Y., Payne, B. K., & Wu, Y. (2008). The impact of situational factors,


http://doi.org/10.1177/109861102250584

http://doi.org/10.1111/1745-9133.12193


http://doi.org/10.1016/j.neubiorev.2008.08.004

http://doi.org/10.1016/j.neubiorev.2011.11.009

http://doi.org/10.1007/s12160-009-9101-z


Thomason, M. E., Chang, C. E., Glover, G. H., Gabrieli, J. D. E., Greicius, M.


Thorndike, E. L., & Woodworth, R. S. (1901c). The influence of improvement in one mental function upon the efficiency of other functions III: Functions involving attention, observation and discrimination. *Psychological Review, 8*(6), 553–564.


van Rijn, S., & Swaab, H. (2015). Executive dysfunction and the relation with...


Belt Books.


http://doi.org/10.1016/S0065-3454(08)60192-8


http://doi.org/10.1080/02699930903205698


http://doi.org/10.1016/S0165-0173(02)00248-5

http://doi.org/10.1016/S1350-4533(01)00112-6


Appendices
Appendix A: Ethical Approval

The Research Ethics Committee of the University of Liverpool provided ethical approval for the work in the current thesis.

Ethical approval was granted under the following reference numbers:

RETH000297 Cognitve testing in the context of police use of force training
IPHS-1213-LB-100 Threat-related attentional bias and police use of force
IPHS-1314-346 Ego depletion and police use of force
Appendix B: Consent Form – Study 1 and 2

English Version:

CONSENT FORM

Project title: Executive Functioning and Police Use of Force

Material gathered during this research will be treated as confidential and securely stored. Please answer each statement concerning the collection and use of the research data by putting your initials in the box.

1. I have read and understood the information sheet.
2. I have been given the opportunity to ask questions about the study.
3. I have had my questions answered satisfactorily.
4. I understand that I can withdraw from the study at any time without having to give an explanation.
5. I understand that none of my personal details will be recorded and that my responses are anonymous.
6. I agree not to discuss this with anyone outside of the study group.
7. I agree to take part in the above study.

Participant
Name (printed) ____________________________________________
Signature __________________________ Date_______________

Researcher
Name (printed) ____________________________________________
Signature __________________________ Date_______________

Feel free to contact us if you have any further questions.
The name of the main investigator is Dr Jon Cole (0151 794 2175 or joncole@liverpool.ac.uk)

1 copy for participant, 1 copy for researcher
EINVERSTÄNDNISERKLÄRUNG

Studienbeschreibung:

Auswirkungen von bewaffneten Auseinandersetzungen auf die Arbeitsweise des Gehirns

Sämtliches im Verlauf der Studie gesammeltes Material wird höchst vertraulich behandelt und sicher aufbewahrt. Bitte bestätigen Sie jede Aussage bezüglich der Erhebung und der Nutzung der Forschungsdaten, indem Sie Ihre Initialen in das Kästchen schreiben.

8. Ich habe das Informationsblatt gelesen und verstanden.

9. Mir wurde die Möglichkeit gegeben, Fragen zur Studie zu stellen.

10. Alle meine Fragen wurden zufriedenstellend beantwortet.


12. Ich verstehe, dass keine persönlichen Details von mir erhoben werden und dass meine erhobenen Daten und gemessenen Werte anonym gespeichert werden.

13. Ich stimme zu, dass ich die Inhalte der Studie nicht mit anderen Personen außerhalb der Probandengruppe bespreche.


Teilnehmer Name (Druckbuchstaben)________________________________________

Unterschrift ___________________________ Date__________

Forscher

Name (Druckbuchstaben)________________________________________

Unterschrift ___________________________ Date__________

Bitte kontaktieren Sie uns, wenn Sie weitere Fragen haben.
Der Name des Forschungsleiters ist: Dr. Jon Cole
Kontaktdaten: +44 151 794 2175 oder joncole@liverpool.ac.uk
Appendix C: Online Consent Form – Study 3

English version:

The current study investigates the influence of various factors while identifying threatening stimuli in the environment. The experiment is carried out by Canadian, German and UK Authorized Firearms Officers.

Your data will be stored anonymously and securely whilst being analyzed in order to be used in a psychological research paper.

Thank you for your time and support.

If you have any queries about the experiment or have any issues with the experiment please contact Prof. Jon C. Cole.

Prof. Dr. Jon Cole
Tactical Decision Making Research Group
Department of Psychological Sciences
University of Liverpool

Contact:
J.C.Cole@liverpool.ac.uk
+44 (0) 151 794 2175

Das Experiment dauert ca. 30-40 Minuten und besteht aus mehreren Einzeltests. Bitte stellen Sie sicher, dass Sie in dieser Zeit nicht gestört oder abgelenkt werden, da auch Ihre Reaktionszeit gemessen wird.

Ihre Daten werden anonym gespeichert und ausgewertet, so dass kein Rückschluss auf Ihre Identität möglich ist.

