Aus der Klinik für Psychosomatik und Psychotherapeutische Medizin
des Zentralinstituts für seelische Gesundheit
der Medizinischen Fakultät Mannheim
der Universität Heidelberg
(Direktor: Prof. Dr. med. Christian Schmahl)

Amygdala Neurofeedback Training in Borderline Personality Disorder: Capturing Improvements in Emotion Regulation

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.) der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Jenny Katharina Zähringer

aus
Freiburg i.Br.

2020

Dekan: Prof. Dr. med. Sergij Goerdt

Referent: Prof. Dr. med. Christian Schmahl

TABLE OF CONTENTS

Page

ABBREVIATIONS

ADHD	Attention deficit hyperactivity disorder
ACC	Anterior cingulate cortex
ALS	Affective Lability Scale
ANOVA	Analysis of Variance
AUC	Area under the curve
BMT	Backward Masking Task
BOLD	Blood-oxygen-level dependent
BPD	Borderline personality disorder
CERQ	Cognitive Emotion Regulation Questionnaire
CI	Confidence interval
dACC	Dorsal ACC
DERS	Difficulties in Emotion Regulation Scale
DBT	Dialectical Behavior Therapy
dlPFC	Dorsolateral prefrontal cortex
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, fifth edition
DSS-4	Dissociation Stress Scale, 4 item version
EEG	Electroencephalography
EMA	Ecological Momentary Assessment
EMG	Electromyography
EPI	Echo-planar-imaging
EWMT	Emotional working memory task
fMRI	Functional magnetic resonance imaging
GLM	General Linear Model
HRV	Heart rate variability
IAPS	International Affective Picture System
IMMO	Implementation and Maintenance Model
IPDE	International Personality Disorder Examination
KDEF	Karoliska Directed Emotional Faces
MeSH	Medical subject headings
MNI	Montreal Neurological Institute
MSSD	Mean squared for successive differences

NA	Negative affect
OFC	Orbitofrontal cortex
PA	Positive affect
PTSD	Posttraumatic stress disorder
RCT	Randomized controlled trial
ROI	Region of interest
SAM	Self-Assessment-Manikin
SEK-27	Selbsteinschaetzung Emotionaler Kompetenzen
SKID	Strukturiertes Klinisches Interview
SMA	Supplementary motor area
SRI	Stochastic Regression Analysis
TAS-26	Toronto Alexithymia Scale
TAU	Treatment as usual
UPPS	Impulsivity Scale (“urgency”, “[lack of] premeditation”, “[lack of] perseverance”, “sensation seeking”)
VDM	Voxel displacement map
vlPFC	Ventrolateral prefrontal cortex
vmPFC	Ventromedial prefrontal cortex
ZAN-BPD	Zanarini Borderline Personality Disorder Interview

1 INTRODUCTION

1.1 Theoretical Background

The way we regulate emotions is a powerful determinant of behavior and directly impacts our affect and physiology (Gross, 2002). The adaptive use of emotion regulation in the long term has been shown to be an essential factor of mental well-being and for establishing healthy relationships (Gross & John, 2003). Therefore, it is not surprising that emotion regulation has been a popular topic in psychological research for the past decades (Gross, 2014; McRae & Gross, 2020). At the same time, emotion regulation is of considerable interest for clinical research, as many mental disorders such as Borderline Personality Disorder (BPD) are in large part disorders of emotion regulation (Aldao, Nolen-Hoeksema, & Schweizer, 2010; Campbell-Sills & Barlow, 2007; Glenn & Klonsky, 2009; Joormann & Gotlib, 2010; Schmahl et al., 2014; Sloan et al., 2017). Because of its important role in mental health, research has endeavored to improve methods to assess emotion regulation in the laboratory, to understand its biological underpinnings and to create trainings and specific clinical programs that aim to augment the ability to regulate emotions in a healthy and adaptive way.

An important question within the field of emotion regulation research has been the adequate assessment of emotion and emotion regulation in the laboratory. Studies have heavily benefitted from psychophysiological measures, as they move beyond self-report and are cost-efficient and relatively easy to implement as compared to functional magnetic resonance imaging (fMRI). However, psychophysiological effects of emotion regulation have been inconsistent across studies and the causes for these inconsistencies remain unknown. The assessment of psychophysiological effects of emotion regulation is not only important for laboratory research wishing to understand the biological mechanisms of emotion regulation but is also critical for studies aiming to assess training-related changes of emotion regulation in clinical populations (De Witte, Sutterlin, Braet, & Mueller, 2017; Svaldi, Tuschen-Caffier, Lackner, Zimmermann, & Naumann, 2012).

Several psychotherapeutic approaches are available for patients with emotion regulation problems (Bateman & Fonagy, 2007; Linehan, 1993) that have been shown to be effective to improve BPD symptoms (Stoffers et al., 2012). Dialectical behavior therapy (DBT; Linehan, 1993), originally developed for patients with borderline personality disorder, has a primary focus to teach patients how to regulate their emotions. DBT has shown medium effect sizes in symptomatology compared to treatment as usual (TAU; Cristea et al., 2017; Stoffers et al.,

2012), but besides the convincing success, there is still a demand to further improve therapy and provide targeted therapeutic options, tailored to the needs of BPD. A new technique to potentially improve emotion regulation is amygdala neurofeedback training (Linhartová et al., 2019). Because amygdala hyperreactivity is assumed to underlie emotion regulation problems in BPD (Schulze, Schmahl, & Niedtfeld, 2016), neurofeedback to downregulate the amygdala might be a potential new training to complement therapy in this clinical population. Amygdala neurofeedback has already been shown to be feasible in BPD (Paret et al., 2016) but there are several questions that are still open for debate. For example, little is known about what aspects of BPD symptomatology and emotion dysregulation improve with normalization of amygdala activation. In addition, the mechanisms of amygdala neurofeedback to change emotion are not fully understood. These questions are critical for the selection of primary outcome measures urgently needed to conduct randomized controlled trials (RCT). Such trials will ultimately answer the question whether amygdala neurofeedback in BPD has an effect above and beyond placebo.

It is because of these reasons that the present thesis is dedicated to identifying suitable measures of emotion regulation changes after amygdala neurofeedback training in BPD and to advancing psychophysiological assessment of emotion regulation in general. As such, the present thesis provides the groundwork for future RCTs in amygdala neurofeedback in BPD. In addition, the present thesis also provides the empirical base for studies wishing to assess emotion regulation effects via psychophysiology in the laboratory.

In the following sections, the concept of emotion regulation will be introduced and its manifestation on behavioral, psychophysiological and neural measures will be reviewed. Following this, previous research on the different manifestations of emotion dysregulation in BPD will be outlined and amygdala neurofeedback training as new technique to tackle emotion regulation problems will be introduced. From significant gaps in each of these fields of literature, the central research questions will be derived and outlined in detail.

1.1.1 Emotion Regulation

Emotion regulation allows us to control “what kind of emotions we have, when we have them and how we experience and express these emotions” (Gross, 1998b; p. 275). Individuals can choose from a variety of different regulation strategies to modify emotions. In an attempt to classify these strategies into broader categories, researchers have developed a number of theoretical models (Gross, 1998a; Gyurak, Gross, & Etkin, 2011; Koole, 2009; Larsen, 2000; Parkinson & Totterdell, 1999).

The arguably most prominent approach is the process model of emotion regulation (Gross, 1998a, 2015). It assumes that the key dimension on which emotion regulation strategies differ is their timing, that is, when they are being applied during the emotion generative process. As such, the process model divides emotion regulation strategies into antecedent-focused emotion regulation strategies, i.e. strategies that are applied before the emotional response has fully unfolded, and response-focused emotion regulation strategies, i.e. emotion regulation when the emotional response is already under way. Antecedent-focused emotion regulation strategies are *situation selection* or *situation modification*, *attentional deployment*, and *cognitive change*, with attentional deployment and cognitive change being the most common antecedent-focused strategies analyzed in emotion regulation studies. Attentional deployment involves directing attention away from certain aspects of a situation and is also rereferred to as distraction. Cognitive change involves reappraising the situation by altering its meaning or to distance oneself from the situation, oftentimes referred to as reappraisal. In contrast to these antecedent-focused strategies, response-focused strategies refer to suppression or exaggeration of physiological, cognitive, and behavioral processes of the emotional response.

A recent extension of the process model, the *extended process model* (Gross, 2015) complements the previous version by defining when emotion regulation starts and stops. In brief, it assumes that emotions are a result from valuating perceptions as "good or bad for me". If the valuation is "bad for me," a goal to modify is activated, which can be achieved with the five emotion regulation strategies mentioned above. Each emotion regulation strategy has different stages, namely identification (whether to regulate the emotion), selection (what strategy to use), and implementation (implementing a specific strategy suited to the present situation). Each of these stages is further divided into sub-stages of perceiving, valuating and acting. If the initially activated goal corresponds to the perceived emotional experience, the cycle stops. Otherwise, it continues in a dynamic fashion by either maintaining or switching a strategy. At each of the stages, impairments can lead to emotion dysregulation (Sheppes, Suri, & Gross, 2015).

Another model categorizes emotion regulation strategies on a spectrum ranging between explicit and implicit emotion regulation (Gyurak et al., 2011). Within this framework, explicit emotion regulation has been defined as processes that are conscious, effortful and require monitoring during implementation. In other words, individuals are aware that they regulate. Implicit emotion regulation on the other hand has been defined as a process that is initiated rather automatically and does not require active monitoring. During implicit emotion regulation, individuals are not necessarily aware that they are regulating emotions. Although explicit strategies are cognitively more demanding than implicit strategies, they are useful for adjusting emotional

reactions in specific contexts, e.g. suppressing an anger reaction to a colleague (Gyurak & Etkin, 2014).

1.1.1.1 Affective Consequences of Emotion Regulation

The process model assumes that antecedent-focused emotion regulation strategies such as reappraisal should be more effective and less effortful to modify emotions than response-focused strategies such as suppression (Gross, 1998a, 2002). This assumption was supported by studies comparing the effects of reappraisal with the effects of suppression. They demonstrated that reappraisal was more successful than suppression in downregulating subjective negative (Ehring, Tuschen-Caffier, Schnulle, Fischer, & Gross, 2010; Gross, 1998a; Hofmann, Heering, Sawyer, & Asnaani, 2009) and positive emotional experiences (Kalokerinos, Greenaway, & Denson, 2015). They also showed that the frequent use of reappraisal was associated with a more adaptive profile of emotion experience and cardiovascular responding (Mauss, Cook, Cheng, & Gross, 2007), appeared to be less cognitively effortful than suppression (Richards & Gross, 2000) and was less associated to self-reported stress than suppression (Moore, Zoellner, & Mollenholt, 2008). Studies comparing reappraisal with distraction indicate that reappraisal could be more successful in downregulating subjective emotional experience emotions than distraction (Schonfelder, Kanske, Heissler, & Wessa, 2014).

More recently however, research has departed from dividing emotion regulation strategies into adaptive and maladaptive categories but rather emphasizes the context in which emotion regulation strategies are applied (McRae, 2016). A growing body of evidence suggests that contextually appropriate and flexible use of emotion regulation may be a marker of mental health. Better performance during an emotion suppression task for example has been related to higher well-being and socio-economic status (Côté, Gyurak, & Levenson, 2010). Moreover, suppression is more frequently used when facing high as opposed to low intensity stimuli (Dixon-Gordon, Aldao, & De Los Reyes, 2015). Similarly, both healthy controls and patients with borderline personality disorder (BPD), choose distraction over reappraisal when dealing with high intensity stimuli but preferred reappraisal when dealing with low intensity stimuli (Sauer et al., 2016), implying that stimulus intensity is an important determinant of adaptive emotion regulation.

1.1.1.2 Neural Correlates of Emotion Regulation

The brain system supporting emotion regulation has been extensively studied in the past. There are five meta-analyses (Buhle et al., 2014; Diekhof, Geier, Falkai, & Gruber, 2011; Frank et

al., 2014; Kohn et al., 2014; Morawetz, Bode, Derntl, & Heekeren, 2017) showing that the fronto-parietal network (i.e. ventrolateral prefrontal cortex (vlPFC)/dorsolateral prefrontal cortex (dlPFC)) extending to the dorsal anterior cingulate (dACC) and inferior and superior parietal cortex, and the amygdala are involved during emotion regulation. More specifically, Kohn et al. (2014) reported co-activation patterns of reappraisal downregulation in the dlPFC, vlPFC, dACC, supplementary motor area (SMA) and parietal cortex. Buhle et al. (2014) assessed the neural correlates of both reappraisal up- and downregulation and identified a similar regulatory network (i.e. dlPFC, vlPFC, dACC, SMA and parietal cortex) and a modulatory influence of these regions on the amygdala. Frank et al. (2014) assessed emotion up- and downregulation and found the bilateral amygdala and parahippocampal gyrus to be decreased during downregulation and increased during upregulation, and a similar regulation network including the dlPFC, dACC and premotor cortex during both up- and downregulation. Finally, Morawetz et al. (2017) analyzed different emotion regulation strategies and found that the left vlPFC, the anterior insula and the SMA were activated independent of the regulation strategy. In addition, vlPFC and posterior cingulate cortex were the main regions consistently found to be recruited during the upregulation as well as the downregulation of emotion.

Together with the findings in recent meta-analyses, previous research points to a specific amygdala–frontal circuit of emotion generation and regulation (Banks, Eddy, Angstadt, Nathan, & Phan, 2007; Etkin, Büchel, & Gross, 2015; Ochsner & Gross, 2005). That is, during cognitive emotion regulation, the fronto-parietal network including the anterior cingulate cortex are assumed to downregulate amygdala activity (Banks et al., 2007; Urry et al., 2006). Although interactions within the fronto-limbic network are rather complex, research widely agrees on the fact that the amygdala plays a crucial role for processing, expressing, and experiencing emotions (Buhle et al., 2014; LeDoux, 2007; Phelps & LeDoux, 2005). Hence, therapeutic interventions that involve the amygdala might have the potential to improve emotion regulation.

1.1.1.3 Psychophysiological Consequences of Emotion Regulation

Emotions initiate a set of psychophysiological reactions including autonomic responses, facial behavior and somatic reflexes (Bradley, Codispoti, Cuthbert, & Lang, 2001; Kreibig, 2010; Siegel et al., 2018). Autonomic responses of emotions include cardiovascular, electrodermal and respiratory responses (Kreibig, 2010; Siegel et al., 2018; Stemmler, 2004). The autonomic nervous system is divided into the excitatory sympathetic nervous system, which can elicit defensive behavior and the inhibitory parasympathetic nervous system. The sympathetic and parasympathetic system run antagonistically, which then causes changes in physiological arousal

(Ulrich-Lai & Herman, 2009). Each autonomic measure differs with respect to its sympathetic and parasympathetic influence. For example, pre-ejection period and skin conductance are commonly regarded as an index of sympathetic activity, whereas heart rate and blood pressure reflect a blend of sympathetic and parasympathetic activity (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Schächinger, Weinbacher, Kiss, Ritz, & Langewitz, 2001). In contrast, high-frequency components of heart rate variability primarily reflect an index of cardiac vagal tone, which can be interpreted as the influence of parasympathetic activity to cardiac regulation (Malik et al., 1996).

Electromyography is a technique to measure electric activity of sceletar muscles. In emotion research, electromyography has been primarily used to measure facial muscle activity, for example the *orbicularis oculi* muscle to assess the startle response (Lang, Bradley, & Cuthbert, 1990), and the zygomaticus major and the corrugator supericilii muscle groups which are associated with smiling and frowning. Previous research has demonstrated that activity of the corrugator and zygomatic muscle vary inversely with the emotional valence of emotional stimuli (Bradley, Codispoti, Cuthbert, et al., 2001). Similarly, a substantial body of work (Davis, Campeau, Kim, & Falls, 1995; Lang et al., 1990) demonstrated that startle amplitudes are increased and decreased in negative and positive emotional states, respectively (i.e. *emotion-modulated* startle), with responses being more potentiated and inhibited during viewing of highly arousing stimuli (Bradley, Codispoti, Cuthbert, et al., 2001; Vrana, Spence, & Lang, 1988).

Much evidence has accumulated suggesting that suppression is related to an increase in sympathetic nervous system activity but does not significantly change self-report to negative stimuli (Gross & Levenson, 1993, 1997; Richards & Gross, 1999). The enhanced sympathetic activation following suppression has led researchers to conclude that suppression "exacts a palpable physiological cost" (Gross & Levenson, 1997, p. 101). In other words, because response-focused strategies involve an active modulation of expressive behavior, increased sympathetic activation might be the result of that effort (Butler et al., 2003). In contrast, past literature has proposed that reappraisal has little impact on sympathetic and cardiovascular measures (Gross, 1998a). A meta-analysis studying the overall physiological effect of different emotion regulation strategies confirmed this general pattern: cognitive change had a smaller effect on physiology than response modulation (Webb, Miles, & Sheeran, 2012).

Both emotion-modulated startle and facial electromyography can also be used to measure emotion regulation. The emotion-modulated startle response is inhibited when down-regulating and potentiated when up-regulating negative emotions (Adolph & Pause, 2012; Bernat, Cadwallader, Seo, Vizueta, & Patrick, 2011; Conzelmann, McGregor, & Pauli, 2015; Dillon &

LaBar, 2005; Driscoll, Tranel, & Anderson, 2009; Grillon, Quispe-Escudero, Mathur, & Ernst, 2015; Jackson, Malmstadt, Larson, & Davidson, 2000; Lee, Shackman, Jackson, & Davidson, 2009; Lissek et al., 2007; Piper & Curtin, 2006). Similarly, studies have shown that corrugator activity is reduced when downregulating negative emotions and increased when upregulating negative emotions (Jackson et al., 2000).

Several questions, however, remain unanswered. First, when looking at individual psychophysiological measures, findings are inconsistent with respect to the effects of emotion regulation on autonomic physiology. Studies show that reappraisal instructions for example have no effect on (Goldin, Moodie, & Gross, 2019; Gross, 1998a; Kalisch et al., 2005), increase (Lohani & Isaacowitz, 2014; Sheppes, Catran, & Meiran, 2009), or decrease skin conductance (Urry, van Reekum, Johnstone, & Davidson, 2009; Wolgast, Lundh, & Viborg, 2011). These inconsistencies may be due to the large heterogeneity between studies, which can substantially affect the magnitude of the physiological responses. The contradictory pattern of results across the literature does not allow a straightforward interpretation and the causes for these inconsistencies are not well understood.

Second, and with respect to the emotion-modulated startle, effects of emotion regulation appear to be more consistent, yet it remains unclear whether the startle response may change as a function of probe timing. In previous studies, small decreases of the startle amplitudes were observed when probes were delivered 3 seconds into the regulation phase, but large decreases were observed when probes were delivered 8-11 seconds into the regulation phase (Dillon & LaBar, 2005; Jackson et al., 2000). However, these studies did not directly test whether startle inhibition during emotion down-regulation was significantly different at early versus late probes. Another study delivered the startle probe 2 seconds into the reappraisal phase and reported non-significant startle inhibition (Eippert et al., 2007). Most pronounced amygdala down-regulation in this study was observed after probe presentation, suggesting that the probe might have been given too early to reliably detect reappraisal effects (Eippert et al., 2007). According to the implementation- maintenance model (IMMO; Kalisch, 2009; Paret et al., 2011) reappraisal is divided into two phases: In the early phase, participants choose and implement a strategy, whereas in the late phase they maintain the strategy in working memory and monitor its success.

Together with results from previous studies, this suggests that startle modulation may become more pronounced as soon as the maintenance of reappraisal predominates and therefore reappraisal might need several seconds until it effectively reduces negative emotions.

1.1.1.4 Assessment of Explicit and Implicit Emotion Regulation

As reviewed above, emotion regulation can be conceptualized as a multicomponent process including affective, behavioral and physiological consequences and these components may be assessed in experimental studies on various system levels.

First of all, questionnaires provide an easy way to assess emotion experience and self-reported behavior, and can be used cost-effectively to measure both trait and state features of emotion regulation. With trait questionnaire such as the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004), or the Emotion Regulation Skills Questionnaire (SEK-27; Berking & Znoj, 2008)) one can assess the subjective use of adaptive emotion regulation.

In addition, state questionnaires such as the Self-Assessment-Manikin (SAM; Bradley & Lang, 1994) can be used in experimental paradigms of emotion regulation, where participants have to rate their current affective state after each emotional stimulus presented.

The laboratory experimental approach may complement questionnaire data, as it offers systematic and reliable information about the relations between specific emotional stimuli and their experiential, physiological and behavioral responses under well-controlled conditions. In a typical laboratory experiment to assess explicit emotion regulation, participants view emotional stimuli such as films or pictures and are either asked to apply a specific emotion regulation strategy (e.g., suppression) or to simply view the stimulus without trying to regulate the emotions that arise in response to the stimulus. By including a "no regulation" condition as a control condition, it is possible to compare each regulation strategy to a condition where participants are told not to regulate their emotions. Successful emotion regulation in these paradigms is usually defined as significant short-term changes in affective (e.g., SAM rating), behavioral (e.g., facial expressions) or other physiological outcomes (e.g., neural or psychophysiological correlates) by contrasting the regulation with the control condition. The larger the difference in these domains, the more successful the regulation. The laboratory paradigm qualifies as an explicit emotion regulation task because participants are explicitly told to use a certain emotion regulation strategy and employ them in a conscious and deliberate way.

A typical test to assess implicit emotion regulation (for definition see section 1.1.1) is the emotional working memory task (EWMT; Krause-Utz et al., 2012; Oei et al., 2012). In this task, participants are shown a set of letters that they have to keep in mind while getting distracted by a picture. A second set of letters is presented afterwards, and the participants have to indicate whether one of the letters in the second set match with the letters shown in the first set. Implicit emotion regulation is thus quantified by contrasting behavioral measures (i.e. response times and accuracy of responses) between trials with negative distractors and a control condition

(usually a neutral distractor). In addition, neural activation during the negative distractors might be contrasted to the control condition in order to measure implicit emotion regulation. The task qualifies as an implicit emotion regulation paradigm, because the distracting negative stimulus requires behavioral adjustments (e.g., emotion regulation) in order to achieve the goal of responding fast and accurately, but participants are not explicitly told to engage in emotion regulation.

Both the laboratory experiment and the EWMT described above exhibit a high degree of standardization, but usually these laboratory experiments are experienced as artificial, which results in low external validity. New approaches such as ecological momentary assessment (EMA; Santangelo, Bohus, & Ebner-Priemer, 2014) try to overcome this problem. With EMA, mean levels of positive and negative affect as well as fluctuations in affect over a certain period of time for example can be measured in daily life (Ebner-Priemer et al., 2007). With that, the EMA approach is less prone to subjective distortion.

The paradigms outlined above represent a selection of the many ways to quantify changes in emotions via emotion regulation. It should be highlighted that each of the paradigms have very different "readout" measures (e.g., neural change, psychophysiological change, self-report, behavioral measures) and consider the components of emotion change to different degrees. As such it becomes clear that there is no "gold-standard" outcome measure of emotion regulation. Rather, the literature suggests a multi-component approach to measuring emotional responses (Mauss & Robinson, 2009). The variety of paradigms however makes it difficult to determine a primary outcome measure of emotion regulation success in clinical research trials.

1.1.2 Borderline Personality Disorder

BPD is a serious mental disorder characterized by an instability in affect, identity, interpersonal relationships, and behavioral dysregulation (American Psychiatric Association, 2013). In addition, BPD patients often show chronic feelings of emptiness, stress-related paranoid ideation or dissociative symptoms, intense states of anger or problems controlling anger, self-injurious behavior and suicidal tendencies, and inadequate efforts to avoid abandonment (Lieb, Zanarini, Schmahl, Linehan, & Bohus, 2004; Oldham, 2006). It is believed that severe emotion dysregulation lies at the core of BPD (Glenn & Klonsky, 2009; Sanislow et al., 2002; Schmahl et al., 2014).

With 10 up to 25% in clinical samples BPD is the most common personality disorder in clinical settings (Gunderson, 2009; Leichsenring, Leibing, Kruse, New, & Leweke, 2011; Torgersen, Kringlen, & Cramer, 2001). In the general population, its prevalence ranges between 0.5 and

5.9 % (Grant et al., 2008; Lenzenweger, Lane, Loranger, & Kessler, 2007; Torgersen et al., 2001; Trull, Jahng, Tomko, Wood, & Sher, 2010). Functional impairment in BPD patients is considerable compared to other personality disorders (Ansell, Sanislow, McGlashan, & Grilo, 2007). For example, completed suicide ranges between 3% and 10% of all BPD patients (Black, Blum, Pfohl, & Hale, 2004; Bohus & Schmahl, 2007; Temes, Frankenburg, Fitzmaurice, & Zanarini, 2019). Suicide attempts range between 60-80% (Black et al., 2004; Goodman et al., 2017; Lieb et al., 2004; Oldham, 2006). These numbers emphasize the severe impairments in psychological well-being and social functioning these patients are suffering from. Given the severity of symptoms and that patients with BPD are such a large subset of psychiatric patients, it is not surprising that they consume considerably more mental health resources than most other psychiatric groups (Soeteman, Hakkaart-van Roijen, Verheul, & Busschbach, 2008; Zanarini, Frankenburg, Khera, & Bleichmar, 2001).

1.1.2.1 Emotion Dysregulation and its Manifestation in BPD

Emotion dysregulation has been conceptualized as maladaptive alterations in emotion generation and regulation (Sheppes et al., 2015). Theoretical approaches to emotion dysregulation differ with respect to the number of emotion dysregulation components and how they are termed (Carpenter & Trull, 2013; D'Agostino, Covanti, Rossi Monti, & Starcevic, 2017). As such, the construct “emotion dysregulation” goes beyond impairments in the implementation of explicit and implicit emotion regulation as reviewed above.

According to the biosocial theory (Linehan, 1993), emotion dysregulation in BPD is developed through an interplay between biological vulnerabilities (e.g. genetic or intrauterine factors) and negative experiences in childhood or adolescence (e.g., interpersonal violence, emotional neglect, invalidation). Because BPD patients are more sensitive to emotional stimuli from birth, they experience heightened negative affect and affective instability when encountering an emotional stimulus, which then leads to an increase in dysfunctional emotion regulation strategies. This in turn may reinforce attention towards negatively valenced stimuli hence resulting in a vicious cycle of emotion dysregulation. Following this, the multi-component model of emotion dysregulation in BPD differs several components of emotion dysregulation, that is, heightened sensitivity towards negative stimuli, impairments in the implementation and maintenance of adaptive and appropriate emotion regulation strategies, and heightened and labile negative affect (for a review see Carpenter & Trull, 2013).

The components of emotion dysregulation in BPD have been tested in a variety of experimental studies (Carpenter & Trull, 2013). These studies approached emotion dysregulation on several system levels such as self-report, neural activation and psychophysiology.

1.1.2.1.1 Self-report

Self-report studies demonstrate that BPD patients report greater lability in anger and anxiety compared to participants with other personality disorders (Koenigsberg et al., 2002). Problems with emotion regulation assessed with the DERS (Gratz & Roemer, 2004) was associated with BPD symptoms when accounting for negative affect, suggesting that emotion dysregulation is a crucial contributor to BPD (Salsman & Linehan, 2012). Moreover, a recent meta-analysis (Daros & Williams, 2019) revealed that symptoms of BPD were associated with less frequent use of emotion regulation strategies that would be considered more effective at reducing negative affect (i.e., cognitive reappraisal, problem solving, and acceptance) and more frequent use of emotion regulation strategies considered less effective at reducing negative affect (i.e., suppression, rumination, and avoidance).

Studies utilizing the ecological momentary assessment approach (Ebner-Priemer, Eid, Kleindienst, Stabenow, & Trull, 2009; Stone & Shiffman, 1994; Trull & Ebner-Priemer, 2009) revealed that BPD patients report negative affect more frequently and oscillate rapidly between emotions which results in high affective instability (Ebner-Priemer et al., 2007; Nica & Links, 2009; Santangelo et al., 2014; Trull et al., 2008).

1.1.2.1.2 Neural Activation

Neuroimaging has become one of the most important methods to detect biological markers in patients with BPD that differentiate them from the healthy population. Studies have shown a reduced amygdala and hippocampal volume in BPD (Niedtfeld et al., 2013; Nunes et al., 2009; Ruocco, Amirthavasagam, & Zakzanis, 2012; Schmahl, Vermetten, Elzinga, & Douglas Bremner, 2003). The most prominent and consistent finding in BPD patients compared to healthy controls is a heightened activation of limbic regions such as the amygdala (Herpertz et al., 2001; Koenigsberg, Siever, et al., 2009; Minzenberg, Fan, New, Tang, & Siever, 2007; Schulze et al., 2011; Schulze et al., 2016; Schulze, Schulze, Renneberg, Schmahl, & Niedtfeld, 2019) and insula (Krause-Utz et al., 2012; Niedtfeld et al., 2010; Ruocco, Amirthavasagam, Choi-Kain, & McMain, 2013; Schulze et al., 2011; van Zutphen et al., 2018) during the processing of negative stimuli. BPD subjects moreover show sustained blood-oxygen-level dependent (BOLD) responses of the amygdala in reaction to emotional stimuli (Hazlett et al., 2012). Sustained amygdala BOLD response could indicate that the amygdala needs more time

to return to baseline and a failure to downregulate the amygdala in response to repeated presentations of emotional pictures (Hazlett et al., 2012). Hyper-arousal of the amygdala is clinically important, given its role in appraising the affective salience of stimuli (Pessoa & Adolphs, 2010), especially the appraisal of perceived threat and mediation of fear responses (LeDoux, 2007). In addition to increased activation of limbic regions, BPD patients also show a hypoactivation of frontal regions (Krause-Utz, Winter, Niedtfeld, & Schmahl, 2014) such as the anterior cingulate cortex (ACC; Minzenberg et al., 2007) and the dlPFC (Schulze et al., 2016) during the processing of negative stimuli. Overall, alterations in limbic and prefrontal regions might underlie the emotional disturbances BPD patients have (Schulze et al., 2016).

To investigate the neural correlates of explicit emotion regulation in BPD, fMRI studies utilized emotion regulation paradigms established in general emotion regulation research, which have been described in section 1.1.1.4. Studies found that during regulation of negative stimuli, BPD patients showed less activity in regions associated with cognitive emotion regulation such as the dlPFC, the orbitofrontal cortex (OFC) and the ACC than healthy controls (Koenigsberg, Fan, et al., 2009; Lang et al., 2012; Schulze et al., 2011; van Zutphen et al., 2018). At the same time BPD patients showed increased activation of limbic regions such as the insula (Schulze et al., 2011) and the amygdala (Koenigsberg, Fan, et al., 2009) during emotion regulation compared to healthy controls. To sum up at this point, the findings of the neural correlates of emotion processing and regulation underpin the assumption of dysfunctions within the fronto-limbic network in BPD, involving ACC, OFC, dlPFC, and - most importantly - the amygdala.

1.1.2.1.3 Psychophysiological Alterations

Studies assessing the biological underpinnings of emotion dysregulation in BPD have used psychophysiological measures such as the emotion-modulated startle reflex (Barnow et al., 2012; Baskin-Sommers, Curtin, et al., 2012; Ebner-Priemer et al., 2005; Hazlett et al., 2007; Herpertz & Koetting, 2005; Herpertz, Kunert, Schwenger, & Sass, 1999; Limberg, Barnow, Freyberger, & Hamm, 2011; Thompson, Allen, Chong, & Chanen, 2018; Vitale & Newman, 2012), skin conductance response (Kuo, Fitzpatrick, Metcalfe, & McMain, 2015; Kuo & Linehan, 2009; Schmahl et al., 2004) and heart rate variability (Weinberg, Klonsky, & Hajcak, 2009) to study baseline psychophysiological arousal and reactivity to emotion stimuli compared to healthy controls or individuals suffering from other mental disorders.

With regard to the emotion-modulated startle, some studies have found baseline heightened startle response in BPD versus healthy controls (Ebner-Priemer et al., 2005) and increased startle response to aversive disorder-specific scripts (Limberg et al., 2011), negative words (Hazlett et al., 2007), and conditioned threat stimuli (Baskin-Sommers, Vitale, Maccoon, & Newman,

2012) in comparison to healthy controls. Other studies however did not support the notion that BPD patients show heightened startle response (Barnow et al., 2012; Herpertz & Koetting, 2005; Herpertz et al., 1999; Vitale & Newman, 2012). Rather, these studies demonstrate that BPD patients exhibit a comparable potentiation of the startle amplitude to negative stimuli (Herpertz et al., 1999), as it is reported in many studies of healthy individuals (Jackson et al., 2000). So far, only one study assessed the emotion-modulated startle reflex during emotion regulation in patients with BPD, but did not find diminished ability to regulate with respect to the modulation of the startle reflex compared to healthy controls (Thompson et al., 2018). Similarly, effects of autonomic responses to emotional stimuli are mixed as well as in studies with healthy population. Some studies report significant differences in heart rate and skin conductance (Kuo & Linehan, 2009; Limberg et al., 2011), whereas others do not (Herpertz et al., 1999; Schmahl et al., 2004). Rosenthal et al. (2008) conclude that many studies provide at least some evidence that BPD patients exhibit greater emotional responding to emotionally evocative stimuli across certain psychophysiological indices, although findings are mixed. They discuss dissociation, which may dampen the psychophysiological response (Barnow et al., 2012; Ebner-Priemer et al., 2005), and the heterogeneity of BPD symptoms as potential contributors to the mixed findings (Rosenthal et al., 2008).

In sum and as outlined in the paragraphs above, there are many tools to assess emotion dysregulation on different system levels in BPD. D'Agostino et al. (2017) however highlight that the instruments available are diverse and capture only a fraction of it, rather than measuring the whole construct of emotion dysregulation.

1.1.2.2 Treatment of BPD

Several methods of psychotherapy are available for patients with borderline personality disorder. Dialectical behavior therapy (DBT; Linehan, 1993) and Mentalization-based Therapy (Bateman & Fonagy, 2007) have both been shown effective for the treatment of BPD (Cristea et al., 2017; Stoffers et al., 2012). To date, DBT remains the most frequently studied treatment for BPD (Cristea et al., 2017) and its effects to reduce BPD symptoms have been shown to be stable over the course of 30 months after in-patient treatment (Fassbinder et al., 2007). When treated adequately, the remission rate after 10 years is at 86%, however, remission in BPD is considerably slower than in other mental disorders like major depression or other personality disorders (Gunderson et al., 2011). Despite the convincing success of DBT, there is still a demand to further improve therapy and provide targeted therapeutic options, tailored to the needs of this patient population. Bohus et al. (2016) highlights a low coverage of treatment for BPD

patients in Germany, which has been attributed to a lack of therapists and in-patient treatment that are specialized on BPD. In addition, some BPD patients profit more from psychotherapy than others. Because BPD comprises a set of diverse symptoms, some patients for example may suffer more from high impulsivity, whereas others may suffer more from emotion dysregulation. As such, individualized treatments may help to address specific problems that patients undergo.

1.1.3 fMRI Neurofeedback

Neurofeedback has a long tradition within the field of electroencephalography (EEG), especially in its application to attention deficit hyperactivity disorder (ADHD) treatment (Enriquez-Geppert, Smit, Pimenta, & Arns, 2019; Gevensleben, Rothenberger, Moll, & Heinrich, 2012). In the recent decade however, fMRI based neurofeedback has become increasingly popular as well (Sulzer et al., 2013; Thibault, Lifshitz, & Raz, 2018). fMRI-based neurofeedback is a form of biofeedback in which real-time online fMRI signals can be used for self-regulating brain function (Cox, Jesmanowicz, & Hyde, 1995; Weiskopf et al., 2003). In particular, information about the activity of a specific brain region (Weiskopf, 2012) or connectivity between brain regions (Koush et al., 2013) is fed back to the subject in real-time. It is assumed that one can train change and regulate one's own neural activity with neurofeedback (Scharnowski & Weiskopf, 2015).

fMRI detects the concentration of oxygenated and deoxygenated hemoglobin in the neural vasculature, which indicates the metabolic demands of the underlying brain activity (Attwell & Iadecola, 2002). fMRI neurofeedback has a higher spatial resolution than EEG neurofeedback and allows to noninvasively tackle subcortical structures such as the amygdala and other limbic regions. On the other hand, fMRI has a lower temporal resolution than EEG techniques. There is a hemodynamic delay of about 4-6 seconds that participants need to consider when engaging in fMRI neurofeedback (Thibault, Lifshitz, & Raz, 2016).

The progress of neuroimaging has provided valuable data on neuronal networks and their relation to mental disorders, which has important implications for therapeutic approaches of these disorders (Linden, 2006; Linden et al., 2012). With neurofeedback we can make use of this information in order to target disturbed neural mechanisms in patient populations. There is accumulating evidence that aberrant neural activity which directly underlies dysfunctional behaviors and mental states may be changed with neurofeedback and therefore neurofeedback might be a promising therapeutic intervention (Lubianiker et al., 2019; Sitaram et al., 2017). Healthy

adults for example showed the ability to self-regulate brain activity in neural structures associated with affect such as the insula, amygdala and ACC (Caria et al., 2007; Hamilton, Glover, Hsu, Johnson, & Gotlib, 2011; Hellrung et al., 2018; Paret et al., 2014; Zotev et al., 2011). In addition, it has been shown that fMRI-based neurofeedback is feasible in various psychiatric conditions such as depression (Linden et al., 2012; Young, Misaki, et al., 2017; Young, Siegle, et al., 2017; Young et al., 2014), borderline personality disorder (Paret et al., 2016), posttraumatic stress disorder (Nicholson et al., 2017), obesity (Kohl et al., 2019; Spetter et al., 2017), tinnitus (Haller, Birbaumer, & Veit, 2010), schizophrenia (Ruiz et al., 2013), and Parkinson's disease (Subramanian et al., 2011). The majority of fMRI-based neurofeedback studies show that neural changes become visible after 30 minutes of neurofeedback training and some studies even show that these changes can sustain up to 14 months (cf. Thibault et al., 2016).

1.1.3.1 fMRI Neurofeedback to Change Emotion

One of the most broadly studied transdiagnostic phenomena with fMRI neurofeedback is emotion regulation (Linhartová et al., 2019). From a theoretical point of view it has been assumed that emotion regulation and brain self-regulation via neurofeedback underlies similar psychological processes (Paret & Hendler, 2020). At the same time, brain regions activated during processes of neurofeedback such as the ACC and the vlPFC (Emmert et al., 2016) largely overlap with the emotion regulation network (Paret & Hendler, 2020). As such it is probable that amygdala neurofeedback training can impact emotion regulation.

In neurofeedback studies aiming to improve emotion regulation, the amygdala represents the most common target in these studies (Linhartová et al., 2019). Neurofeedback studies have both instructed participants to upregulate (Misaki et al., 2019; Young, Misaki, et al., 2017; Young, Siegle, et al., 2017; Young et al., 2014; Zotev et al., 2018) or to downregulate their amygdala response (Brühl et al., 2014; Herwig et al., 2019; Nicholson et al., 2017; Paret et al., 2014; Paret et al., 2016; Paret et al., 2018), indicating that neuromodulation of amygdala activation in both directions is possible with neurofeedback.

Besides neuromodulation, the feasibility to change emotion with fMRI neurofeedback has also been demonstrated by a number of studies showing that symptoms of emotion dysregulation in clinical populations were improved after fMRI neurofeedback treatment (Gerin et al., 2016; Hamilton et al., 2016; Linden et al., 2012; Paret et al., 2016; Young, Siegle, et al., 2017; Young et al., 2014). Four controlled trials demonstrated greater improvement in measures of negative emotion with treatment, compared to a control group (Hamilton et al., 2016; Linden et al., 2012; Young, Siegle, et al., 2017; Young et al., 2014). Two studies which employed control groups

in a double-blind design showed that amygdala neurofeedback, compared to placebo-neurofeedback, allows patients suffering from depression to regulate their amygdala and to improve their mood (Young, Siegle, et al., 2017; Young et al., 2014). However, both studies included small sample sizes, thereby limiting the interpretation of the effects.

Symptom reduction after amygdala neurofeedback training has also been reported in patients with posttraumatic stress disorder (PTSD; Nicholson et al., 2017) and BPD (Paret et al., 2016). Because these studies did not assess a control group, specific effects of neurofeedback above and beyond placebo are yet to be shown in future RCTs (Linhartová et al., 2019).

In addition, the mechanisms of amygdala neurofeedback to change emotion and emotion regulation are not fully understood. Theoretical models of neurofeedback learning for example (Birbaumer, Ruiz, & Sitaram, 2013; Gevensleben, Moll, Rothenberger, & Heinrich, 2014) differentiate between the *conditioning-and-repairing model*, which assumes that neurofeedback repairs a neural dysfunctions by implicit operant conditioning processes resulting in a decrease in related symptoms, and the *skill-acquisition model*, which assumes that during neurofeedback, effortful, explicit learning allow the acquisition of specific skills through cognitive-behavioral strategies. According to the skill acquisition model, improved brain self-regulation and reduction of the severity of symptoms do not necessarily correlate. There is a lack of experimental studies directly testing these learning models during amygdala neurofeedback (Paret & Hendler, 2020). As such, it remains unknown what aspects of emotion regulation might be improved with amygdala neurofeedback.