Vielen Dank für Ihre Unterstützung.

Mario Staller & Jon Cole

Dr. Mario Staller & Prof. Dr. Jon Cole
Tactical Decision Making Research Group
Department of Psychological Sciences
University of Liverpool

Kontakt:
mario.staller@liverpool.ac.uk
+49 (0)151-22289759
### Appendix D: Stait-Trait Anxiety Inventory

**English version:**

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you feel right now. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feeling best.

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>almost never</th>
<th>sometimes</th>
<th>often</th>
<th>almost always</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I feel calm</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>I feel secure</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>I am tense</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>I feel strained</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>I feel at ease</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>I feel upset</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>I am presently worrying over possible misfortunes</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>I feel satisfied</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>I feel frightened</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>I feel comfortable</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>I feel self-confident</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>I feel nervous</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>I am jittery</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>I feel indecisive</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>I am relaxed</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>I feel content</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>I am worried</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>I feel confused</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>I feel steady</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>I feel pleasant</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you **generally** feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you **generally** feel.

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>21  I feel pleasant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22  I feel nervous and restless</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23  I feel satisfied with myself</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24  I wish I could be as happy as others seem to be</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25  I feel like a failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26  I feel rested</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27  I am „calm, cool, and collected“</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28  I feel that difficulties are piling up so that I cannot overcome them</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29  I worry too much over something that really doesn’t matter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30  I am happy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31  I have disturbing thoughts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32  I lack self-confidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33  I feel secure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34  I make decisions easily</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35  I feel inadequate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36  I am content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37  Some unimportant thought runs through my mind and bothers me</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38  I take disappointments so keenly that I can’t put them out of my mind</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39  I am a steady person</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40  I get in a state of tension or turmoil as I think over my recent concerns and interests</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
German version:

Im folgenden Fragebogen finden Sie eine Reihe von Feststellungen, mit denen man sich selbst beschreiben kann. Bitte lesen Sie jede Feststellung durch und wählen Sie aus den vier Antworten diejenige aus, die angibt, wie Sie sich JETZT, d.h. IN DIESEM MOMENT, fühlen. Kreuzen Sie bitte bei der Feststellung die Zahl unter der von Ihnen gewählten Antwort an. Es gibt keine richtigen oder falschen Antworten. Überlegen Sie bitte nicht lange und denken Sie daran, diejenige Antwort auszuwählen, die ihren AUGENBLICKLICHEN Gefühlszustand am besten beschreibt.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Überhaupt nicht</th>
<th>Ein wenig</th>
<th>ziemlich</th>
<th>sehr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ich bin ruhig</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Ich fühle mich geborgen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Ich fühle mich angespannt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Ich bin bekümmert</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Ich bin gelöst</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Ich bin aufgeregt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Ich bin besorgt, dass etwas schief gehen könnte</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Ich fühle mich ausgeruht</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>Ich bin beunruhigt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Ich fühle mich nicht wohl</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>Ich fühle mich selbstsicher</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>Ich bin nervös</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>Ich bin zappelig</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>Ich bin verkrampft</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>Ich bin entspannt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>Ich bin zufrieden</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>Ich bin besorgt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>Ich bin überreizt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>Ich bin froh</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>Ich bin vergnügt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Im folgenden Fragebogen finden Sie eine Reihe von Feststellungen, mit denen man sich selbst beschreiben kann. Bitte lesen Sie jede Feststellung durch und wählen Sie aus den vier Antworten diejenige aus, die angibt, wie Sie sich **IM ALLGEMEINEN** fühlen. Kreuzen Sie bitte bei der Feststellung die Zahl unter der von Ihnen gewählten Antwort an. Es gibt keine richtigen oder falschen Antworten. Überlegen Sie bitte nicht lange und denken Sie daran, diejenige Antwort auszuwählen, die am besten beschreibt, wie sie sich im **ALLGEMEINEN** fühlen.

<table>
<thead>
<tr>
<th>Feststellung</th>
<th>Fast nie</th>
<th>Manchmal</th>
<th>Oft</th>
<th>Fast immer</th>
</tr>
</thead>
<tbody>
<tr>
<td>21 Ich bin vergnügt</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>22 Ich werde schnell müde</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>23 Mir ist zum Weinen zumute</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>24 Ich glaube, mir geht es schlechter als den anderen Leuten</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>25 Ich verpasse günstige Gelegenheiten, weil ich mich nicht schnell genug entscheiden kann</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>26 Ich fühle mich ausgeruht</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>27 Ich bin ruhig und gelassen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>28 Ich glaube, dass mir meine Schwierigkeiten über den Kopf wachsen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>29 Ich mache mir zu viel Gedanken über unwichtige Dinge</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>30 Ich bin glücklich</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>31 Ich neige dazu, alles schwer zu nehmen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>32 Mir fehlt es an Selbstvertrauen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>33 Ich fühle mich geborgen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>34 Ich mache mir Sorgen über mögliches Missgeschick</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>35 Ich fühle mich niedergeschlagen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>36 Ich bin zufrieden</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>37 Unwichtige Gedanken gehen mir durch den Kopf und bedrücken mich</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>38 Enttäuschungen nehme ich so schwer, dass sich sie nicht vergessen kann</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>39 Ich bin ausgeglichen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>40 Ich werde nervös und unruhig, wenn ich an meine derzeitigen Angelegenheiten denke</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
## Appendix E: Words Used in the Dot Probe Task

<table>
<thead>
<tr>
<th>Threat Words</th>
<th>Neutral Words</th>
<th>Threat Words</th>
<th>Neutral Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>PANICKY</td>
<td>CLARETS</td>
<td>KOPFLOS</td>
<td>ROTWEIN</td>
</tr>
<tr>
<td>DESPERATE</td>
<td>VARIABLE</td>
<td>VERWZEIFELT</td>
<td>VARIABLE</td>
</tr>
<tr>
<td>CRINGING</td>
<td>NECKLACE</td>
<td>ZURÜCKSCHRECKE</td>
<td>HALSKETTE</td>
</tr>
<tr>
<td>CAUTIOUS</td>
<td>EDITIONS</td>
<td>VORSICHTIG</td>
<td>VERSION</td>
</tr>
<tr>
<td>SCARED</td>
<td>PLANET</td>
<td>ERSCHROCKEN</td>
<td>PLANET</td>
</tr>
<tr>
<td>CANCER</td>
<td>SADDLE</td>
<td>KREBS</td>
<td>SATTEL</td>
</tr>
<tr>
<td>DANGER</td>
<td>EAGUE</td>
<td>GEFAHR</td>
<td>LIGA</td>
</tr>
<tr>
<td>ALARMED</td>
<td>SHELVES</td>
<td>ALAMIERT</td>
<td>REGALE</td>
</tr>
<tr>
<td>CORONARY</td>
<td>RECEIPTS</td>
<td>INFARKT</td>
<td>EINNAHMEN</td>
</tr>
<tr>
<td>TRAUMA</td>
<td>ENJOIN</td>
<td>TRAUMA</td>
<td>BEFEHLEN</td>
</tr>
<tr>
<td>EDGTY</td>
<td>RINK</td>
<td>GEREIZT</td>
<td>EISBAHN</td>
</tr>
<tr>
<td>EMERGENCY</td>
<td>LISTENING</td>
<td>NOTFALL</td>
<td>HÖREN</td>
</tr>
<tr>
<td>DISTRESS</td>
<td>CREATURE</td>
<td>BEDRÄNGNIS</td>
<td>KREATUR</td>
</tr>
<tr>
<td>UNSETTLED</td>
<td>ADVOCATED</td>
<td>UNGEKLÄRT</td>
<td>BEFÜRWORTEND</td>
</tr>
<tr>
<td>HOSTILE</td>
<td>ROLLING</td>
<td>FEINDLICH</td>
<td>ROLLEN</td>
</tr>
<tr>
<td>CONFUSED</td>
<td>ARRANGED</td>
<td>VERWIRRT</td>
<td>GEORDNET</td>
</tr>
<tr>
<td>TOUCHY</td>
<td>TINTED</td>
<td>EMPFINDLICH</td>
<td>GETÖNT</td>
</tr>
<tr>
<td>FRIGHT</td>
<td>SIPPED</td>
<td>SCHRECKEN</td>
<td>SCHLÜRFEND</td>
</tr>
<tr>
<td>WORRIED</td>
<td>CONTEXT</td>
<td>BESORGT</td>
<td>KONTEXT</td>
</tr>
<tr>
<td>NAUSEA</td>
<td>CONFER</td>
<td>ÜBELKEIT</td>
<td>VERLEIHEN</td>
</tr>
<tr>
<td>DISEASE</td>
<td>REMARKS</td>
<td>KRANKHEIT</td>
<td>ANMERKUNGEN</td>
</tr>
<tr>
<td>WARY</td>
<td>FOLD</td>
<td>VORSICHTIG</td>
<td>GEFALTET</td>
</tr>
<tr>
<td>SINNER</td>
<td>INTEGRAL</td>
<td>UNHEIMLICH</td>
<td>FALTE</td>
</tr>
<tr>
<td>MUTILATED</td>
<td>DECANTING</td>
<td>VERSTÜMMELT</td>
<td>UMGEFÄLLT</td>
</tr>
<tr>
<td>NEARVOS</td>
<td>OUTCOME</td>
<td>NERVÖS</td>
<td>ERGEBNIS</td>
</tr>
<tr>
<td>ASSAULT</td>
<td>BOTTLES</td>
<td>ÜBERFALL</td>
<td>FLASCHEN</td>
</tr>
<tr>
<td>SUCCOTATING</td>
<td>CONSTITUENT</td>
<td>ERSTICKEN</td>
<td>BESTANDTEIL</td>
</tr>
<tr>
<td>CATASTROPHE</td>
<td>APPROXIMATE</td>
<td>KATASTROPHE</td>
<td>ANNÄHERND</td>
</tr>
<tr>
<td>LOST</td>
<td>READ</td>
<td>VERLOREN</td>
<td>LESEN</td>
</tr>
<tr>
<td>HARM</td>
<td>NOON</td>
<td>SCHADEN</td>
<td>MITTAG</td>
</tr>
<tr>