1.1.3.2 Emotion Dysregulation, BPD and Amygdala Neurofeedback

As outlined above, it is assumed that aberrant amygdala activity underlies emotion dysregulation in BPD (Schulze et al., 2016) and thus neurofeedback to downregulate the amygdala may be a candidate training to improve the neural dysfunctions underlying emotion regulation problems in BPD. Studies demonstrate associations between BPD diagnosis and amygdala hyperactivation in response to emotion and emotion regulation (Schulze et al., 2019; Zilverstand, Parvaz, & Goldstein, 2017), show amygdala decrease after self-injury which is characteristic for BPD (Reitz et al., 2015), and report amygdala normalization after response to psychotherapy (Goodman et al., 2014). However, whether the training can cause actual improvements in emotion dysregulation remains unknown. Very little knowledge exists on how amygdala hyperactivation maps on behavioral correlates of emotion dysregulation and how it causes symptoms in BPD. In addition, we do not know how improvements in emotion regulation after amygdala neurofeedback training might be expressed in BPD. It could for example result in new or

strengthened skills to regulate emotions or in normalization of amygdala hyperreactivity in transfer tasks without feedback.
This lack of knowledge impedes the informed selection of primary outcome measures for RCTs aiming to test the effectivity of amygdala neurofeedback in BPD.

1.2 Research Questions

The present work provides the groundwork for future RCTs of amygdala neurofeedback as a potential training of emotion dysregulation and extends knowledge about the assessment of emotion regulation in the laboratory using autonomic and electromyographic responses, in particular the emotion-modulated startle.
Given the severity of BPD and its burden to the mental health care system, targeted therapeutic options tailored to the needs of this patient population are needed to complement therapy of BPD. Amygdala neurofeedback training is a candidate training to tackle emotion dysregulation in BPD. As outlined above, little knowledge exists on how amygdala hyperactivation maps on behavioral correlates of emotion dysregulation and how it causes symptoms in BPD. Empirical evidence for neurofeedback learning models is lacking and as such it is difficult to predict how amygdala neurofeedback maps on neural and behavioral measures. In addition, emotion (dys)regulation in both healthy subjects and BPD patients can be assessed at multiple levels of analysis, i.e. self-report, behavior, and biological correlates such as neural or psychophysiological activity and researcher have a range of tools at hand to measure them. Each of these tools capture only a fraction of the construct of emotion dysregulation. Therefore, several questions need to be answered first, in order to be able to select a primary outcome measures for amygdala neurofeedback RCTs.
Psychophysiological measures are cost-effective and move beyond self-report yet bear difficulties in assessment, as previous research revealed highly inconsistent results and important moderators are not identified yet. As such, it is very hard to predict emotion regulation effects on psychophysiology. In contrast to autonomic measures, the emotion-modulated startle seems to be a promising candidate to capture effects of emotion regulation. As reviewed above, previous studies demonstrated greater startle potentiation for probes delivered later compared to those delivered earlier during a 6-12s period of emotional picture viewing (Bradley, Codispoti, Cuthbert, et al., 2001; Sutton, Davidson, Donzella, Irwin, & Dottl, 1997). Studies assessing earlier probes did not reveal significant effects (Dillon & LaBar, 2005; Eippert et al., 2007). As such, it remains unclear whether the startle response may change as a function of probe timing.
In light of the above, the first two studies were dedicated to identifying the effects of emotion

regulation on psychophysiological measures, to identify whether important aspects of the study design moderate these effects, and whether the timing of the startle probe significantly influences the startle during emotion regulation.

Study I quantitively summarized the past literature on the effects of emotion regulation on common psychophysiological measures in healthy individuals using a meta-analytic approach. In total, 1353 studies were screened and $k = 78$ studies were identified as relevant. These studies each contributed multiple sub-samples and 23 meta-analyses for combinations of emotion regulation strategy and psychophysiological measure were conducted. In addition, important moderating variables potentially causing some of the inconsistencies such as trial duration, nature of control instruction and emotion induction were tested and discussed as well. As such, study I provides an important contribution to the field as it identifies psychophysiological measures that are suited to quantify emotion regulation success in the laboratory as well as important moderating factors.

Study II systematically tested effects of the startle probe timing on startle responses during emotion regulation in 47 healthy individuals in order to optimize the way we can track emotion regulation with the emotion-modulated startle and to avoid misinterpretation of non-significant results. In particular, it was hypothesized that later startle probes would achieve larger startle inhibition during reappraisal than earlier probes.

To provide a profound methodological foundation of the assessment of emotion regulation via psychophysiology, the first two studies of the present thesis addressed a healthy population. The results of these studies were then used to address the last and most important aim of the present thesis, that is, identifying ways to test training-related changes of emotion dysregulation after amygdala neurofeedback training in BPD (**Study III**). A four-session amygdala neurofeedback training was tested in 24 female BPD patients. Before and after the neurofeedback training, as well as at a 6-week follow-up assessment, measures of emotion dysregulation and BPD psychopathology were tested at diverse levels of analysis, including the emotion-modulated startle piloted in Study II. It was hypothesized that with amygdala neurofeedback training, BPD patients would improve in general BPD psychopathology, decrease their amygdala response to negative pictures, improve emotion regulation abilities measured with self-report and the startle response, and increased resting heart-rate variability.

For an overview of the research structure and the relationship between the experimental studies, see Figure 1.1.

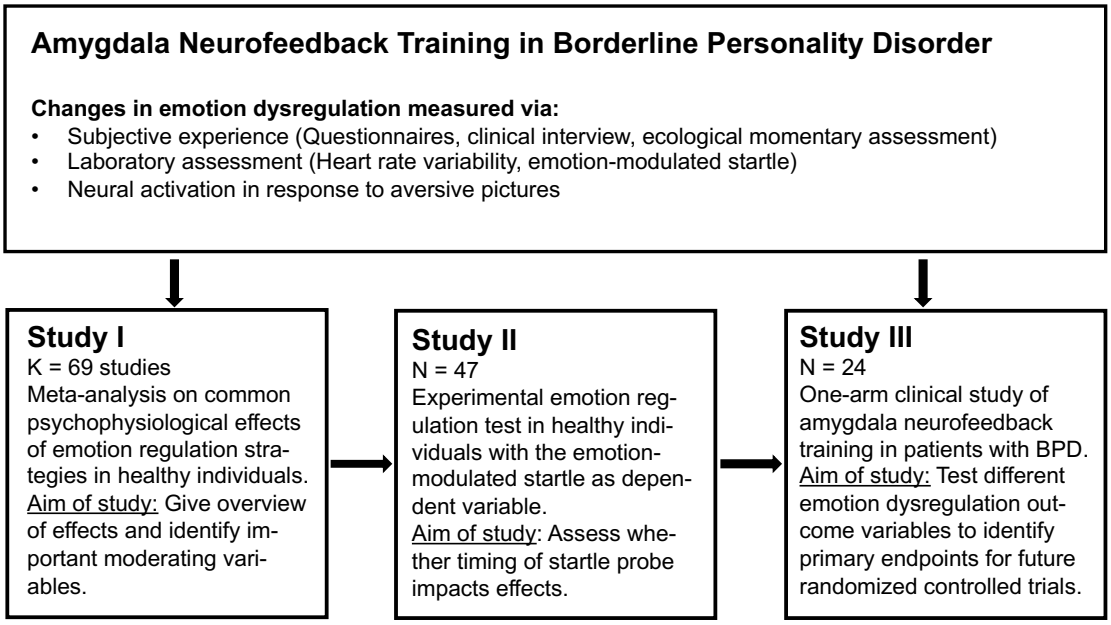

Figure 1.1. Research structure and the relationship between experimental studies.

2 STUDY I: PSYCHOPHYSIOLOGICAL EFFECTS OF DOWNREGULATING NEGATIVE EMOTIONS: INSIGHTS FROM A META-ANALYSIS IN HEALTHY ADULTS

An adapted version of this chapter has been published as ‘Zaehringer, J., Jennen-Steinmetz, C., Schmahl, C., Ende, G., & Paret, C. (2020). Psychophysiological Effects of Downregulating Negative Emotions: Insights From a Meta-Analysis of Healthy Adults. *Frontiers in Psychology, 11*(470). doi:10.3389/fpsyg.2020.00470‘

2.1 Abstract

Assessing psychophysiological responses of emotion regulation is a cost-efficient way to quantify emotion regulation and to complement subjective report that may be biased. Previous studies have revealed inconsistent results complicating a sound interpretation of these findings. In the present study, we summarized the existing literature through a systematic search of articles. Meta-analyses were used to evaluate effect sizes of instructed downregulation strategies on common autonomic (electrodermal, respiratory, cardiovascular and pupillometric) and electromyographic (corrugator activity, emotion-modulated startle) measures. Moderator analyses were conducted, with moderators including study design, emotion induction, control instruction and trial duration. We identified k = 78 studies each contributing multiple sub-samples and performed 23 meta-analyses for combinations of emotion regulation strategy and psychophysiological measure. Overall, results showed that effects of reappraisal and suppression on autonomic measures were highly inconsistent across studies with rather small mean effect sizes. Electromyography (startle and corrugator activity) showed medium effect sizes that were consistent across studies. Our findings highlight the diversity as well as the low level of standardization and comparability of research in this area. Significant moderation of effects by study design, trial duration, and control condition emphasizes the need for better standardization of methods. In addition, the small mean effect sizes resulting from our analyses on autonomic measures should be interpreted with caution. Findings corroborate the importance of multichannel approaches.

2.2 Introduction

Emotion regulation is a vital part of our daily lives. It permits individuals to control the occurrence, intensity, type, and duration of emotions (Gross & Thompson, 2007). Strategies to regulate emotions not only alter the subjective experience of emotions (Gross, 1998a), but also map onto bodily responses such as changes in measures of the autonomic nervous system (Gross, 2002; Webb et al., 2012), emotion-expressive behavior (Dan-Glauser & Gross, 2011, 2015), somatic reflexes such as the emotion-modulated startle (Jackson et al., 2000) or neural activation (Buhle et al., 2014; Ochsner et al., 2004). The habitual use of adaptive emotion regulation strategies is a hallmark of successful functioning and is associated with increased well-being, whereas difficulties with regulating emotions have been linked to many psychopathologies (Aldao et al., 2010; Joormann & Vanderlind, 2014; Schmahl et al., 2014). In light of the significance of emotion regulation, appropriate experimental paradigms are required that are suitable for research involving large sample sizes and patient populations. In a typical emotion regulation study, emotions are experimentally induced using affective stimuli such as films (Gross & Levenson, 1995) or pictures (e.g., International affective picture system; Lang, Bradley, & Cuthbert, 2008). Participants are instructed to regulate their emotional experience or to respond naturally without regulating their emotions (i.e. the control condition). By comparing the regulation with the control condition it is possible to determine the effect of regulation, which has been used as an indirect measure of emotion regulation effectiveness (Webb et al., 2012). Assessing psychophysiological correlates has several important advantages. They move beyond on-line self-reports and retrospective assessments, as physiological responding is regarded as automatic, relatively unconscious, and fast (Bradley, Lang, & Cuthbert, 1993; Edelmann & Baker, 2002; Lapate, Rokers, Li, & Davidson, 2014; Öhman & Soares, 1994; Olsson & Phelps, 2004). Research focusing on the direct effects of emotion regulation has found significant psychophysiological changes even when subjective experience remained unaffected (Gross & Levenson, 1993, 1997). Hence, psychophysiological measures can offer important insights into internal emotional experiences that are not available by assessing self-report. In addition, psychophysiological responses are easier to assess than neural physiological measures (e.g., functional magnetic resonance imaging) and are thus cost-efficient methods for quantifying differences in emotion regulation.

2.2.1 Conceptual Foundations of Emotion Regulation

There have been multiple attempts to classify emotion regulation strategies (Gross, 1998a, 1998b; Koole, 2009; Larsen, 2000). One of the most influential models is the *process model of*

emotion regulation (Gross, 1998a, 1998b, 2015), which broadly categorizes strategies as either being antecedent-focused, i.e. strategies are implemented before the emotional response has fully unfolded, and as response-focused, i.e. strategies are implemented after the emotional response has already been generated. The process model distinguishes five major emotion regulation processes: situation selection (i.e. attempts to change a future emotional response), situation modification (i.e. changing the situation in order to modify its emotional effect), attentional deployment (i.e. distraction away from or concentration on an emotional stimulus to modify the emotion itself), cognitive change (i.e. reappraise a situation or to change the perspective so that the emotional experience is modulated) and response modulation (i.e. strategies to suppress expressive behavior, thoughts, or emotions). Situation selection, situation modification, attentional deployment and cognitive change are regarded as antecedent-focused and response modulation is regarded as a response-focused process.

A majority of past emotion regulation studies have instructed participants to distract themselves from, reappraise or suppress[1] a target stimulus in order to downregulate emotions. These strategies correspond to attentional deployment, cognitive change and response modulation, respectively, In addition, a considerable number of studies allowed participants to use a strategy of their own choice (Baur, Conzelmann, Wieser, & Pauli, 2015; Conzelmann et al., 2015; Dillon & LaBar, 2005; Driscoll et al., 2009; Golkar et al., 2014; Grillon et al., 2015; Jackson et al., 2000; Lee et al., 2009; Lissek et al., 2007; Piper & Curtin, 2006)

The present meta-analysis thus focuses on these four major types of downregulation instructions, that is distraction, reappraisal, suppression and downregulation instructions that allowed participants to choose their own strategy. Other strategies were out of the scope. For a comprehensive overview see Table 2.1.

2.2.2 Psychophysiological Responses of Emotions and Emotion Regulation

There is great interest in understanding the relationship between emotions and psychophysiological responses including responses of the autonomic nervous system (i.e. cardiovascular, electrodermal, respiratory, pupillometric) and responses measured with the electromyogram (EMG) such as facial muscle activity (e.g., corrugator supercilii activity) and somatic reflexes

[1] Acceptance has become increasingly popular across the emotion regulation literature too, yet there has been a debate as to whether it belongs to antecedent (Webb et al., 2012) or response-focused processes (Hofmann & Asmundson, 2008) and as to whether it is a strategy or rather a function of different strategies. Given that very few studies on acceptance assessed psychophysiological responses, it is not included in the present review.

(e.g., emotion-modulated startle). The interested reader is directed to detailed reviews by Cacioppo et al. (2000), Kreibig (2010), Siegel et al. (2018) and Stemmler (2004).

Table 2.1

Emotion downregulation processes and their strategies considered in this meta-analysis.

Process	Strategy	Subtype	Example
Emotion regulation instructions			
Attentional deployment	Distraction	Active distraction	Participants are instructed to think about something positive or neutral that is unrelated to the target emotion/stimulus
Cognitive change	Reappraisal	Reinterpret the emotional stimulus	Participants are instructed to reinterpret the emotional stimulus to decrease the target emotion
		Reappraise via perspective taking, i.e. distancing	Participants are instructed to alter the impact of a stimulus by adopting a more objective perspective
		Reappraisal mixed	A mixture of reappraisal instructions
Response Modulation	Suppression	Suppress the expression of emotion	Participants are instructed to hide the way they are feeling, e.g. not to smile
		Suppress the experience of an emotion	Participants are instructed to suppress their emotional experience
		Suppress thoughts of the emotion eliciting event	Participants are instructed to suppress thoughts about the emotion-eliciting event
		Suppression mixed	A mixture of suppression instructions
Downregulation unspecified	Own choice	Own choice	Participants are free to choose a strategy that works best for them. They are not allowed to create a different emotion or think of something unrelated to the stimulus

Table 2.1 *(continued)*

Process	Strategy	Subtype	Example
Control instructions			
		No instruction (C1)	No instructions are given
		Instructions not to regulate (C2)	Participants are told that they should not use a regulation strategy
		Instructions to maintain (C3)	Participants are instructed to maintain the target emotion
		Instructions to experience naturally (C4)	Participants are instructed to respond naturally without regulating it
		Control mixed (C5)	A mixture of control instructions

See Table 2.2 for an overview of relevant psychophysiological measures within the emotion regulation literature. Such relations have most commonly been studied in terms of two affective dimensions, that is valence (positive-negative) and arousal (high-low) (Bradley, Codispoti, Cuthbert, et al., 2001; Lang, 1995). Some measures such as heart rate, emotion-modulated startle and facial activity are specific to the valence of the emotion (Bradley, Codispoti, Cuthbert, et al., 2001) and others such as skin conductance and pupil dilation are more specific to the arousal dimension (Bradley, Codispoti, Cuthbert, et al., 2001; Greenwald, Cook, & Lang, 1989; VanOyen Witvliet & Vrana, 1995). Past research has also put a lot of effort into answering the question whether different emotion categories (e.g., disgust, sadness, fear) produce distinct physiological response patterns. In a recent meta-analysis the hypothesis could not be confirmed (Siegel et al., 2018). Rather, emotions seem to elicit an unspecific set of psychophysiological changes.

When it comes to the *regulation* of emotions, much evidence has accumulated suggesting that suppression is related to an increase in sympathetic nervous system activity but no difference in self-report to negative stimuli (Gross & Levenson, 1993, 1997; Richards & Gross, 1999). The enhanced sympathetic activation following suppression has led researchers to conclude that suppression "exacts a palpable physiological cost" (Gross & Levenson, 1997, p. 101). In other words, because response-focused strategies involve an active modulation of expressive behavior, increased sympathetic activation might be the result of that effort (Butler et al., 2003). In contrast, past literature has proposed that reappraisal has little impact on sympathetic and cardiovascular measures (Gross, 1998a). A meta-analysis studying the overall physiological

effect of different emotion regulation strategies confirmed this general pattern: cognitive change had a smaller effect on physiology than response modulation (Webb et al., 2012). However, as noted earlier, there is a vast range of different psychophysiological outcome measures ranging from cardiovascular, electrodermal, respiratory, pupillometric and electromyographic response systems and it has been shown that the nature of the relationship between cognitive emotion regulation and different psychophysiological responses can vary largely (Bernat et al., 2011). By simply combining all psychophysiological measures to a composite score is helpful in looking at the overall effectiveness of an emotion regulation strategy (as has been done in the meta-analysis by Webb et al. (2012), but it does not reveal which of the individual psychophysiological responses change or do not change with an emotion regulation strategy.

When looking at individual psychophysiological measures, findings are mixed with respect to the effects of emotion regulation on autonomic physiology. Reappraisal instructions focusing on decreasing negative emotions compared to a control condition have been shown to have no effect on (Goldin et al., 2019; Gross, 1998a; Kalisch et al., 2005), increase (Lohani & Isaacowitz, 2014; Sheppes et al., 2009), or decrease (Urry et al., 2009; Wolgast et al., 2011) skin conductance and to increase (Urry et al., 2006; van Reekum et al., 2007) or decrease (Bebko, Franconeri, Ochsner, & Chiao, 2011) pupil diameter. Contradictory patterns can also be found for suppression strategies. For example, individuals' heart rate was significantly increased (Ben-Naim, Hirschberger, Ein-Dor, & Mikulincer, 2013; Butler, Wilhelm, & Gross, 2006; Hagemann, Levenson, & Gross, 2006), decreased (Gross & Levenson, 1993; Robinson & Demaree, 2009), or stayed the same (Gross, 1998a) when individuals suppressed negative emotions compared to a control condition. These inconsistencies may be due to the large heterogeneity between studies, which can substantially affect the magnitude of the physiological responses. The contradictory pattern of results across the literature does not allow a straightforward interpretation. The causes for these inconsistencies are, however, not well understood, and this inevitably obscures the detection of common trends.

Table 2.2

Common psychophysiological measures of emotion regulation studies.

Measurement	Abbreviation	Measurement system (units)	Description
Cardiovascular			
Cardiac output	CO	l/min	Blood volume pumped by the heart per minute.
Diastolic blood pressure	**DBP**	**mmHg**	Lowest blood pressure of circulating blood on the walls of blood vessels in between two heartbeats, measured in millimeters of mercury.
Ear pulse transit time	**EPTT**	**ms**	Time interval between the R-wave of the electrocardiogram to the pulse wave arrival at the ear.
Finger pulse amplitude	**FPA**	**Arbitrary**	Amplitude of the pulse waveform measured in the finger. Indicator of dilation and constriction of the blood vessels.
Finger pulse transit time	**FPTT**	**ms**	Time interval between the R-wave of the electrocardiogram to the pulse wave arrival at the finger.
Heart rate/interbeat interval/heart period	**HR/HP**	**bpm/ms/ms**	Number of beats per unit of time/time between heart beats (inverse of heart rate).
Heart rate variability	**HRV**	**Units vary by method**	Variation in heart rate. Refers specifically to the high-frequency HRV (also called respiratory sinus arrhythmia (RSA)).
Low frequency HRV	LF	Units vary by method	Variation in heart rate. Refers specifically to the low-frequency HRV.
Mean arterial pressure	**MAP**	**mmHg**	Mean blood pressure of circulating blood on the walls of blood vessels in between two heartbeats, measured in millimeters of mercury.
Pre-ejection period	PEP	ms	Period between the beginning of electrical stimulation of the heart to the opening of the aortic valve. Indicator of the cardiac contractile force (i.e. how hard the heart is beating).
Stroke volume	SV	mL	Volume of blood pumped from the left ventricle per beat.