<td>TENSE</td>
<td>BACKS</td>
<td>SPANNUNG</td>
<td>RÜCKSEITE</td>
</tr>
<tr>
<td>JUMPY</td>
<td>LOTUS</td>
<td>SCHRECKHAFT</td>
<td>LOTUS</td>
</tr>
<tr>
<td>DEFENSIVE</td>
<td>GEOMETRIC</td>
<td>ABWEHR</td>
<td>GEOMETRISCH</td>
</tr>
<tr>
<td>EMBARRASSED</td>
<td>MICROSCOPIC</td>
<td>PEINLICH</td>
<td>MIKROSKOPISCH</td>
</tr>
<tr>
<td>LETHAL</td>
<td>RACKET</td>
<td>TÖDLICH</td>
<td>TENNISCHLÄGER</td>
</tr>
<tr>
<td>APPREHENSION</td>
<td>INSTRUMENTAL</td>
<td>FESTNAHME</td>
<td>HILFREICH</td>
</tr>
<tr>
<td>AMBULANCE</td>
<td>MOMENTARY</td>
<td>KRANKENWAGEN</td>
<td>VORÜBERGEHEND</td>
</tr>
<tr>
<td>ATTACK</td>
<td>SEASON</td>
<td>ANGRIFF</td>
<td>JAHRESZEIT</td>
</tr>
<tr>
<td>HORROR</td>
<td>WAGONS</td>
<td>HORROR</td>
<td>WAGON</td>
</tr>
<tr>
<td>STRANGLING</td>
<td>SIGNATURE</td>
<td>ERWÜRGEN</td>
<td>UNTERSCHRIFT</td>
</tr>
<tr>
<td>AGITATION</td>
<td>FIREPLACE</td>
<td>UNRUHE</td>
<td>FEUERPLATZ</td>
</tr>
<tr>
<td>INSECURE</td>
<td>FETCHING</td>
<td>UNSICHER</td>
<td>BEZAUBERND</td>
</tr>
<tr>
<td>INCURABLE</td>
<td>RECLAIMED</td>
<td>UNHEILBAR</td>
<td>REGENERIERT</td>
</tr>
<tr>
<td>SUICIDE</td>
<td>SUMMERS</td>
<td>ELBSTMORD</td>
<td>SOMMER</td>
</tr>
<tr>
<td>FUNERAL</td>
<td>HUNTING</td>
<td>BEERDIGUNG</td>
<td>JAGD</td>
</tr>
<tr>
<td>GRIEVING</td>
<td>HALLMARK</td>
<td>TRAUERND</td>
<td>KENNZEICHEN</td>
</tr>
<tr>
<td>TRAGEDY</td>
<td>REQUEST</td>
<td>TRAGÄ–DIE</td>
<td>BITTE</td>
</tr>
<tr>
<td>WORTHLESS</td>
<td>BATTERIES</td>
<td>WERTLOS</td>
<td>BATTERIEN</td>
</tr>
<tr>
<td>UNPOPULAR</td>
<td>SHORELINE</td>
<td>UNBELIEBT</td>
<td>UFER</td>
</tr>
<tr>
<td>SAD</td>
<td>PAT</td>
<td>TRAURIG</td>
<td>KLAPS</td>
</tr>
<tr>
<td>Threat Words</td>
<td>Neutral Words</td>
<td>Threat Words</td>
<td>Neutral Words</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td>POINTLESS</td>
<td>LUNCHROOM</td>
<td>SINNLOS</td>
<td>KANTINE</td>
</tr>
<tr>
<td>GLUM</td>
<td>KITE</td>
<td>BEDRÜCKT</td>
<td>DRACHEN</td>
</tr>
<tr>
<td>REJECTED</td>
<td>QUANTITY</td>
<td>ABGELEHN</td>
<td>MENCE</td>
</tr>
<tr>
<td>POWERLESS</td>
<td>MULTITUDE</td>
<td>MACHTLOS</td>
<td>VIELZAHN</td>
</tr>
<tr>
<td>SICKLY</td>
<td>TOKENS</td>
<td>KRÄNKLICH</td>
<td>ZEICHEN</td>
</tr>
<tr>
<td>DEVASTATED</td>
<td>STAGECOACH</td>
<td>VERWÜSTET</td>
<td>POSTKUTSCHEN</td>
</tr>
<tr>
<td>INFERIOR</td>
<td>SHEARING</td>
<td>UNTERLEGEN</td>
<td>SCHUR</td>
</tr>
<tr>
<td>FAIL</td>
<td>EARS</td>
<td>SCHEITERN</td>
<td>OHREN</td>
</tr>
<tr>
<td>DEFEAT</td>
<td>MUSEUM</td>
<td>NIEDERLAGEN</td>
<td>MUSEUM</td>
</tr>
<tr>
<td>LONELY</td>
<td>JERSEY</td>
<td>EINSAM</td>
<td>TRIKOT</td>