Table 2.2 (*continued)*

Measurement	Abbreviation	Measurement system (units)	Description
Systolic blood pressure	**SBP**	**mmHg**	Maximum blood pressure of circulating blood on the walls of blood vessels in between two heartbeats, measured in millimeters of mercury.
Total peripheral resistance	TPR	Unity vary by method	Overall resistance that must be overcome to push blood through the whole circulatory system (i.e. all major arterial trees).
Electrodermal			
Skin conductance response	**SCR**	**MicroSiemens**	Peak amplitude, magnitude or local maximum of the skin conductance response. Includes non-specific skin conductance responses during longer periods of time if reported as amplitude.
Skin conductance level	**SCL**	**MicroSiemens**	Mean change of skin conductance over a specific period of time. Operationalized as simple average, change from baseline, area under the curve or integrated signal.
Number of skin conductance responses	nSCR	n	Number of skin responses per unit of time (e.g., per minute).
Respiratory			
Inspiration/expiration time	IT/ET	sec	Average inhalation/exhalation time per respiratory cycle.
Respiration amplitude	**RA**	**mL**	Difference in volts between the point of maximum inspiration and the point of maximum expiration.
Respiration rate	RR	c/min	Number of breaths per minute.
Tidal volume	TV	mL	Air volume that moves into or out of the lungs while breathing quietly.
Pupillometric			
Pupil dilation	**PD**	**mm**	Average diameter of pupil in millimeter during a specific period of time.

Table 2.2 (*continued)*

Measurement	Abbre-viation	Measurement system (units)	Description
Electromyographic			
Emotion-modulated startle	**Startle**	**MicroVolt**	Amplitude of the startle eyeblink response (orbicularis oculi) in response to affective stimuli.
Corrugator supercilii activity	**cEMG**	**MicroVolt**	Muscular activity of the corrugator supercilii responsible for furrowing of the brow.
Zygomaticus major activity	zEMG	MicroVolt	Muscular activity of the zygomaticus major responsible for smiling.
Other			
Finger temperature	**FT**	**F/C°**	Temperature of the finger, in Fahrenheidt (F) or Celcius (C°).

Note. The measures in bold were included in our meta-analysis; for the other measures the number of studies was insufficient (k < 5 studies per cell). Because heart rate (HR) and interbeat interval (IBI) are inversely related, we switched the direction of the effect sizes when IBI was extracted (instead of HR). Descriptions derived and adapted from (Berntson, Quigley, Norman, & Lozano, 2016; Blumenthal et al., 2005; Cacioppo et al., 2000; Dawson, Schell, & Filion, 2016; Siegel et al., 2018).

2.2.3 Factors Related to the Impact of Emotion Regulation on Psychophysiology

2.2.3.1 Study Design

Studies using within-study designs found larger effects of emotion regulation on experiential, behavioral and physiological outcomes than did studies employing between-study designs (cf. Webb et al., 2012). Employing within-study designs reduces sampling error thereby increasing power. On the other hand, within-study designs may also increase task difficulty because participants are required to engage in more than just one emotion regulation strategy. In event-related designs typical for within-subject studies, participants may even shift continuously between different strategies.

2.2.3.2 Emotion Induction

Emotion regulation studies have used a variety of different emotional stimuli, including pictures (e.g., the International Affective Picture System; IAPS; Lang et al., 2008), film clips (Gross & Levenson, 1995), stressful tasks (e.g., the Trier Social Stress Test; Kirschbaum, Pirke, &

Hellhammer, 1993), dyadic interactions (Levenson & Gottman, 1983) or threat of shock paradigms (Delgado, Nearing, LeDoux, & Phelps, 2008). Each type of stimulus provides a reliable method to generate emotions. However, a key dimension on which induction methods differ is whether they require participants to sit passively in front of a monitor or whether they employ a stressful task or conversation with a (romantic) partner. Somatic activity has a significant influence on autonomic response measures, especially on heart rate (Obrist, 1981). In addition, stressful tasks such as giving a speech alter the sympathetic nervous system to a stronger degree than picture viewing (Fechir et al., 2008). When it comes to potential differences between films and pictures, findings are mixed. Studies on emotion processing have been shown that e.g. heart rate returns to baseline if the picture remains still, but further slows down if the picture involves motion (Detenber, Simons, & Bennett Jr, 1998; Simons, Detenber, Roedema, & Reiss, 1999). However, a recent study on emotion regulation reported that films and pictures did not differently affect the emotion regulation process on a physiological level, although films elicited a stronger absolute skin conductance response than pictures (Morawetz, Bode, Baudewig, Jacobs, & Heekeren, 2016a). We are not aware of any other study directly assessing the impact of the emotion induction method on psychophysiological effects in the context of emotion regulation and thus we will address this question in the present analysis[2].

2.2.3.3 Control Instruction

Effects of emotion regulation strategies on psychophysiological measures can be determined by contrasting the emotion regulation instruction against different control instructions. For example, participants can be instructed to "maintain" the emotion they feel (Jackson et al., 2000), to "view" the emotional stimulus (Gross & Levenson, 1993), or to "respond naturally" (Shiota & Levenson, 2009). Previous literature has shown that differences in neural activation depend on the control condition instruction (Schaefer et al., 2002), with higher amygdala activation reported for "maintain" than for "view" instructions. The terminology used as control instructions (e.g., maintain vs. view) has not been systematically explored in psychophysiological studies of emotion regulation yet. However, it could have important influences on physiological processes as shown by an fMRI study (Diers, Weber, Brocke, Strobel, & Schönfeld, 2014).

[2] It should be noted that there might be more aspects of visual stimuli that could possibly influence effect sizes. For example, within the field of visual perception, studies show that faces are not as evocative as scenes (Alpers, Adolph, & Pauli, 2011; Wangelin, Bradley, Kastner, & Lang, 2012). A fine-grained moderator analysis of different aspects of picture and film stimuli however was not possible due to the small number of studies available and because most studies included in the present analysis used a blend of negative scenes and faces as stimuli.

Similarly, Webb et al. (2012) found that the control condition moderated the physiological effects of emotion regulation (Webb et al., 2012).

2.2.3.4 Trial duration

Another important aspect of the study design which varies largely across studies is the trial duration of the regulation period. According to the implementation and maintenance model (IMMO; Kalisch, 2009; Paret et al., 2011), reappraisal for example is divided into two phases: In the early phase, participants choose and implement a regulation strategy, whereas in the late phase they maintain the strategy in working memory and monitor its success. Hence, reappraisal might need several seconds until it effectively reduces negative emotions. Thus, the effect of reappraisal might become larger with increasing trial duration, which might also affect physiology.

2.2.4 Aim of Study

The primary aim of the present study was to quantitatively summarize the relation between popular emotion downregulation instructions (distraction, reappraisal, suppression, own choice) and common psychophysiological measures (i.e. cardiovascular, electrodermal, respiratory, pupillometric, electromyographic) in healthy adults. In light of the contradictory pattern of psychophysiological effects in the emotion regulation literature we aimed to answer the following questions: a) What are the effects of distraction, reappraisal, suppression and downregulation where participants choose a strategy that works best for them on individual psychophysiological response measures? b) How consistent are these effects across studies? and c) What aspects of the study design moderate the effects? In light of the hypothesis that psychophysiological measures are somewhat sensitive to the valence of the induced emotion and because the majority of studies on emotion regulation and psychophysiology induced negative emotions, the present meta-analysis focuses on the downregulation of negative stimuli (for an overview of studies employing positive stimuli see supplement Table 2.9).

We first systematically searched for emotion regulation studies that instructed participants to use emotion regulation strategies and that assessed psychophysiological measures of our interest as dependent variable. To advance current knowledge, we performed meta-analyses to separately quantify the effects for each of these measures during emotion regulation. In addition, we performed moderator analyses to explore the impact of study characteristics on the effect sizes. Moderators of interest were study design, trial duration, control instruction, and emotion induction method. It is important to note that our ability to identify the effects of cognitive

emotion regulation strategies on psychophysiological variables and potential moderators is limited by the published studies available for meta-analysis.

2.3 Methods

2.3.1 Selection of Studies

Studies were identified through a systematic literature search of articles using the PubMed, Web of Science, and PsychINFO databases. The search strategy was developed to maximize the sensitivity of article identification by combining individual words and medical subject headings (MeSH). We searched for the keywords *emotion regulation or emotional regulation* cross referenced with *psychophysiology* [MeSH]*, psychophysiologic*, autonomic, parasympathetic, sympathetic, respiration* [MeSH], *cardiovascular, electrocardiography* [MeSH]*, respiratory sinus arrhythmia* [MeSH], *blood pressure* [MeSH]*, heart rate* [MeSH], *startle, startle reflex* [MeSH], *electromyography* [MeSH], *pupil diameter, pupil dilation, electrodermal or skin conductance,* and *galvanic skin response* [MeSH] cross referenced with *stimulus, stimuli, film*, picture*, image*, script*, anxiety, fear*, threat**, and *video**. Additionally, reference lists from identified studies that met the inclusion criteria (see the next section for criteria) as well as relevant articles in the authors' library were reviewed for titles that might have been previously missed. Subsequently, studies identified in this manner (n = 13) were collected for inclusion.
The search process described above yielded a total of 1353 potentially relevant articles on July 18, 2019 (after duplicates were removed)[3]. The first author and another independent reviewer (Stephanie Mall, research assistant) systematically examined titles and relevant abstracts using the Covidence website (www.covidence.org) to determine whether an article would be subsequently reviewed in full-text format. The following criteria were applied: The study presented original empirical results, was published in a peer-reviewed journal, was written in English or German, included adult healthy participants, and an explicit emotion regulation paradigm was assessed where participants are explicitly told to use emotion regulation strategies to modulate an emotion. We discarded studies that did not assess a psychophysiological measure of interest (e.g., electroencephalography studies) at this point. Based on these criteria, the same two reviewers independently reviewed 157 studies in full-text format.

[3] The search process was updated two times in total. The first search yielded a total of 848 potentially relevant articles on January 22, 2016 (after duplicates were removed). A second search one year later (on February 8, 2017) yielded an additional 210 potentially relevant articles (after duplicates were removed). A third search two years later (on July 18, 2019) yielded an additional 295 potentially relevant articles (after duplicates were removed).

2.3.2 Inclusion/Exclusion Criteria

The 157 studies were examined to determine if they met the following inclusion criteria of our analysis: The study (1) included a control condition in which participants were confronted with emotional contents but did not regulate emotions (see Table 2.1 for definitions of possible control instructions), (2) sampled a psychophysiological measure throughout the regulation phases, (3) did not assess an experimental intervention before the emotion regulation task that may influence the performance of emotion regulation, (4) provided sufficient information to compute the effect size, (5) induced negative emotions, (6) instructed participants to use one or more of the strategies provided in Table 2.1. If studies met inclusion criteria (1) to (6) but did not provide adequate information for effect size computation, we asked the authors for the needed information via e-mail.

Finally, a total of n = 78 studies fulfilled all inclusion criteria. Of those, n = 68 entered our quantitative synthesis (for an overview see Table 2.3). The remaining 10 studies (Baur et al., 2015; Delgado et al., 2008; Driscoll et al., 2009; Jamieson, Mendes, & Nock, 2013; Jamieson, Nock, & Mendes, 2012; Kotwas et al., 2019; Peters & Jamieson, 2016; Peters, Overall, & Jamieson, 2014; Reinecke et al., 2015; Zaehringer, Schmahl, Ende, & Paret, 2018) were not considered, as a meta-analysis on the respective combination of emotion regulation strategy and psychophysiological measure was not possible because the number of studies was too small. See Figure 2.1 for a PRISMA flowchart depiction of the screening and selection of studies.

2.3.3 Data Extraction

The first author coded the sample sizes, group means, standard deviations, t and p-values for tests on group effects and participants' mean age of the eligible studies. Another person independently coded 50% of the included studies to evaluate reliability. Correlation analysis confirmed high interrater-reliability (mean r = .95, range = .66-1.0). In addition, inconsistencies between raters were identified and subsequently corrected. Additionally, the psychophysiological measure, and the specific emotion regulation strategy (distraction, reappraisal, suppression, own choice) were coded. When comparing emotion regulation studies, a major problem arises from inconsistencies in the way emotion regulation instructions are labeled. For example, studies that labeled a condition as "suppression" either instructed participants to use reappraisal (Bernat et al., 2011; Eippert et al., 2007) or to suppress thoughts or facial expressions (Gross & Levenson, 1993; Ohira et al., 2006). To prevent confusion, we specifically evaluated the particular emotion regulation instructions as reported in the articles and coded them according to

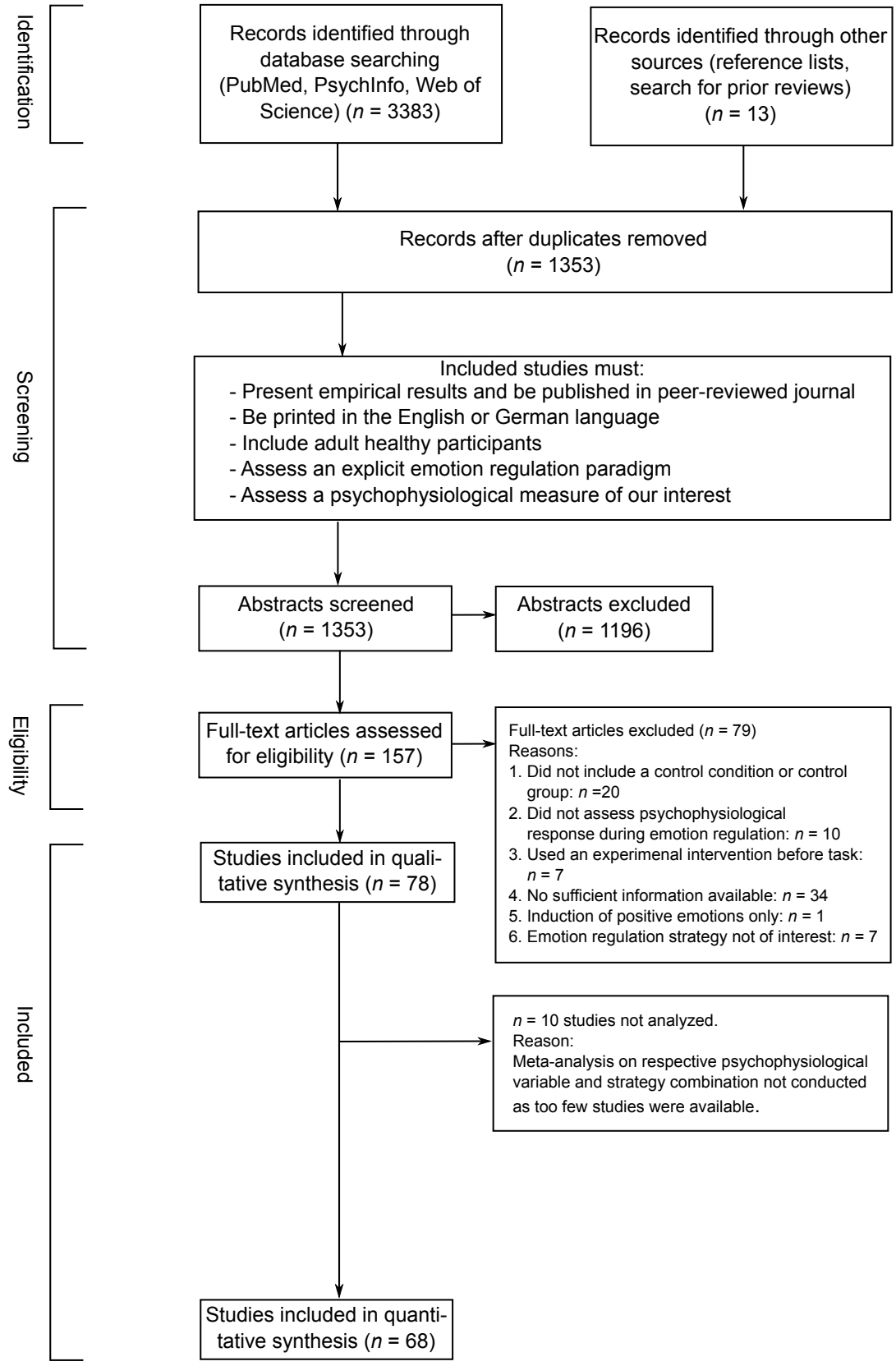

Figure 2.1. PRISMA flowchart of the literature search process.

the taxonomy adapted from Webb et al. (2012). See Table 2.1 for definitions and examples. For this meta-analysis, we also subdivided the control strategies into five types (classifications can be derived from Table 2.1; adapted from Webb et al., 2012): no instruction at all (i.e. "view"), instruction "not to regulate in a certain manner", instructions to "respond naturally", instructions to "maintain" the target emotion or a combination of the above instructions. Furthermore, the researcher(s) also coded whether a study used a between-subject design with two independent groups for the control and the experimental group or a within-subject design with a single group undergoing both regulation and control conditions. In addition, the nature of emotion induction if applicable (pictures, film, music, dyadic interaction, past experience or personally relevant thought, threat of shock (ToS), stressor task, anger task) was also coded. Finally, we coded the trial duration (i.e. the length of the regulation period of a trial, in seconds). We defined the length of a regulation period as the length of one regulation attempt. In event-related designs a regulation attempt thus corresponds to one trial (i.e. after instruction until picture offset), whereas in studies presenting films or stress tasks, a regulation attempt corresponds to the whole film viewing period or task period (i.e. after instruction until end of film/task).
Regarding electrodermal activity, there was great variability in the quantification of skin conductance across studies. We developed a taxonomy by which we divided electrodermal activity measures in skin conductance level, skin conductance response and number of skin conductance responses (see Table 2.2 for a detailed description of the taxonomy and a table summarizing all included studies on electrodermal responses with information about the categorization can be found in the supplement (2.6.4).

2.3.4 Statistical Analysis

Cohen's d was used as the effect size measure in the meta-analyses. For between-subject studies, effect sizes were calculated from the means and standard deviations of the control and experimental (regulation) groups. For within-subject studies, we used the means and standard deviations of the control and experimental (regulation) conditions. If these values were not available, effect sizes were calculated using t-values. Furthermore, the variances of the effect sizes were determined. In within-subject designs, the variance of the effect size estimate depends on the correlation between the paired measurements. If the correlation was not available from the original data, the median correlation from the other studies entering the meta-analysis was used. Effect sizes were interpreted based on Cohen's guidelines (Cohen, 1988). Therefore, effects at the 0.2, 0.5, and 0.8 levels were considered as small, medium, and large, respectively.

Since the experimental conditions of the studies differ in many ways, it is unlikely that the studies share a common effect size. Fixed-effect models are therefore implausible. Following recommendations of Borenstein, Hedges, Higgins, and Rothstein (2010) we conducted random effects meta-analyses. We calculated average effect sizes and 95% confidence intervals (CI). Heterogeneity of effect sizes was assessed with the I^2-statistic which represents the proportion of total variation in the estimated effect sizes that is due to heterogeneity between studies (Higgins & Thompson, 2002). The analyses were performed separated by psychophysiological measure and emotion regulation strategy. Meta-analyses were only conducted when five or more independent samples were available[4].

For each significant meta-analysis we constructed a funnel plot with the effect sizes on the horizontal axis and their standard errors on the vertical axis. Egger's tests (Egger, Davey Smith, Schneider, & Minder, 1997) were applied to evaluate asymmetry in funnel plots which may be caused by publication bias.

Several studies included two or three assessments within a given measure (e.g., skin conductance level during the regulation of sad and disgusting stimuli) so that there was more than one effect size reported for a specific sample. In these cases, we used the mean of the multiple effect sizes. To calculate the variance of this mean effect size, we assumed that the correlation between the effect sizes was 0.5. If studies reported sufficient results from multiple independent samples (e.g., men and women, prone to disgust vs. not prone to disgust), each of them entered the analysis. Effect sizes for interbeat interval and heart rate were included in the same analyses. To align to polarity of the effect sizes, the parameter for interbeat interval was multiplied by minus one. Thus, a negative size of interbeat interval corresponds to decreased heart rate.