</tr>
<tr>
<td>FATIGUE</td>
<td>AVOCADO</td>
<td>MÜDIGKEIT</td>
<td>AVOKADO</td>
</tr>
<tr>
<td>IGNORED</td>
<td>LIGHTED</td>
<td>VERNACHLÄSSIGT</td>
<td>ERELEUTCHT</td>
</tr>
<tr>
<td>SLUGGISH</td>
<td>TEXTURED</td>
<td>TRÄGE</td>
<td>STRUKTURIERT</td>
</tr>
<tr>
<td>GROAN</td>
<td>FLUTE</td>
<td>STÖHNNEN</td>
<td>FLÖTE</td>
</tr>
<tr>
<td>MISERABLE</td>
<td>STATEWIDE</td>
<td>ELENDE</td>
<td>LANDESWEITE</td>
</tr>
<tr>
<td>BROODING</td>
<td>PROTEINS</td>
<td>GRÜBELEI</td>
<td>PROTEINE</td>
</tr>
<tr>
<td>UNFORTUNATE</td>
<td>COMMODITIES</td>
<td>UNGLÜCKLICH</td>
<td>ROHSTOFFE</td>
</tr>
<tr>
<td>DESPISED</td>
<td>TOMATOES</td>
<td>VERACHTET</td>
<td>TOMATEN</td>
</tr>
<tr>
<td>PATHETIC</td>
<td>CLEANERS</td>
<td>ARMSELIG</td>
<td>REINIGER</td>
</tr>
<tr>
<td>HUMILIATED</td>
<td>WATERPROOF</td>
<td>ERNIEDRIGT</td>
<td>WASSERFEST</td>
</tr>
<tr>
<td>HAZARD</td>
<td>BALLOT</td>
<td>GEFARH</td>
<td>ABSTIMMUNG</td>
</tr>
<tr>
<td>INTIMIDATED</td>
<td>COEFFICIENT</td>
<td>EINGESCHÜCHTERT</td>
<td>KOEFFIZIENT</td>
</tr>
<tr>
<td>QUAKING</td>
<td>ANAGRAM</td>
<td>ZITTERND</td>
<td>ANAGRAM</td>
</tr>
<tr>
<td>TIMID</td>
<td>SATIN</td>
<td>ÄNGSTLICH</td>
<td>SATIN</td>
</tr>
<tr>
<td>TEASE</td>
<td>AISLE</td>
<td>ÄRGERN</td>
<td>GANG</td>
</tr>
<tr>
<td>DISMAL</td>
<td>MIDWAY</td>
<td>DÜSTER</td>
<td>MITEN</td>
</tr>
<tr>
<td>HOPELESS</td>
<td>FEATHERS</td>
<td>HOFFNUNGSLOS</td>
<td>GEFIEDER</td>
</tr>
<tr>
<td>AWFUL</td>
<td>TRACT</td>
<td>SCHRECKLICH</td>
<td>STRECKE</td>
</tr>
<tr>
<td>DESERTED</td>
<td>MARCHING</td>
<td>VERLASSEN</td>
<td>MARSCHIEREND</td>
</tr>
<tr>
<td>FUTILE</td>
<td>ATTIRE</td>
<td>VERGEBLICH</td>
<td>BEKLEIDUNG</td>
</tr>
<tr>
<td>MOURN</td>
<td>SCANS</td>
<td>TRAURN</td>
<td>SCAN</td>
</tr>
<tr>
<td>DISCOURAGED</td>
<td>CONNECTIONS</td>
<td>ENTMITUTIG</td>
<td>VERBINDUNGEN</td>
</tr>
<tr>
<td>INADEQUATE</td>
<td>TRANSITION</td>
<td>UNZUREICHEND</td>
<td>ÜBERGANG</td>
</tr>
<tr>
<td>DISCONTENDED</td>
<td>HOUSEHOLDERS</td>
<td>UNZUFRIEDEN</td>
<td>HAUSHERR</td>
</tr>
<tr>
<td>INFORM</td>
<td>DEPOTS</td>
<td>BELEHREN</td>
<td>LAGERHAUS</td>
</tr>
<tr>
<td>MISUNDERSTOOD</td>
<td>MANIFESTATION</td>
<td>VERKANNT</td>
<td>ERSCHEINUNG</td>
</tr>
<tr>
<td>GLOOMY</td>
<td>PASTEL</td>
<td>DÜSTER</td>
<td>PASTELL</td>
</tr>
<tr>
<td>USELESS</td>
<td>FLOWING</td>
<td>NUTZLOS</td>
<td>FLÜSSIG</td>
</tr>
<tr>
<td>FORDLORN</td>
<td>KEYHOLE</td>
<td>VERLOREN</td>
<td>SCHLÜSSELLOCH</td>
</tr>
<tr>
<td>TORMENTED</td>
<td>MYTHOLOGY</td>
<td>GEFOLTERT</td>
<td>MYTHOLOGIE</td>
</tr>
<tr>
<td>DEATHBED</td>
<td>SOFTENER</td>
<td>STERBEBETT</td>