As physiological measures have been shown to discriminate between negative and positive emotional states (Bradley & Lang, 2000; Kreibig, 2010; Levenson, Ekman, & Friesen, 1990), we aimed for distinguishing between positive and negative target emotions in our analyses. Only 13 studies in total (Baur et al., 2015; Conzelmann et al., 2015; Dan-Glauser & Gross, 2011, 2015; Demaree, Schmeichel, Robinson, & Everhart, 2004; Driscoll et al., 2009; Giuliani, McRae, & Gross, 2008; Gomez, Scholz, & Danuser, 2015; Gross & Levenson, 1997; Gruber, Hay, & Gross, 2014; Kotwas et al., 2019; Ohira et al., 2006; Wu, Liang, Wang, Zhao, & Zhou, 2016) induced positive emotions. Combinations of psychophysiological measure and emotion regulation strategy resulted in a maximum of three studies. Therefore, meta-analyses on the

[4] Some studies included several independent samples. The minimum number of independent studies required to conduct a meta-analysis was accepted as three.

regulation of positive emotions were not computed in the present study. See an overview of studies using positive emotions in the supplement Table 2.9.

We conducted moderator analyses to test whether features of the experimental context influenced the effect sizes. We used four moderator variables in our analyses: study design (within-subject vs. between-subject), nature of control condition (instruction to respond naturally vs. no instruction), nature of emotion induction (films vs. pictures) and trial duration (i.e. length of a regulation trial, in seconds), as far as there were enough studies for statistical comparison. To evaluate the effects of moderators we used meta-regression analyses and present the regression coefficients.

Statistical analyses were conducted with the metaphor package from R (version 3.2) and SAS 9.4 (SAS Institute Inc., Cary, NC, USA). Statistical significance was defined at the 5% level.

2.3.4.1 Heterogeneity

We investigated whether the variance between the observed effect sizes was larger than what would be expected on the basis of sampling variance alone (Hedges, 1982; Rosenthal & Rubin, 1982). If the effect sizes are heterogeneous it means that the mean effect size does not represent individual effect sizes for studies within the population in that moderators of the effect sizes may be present (e.g., nature of emotion induction). In an analysis with a small number of effect sizes, especially if they are based on small sample size studies, the *Q*-statistic may be nonsignificant even when there is considerable variability among the effect sizes. Therefore, we computed the percent of variability in effect sizes due to heterogeneity using the I^2 statistic (Higgins & Thompson, 2002). *I* represents the amount of variability in effect sizes that is accounted for by heterogeneity as a proportion of the total variability. According to Higgins and Thompson's (2002) general guidelines, mild heterogeneity would be suggested by an I^2 = 30% of the variability in effect sizes, moderate heterogeneity by an I^2 between 30% and 50%, and notable heterogeneity when I^2 is > 50% of the variability.

2.3.4.2 Moderator Analyses

We conducted moderator analyses to test whether features of the experimental context influenced the observed effect sizes. We used four moderator variables in our analyses: study design

(within-subject vs. between-subject), nature of control condition (instruction to respond naturally vs. no instruction)[5], nature of emotion induction (films vs. pictures)[6], trial duration (i.e. length of a regulation trial, in seconds), as far as there were sufficient cases for statistical comparison. We used meta-regression (Thompson & Sharp, 1999) to evaluate moderators. The advantage of meta-regression is that continuous moderators (e.g., trial duration) can be evaluated alongside categorical moderators (e.g., within- versus between-participants designs). For the meta- regressions, β is the beta weight or coefficient assigned to the predictor; t (and the associated p value) tests whether the beta weight is significantly different from zero.

2.4 Results

2.4.1 Descriptive Analyses

Across the 78 studies that were initially considered in our qualitative analysis, heart rate and skin conductance level was measured most frequently, with three times as many effect sizes as for any other measure (see Figure 2.2 for an overview). Thus, emotion regulation strategies and psychophysiological measures were not evenly represented in the published literature. Certain combinations of emotion regulation strategy and psychophysiological measures occurred frequently in published experiments (e.g., reappraisal and measuring heart rate) whereas other combinations were rare or nonexistent (e.g., suppression while measuring stroke volume).
69 individual studies entered our quantitative analyses (for a flowchart of the selection and screening process see Figure 2.1). Study characteristics of these studies are presented in Table 2.3. There are n = 4,474 unique individuals across all of the 68 included studies (meaning that this is the total n across all studies) with many individuals contributing data to more than one effect size for a total of n = 13,380 data points across all meta-analytic comparisons. Because not all studies reported demographic statistics, reported information about age and sex is only an estimated number.

[5] We were unable to test other types of control instructions as there were too few studies available.

[6] We were unable to test other types of emotion inductions (i.e. music, dyadic interaction, past experience or personally relevant thought, threat of shock, stressor task, anger task) as there were too few studies available.

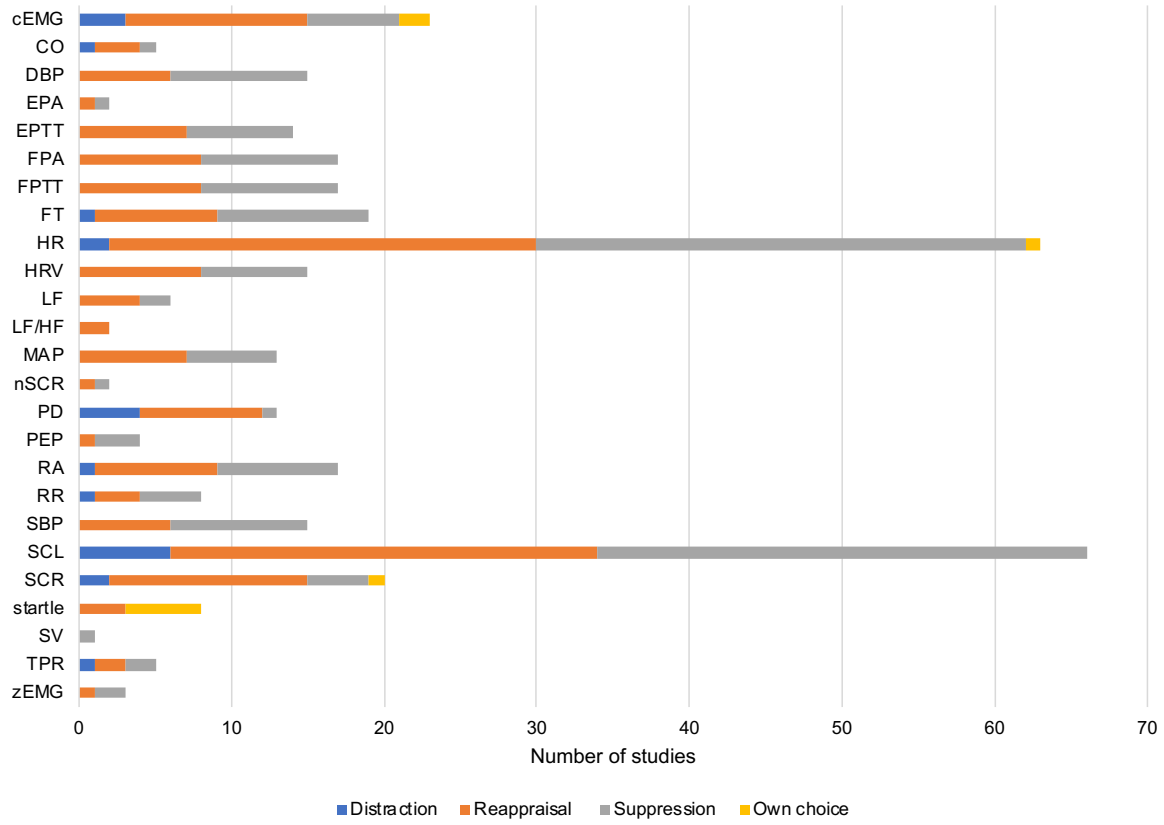

Figure 2.2. Number of available effect sizes for each measure as a function of emotion regulation strategy (distraction, reappraisal, suppression, own choice). Note that the statistic refers to the *k* = 78 studies initially identified in our qualitative analysis. cEMG = corrugator activity; CO = cardiac output; DBP = diastolic blood pressure; EPA = ear pulse amplitude; EPTT = ear pulse transit time; FPA = finger pulse amplitude; FPTT = finger pulse transit time; FT = finger temperature; HR = heart rate; HRV = heart rate variability; LF = low frequency HRV; LF/HF = ratio between low and high frequency HRV; MAP = mean arterial pressure; nSCR = number of skin conductance responses; PD = pupil dilation; PEP = pre-ejection period; RA = respiration amplitude; RR = respiration rate; SBP = systolic blood pressure; SCL = skin conductance level; SCR = skin conductance response; SV = stroke volume; TPR = total peripheral resistance; zEMG = zygomatic activity.

Table 2.3

Characteristics and effect sizes for studies included in the meta-analyses.

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	$\%_{women}$	Age (*M*)	*N*	ES
Ajaya, Peckham, and Johnson (2016)	Reappraisal	HRV	Anger	B	120	A	C1	66	60.61	20.62	40	-0.10
Aldao and Mennin (2012)	Reappraisal	HRV	Disgust, fear, sadness	B	62	F	C1	58	56.90	29.57	38	0.75
Azbel-Jackson, Butler, Ellis, and van Reekum (2015), study 1	Suppression	HR	Negative	B	7	I	C2	60	70.00	21.50	60	-0.22
Azbel-Jackson et al. (2015), study 1	Suppression	SCL	Negative	B	7	I	C2	60	70.00	21.50	60	-0.04
Azbel-Jackson et al. (2015), study 2	Suppression	HR	Negative	B	7	I	C2	80	85.00	22.20	40	0.40
Azbel-Jackson et al. (2015), study 2	Suppression	SCL	Negative	B	7	I	C2	80	85.00	22.20	40	0.73
Bebko et al. (2011)	Reappraisal	PD	Negative	W	10	I	C4	84	47.62	19.67	40	-0.09
Ben-Naim et al. (2013)	Reappraisal	FPA	Negative	B	900	D	C1	254	50.00	24.00	86	-1.52
Ben-Naim et al. (2013)	Reappraisal	FPTT	Negative	B	900	D	C1	254	50.00	24.00	86	-0.18
Ben-Naim et al. (2013)	Reappraisal	HR	Negative	B	900	D	C1	254	50.00	24.00	86	0.33
Ben-Naim et al. (2013)	Reappraisal	SCL	Negative	B	900	D	C1	254	50.00	24.00	86	0.16
Ben-Naim et al. (2013)	Reappraisal	SCR	Negative	B	900	D	C1	254	50.00	24.00	86	-0.39
Ben-Naim et al. (2013)	Suppression	EPTT	Negative	B	900	D	C1	254	50.00	24.00	85	0.09
Ben-Naim et al. (2013)	Suppression	FPA	Negative	B	900	D	C1	254	50.00	24.00	85	-0.66
Ben-Naim et al. (2013)	Suppression	FPTT	Negative	B	900	D	C1	254	50.00	24.00	85	-0.32
Ben-Naim et al. (2013)	Suppression	HR	Negative	B	900	D	C1	254	50.00	24.00	85	0.35
Ben-Naim et al. (2013)	Suppression	SCL	Negative	B	900	D	C1	254	50.00	24.00	85	0.04
Braams, Blechert, Boden, and Gross (2012)	Suppression	HR	Fear	B	17	ToS	C1	123	46.34	21.70	62	-0.04
Bulut, Würz, Küpeli, Bulut, and Sungur (2018), study 1	Reappraisal	HRV	Negative	B	300	I	C4	28	67.86	23.67	28	0.47
Butler et al. (2003), study1	Suppression	MAP	Negative	B		D	C1	72	100.00	20.30	60	-0.09
Butler et al. (2006)	Reappraisal	HR	Negative	B	591	D	C1	190	100.00	20.00	62	-0.24
Butler et al. (2006)	Reappraisal	HRV	Negative	B	591	D	C1	190	100.00	20.00	62	0.51

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	$\%_{women}$	Age (*M*)	*N*	ES
Butler et al. (2006)	Reappraisal	RA	Negative	B	591	D	C1	190	100.00	20.00	62	0.12
Butler et al. (2006)	Suppression	HR	Negative	B	571	D	C1	190	100.00	20.00	69	0.10
Butler et al. (2006)	Suppression	HRV	Negative	B	571	D	C1	190	100.00	20.00	69	0.39
Butler et al. (2006)	Suppression	RA	Negative	B	571	D	C1	190	100.00	20.00	69	-0.76
Butler, Gross, and Barnard (2014)	Reappraisal	SCL	Negative	B	591	D	C1	190	14.74	20.10	61	-0.28
Butler et al. (2014)	Suppression	SCL	Negative	B	571	D	C1	190	14.74	20.10	68	-0.26
Chu et al. (2019)	Reappraisal	HR	Anger	B	10	A	C1	68	54.41	40.00	68	-0.14
Colby, Lanzetta, and Kleck (1977)	Suppression	SCL	Fear	W	6	ToS	C4	10	0.00		10	-0.11
Conzelmann et al. (2015)	Own choice	Startle	Negative	W	8	I	C3	31	48.39	22.00	31	-0.60
Dan-Glauser and Gross (2011)	Suppression	FT	Negative	W	8	I	C4	37	100.00	20.20	37	-0.16
Dan-Glauser and Gross (2011)	Suppression	HR	Negative	W	8	I	C4	37	100.00	20.20	37	-0.57
Dan-Glauser and Gross (2011)	Suppression	MAP	Negative	W	8	I	C4	37	100.00	20.20	37	-0.07
Dan-Glauser and Gross (2011)	Suppression	RA	Negative	W	8	I	C4	37	100.00	20.20	37	-0.82
Dan-Glauser and Gross (2015)	Suppression	FPA	Negative	W	8	I	C4	37	100.00	20.20	37	0.42
Dan-Glauser and Gross (2015)	Suppression	FPTT	Negative	W	8	I	C4	37	100.00	20.20	37	-0.13
Dan-Glauser and Gross (2015)	Suppression	FT	Negative	W	8	I	C4	37	100.00	20.20	37	-0.16
Dan-Glauser and Gross (2015)	Suppression	HR	Negative	W	8	I	C4	37	100.00	20.20	37	-0.71
Dan-Glauser and Gross (2015)	Suppression	MAP	Negative	W	8	I	C4	37	100.00	20.20	37	-0.50
Dan-Glauser and Gross (2015)	Suppression	RA	Negative	W	8	I	C4	37	100.00	20.20	37	-0.45
Demaree et al. (2006)	Suppression	HR	Disgust	B	120	F	C4	69	52.17	19.32	35	0.09
Demaree et al. (2006)	Suppression	HRV	Disgust	B	120	F	C4	69	52.17	19.32	35	0.21
Demaree et al. (2006)	Suppression	RA	Disgust	B	120	F	C4	69	52.17	19.32	35	0.43
Demaree et al. (2006)	Suppression	SCL	Disgust	B	120	F	C4	69	52.17	19.32	35	0.12

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	%$_{women}$	Age (*M*)	*N*	ES
Denson, Creswell, Terides, and Blundell (2014), study 1	Reappraisal	HR	Fear	B	600	S	C1	90	52.22	20.54	90	-0.09
Denson et al. (2014), study 1	Reappraisal	HR	Fear	B	300	S	C1	90	52.22	20.54	86	-0.07
Denson, Grisham, and Moulds (2011)	Reappraisal	HRV	Anger	B	180	F	C1	131	100.00	20.23	86	0.37
Denson et al. (2011)	Suppression	HR	Anger	B	180	F	C1	131	100.00	20.23	89	0.25
Denson et al. (2011)	Suppression	HRV	Anger	B	180	F	C1	131	100.00	20.23	89	0.17
Deveney and Pizzagalli (2008)	Reappraisal	cEMG	Negative	W	5	I	C3	32	78.13	23.97	26	-0.09
Di Simplicio et al. (2012), sample 1	Reappraisal	HR	Negative	W	4	I	C4	30	53.33	28.59	20	0.00
Di Simplicio et al. (2012), sample 1	Reappraisal	HRV	Negative	W	4	I	C4	30	53.33	28.59	20	0.05
Di Simplicio et al. (2012), sample 2	Reappraisal	HR	Negative	W	4	I	C4	30	53.33	28.59	10	0.09
Di Simplicio et al. (2012), sample 2	Reappraisal	HRV	Negative	W	4	I	C4	30	53.33	28.59	10	-0.15
Dillon and LaBar (2005), sample 1	Own choice	Startle	Negative	W	12	I	C3	48	77.08	22.00	12	-0.09
Dillon and LaBar (2005), sample 2	Own choice	Startle	Negative	W	12	I	C3	48	77.08	22.00	12	-0.75
Efinger, Thuillard, and Dan-Glauser (2019)	Reappraisal	HR	Negative	W	8	I	C4	77	100.00	20.70	77	-0.27
Efinger et al. (2019)	Reappraisal	RA	Negative	W	8	I	C4	77	100.00	20.70	77	0.06
Efinger et al. (2019)	Reappraisal	SCL	Negative	W	8	I	C4	77	100.00	20.70	77	-0.19
Efinger et al. (2019)	Distraction	SCL	Negative	W	8	I	C4	77	100.00	20.70	77	-0.27
Fitzpatrick and Kuo (2016)	Distraction	SCL	Negative	W	10	I		30	66.67	30.07	30	0.00
Fuentes-Sánchez, Jaén, Escrig, Lucas, and Pastor (2019)	Reappraisal	SCR	Negative	W	8	I	C4	122	59.02	25.10	106	-0.01
Goldin et al. (2019)	Reappraisal	HR	Negative	W	12	SB	C4	35	57.14	32.20	35	-0.03
Goldin et al. (2019)	Reappraisal	SCL	Negative	W	12	SB	C4	35	57.14	32.20	35	-0.01
Golkar et al. (2014)	Own choice	Startle	Negative	W	5	I	C2	61	54.10	30.90	61	-0.47
Gomez et al. (2015)	Reappraisal	SCR	Disgust	B	10	I	C1	81	64.20	28.15	40	-0.11
Gross and Levenson (1993), study 1	Suppression	EPPT	Disgust	B	64	F	C1	42	0.00	19.30	43	0.07

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	%$_{women}$	Age (*M*)	*N*	ES
Gross and Levenson (1993), study 1	Suppression	FPA	Disgust	B	64	F	C1	42	0.00	19.30	43	-0.38
Gross and Levenson (1993), study 1	Suppression	FPTT	Disgust	B	64	F	C1	42	0.00	19.30	43	-0.24
Gross and Levenson (1993), study 1	Suppression	FT	Disgust	B	64	F	C1	42	0.00	19.30	43	-0.30
Gross and Levenson (1993), study 1	Suppression	HR	Disgust	B	64	F	C1	42	0.00	19.30	43	-0.53
Gross and Levenson (1993), study 1	Suppression	RA	Disgust	B	64	F	C1	42	0.00	19.30	43	-0.18
Gross and Levenson (1993), study 1	Suppression	SCL	Disgust	B	64	F	C1	42	0.00	19.30	43	0.24
Gross and Levenson (1993), study 2	Suppression	EPPT	Disgust	B	64	F	C1	43	100.00	19.20	42	-0.55
Gross and Levenson (1993), study 2	Suppression	FPA	Disgust	B	64	F	C1	43	100.00	19.20	42	-0.81
Gross and Levenson (1993), study 2	Suppression	FPTT	Disgust	B	64	F	C1	43	100.00	19.20	42	0.21
Gross and Levenson (1993), study 2	Suppression	FT	Disgust	B	64	F	C1	43	100.00	19.20	42	-0.96
Gross and Levenson (1993), study 2	Suppression	HR	Disgust	B	64	F	C1	43	100.00	19.20	42	-0.21
Gross and Levenson (1993), study 2	Suppression	RA	Disgust	B	64	F	C1	43	100.00	19.20	42	0.11
Gross and Levenson (1993), study 2	Suppression	SCL	Disgust	B	64	F	C1	43	100.00	19.20	42	0.46
Gross and Levenson (1997)	Suppression	SCL	sadness	B	210	F	C1	180	100.00		180	0.29
Gross (1998a)	Reappraisal	FPA	Disgust	B	64	F	C1	120	50.00	21.00	80	0.12
Gross (1998a)	Reappraisal	FT	Disgust	B	64	F	C1	120	50.00	21.00	80	-0.33
Gross (1998a)	Reappraisal	HR	Disgust	B	64	F	C1	120	50.00	21.00	80	-0.09
Gross (1998a)	Reappraisal	SCL	Disgust	B	64	F	C1	120	50.00	21.00	80	-0.19
Gross (1998a)	Suppression	FPA	Disgust	B	64	F	C1	120	50.00	21.00	80	-0.60
Gross (1998a)	Suppression	FT	Disgust	B	64	F	C1	120	50.00	21.00	80	-1.04
Gross (1998a)	Suppression	HR	Disgust	B	64	F	C1	120	50.00	21.00	80	0.02
Gross (1998a)	Suppression	SCL	Disgust	B	64	F	C1	120	50.00	21.00	80	0.41
Hagemann et al. (2006)	Suppression	EPPT	Negative	B	5	ToS, I	C1	252	51.98	20.50	168	-0.38

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	%$_{women}$	Age (*M*)	*N*	ES
Hagemann et al. (2006)	Suppression	FPA	Negative	B	5	ToS, I	C1	252	51.98	20.50	168	-0.25
Hagemann et al. (2006)	Suppression	FPTT	Negative	B	5	ToS, I	C1	252	51.98	20.50	168	-0.39
Hagemann et al. (2006)	Suppression	FT	Negative	B	5	ToS, I	C1	252	51.98	20.50	168	-0.55
Hagemann et al. (2006)	Suppression	HR	Negative	B	5	ToS, I	C1	252	51.98	20.50	168	0.73
Hagemann et al. (2006)	Suppression	HRV	Negative	B	20	ToS, I	C1	252	51.98	20.50	168	-0.34
Hagemann et al. (2006)	Suppression	SCL	Negative	B	5	ToS, I	C1	252	51.98	20.50	168	0.49
Hallam et al. (2015)	Reappraisal	SCL	Negative	W	10	I	C4	40	50.00	20.00	26	0.00
Hallam et al. (2015)	Suppression	SCL	Negative	W	10	I	C4	40	50.00	20.00	26	-0.01
Jackson et al. (2000)	Own choice	Startle	Negative	W	14	I	C3	48	68.75	20.50	44	-1.04
Kim and Hamann (2012)	Reappraisal	cEMG	Negative	W	24	I	C4	36	50.00	20.19	33	-0.30
Kim and Hamann (2012)	Reappraisal	SCR	Negative	W	24	I	C4	36	50.00	20.19	32	0.11
Kinner et al. (2017)	Reappraisal	PD	Negative	W	5	I	C4	30	100.00	24.40	28	0.26
Kinner et al. (2017)	Reappraisal	SCR	Negative	W	5	I	C4	30	100.00	24.40	25	0.00
Kunzmann, Kupperbusch, and Levenson (2005)	Suppression	HR	Disgust	W	117	F	C1	95	49.47	46.00	47	-0.26
Kunzmann et al. (2005)	Suppression	SCL	Disgust	W	117	F	C1	95	49.47	46.00	47	0.15
Leiberg, Eippert, Veit, and Anders (2012)	Reappraisal	SCR	Negative	W	6	I	C4	24	100.00	24.10	24	0.17
Lohani and Isaacowitz (2014), sample 1	Distraction	SCL	Sadness	W	300	F	C1	42	73.81	18.50	40	0.48
Lohani and Isaacowitz (2014), sample 1	Reappraisal	cEMG	Sadness	W	300	F	C1	48	79.17	71.42	42	-0.17
Lohani and Isaacowitz (2014), sample 1	Reappraisal	SCL	Sadness	W	300	F	C1	42	73.81	18.50	40	0.56
Lohani and Isaacowitz (2014), sample 1	Suppression	SCL	Sadness	W	300	F	C1	42	73.81	18.50	40	0.52
Lohani and Isaacowitz (2014), sample 2	Distraction	SCL	Sadness	W	300	F	C1	48	79.17	71.42	44	0.24
Lohani and Isaacowitz (2014), sample 2	Reappraisal	cEMG	Sadness	W	300	F	C1	42	73.81	18.50	40	-0.30
Lohani and Isaacowitz (2014), sample 2	Reappraisal	SCL	Sadness	W	300	F	C1	48	79.17	71.42	44	0.09

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	$\%_{women}$	Age (*M*)	*N*	ES
Lohani and Isaacowitz (2014), sample 2	Suppression	SCL	Sadness	W	300	F	C1	48	79.17	71.42	44	0.13
Low, Stanton, and Bower (2008)	Reappraisal	HR	Negative	B	600	S	C3	81	58.02	20.60	56	0.29
Martins, Florjanczyk, Jackson, Gatz, and Mather (2018)	Reappraisal	PD	Negative	W	7	I	C4	48	68.75	69.10	48	0.06
Martins et al. (2018)	Reappraisal	PD	Negative	W	7	I	C4	48	60.42	21.06	48	0.06
Morawetz et al. (2016a)	Reappraisal	SCR	Negative	W	8	I, F	C4	59	33.90	32.47	47	0.08
Morawetz et al. (2016b)	Reappraisal	SCR	Negative	W	8	I	C4	23	52.17	25.70	16	-0.19
Morawetz, Bode, Baudewig, Kirilina, and Heekeren (2016c)	Reappraisal	SCR	Negative	W	8	F	C4	23	65.22	22.95	22	-0.03
Ohira et al. (2006)	Suppression	HR	Negative	W	60	I	C4	10	100.00	24.22	9	0.04
Opitz, Lee, Gross, and Urry (2014), sample 1	Reappraisal	cEMG	Sadness	W	8	I	C4	30	53.33	61.90	29	-0.43
Opitz et al. (2014), sample 1	Reappraisal	HR	Sadness	W	8	I	C4	30	63.33	19.45	28	-0.02
Opitz et al. (2014), sample 1	Reappraisal	SCL	Sadness	W	8	I	C4	30	63.33	19.45	27	-0.02
Opitz et al. (2014), sample 2	Reappraisal	cEMG	Sadness	W	8	I	C4	30	63.33	19.45	28	-1.07
Opitz et al. (2014), sample 2	Reappraisal	HR	Sadness	W	8	I	C4	30	53.33	61.90	29	-0.14
Opitz et al. (2014), sample 2	Reappraisal	SCL	Sadness	W	8	I	C4	30	53.33	61.90	29	-0.27
Ortner (2015)	Reappraisal	SCR	Negative	B	8	I	C1	120	75.83		76	0.01
Plieger et al. (2017)	Reappraisal	SCL	Negative	W	5	I	C1	91	82.42	24.53	91	-0.28
Richards and Gross (1999), study 2	Suppression	DBP	Negative	B	84	I	C1	85	100.00	18.80	74	0.36
Richards and Gross (1999), study 2	Suppression	FT	Negative	B	84	I	C1	85	100.00	18.80	74	-0.37
Richards and Gross (1999), study 2	Suppression	HR	Negative	B	84	I	C1	85	100.00	18.80	74	-0.11
Richards and Gross (1999), study 2	Suppression	SBP	Negative	B	84	I	C1	85	100.00	18.80	74	0.27
Richards and Gross (1999), study 2	Suppression	SCL	Negative	B	84	I	C1	85	100.00	18.80	74	-0.14
Roberts, Levenson, and Gross (2008), sample 1	Suppression	DBP	Disgust	B	62	F	C1	40	60.00	20.80	40	0.91
Roberts et al. (2008), sample 1	Suppression	HR	Disgust	B	62	F	C1	40	60.00	20.80	40	-0.23

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	%$_{women}$	Age (*M*)	*N*	ES
Roberts et al. (2008), sample 1	Suppression	SBP	Disgust	B	62	F	C1	40	60.00	20.80	40	0.60
Roberts et al. (2008), sample 1	Suppression	SCL	Negative	B	62	F	C1	40	60.00	20.80	40	0.00
Roberts et al. (2008), sample 2	Suppression	DBP	Disgust	B	62	F	C1	40	60.00	20.80	40	0.84
Roberts et al. (2008), sample 2	Suppression	HR	Disgust	B	62	F	C1	40	60.00	20.80	40	0.08
Roberts et al. (2008), sample 2	Suppression	SBP	Disgust	B	62	F	C1	40	60.00	20.80	40	0.66
Roberts et al. (2008), sample 2	Suppression	SCL	Negative	B	62	F	C1	40	60.00	20.80	40	0.35
Roberts et al. (2008), sample 3	Suppression	DBP	Disgust	B	62	F	C1	40	60.00	20.80	40	-0.31
Roberts et al. (2008), sample 3	Suppression	HR	Disgust	B	62	F	C1	40	60.00	20.80	40	-0.61
Roberts et al. (2008), sample 3	Suppression	SBP	Disgust	B	62	F	C1	40	60.00	20.80	40	0.01
Roberts et al. (2008), sample 3	Suppression	SCL	Negative	B	62	F	C1	40	60.00	20.80	40	0.62
Roberts et al. (2008), sample 4	Suppression	DBP	Disgust	B	62	F	C1	40	60.00	20.80	40	0.12
Roberts et al. (2008), sample 4	Suppression	HR	Disgust	B	62	F	C1	40	60.00	20.80	40	0.26
Roberts et al. (2008), sample 4	Suppression	SBP	Disgust	B	62	F	C1	40	60.00	20.80	40	0.11
Roberts et al. (2008), sample 4	Suppression	SCL	Negative	B	62	F	C1	40	60.00	20.80	40	0.30
Robinson and Demaree (2009)	Suppression	HR	Sadness	W	120	F	C4	102	50.98	19.75	102	-0.23
Robinson and Demaree (2009)	Suppression	HRV	Sadness	W	120	F	C4	102	50.98	19.75	102	0.41
Robinson and Demaree (2009)	Suppression	SCL	Sadness	W	120	F	C4	102	50.98	19.75	102	0.26
Rohrmann, Hopp, Schienle, and Hodapp (2009), sample 1	Reappraisal	HR	Disgust	B	60	F	C1	120	0.00	25.47	36	0.22
Rohrmann et al. (2009), sample 1	Reappraisal	HR	Disgust	B	60	F	C1	120	0.00	25.47	36	-0.34
Rohrmann et al. (2009), sample 1	Suppression	HR	Disgust	B	60	F	C1	120	0.00	25.47	36	0.47
Rohrmann et al. (2009), sample 1	Suppression	HR	Disgust	B	60	F	C1	120	0.00	25.47	36	-0.66
Rohrmann et al. (2009), sample 2	Reappraisal	SCL	Disgust	B	60	F	C1	120	0.00	25.47	36	0.35
Rohrmann et al. (2009), sample 2	Reappraisal	SCL	Disgust	B	60	F	C1	120	0.00	25.47	36	-0.57

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	%$_{women}$	Age (*M*)	*N*	ES
Rohrmann et al. (2009), sample 2	Suppression	SCL	Disgust	B	60	F	C1	120	0.00	25.47	36	0.85
Rohrmann et al. (2009), sample 2	Suppression	SCL	Disgust	B	60	F	C1	120	0.00	25.47	36	-0.23
Roth et al. (2014), study 2	Suppression	SCL	Fear	B	197	F	C1	116	60.34	24.90	65	-0.04
Roth et al. (2014), study 2	Distraction	SCL	Fear	B	197	F	C1	116	60.34	24.90	67	-0.77
Sheppes et al. (2009)	Reappraisal	FT	Sadness	B	190	F	C5	45	100.00	22.90	29	0.22
Sheppes et al. (2009)	Reappraisal	SCL	Sadness	B	190	F	C5	45	100.00	22.90	29	1.13
Sheppes et al. (2009)	Distraction	SCL	Sadness	B	190	F	C5	45	100.00	22.90	29	0.23
Shermohammed et al. (2017)	Reappraisal	HR	Negative	W	8	I	C1	25	48.00	20.89	19	0.65
Shermohammed et al. (2017)	Reappraisal	SCR	Negative	W	8	I	C1	25	48.00	20.89	17	0.12
Shiota and Levenson (2009, 2012), sample 1	Suppression	DBP	Disgust	W	180	F	C4	76	50.00	25.50	73	-0.66
Shiota and Levenson (2009, 2012), sample 1	Suppression	EPPT	Disgust	W	180	F	C4	76	50.00	25.50	74	0.33
Shiota and Levenson (2009, 2012), sample 1	Suppression	FPA	Disgust	W	180	F	C4	76	50.00	25.50	75	0.49
Shiota and Levenson (2009, 2012), sample 1	Suppression	FPTT	Disgust	W	180	F	C4	76	50.00	25.50	75	-0.12
Shiota and Levenson (2009, 2012), sample 1	Suppression	FT	Disgust	W	180	F	C4	76	50.00	25.50	76	-0.24
Shiota and Levenson (2009, 2012), sample 1	Suppression	HR	Disgust	W	180	F	C4	76	50.00	25.50	75	-0.40
Shiota and Levenson (2009, 2012), sample 1	Suppression	MAP	Disgust	W	180	F	C4	76	50.00	25.50	73	-0.66
Shiota and Levenson (2009, 2012), sample 1	Suppression	RA	Disgust	W	180	F	C4	76	50.00	25.50	72	-0.29
Shiota and Levenson (2009, 2012), sample 1	Suppression	SBP	Disgust	W	180	F	C4	76	50.00	25.50	73	-0.69
Shiota and Levenson (2009, 2012), sample 1	Suppression	SCL	Disgust	W	180	F	C4	76	50.00	25.50	73	-0.42
Shiota and Levenson (2009, 2012), sample 2	Reappraisal	FPA	Disgust, sadness	W	180	F	C4	22	50.00	25.50	23	0.37
Shiota and Levenson (2009, 2012), sample 2	Reappraisal	FPTT	Disgust, sadness	W	180	F	C4	22	50.00	25.50	23	0.47
Shiota and Levenson (2009, 2012), sample 2	Reappraisal	FT	Disgust, sadness	W	180	F	C4	22	50.00	25.50	23	0.36
Shiota and Levenson (2009, 2012), sample 2	Reappraisal	HR	Disgust, sadness	W	180	F	C4	22	50.00	25.50	23	-0.29

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	$\%_{women}$	Age (*M*)	*N*	ES
Shiota and Levenson (2009, 2012), sample 2	Reappraisal	RA	Disgust, sadness	W	180	F	C4	22	50.00	25.50	22	-0.34
Shiota and Levenson (2009, 2012), sample 2	Reappraisal	SCL	Disgust, sadness	W	180	F	C4	22	50.00	25.50	23	-0.27
Shiota and Levenson (2009, 2012), sample 3	Reappraisal	FPA	Disgust, sadness	W	180	F	C4	26	50.00	25.30	25	0.14
Shiota and Levenson (2009, 2012), sample 3	Reappraisal	FPTT	Disgust, sadness	W	180	F	C4	26	50.00	25.30	25	0.02
Shiota and Levenson (2009, 2012), sample 3	Reappraisal	FT	Disgust, sadness	W	180	F	C4	26	50.00	25.30	26	0.12
Shiota and Levenson (2009, 2012), sample 3	Reappraisal	HR	Disgust, sadness	W	180	F	C4	26	50.00	25.30	25	-0.11
Shiota and Levenson (2009, 2012), sample 3	Reappraisal	RA	Disgust, sadness	W	180	F	C4	26	50.00	25.30	24	-0.10
Shiota and Levenson (2009, 2012), sample 3	Reappraisal	SCL	Disgust, sadness	W	180	F	C4	26	50.00	25.30	24	0.10
Shiota and Levenson (2009, 2012), sample 4	Suppression	DBP	Disgust	W	180	F	C4	72	50.00	44.70	64	-0.27
Shiota and Levenson (2009, 2012), sample 4	Suppression	EPPT	Disgust	W	180	F	C4	72	50.00	44.70	71	-0.06
Shiota and Levenson (2009, 2012), sample 4	Suppression	FPA	Disgust	W	180	F	C4	72	50.00	44.70	71	0.27
Shiota and Levenson (2009, 2012), sample 4	Suppression	FPTT	Disgust	W	180	F	C4	72	50.00	44.70	72	0.11
Shiota and Levenson (2009, 2012), sample 4	Suppression	FT	Disgust	W	180	F	C4	72	50.00	44.70	72	-0.03
Shiota and Levenson (2009, 2012), sample 4	Suppression	HR	Disgust	W	180	F	C4	72	50.00	44.70	72	-0.30
Shiota and Levenson (2009, 2012), sample 4	Suppression	MAP	Disgust	W	180	F	C4	72	50.00	44.70	64	-0.28
Shiota and Levenson (2009, 2012), sample 4	Suppression	RA	Disgust	W	180	F	C4	72	50.00	44.70	66	-0.07
Shiota and Levenson (2009, 2012), sample 4	Suppression	SBP	Disgust	W	180	F	C4	72	50.00	44.70	64	-0.32
Shiota and Levenson (2009, 2012), sample 4	Suppression	SCL	Disgust	W	180	F	C4	72	50.00	44.70	69	-0.39
Shiota and Levenson (2009, 2012), sample 5	Reappraisal	FPA	Disgust, sadness	W	180	F	C4	22	50.00	44.70	23	0.23
Shiota and Levenson (2009, 2012), sample 5	Reappraisal	FPTT	Disgust, sadness	W	180	F	C4	22	50.00	44.70	24	-0.28
Shiota and Levenson (2009, 2012), sample 5	Reappraisal	FT	Disgust, sadness	W	180	F	C4	22	50.00	44.70	24	0.00
Shiota and Levenson (2009, 2012), sample 5	Reappraisal	HR	Disgust, sadness	W	180	F	C4	22	50.00	44.70	24	-0.31
Shiota and Levenson (2009, 2012), sample 5	Reappraisal	RA	Disgust, sadness	W	180	F	C4	22	50.00	44.70	23	-0.18

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	$\%_{women}$	Age (*M*)	*N*	ES
Shiota and Levenson (2009, 2012), sample 5	Reappraisal	SCL	Disgust, sadness	W	180	F	C4	22	50.00	44.70	22	-0.10
Shiota and Levenson (2009, 2012), sample 6	Reappraisal	FPA	Disgust, sadness	W	180	F	C4	26	50.00	43.20	26	0.17
Shiota and Levenson (2009, 2012), sample 6	Reappraisal	FPTT	Disgust, sadness	W	180	F	C4	26	50.00	43.20	26	0.21
Shiota and Levenson (2009, 2012), sample 6	Reappraisal	FT	Disgust, sadness	W	180	F	C4	26	50.00	43.20	26	0.23
Shiota and Levenson (2009, 2012), sample 6	Reappraisal	HR	Disgust, sadness	W	180	F	C4	26	50.00	43.20	26	-0.06
Shiota and Levenson (2009, 2012), sample 6	Reappraisal	RA	Disgust, sadness	W	180	F	C4	26	50.00	43.20	24	-0.10
Shiota and Levenson (2009, 2012), sample 6	Reappraisal	SCL	Disgust, sadness	W	180	F	C4	26	50.00	43.20	25	-0.09
Shiota and Levenson (2009, 2012), sample 7	Suppression	DBP	Disgust	W	180	F	C4	72	50.00	64.80	69	-0.30
Shiota and Levenson (2009, 2012), sample 7	Suppression	EPPT	Disgust	W	180	F	C4	72	50.00	64.80	68	-0.01
Shiota and Levenson (2009, 2012), sample 7	Suppression	FPA	Disgust	W	180	F	C4	72	50.00	64.80	65	0.23
Shiota and Levenson (2009, 2012), sample 7	Suppression	FPTT	Disgust	W	180	F	C4	72	50.00	64.80	65	0.16
Shiota and Levenson (2009, 2012), sample 7	Suppression	FT	Disgust	W	180	F	C4	72	50.00	64.80	72	0.11
Shiota and Levenson (2009, 2012), sample 7	Suppression	HR	Disgust	W	180	F	C4	72	50.00	64.80	69	-0.12
Shiota and Levenson (2009, 2012), sample 7	Suppression	MAP	Disgust	W	180	F	C4	72	50.00	64.80	69	-0.30
Shiota and Levenson (2009, 2012), sample 7	Suppression	RA	Disgust	W	180	F	C4	72	50.00	64.80	66	-0.26
Shiota and Levenson (2009, 2012), sample 7	Suppression	SBP	Disgust	W	180	F	C4	72	50.00	64.80	69	-0.27
Shiota and Levenson (2009, 2012), sample 7	Suppression	SCL	Disgust	W	180	F	C4	72	50.00	64.80	69	-0.46
Shiota and Levenson (2009, 2012), sample 8	Reappraisal	FPA	Disgust, sadness	W	180	F	C4	24	50.00	64.80	23	-0.08
Shiota and Levenson (2009, 2012), sample 8	Reappraisal	FPTT	Disgust, sadness	W	180	F	C4	24	50.00	64.80	23	0.03
Shiota and Levenson (2009, 2012), sample 8	Reappraisal	FT	Disgust, sadness	W	180	F	C4	24	50.00	64.80	24	0.10
Shiota and Levenson (2009, 2012), sample 8	Reappraisal	HR	Disgust, sadness	W	180	F	C4	24	50.00	64.80	23	-0.19
Shiota and Levenson (2009, 2012), sample 8	Reappraisal	RA	Disgust, sadness	W	180	F	C4	24	50.00	64.80	20	-0.19
Shiota and Levenson (2009, 2012), sample 8	Reappraisal	SCL	Disgust, sadness	W	180	F	C4	24	50.00	64.80	23	-0.11