<td>WEICHMACHER</td>
</tr>
<tr>
<td>UNHAPPY</td>
<td>BRIDGES</td>
<td>UNGLÜCKLICH</td>
<td>BRÜCKEN</td>
</tr>
<tr>
<td>DULL</td>
<td>FLEW</td>
<td>STUMPF</td>
<td>FLIEGEN</td>
</tr>
<tr>
<td>DREADFUL</td>
<td>COMPUTER</td>
<td>FURCHTBAR</td>
<td>COMPUTER</td>
</tr>
<tr>
<td>COFFIN</td>
<td>EDITED</td>
<td>SARG</td>
<td>BEARBEITET</td>
</tr>
<tr>
<td>MISTAKE</td>
<td>QUARTER</td>
<td>FEHLER</td>
<td>VIERTEL</td>
</tr>
<tr>
<td>Aggression-themed words</td>
<td>Positive emotion words</td>
<td>Negative emotion words</td>
<td>Neutral words</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>English</td>
<td>German</td>
<td>English</td>
<td>German</td>
</tr>
<tr>
<td>rage</td>
<td>Wut</td>
<td>devotion</td>
<td>Hingabe*</td>
</tr>
<tr>
<td>anger</td>
<td>Zorn</td>
<td>affection</td>
<td>Zuneigung</td>
</tr>
<tr>
<td>tear</td>
<td>zerreißen</td>
<td>admire</td>
<td>bewundern</td>
</tr>
<tr>
<td>assault</td>
<td>Angriff</td>
<td>euphoric</td>
<td>euphorisch</td>
</tr>
<tr>
<td>kick</td>
<td>Trit</td>
<td>fond</td>
<td>zärtlich</td>
</tr>
<tr>
<td>shout</td>
<td>Schrei</td>
<td>grateful</td>
<td>dankbar</td>
</tr>
<tr>
<td>punch</td>
<td>Schlag</td>
<td>tolerant</td>
<td>tolerant</td>
</tr>
<tr>
<td>hate</td>
<td>Hass</td>
<td>affectionate</td>
<td>herzlich</td>
</tr>
<tr>
<td>argue</td>
<td>streiten</td>
<td>amused</td>
<td>amüsier</td>
</tr>
<tr>
<td>temper</td>
<td>Wut</td>
<td>love</td>
<td>Liebe</td>
</tr>
<tr>
<td>fight</td>
<td>Kampf</td>
<td>joy</td>
<td>Freude</td>
</tr>
<tr>
<td>kill</td>
<td>töten</td>
<td>proud</td>
<td>stolz</td>
</tr>
<tr>
<td>punish</td>
<td>bestrafen</td>
<td>fond</td>
<td>liebevoll</td>
</tr>
<tr>
<td>annoyed</td>
<td>verärgert</td>
<td>funny</td>
<td>lustig</td>
</tr>
<tr>
<td>guilt</td>
<td>Schuld</td>
<td>glad</td>
<td>froh</td>
</tr>
<tr>
<td>lust</td>
<td>Verlangen</td>
<td>comfortable</td>
<td>komfortabel</td>
</tr>
<tr>
<td>crush</td>
<td>vernichten</td>
<td>beloved</td>
<td>geliebte</td>
</tr>
<tr>
<td>slash</td>
<td>Hieb</td>
<td>calm</td>
<td>Ruhe</td>
</tr>
<tr>
<td>rape</td>
<td>Vergewaltigung</td>
<td>peace</td>
<td>Frieden</td>
</tr>
<tr>
<td>cut</td>
<td>Schnitt</td>
<td>daring</td>
<td>Kühneheit</td>
</tr>
<tr>
<td>wound</td>
<td>Wunde</td>
<td>cheerful</td>
<td>heiter</td>
</tr>
<tr>
<td>injure</td>
<td>verletzen</td>
<td>warm</td>
<td>warm</td>
</tr>
<tr>
<td>threaten</td>
<td>bedrohen</td>
<td>protective</td>
<td>schützend</td>
</tr>
<tr>
<td>knife</td>
<td>Messer</td>
<td>hope</td>
<td>Hoffnung</td>
</tr>
<tr>
<td>slap</td>
<td>Ohrfeige</td>
<td>lively</td>
<td>lebhaft</td>
</tr>
</tbody>
</table>