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	%$_{women}$	Age (*M*)	*N*	ES
Shiota and Levenson (2009, 2012), sample 9	Reappraisal	FPA	Disgust, sadness	W	180	F	C4	24	50.00	64.50	22	0.40
Shiota and Levenson (2009, 2012), sample 9	Reappraisal	FPTT	Disgust, sadness	W	180	F	C4	24	50.00	64.50	22	-0.12
Shiota and Levenson (2009, 2012), sample 9	Reappraisal	FT	Disgust, sadness	W	180	F	C4	24	50.00	64.50	23	0.58
Shiota and Levenson (2009, 2012), sample 9	Reappraisal	HR	Disgust, sadness	W	180	F	C4	24	50.00	64.50	22	-0.10
Shiota and Levenson (2009, 2012), sample 9	Reappraisal	RA	Disgust, sadness	W	180	F	C4	24	50.00	64.50	22	-0.26
Shiota and Levenson (2009, 2012), sample 9	Reappraisal	SCL	Disgust, sadness	W	180	F	C4	24	50.00	64.50	22	-0.64
Soto, Lee, and Roberts (2016)	Suppression	HR	Disgust	W	58	F	C1	59	54.24	19.51	48	-0.19
Soto et al. (2016)	Suppression	SCL	Disgust	W	58	F	C1	59	54.24	19.51	47	-0.15
Stiller, Kattner, Gunzenhauser, and Schmitz (2019)	Reappraisal	HR	Negative	B	165	F	C2	61	73.77	24.30	41	0.15
Stiller et al. (2019)	Reappraisal	SCL	Negative	B	165	F	C2	61	73.77	24.30	41	0.49
Stiller et al. (2019)	Suppression	HR	Negative	B	165	F	C2	61	73.77	24.30	40	0.58
Stiller et al. (2019)	Suppression	SCL	Negative	B	165	F	C2	61	73.77	24.30	40	0.35
Strauss, Ossenfort, and Whearty (2016)	Reappraisal	PD	Negative	W	5	I	C4	25	64.00	19.80	25	0.14
Svaldi, Caffier, and Tuschen-Caffier (2010)	Reappraisal	FPTT	Sadness	W	125	F	C1	25	100.00	38.30	21	-0.11
Svaldi et al. (2010)	Reappraisal	HR	Sadness	W	125	F	C1	25	100.00	38.30	25	-0.32
Svaldi et al. (2010)	Reappraisal	HRV	Sadness	W	125	F	C1	25	100.00	38.30	21	-0.67
Svaldi et al. (2010)	Reappraisal	SCL	Sadness	W	125	F	C1	25	100.00	38.30	23	0.10
Svaldi et al. (2010)	Suppression	FPTT	Sadness	W	211	F	C1	25	100.00	38.30	21	-0.68
Svaldi et al. (2010)	Suppression	HR	Sadness	W	211	F	C1	25	100.00	38.30	25	-0.16
Svaldi et al. (2010)	Suppression	HRV	Sadness	W	211	F	C1	25	100.00	38.30	21	-0.18
Svaldi et al. (2010)	Suppression	SCL	Sadness	W	211	F	C1	25	100.00	38.30	23	0.50
Urry et al. (2006)	Reappraisal	PD	Negative	W	5	I	C3	17	52.94	62.90	14	0.43
Urry et al. (2009)	Reappraisal	PD	Negative	W	8	I	C3	26	57.69	64.80	26	0.46

Table 2.3 *(continued)*

Study name	Strategy	Measure	Emotion	D	TD (s)	EI	CTL	N_{total}	%$_{women}$	Age (*M*)	*N*	ES
Urry et al. (2009)	Reappraisal	SCL	Negative	W	8	I	C3	26	57.69	64.80	26	-0.42
Urry (2009)	Reappraisal	cEMG	Negative	W	8	I	C2	41	63.41	20.00	40	0.03
Urry (2009)	Reappraisal	HR	Negative	W	8	I	C2	41	63.41	20.00	40	-0.14
Urry (2009)	Reappraisal	SCR	Negative	W	8	I	C2	41	63.41	20.00	39	0.18
Urry (2010)	Reappraisal	cEMG	Negative	W	4	I	C4	54	48.15	18.80	54	-0.32
Urry (2010)	Reappraisal	HR	Negative	W	4	I	C4	54	48.15	18.80	53	0.03
Urry (2010)	Reappraisal	SCL	Negative	W	4	I	C4	54	48.15	18.80	52	0.03
Uy et al. (2013)	Suppression	HRV	Disgust	W	133	F	C1	7	57.14	29.80	7	0.45
van Reekum et al. (2007)	Reappraisal	PD	Negative	W	8	I	C4	29	62.07	63.00	21	0.53
L. E. Williams, Bargh, Nocera, and Gray (2009), study1	Reappraisal	HR	Fear	B	180	S	C1	39	64.10	20.60	26	-0.15
Williams et al. (2009), study2	Reappraisal	HR	Fear	B	180	S	C1	47	65.96	21.30	30	0.05
Wolgast et al. (2011)	Reappraisal	cEMG	Disgust, fear, sadness	B	153	F	C1	94	51.06	27.40	62	-0.79
Wolgast et al. (2011)	Reappraisal	SCL	Disgust, fear, sadness	B	153	F	C1	94	51.06	27.40	62	-0.87
Wu et al. (2016), study 2	Reappraisal	SCL	Sadness	B	180	F	C1	42	100.00	22.31	42	-0.18
Yuan, Liu, Ding, and Yang (2014)	Suppression	SCL	Anger	B	1800	A	C1	64	0.00	29.52	43	-0.72

Note. cEMG = corrugator electromyography; DBP = diastolic blood pressure; EPTT = ear pulse transit time; FPA = finger pulse amplitude; FPTT = finger pulse transit time; FT = finger temperature; HR = heart rate; HRV = heart rate variability; MAP = mean arterial pressure; PD = pupil dilation; RA = respiration amplitude; SBP = systolic blood pressure; SCL = skin conductance level; SCR = skin conductance response; D = study design; B = between-subject design study; W = within-subject design study; TD = trial duration; EI = nature of emotion induction; F = film; I = images; A = anger task; D = dyadic interaction; S = Stress induction; SB = negative self-beliefs; ToS = Threat of shock; CTL = nature of control induction; C1 = no instruction given ("view"); C2 = instruction not to regulate; C3 = instruction to maintain target emotion; C4 = instruction to respond naturally; C5 = a combination of C1-C4; N_{total} = number of participants in the study; %$_{women}$ = Percent of women in the respective sample of the study. *N* = number of participants in the subsample; ES = effect size (Cohen's *d*).

Table 2.4

Mean computed effect sizes for each emotion regulation strategy and psychophysiological measure.

Strategy	Response system	Measure	*k*	ES	*SE*	CI_{lower}	CI_{upper}	I^2	*p*	Direction of effect
Distraction	Electrodermal	SCL	6	-0.004	0.175	-0.454	0.447	95.53	0.984	-
Reappraisal	Cardiovascular	FPA	8	-0.015	0.215	-0.524	0.495	88.90	0.948	-
		FPTT	8	-0.021	0.074	-0.195	0.153	24.73	0.785	-
		FT	8	0.159	0.091	-0.056	0.373	21.99	0.124	-
		HR	28	-0.092	0.039	-0.171	-0.012	21.91	0.026*	REG < CTL
		HRV	8	0.106	0.164	-0.282	0.494	87.62	0.537	-
	Electrodermal	SCL	26	-0.065	0.069	-0.206	0.077	71.11	0.355	-
		SCR	12	-0.041	0.031	-0.028	0.109	33.01	0.218	-
	Pupillometric	PD	8	0.136	0.071	-0.033	0.305	69.82	0.098	-
	Respiratory	RA	8	-0.097	0.051	-0.218	0.024	00.00	0.101	-
	Electromyographic	cEMG	9	-0.321	0.098	-0.546	-0.096	42.84	0.011*	REG < CTL
Suppression	Cardiovascular	DBP	8	0.039	0.199	-0.431	0.510	83.99	0.849	-
		EPPT	7	-0.048	0.107	-0.309	0.213	54.77	0.670	-
		FPA	9	-0.108	0.165	-0.488	0.272	84.160	0.530	-
		FPTT	9	-0.174	0.100	-0.404	0.057	70.10	0.121	-
		FT	10	-0.327	0.115	-0.586	-0.067	70.03	0.019*	REG < CTL
		HR	29	-0.093	0.067	-0.231	0.045	78.28	0.177	-
		HRV	7	0.126	0.122	-0.174	0.425	78.76	0.344	-
		MAP	6	-0.338	0.084	-0.554	-0.123	16.45	0.010**	REG < CTL
		RA	9	-0.285	0.118	-0.558	-0.012	61.21	0.042*	REG < CTL
		SBP	8	-0.018	0.164	-0.407	0.371	78.32	0.917	-

Table 2.4 *(continued)*

Strategy	Response system	Measure	*k*	ES	*SE*	CI_{lower}	CI_{upper}	I^2	*p*	Direction of effect
	Electrodermal	SCL	31	0.106	0.064	-0.025	0.236	77.57	0.108	-
Own choice	Electromyographic	Startle	5	-0.621	0.145	-1.021	-0.219	47.35	0.013**	REG < CTL

Note. *k* = number of studies; ES = effect size; *SE* = standard error; CI = confidence interval; cEMG = corrugator activity; DBP = diastolic blood pressure; EPTT = ear pulse transit time; FPA = finger pulse amplitude; FPTT = finger pulse transit time; FT = finger temperature; HR = heart rate; HRV = heart rate variability; MAP = mean arterial pressure; PD = pupil dilation; RA = respiration amplitude; SBP = systolic blood pressure; SCL = skin conductance level; SCR = skin conductance response. I^2 = percent of variability in effect sizes that is due to heterogeneity between studies.
*$p \leq .05$, **$p \leq .01$.

2.4.1.1 Electromyographic Responses

When considering studies that instructed participants to choose a strategy that worked best for them only, downregulation of negative emotions had a significant negative effect on the emotion-modulated startle (d = -.62, CI = [-1.02, -.22], p = .01, k = 5, I^2 = 47.35)[7] with a large effect size and moderate heterogeneity (see Table 2.4 and Figure 2.26 for details). This means that the instruction to decrease negative emotions reduced, on average, the startle response compared to the control instruction. Moreover, reappraisal significantly decreased corrugator activity (d = -.32, CI = [-.55, -.10], p = .01, k = 9, I^2 = 42.84) with medium effect size and moderate heterogeneity (see Table 2.4 and Figure 2.5 for details). However, number of studies on the startle (k = 5) and corrugator activity (k =9) was small and thus should be interpreted with caution.

2.4.1.2 Electrodermal Responses

No significant effect was obtained for distraction on skin conductance level compared to the control condition (d = -.004, CI = [.98, .45], p = .45, k = 6, I^2 = 95.35; see Figure 2.4). Similarly, reappraisal had no significant effect on skin conductance level (d = -.07, CI = [-.21, .08], p = .35, k = 26, I^2 = 71.11; see Figure 2.13) and skin conductance response (d = .04, CI = [-.03, .11], p = .11, k = 12, I^2 = 33.01; see Figure 2.14), compared to the control condition.
In addition, suppression did not significantly change the skin conductance level (d = .11, CI = [-.03, .24], p = .11, k = 31, I^2 = 77.57; see Table 2.4 and Figure 2.25).

2.4.1.3 Respiratory Responses

Suppression significantly decreased respiration amplitude (d = -.29, CI [-.56, -.01], p = .04, k = 9, I^2 = 61.21; see Figure 2.23). Sample size was small (k = 9) and thus should be interpreted with caution.

2.4.1.4 Pupillometric responses

On average, reappraisal did not significantly change pupil dilation in response to negative stimuli compared to a control condition (see Table 2.4 and Figure 2.11 for details). Descriptively,

[7] Instructions to downregulate negative emotions (own choice and reappraisal instructions combined) had a significant negative effect on the emotion-modulated startle too (d = -.44, CI = [-.75, -.14], p = .01, k = 8, I^2 = 74.76).

this result might have been driven by one study (Bebko et al., 2011) which found a decrease in pupil size during reappraisal, whereas other studies (Strauss et al., 2016; Urry et al., 2009; van Reekum et al., 2007) found an increase in pupil size during reappraisal. Overall sample size (k = 8) was small and thus should be interpreted with caution.

2.4.2 Evaluation of Publication Bias

For each significant meta-analysis we constructed a funnel plot with the effect sizes on the horizontal axis and their standard errors on the vertical axis. Egger's tests (Egger et al., 1997) were applied to evaluate asymmetry in funnel plots which may be caused by publication bias. Egger's test revealed that there was significant asymmetry only for the effect of reappraisal on heart rate (p = .008). Individual funnel plots are presented in the supplement (Figure 2.27).

2.4.3 Moderator Analyses

We report moderator analyses only for reappraisal and suppression. For distraction and own choice the number of studies was too small or the distributions of the moderators were inadequate.

2.4.3.1 Study Design

Study design (within-subject vs. between-subject) significantly moderated effect sizes of suppression on finger temperature (β = .54, $p \leq .01$), finger pulse amplitude (β = .78, $p \leq 0.001$) and heart rate (β = -.38, $p \leq .01$). See Table 2.5 for details. The effect of suppression on finger temperature were significant for between-subject design studies (d = -.62, $p \leq .001$, k = 5), whereas the effect on heart rate became significant for within-subject designs (d = -.29, $p \leq .001$, k = 10).

2.4.3.2 Nature of Control Instruction

Effect sizes of suppression on finger temperature (β = .54, $p \leq .01$), finger pulse transit time (β = .42, $p < .05$) and finger pulse amplitude (β = .78, $p < .001$) were significantly moderated by the control instruction (instruction to respond naturally vs. no instruction) (see Table 2.6). The effect of suppression on heart rate (β = -.29, $p < .05$) and skin conductance level (β = -.35, $p \leq .01$) was also moderated by the control instruction (instruction to respond naturally vs. no instruction). When studies with no instruction were considered only, suppression significantly increased skin conductance level (d = .19, p ≤ .01, k = 21), decreased finger temperature (d = -

.62, $p \leq .001$, $k = 5$) and finger pulse transit time ($d = -.40$, $p \leq .01$, $k = 5$). Conversely, when studies with instruction to respond naturally were considered only, suppression significantly decreased heart rate ($d = -.32$, $p \leq .01$, $k = 8$).

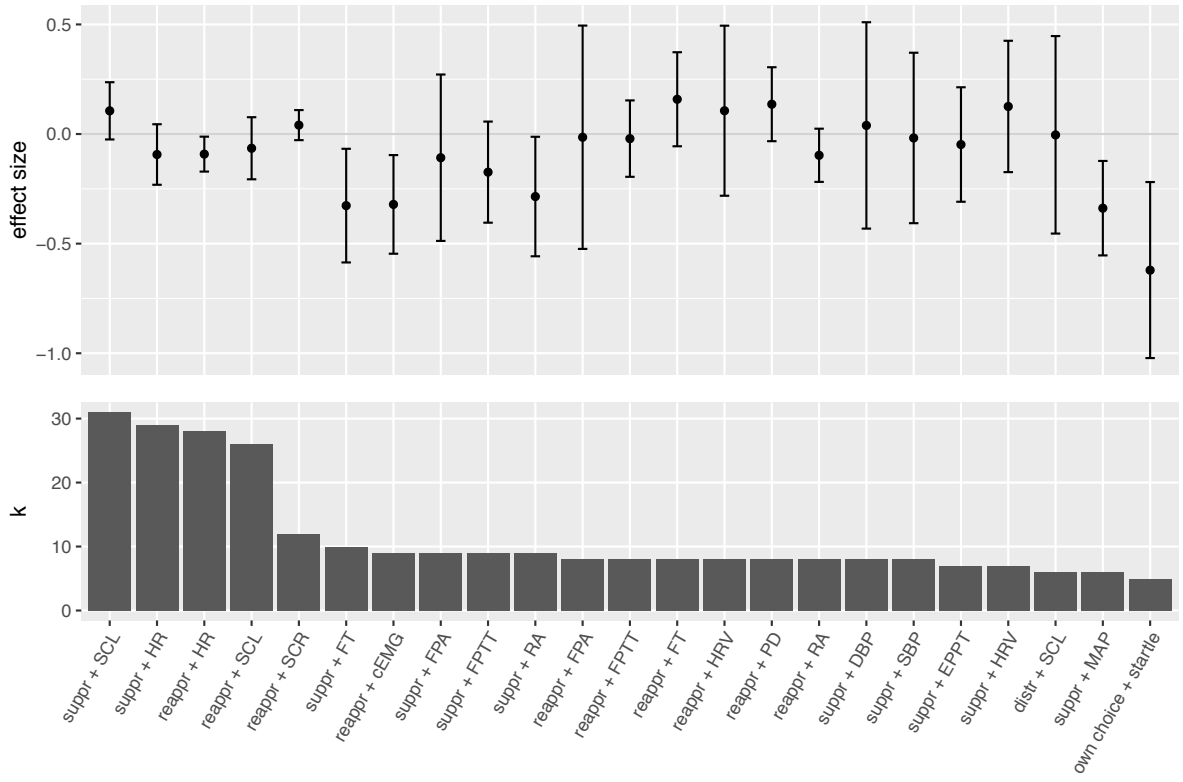

Figure 2.3. Mean effect sizes and confidence intervals for each conducted meta-analysis (upper panel) sorted by number of samples k of the meta-analysis, respectively (lower panel). Suppr = suppression; reappr = reappraisal; distr = distraction. HR = heart rate; SCL = skin conductance level; SCR = skin conductance response; FT = finger temperature; cEMG = corrugator activity; FPA = finger pulse amplitude; FPTT = finger pulse transit time; RA = respiration amplitude; HRV = heart rate variability; PD = pupil dilation; DBP = diastolic blood pressure; SBP = systolic blood pressure; EPTT = ear pulse transit time; MAP = mean arterial pressure.

Table 2.5

Moderator analyses on study design (within-study design vs. between-study design).

Strategy	Measure	*k*	k_{within}	$k_{between}$	N_{total}	β	*SE*	*p*
Reappraisal	SCL	26	17	9	1082	-0.001	0.161	0.997
Reappraisal	HR	28	16	12	1176	-0.131	0.085	0.134
Suppression	SCL	31	11	20	1805	-0.176	0.126	0.174
Suppression	FT	10	5	5	701	0.543	0.138	0.004**
Suppression	FPTT	9	5	4	608	0.080	0.219	0.725
Suppression	FPA	9	4	5	666	0.775	0.115	0.000**
Suppression	RA	9	5	4	467	-0.188	0.263	0.497
Suppression	HR	29	10	19	1640	-0.379	0.113	0.002**

Note. k = number of studies; *SE* = standard error; FPA = finger pulse amplitude; FPTT = finger pulse transit time; FT = finger temperature; HR = heart rate; RA = respiration amplitude; SCL = skin conductance level; β regression coefficient (within vs. between).
**$p \leq .01$

Table 2.6

Moderator analyses on nature of control instruction (instruction to respond naturally vs. no instruction).

Strategy	Measure	*k*	k_{C4}	k_{C1}	N_{total}	β	*SE*	*p*
Reappraisal	SCL	23	12	11	986	-0.063	0.127	0.625
	SCR	11	7	4	491	0.087	0.113	0.460
	HR	25	13	12	1039	-0.033	0.083	0.696
Suppression	SCL	28	7	21	1665	-0.347	0.130	0.012*
	FT	10	5	5	701	0.543	0.138	0.004**
	FPTT	9	4	5	608	0.422	0.156	0.030*
	FPA	9	4	5	666	0.775	0.115	0.000**
	HR	26	8	18	1500	-0.293	0.130	0.034*

Note. k = number of studies; *SE* = standard error; FPA = finger pulse amplitude; FPTT = finger pulse transit time; FT = finger temperature; HR = heart rate; SCL = skin conductance level; SCR = skin conductance response; C4 = instruction to respond naturally; C1 = no instruction; β regression coefficient (respond naturally vs. no instruction).
*$p \leq .05$, **$p \leq .01$

2.4.3.3 Emotion Induction

Moderator analyses of effect sizes were conducted for film vs. picture only, as too few studies employing other emotion induction methods for each strategy and psychophysiological measure combination were available to interpret moderator analyses in a meaningful way. Emotion

induction (films vs. pictures) did not significantly moderate the effect sizes of reappraisal and suppression on skin conductance level and heart rate (see Table 2.7).

Table 2.7

Moderator analysis on emotion induction (films v. images).

Strategy	Measure	*k*	k_{films}	k_{images}	N_{total}	β	*SE*	*p*
Reappraisal	SCL	23	16	7	900	0.126	0.167	0.458
Reappraisal	HR	20	12	8	723	-0.150	0.086	0.101
Suppression	SCL	26	22	4	1431	0.049	0.187	0.795
Suppression	HR	25	19	6	1256	0.145	0.144	0.324

Note. *k* = number of studies; *SE* = standard error; HR = heart rate; SCL = skin conductance level; β = regression coefficient (films vs. pictures).

Table 2.8

Moderator analyses on trial duration (s).

Strategy	Measure	*k*	N_{total}	β	*SE*	*p*
Distraction	SCL	6	287	0.084	0.081	0.354
Reappraisal	HR	28	1176	0.015	0.012	0.209
	HRV	8	305	0.071	0.053	0.232
	PD	8	250	-2.492	1.996	0.258
	SCL	26	1082	0.021	0.021	0.324
	SCR	12	530	-0.028	0.013	0.053*
	cEMG	9	354	0.000	0.051	0.997
Suppression	DBP	8	440	-0.408	0.141	0.028*
	EPPT	7	551	0.022	0.024	0.403
	FPA	9	666	-0.030	0.039	0.464
	FPTT	9	608	-0.011	0.026	0.677
	FT	10	701	0.130	0.086	0.172
	HR	29	1640	0.028	0.022	0.214
	HRV	7	491	0.044	0.047	0.392
	RA	9	467	-0.032	0.048	0.526
	SBP	8	440	-0.387	0.094	0.006**
	SCL	31	1805	-0.026	0.012	0.039*

Note. *k* = number of studies; *SE* = standard error; cEMG = corrugator electromyography; DBP = diastolic blood pressure; EPTT = ear pulse transit time; FPA = finger pulse amplitude; FPTT = finger pulse transit time; FT = finger temperature; HR = heart rate; HRV = heart rate variability; MAP = mean arterial pressure; PD = pupil dilation; RA = respiration amplitude; SBP = systolic blood pressure; SCL = skin conductance level; SCR = skin conductance response; β regression coefficient (refers to one minute change in trial duration).
*$p \leq .05$, **$p \leq .01$

2.4.4 Meta-Analyses

As the 68 studies contributed data to multiple effect sizes, we computed 267 individual effect sizes (see Table 2.3) that entered 24 different meta-analyses (see Table 2.4 and Figure 2.3). Overall, computed individual mean effect sizes for each combination of regulation strategy with measure did not exceed d = .62 (own choice effect on startle; see Table 2.4). Figure 2.3 also highlights that some meta-analyses revealed large confidence intervals and non-significant effect sizes, suggesting that these effects are rather inconsistent (e.g., suppression effect on skin conductance response, ear pulse transit time, diastolic blood pressure and finger pulse amplitude, reappraisal effect on finger pulse amplitude, heart rate variability, and distraction effect on skin conductance level). Largest effect sizes were obtained for electromyographic responses (startle and corrugator activity), followed by suppression effects on some cardiovascular measures (i.e. finger temperature and mean arterial pressure). For many computed mean effect sizes confidence intervals around the mean effect were large (see Figure 2.3), indicating that the accuracy of our analysis to predict the true effect was rather low. Moreover, heterogeneity differed largely across meta-analyses (see Table 2.4). For individual forest plots of each meta-analysis see supplement Figure 2.4 - Figure 2.26.

2.4.4.1 Cardiovascular Responses

Reappraisal significantly decreased heart rate (d = -0.09, CI = [-.17, -.01], p = .03, k = 28, I^2 = 21.90), yet the effect size was very small and direction of effects across individual studies were inconsistent (see Figure 2.9). Reappraisal had no significant effect on all other tested cardiovascular measures (i.e. finger pulse amplitude, finger pulse transit time, finger temperature and heart rate variability) with mean effect sizes ranging between -.02 and .16 (see Table 2.4).
Suppression significantly decreased finger temperature (d = -.33, CI = [-.59, -.07], p = .02, k = 10, I^2 = 70.03; see Figure 2.19), and mean arterial pressure (d = -.34, CI = [-.55, -.12], p = .01, k = 6, I^2 = 16.45; see Figure 2.22), with small to medium effect sizes and mild to notable heterogeneity. Suppression did not significantly change diastolic blood pressure, ear pulse transit time, heart rate, heart rate variability, systolic blood pressure, and skin conductance response (see Table 2.4 for details and statistics).

2.4.4.2 Trial Duration

Trial duration significantly moderated the effect of reappraisal on skin conductance response ($\beta = -.03$, $p = .05$, $k = 12$) and the effect of suppression on skin conductance level ($\beta = -.03$, $p < .05$, $k = 31$), diastolic ($\beta = -.41$, $p < .05$, $k = 8$) and systolic blood pressure ($\beta = -.39$, $p < .01$, $k = 8$) in that the effect became more negative with longer trial durations (see Table 2.8). The moderating effect of trial duration on suppression and skin conductance level was mainly driven by one study (Yuan et al., 2014).

2.5 Discussion

Over the past two decades, emotion regulation has become a vibrant research field. Our literature search corroborates this trend. It revealed an increase of almost 60% of potentially relevant publications for our meta-analysis within the recent three years. The vast growth of literature illustrates a vigorous interest in understanding the psychophysiological mechanisms of emotion regulation.

Previous studies on the psychophysiological responses to emotion regulation revealed inconsistent results. Moreover, distraction and reappraisal strategies appeared to have no or little effect on psychophysiology (Webb et al., 2012), and suppression significantly increased sympathetic arousal (Gross, 1998a; Gross & Levenson, 1993). This meta-analysis provides the first attempt to elucidate common trends with means of a quantitative summary of the effects of common emotion regulation strategies on different cardiovascular, electrodermal, respiratory, pupillometric, and electromyographic measures. We performed a structured literature review and conducted a meta-analysis for each combination of psychophysiological measure and emotion regulation strategy whenever there were enough studies available. In brief, we found that suppression significantly decreased mean arterial pressure, finger temperature, and respiration amplitude, whereas reappraisal led to decreased heart rate and decreased corrugator activity (see Table 2.4 and Figure 2.3 for an overview of effects). When participants were free to choose between emotion regulation strategies, a significant inhibition of the emotion-modulated startle (sometimes referred to as fear-potentiated startle) response could be observed. Due to the limited number of studies on distraction, we were not able to conduct meta-analyses on psychophysiological responses except for skin conductance level, and this meta-analysis revealed no significant effect. Publication bias appeared to have an overall minor effect.

As Figure 2.3 illustrates, aggregated effect sizes from the tested autonomic responses were small in general. We did not compute an overall effect size across all psychophysiological

measures. Yet aggregated effect sizes for each psychophysiological measure correspond with the results reported by Webb et al.'s meta-analysis (Webb et al., 2012). They had reported an overall small negative effect of response modulation (e.g., suppression strategies) on psychophysiology (d = .19, [CI = .14, .01]). Attentional deployment (e.g., distraction strategies) had no significant effect on physiological measures (d = .00, CI = [.14, .15]), and so did cognitive change (e.g., reappraisal) (d = .05, [CI = .07 to .16]) (Webb et al., 2012). We conclude that effects of emotion regulation on autonomic measures – if at all present – seem to be rather small and raise the question whether emotion regulation success can be reliably quantified with autonomic measures. It should however be noted that the psychophysiological measures entering our analysis were limited. Figure 2.2 illustrates that there were a number of measures not included as too few studies were available. For example, measures of cardiac function that can be derived via impedance cardiography have received scant attention in the previous literature but provide promising results: Studies have shown that emotion regulation changed total peripheral resistance with medium to large effect sizes (Jamieson et al., 2013; Jamieson et al., 2012; Peters & Jamieson, 2016; Peters et al., 2014).

Activation of the sympathetic nervous system causes an increase in skin conductivity, pupil dilation, heart rate, pre-ejection period, blood pressure, peripheral vasoconstriction, and increased respiration amplitude and respiration rate. Successful emotion regulation should be accompanied by a reduction of sympathetic activity (McRae & Shiota, 2017). Our study reveals that the effects are not quite that straightforward. Suppression lowered finger temperature (indicative of increased sympathetic activity), yet also decreased mean arterial pressure and respiration amplitude (indicative of lower sympathetic activity). Similarly, reappraisal decreased heart rate (indicative of lower sympathetic activity) but did not change any of the other tested autonomic measures. McRae & Shiota (2017) point out that psychophysiological effects often diverge in patterns that correspond to different psychological states (Kreibig, 2010; Shiota, Neufeld, Yeung, Moser, & Perea, 2011), which can result in misinterpretations about the association between psychophysiological responses and the underlying psychological processes (Cacioppo & Tassinary, 1990; Cacioppo, Tassinary, & Berntson, 2007). Psychophysiological responses are usually influenced by various factors, such as stress, workload, or tiredness, and thus may distort the effects of emotion regulation. Decreased pupil size during reappraisal was observed in one study and has been interpreted to be the result of decreased emotional arousal (Bebko et al., 2011). Alternatively, studies have interpreted larger pupil size during reappraisal as an indicator of higher cognitive effort (Urry et al., 2006; van Reekum et al., 2007). They

infer that pupil size may increase during successful emotion regulation as an indicator of increased cognitive processing. The ambiguity of such effects implies that we need a better understanding of cognitive and emotional processes causing autonomic change, and how these changes relate to emotion regulation success.

Another problem is the inconsistency of direction of effect sizes. Different directions of effect sizes rendered the meta-analyses insignificant and infer that there are important factors not yet understood. For example, the meta-analysis of pupil dilation during reappraisal (see Figure 2.11) revealed that one study (Bebko et al., 2011), which received a strong weight in the analysis, found a significant decrease in pupil diameter during reappraisal, while other studies found an increase in pupil diameter (e.g., Strauss et al., 2016; Urry et al., 2009; van Reekum et al., 2007). Similarly, our meta-analysis on heart rate during suppression (see Figure 2.20) revealed that studies found mean heart rate acceleration in response to suppression (e.g., (Hagemann et al., 2006; Stiller et al., 2019), whereas other studies found a heart rate deceleration (Dan-Glauser & Gross, 2011, 2015; Kunzmann et al., 2005). Therefore, the second aim of the present work was to explore the impact of methodological differences using several moderators (trial duration, nature of emotion induction, nature of control instruction, study design).

Effects of suppression on heart rate, finger temperature and finger pulse amplitude were significantly moderated by study design (within vs. between-subject). Between-subject design studies showed a significant decrease in finger temperature and finger pulse amplitude during suppression whereas studies with a within-subject design revealed no significant effect. Conversely, within-subject design studies showed a significant decrease in heart rate whereas studies with a between-subject design revealed no significant effect. The moderating effect of study design on heart rate might also reflect that between-subject design studies in this particular meta-analysis assessed extremely diverse emotion induction methods. For example, two studies (Ben-Naim et al., 2013; Butler et al., 2006) assessed emotion regulation in dyadic interactions. Hagemann et al. (2006) used startle tones in combination with pictures. Rohrmann et al. (2009), Gross (1998a), Denson et al. (2011) used film stimuli. Within-subject design studies considered in this meta-analysis used films and pictures only. Therefore, the nature of emotion induction may account for some variance in the effect sizes obtained across studies using between-subject designs. When data from more studies will be available in the future, it might be possible to confirm this assumption.

Effects of reappraisal and suppression on several electrodermal and cardiovascular measures (i.e. skin conductance level, finger temperature, finger pulse transit time, finger pulse amplitude and heart rate) were significantly moderated by the nature of control instructions. Except for

finger pulse amplitude, the effects became significant when no instruction (i.e. "view" instruction) was given but did not become significant when the instruction to respond naturally was given. This does not correspond with findings by Webb et al. (2012) who found that emotion regulation strategies in general had smaller effects on experiential, behavioral and physiological measures combined when the control condition required participants to "view" or "not to regulate" and larger effects when the control condition required participants to respond naturally. In contrast to our study, they did not determine the moderating effect of control instruction on physiological effects of emotion regulation but considered the overall effect of psychophysiological, behavioral and experiential measures. Control conditions requiring participants to simply view a negative stimulus might correspond to a physiological baseline condition. However, when receiving the instruction to respond naturally, participants might unconsciously pay more attention to their emotional response, which may be particularly sensitive to psychophysiological responses.

Trial duration significantly moderated effect sizes of suppression on skin conductance level, diastolic and systolic blood pressure, and of reappraisal on skin conductance response in that the effects became more negative with increasing trial length. Studies on electrodermal responses may be difficult to compare within the conducted meta-analyses because trial durations varies largely across studies. This might be especially problematic for skin conductance level, as longer time windows carry the risk that non-specific skin conductance responses occur. If these phasic responses are not separated from the tonic parts, they might influence the absolute skin conductance level (Boucsein et al., 2012). Hence, skin conductance level assessed over several seconds in an event-related design might be different than skin conductance level assessed over several minutes in a block-design. We accounted for this variability in parts by conducting a moderator analysis with trial duration as the moderator. We observed effects in both positive and negative direction. Studies with very short trial duration tend to report an increase in skin conductance, whereas studies with longer or extremely long trial durations tend to report a decrease in skin conductance. However, we acknowledge that our analysis did not allow to differentiate for example between studies that assessed skin conductance averages but eliminated the tonic parts (Hallam et al., 2015; Plieger et al., 2017) and studies that assessed skin conductance level without separating the phasic from the tonic responses. We encourage future researcher to use similar research methodology and terminology as suggested by the committee report on publication recommendations (Boucsein et al., 2012) to make studies more comparable in the future. In total, the varying effects of skin conductance across studies may be in part due to the high variability in assessment and quantification.

Compared to the tested *autonomic* responses (i.e. cardiovascular, electrodermal, pupillometric and respiratory responses), our present analysis revealed that effects of measures assessed with *electromyography* were medium and consistent across individual studies (see Figure 2.5 and Figure 2.26). Regarding the emotion-modulated startle, we found a significant decrease through emotion downregulation with a mean effect size of d = -.62. Corrugator activity significantly decreased with reappraisal of negative emotions with a medium effect size of d = -.32. As both analyses included a rather small number of studies resulting in large confidence intervals, they should be treated with caution (see Figure 2.3). Nevertheless, the results on electromyography showed more consistent results compared to the autonomic measures assessed in the present review and this encourages possible reasons that might have accounted for this consistency.

Studies have shown that both the emotion-modulated startle and corrugator activity are specific to valence: The startle is inhibited in response to pleasant but potentiated in response to unpleasant stimuli with stronger responses for high- than for low-arousing stimuli (Bradley, Cuthbert, & Lang, 1993; Hamm, Cuthbert, Globisch, & Vaitl, 1997; Hawk & Cook, 2000; Schupp, Cuthbert, Bradley, Birbaumer, & Lang, 1997; Vrana et al., 1988). Corrugator supercilii is generally considered to correspond to changes in valence, too (Tassinary, Cacioppo, & Vanman, 2007). The valence-specificity might facilitate to measure the correspondence to changes in valence and hence allows to track the regulation effect more closely, compared with autonomic measures that rather reflect changes in arousal. However, there are also studies showing that in the context of emotion regulation, the startle response is more sensitive to changes in arousal (Dillon & LaBar, 2005; Zaehringer et al., 2018).

Animal studies have shown that the amygdala, a key structure in emotion processing, directly modulates the auditory startle reflex via modulation of midbrain neurons (Davis, 1992; Rosen & Davis, 1988), which has been recently complemented by fMRI work in human subjects (Kuhn et al., 2020). Researcher have argued that the emotional modulation as indexed by the startle reflex may serve as a direct indicator of amygdala activation independent of task demands (Grillon & Baas, 2003). Similarly, the amygdala projects to the facial motor nucleus thereby coupling emotional facial expressions to the motive circuit (Davis, 2000). The amygdala is a robust neural target of emotion regulation (Buhle et al., 2014) and altered amygdala activation with emotion regulation thus likely mediates the modulatory effect on the startle response and corrugator activity. Taken together, the specificity to the valence dimension and the direct modulation via the brain's motivational system may contribute to the findings of emotion regulation effects on emotion modulated startle and corrugator activity.

With regard to the emotion-modulated startle, it is also possible that the emotion regulation instruction might have influenced the obtained effect sizes. Participants in these studies were free to choose an emotion regulation strategy that worked best for them. By allowing participants to choose from different strategies, they might be more successful in regulating their emotions, which could result in larger effects. Moreover, the startle response unfolds within milliseconds, whereas autonomic responses such as pupil dilation, electrodermal responses, and heart rate variability rather unfold over several seconds, or even minutes. Therefore, the startle response may be easier to measure because it is clearly time-locked to the startle probe and all changes can be measured in studies with shorter observation times during the trials, whereas a skin conductance response with a slower response latency to peak may carry over effects to the next trial. In addition, emotion-modulated startle studies largely converge on the measurement and quantification of the startle response, whose setup is known to be relatively simple. In our meta-analysis on the emotion-modulated startle, all studies rectified and integrated the raw EMG signal with a time constant of 20ms, calculated the startle amplitude by subtracting a 20 or 50 milliseconds pre-startle baseline from the peak 20 – 120 or 20 – 150 milliseconds after startle probe onset and finally t- or z-transformed the mean amplitudes (Conzelmann et al., 2015; Dillon & LaBar, 2005; Golkar et al., 2014; Jackson et al., 2000).

In contrast, we observed tremendous variation in the quantification of the autonomic indices. For example, studies on skin conductance level during reappraisal assessed baseline activity during a neutral condition that included the presentation of neutral stimuli (Lohani & Isaacowitz, 2014; Wolgast et al., 2011), right before stimulus onset (e.g., Shiota & Levenson, 2009), right before instruction (Opitz et al., 2014), after instruction (Urry et al., 2009), or reported no baseline assessment (Goldin et al., 2019). These studies then either subtracted mean activity of the respective baseline from mean activity during the regulation period (e.g., Opitz et al., 2014; Shiota & Levenson, 2009), calculated raw means (Goldin et al., 2019), or area under the curve (Urry et al., 2009). It should be noted that these observations remain solely on a descriptive level. We did not conduct a moderator analysis to account for this variation since too few studies were available. Future studies would be helpful to corroborate our considerations.

The meta-analyses we presented in this article suggest that electromyographic measures such as the emotion-modulated startle might be robust options to assess emotion regulation effects, whereas autonomic measures might be context dependent and thus should be selected carefully. Autonomic measures are still important and interesting for emotion regulation research as they allow to track the extended reaction of the body to an emotional event or a series of events,

whereas the emotion-modulated startle is being assessed at one given time and thus does not allow to track the time-course of the regulation period.

2.5.1 Limitations and Future Research

While the present study represents the first meta-analysis of specific psychophysiological effects during distraction, reappraisal, suppression and instructions to choose a downregulation strategy, it is not without limitations. First of all, we emphasize that the number of available studies was small with the exception of heart rate and skin conductance level. In particular, most of the significant meta-analyses in the present study included few studies and these studies often stemmed from an even smaller number of labs (e.g., mean arterial pressure, finger temperature; see Figure 2.3). Thus, we need more research to test whether the effects would become insignificant with increasing number of independent studies. Similarly, absence of significance in meta-analyses with small number of samples should not be taken as evidence that there is no effect at all. Thus, studies that assess less common psychophysiological measures and emotion regulation instructions are urgently needed to increase knowledge about psychophysiological responses during emotion regulation.

Furthermore, no meta-analysis is free of a potential publication bias. The bias refers to the phenomenon that significant findings get published earlier and are more likely than non-significant findings. Statistical analyses indicated that there might be some publication bias, but this seemed not to appreciably impact the results. In addition, psychophysiological measures are usually not the primary outcome of emotion regulation studies, and many published studies have reported negative findings. Thus, we consider the publication bias to be relatively small in this review.

We also highlight the substantial variability in the research methodology used across the emotion regulation studies included in our meta-analysis. We explored the impact of methodological differences using several moderators (trial duration, nature of emotion induction, nature of control instruction, study design) and showed that central design aspects are explaining some differences in the overserved autonomic effect sizes. This raises the question to which degree the studies included in the present review are actually comparable.

Sample size was very small and conducting the meta-analyses and moderator analyses required a large number of separate analyses. In light of this, significant results presented here should be treated with caution as multiple comparisons might have increased the chances of false discovery. More research is needed to confirm our results. We also acknowledge that we assessed a limited sample of potential moderators. As mentioned above, there was tremendous variation

in the quantification of the autonomic indices, which we were not able to account for as there were too few studies available to conduct meaningful moderator analyses. Finally, we highlight that our meta-analysis was limited to the regulation of negative emotions only, mainly focusing on reappraisal and suppression.

In light of these limitations, we need particularly larger and more comparable studies with identical setup to control the moderator variables identified in this meta-analysis (in particular trial duration, comparable control conditions and the same study design