*Note: The asterisk (*) next to "Hingabe" indicates a non-standard or informal German term.*
Appendix G: Consent Form – Study 4

English version:

PARTICIPANT CONSENT FORM

Title of Research Project: Ego depletion and police use of force.

Researcher(s): Mario Staller

1. I confirm that I have read and have understood the information sheet dated 7th July 2014 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my rights being affected. In addition, should I not wish to answer any particular question or questions, I am free to decline.

3. I understand that my verbal responses, blood pressure, heart rate, and core temperature will be recorded during my participation in this study.

4. I understand that once I submit my data it will become anonymised and I will therefore no longer be able to withdraw my data.

5. I agree to take part in the above study.

__________________________________________________________________________
Participant Name Date Signature

__________________________________________________________________________
Name of Person taking consent Date Signature

__________________________________________________________________________
Researcher Date Signature

Principal Investigator: Professor Jon Cole
Department of Psychological Sciences
Eleanor Rathbone Building
Phone: 0151 784 2175
Email: joncole@liv.ac.uk

Student Researcher: Mario Staller
School of Psychology
Eleanor Rathbone Building
Phone: +49 151 222 897 59
Email: mstaller@liv.ac.uk
German version:

EINVERSTÄNDNISERKLÄRUNG

Title of Research: Ego depletion and police use of force.
Project:

Researcher(s): Mario Staller


7. Ich verstehe, dass meine Teilnahme auf freiwilliger Basis erfolgt und dass ich jederzeit meine Teilnahme an der Studie beenden kann.

8. Ich verstehe, dass meine Antworten, Verhalten und Herzfrequenz während der Studie aufgenommen werden.

9. Ich verstehe, dass sobald meine Teilnahme an der Studie beendet ist, meine Daten nicht mehr gelöscht werden können, da sie ab diesem Zeitpunkt nur anonymisiert vorliegen.

10. Ich nehme an der Studie teil.

_________________________        ___________________________        ___________________________
Teilnehmer                   Datum                        Unterschrift

_________________________        ___________________________        ___________________________
Wissenschaftler               Datum                        Unterschrift

Principal Investigator:
Professor Jon Cole
Department of Psychological Sciences
Eleanor Rathbone Building
Phone: 0151 784 2175
Email: joncole@liv.ac.uk

Student Researcher:
Mario Staller
School of Psychology
Eleanor Rathbone Building
Phone: +49 151 222 897 59
Email: mstaller@liv.ac.uk
Appendix H: Information Sheet – Study 4

PARTICIPANT INFORMATION SHEET

Title of Research Project: Ego depletion and police use of force.

Researcher(s): Dr. Mario Staller

My name is Mario, I am a postgraduate student studying at the University Of Liverpool (Department of Psychological Sciences), investigating decision-making in high stress environments under the supervision of Professor Cole and Dr Paul Christiansen. You are invited to participate in a research study. Before you decide whether to participate, it is important for you to understand why this research is being conducted and what it will involve. Please take time to read the following information carefully and feel free to ask if you would like more information or if there is anything that you do not understand. Please also feel free to discuss this with your friends and colleagues if you wish. We would like to stress that you do not have to accept this invitation and should only agree to take part if you want to.

Thank you for taking the time to read this.

Purpose of Study
The aim of this study is to investigate the impact of training on decision making in law enforcement personnel. By examining this area it may be possible to develop and improve training packages to create the best decision-making environment possible. To do this, you are invited to take part in a study which consists of two parts; measurement of your cardiovascular function and some short cognitive tasks.

The cognitive testing comprises tasks which aim to measure your memory and self-control. Your scores from the cognitive tasks and cardiovascular function will be combined to examine the relationship between these measures and your performance on the training exercise.

You have been selected to take part in the study as a qualified Police Officer. The only other criteria being that you are over 18 years of age and are of good health both physically and mentally. Participation in the experiment is voluntary and you are free to withdraw at anytime without any explanation or incurring a disadvantage. When taking part in this study you shall be covered by the University of Liverpool insurance policy.

Participation in this study should not last any more than 30 minutes in total. The testing will be conducted in a secure room within the Hesse State Police Training Academy.

It should be noted that there are minimal perceived risks from taking part in the study. The health and safety of the data collection will be in accordance to normal procedures within the Hesse State Police. If you feel any discomfort or disadvantage due to taking part in the study at any point you must let it be known immediately, as there is no intended benefit from taking part in it.
It must again be emphasised that you are free to withdraw from the study at any time and your results can be destroyed if you so wish. However this can only happen whilst you are with the researcher as the data are anonymous once the data collection is finished. If you are unhappy, or you feel that there is a problem, please feel free to let us know by contacting Mario on +49 151 222 897 59 and we will try to help. If you remain unhappy or have a complaint which you feel you cannot come to us with then you should contact the Research Governance Officer on +44 151 794 8290 (ethics@liv.ac.uk). When contacting the Research Governance Officer, please provide details of the name or description of the study (so that it can be identified), the researcher(s) involved and the details of the complaint you wish to make.

The data collected is anonymised and stored securely on the University secure system. The data will be stored for no longer than 10 years with only the primary researcher (Mario) and supervisors (Professor Cole and Dr Christiansen) having access to it. After this, the data shall be disposed of. It is intended that the results of the study will be published in an appropriate journal. You can be made aware of this publication if you so wish by leaving your contact details. You will be totally anonymous in any publication of data resulting from the experiment.

**Principal Investigator:**
Professor Jon Cole  
Department of Psychological Sciences  
Eleanor Rathbone Building  
Phone: +44 151 794 2175  
Email: joncole@liv.ac.uk

**Student Researcher:**
Dr. Mario Staller  
Department of Psychological Sciences  
Eleanor Rathbone Building  
Phone: +49 151 222 897 59  
Email: mstaller@liv.ac.uk
Appendix I: Debrief Sheet – Study 4

PARTICIPANT DEBRIEF SHEET

Title of Research Project: Ego depletion and police use of force.

Researcher(s): Mario Staller

Thank you!

You have now completed the study. The researchers would like to thank you for taking part and hope that it was an enjoyable experience. The information you have provided will be used to examine the impact of self-control on the police use of force. These findings will help us develop a model of the processes underlying decision making in critical incidents. This model will be used to improve training packages to create the best decision-making environment possible.

Your feedback and reflections on the research study are important in assisting future research. If you have any questions about the current study or would like to provide any feedback, then please contact the supervisor (joncole@liv.ac.uk) or student researcher (mstaller@liv.ac.uk). If you have a complaint which you feel you cannot come to us with, then you should contact the Research Governance Officer on +44 151 794 8290 (ethics@liv.ac.uk). When contacting the Research Governance Officer, please provide details of the name or description of the study (so that it can be identified), the researcher(s) involved and the details of the complaint you wish to make.

It is intended that the results of the study will be published in an appropriate journal. You can be made aware of this publication if you so wish by leaving your contact details with the student researcher. You will be totally anonymous in any publication of data resulting from the experiment.

Thank you for taking part in this study.

Principal Investigator: Professor Jon Cole
Department of Psychological Sciences
Eleanor Rathbone Building
Phone: +44 151 794 2175
Email: joncole@liv.ac.uk

Student Researcher: Mario Staller
Department of Psychological Sciences
Eleanor Rathbone Building
Phone: +49 151 222 897 59
Email: mstaller@liv.ac.uk
Appendix J: Brief Mood Introspection Scale

English version:

INSTRUCTIONS: Circle the response on the scale below that indicates how well each adjective or phrase describes your present mood.

1 = definitely do not feel / 2 = do not feel / 3 = slightly feel / 4 = definitely feel

<table>
<thead>
<tr>
<th>Adjective</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Lively</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Drowsy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Happy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Grouchy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Sad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Peppy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Tired</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Nervous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Caring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Calm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Loving</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Gloomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Fed up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Jittery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall, my mood is:

Very Unpleasant

<table>
<thead>
<tr>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
</tr>
<tr>
<td>-9</td>
</tr>
<tr>
<td>-8</td>
</tr>
<tr>
<td>-7</td>
</tr>
<tr>
<td>-6</td>
</tr>
<tr>
<td>-5</td>
</tr>
<tr>
<td>-4</td>
</tr>
<tr>
<td>-3</td>
</tr>
<tr>
<td>-2</td>
</tr>
<tr>
<td>-1</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

Very Pleasant

<table>
<thead>
<tr>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

438
German version:

BMIS (Mayer & Gaschke, 1998)

Es folgt nun eine Anzahl von Wörtern, die verschiedene Gefühle und Empfindungen beschreiben. Setzen Sie bitte spontan an der Stelle ein Kreuz, die Ihre momentane Stimmung am besten beschreibt.

Sie können zwischen vier Abstufungen wählen, nämlich:

1 („trifft gar nicht zu“)
2 („trifft eher nicht zu“)
3 („trifft etwas zu“) bis
4 („trifft völlig zu“)

<table>
<thead>
<tr>
<th></th>
<th>trifft gar nicht zu</th>
<th>trifft eher nicht zu</th>
<th>Trifft etwas zu</th>
<th>Trifft völlig zu</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>lebhaft</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>schlaftrig</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>glücklich</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>grantig</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>traurig</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>schwungvoll</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>müde</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>nervös</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>fürsorglich</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>ruhig</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>zufrieden</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>liebevoll</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>bedrückt</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>deprimiert</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>kribbelig</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>aktiv</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Im Allgemeinen ist meine Stimmung:

Sehr unangenehm  Sehr angenehm

| -10 | -9 | -8 | -7 | 6 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |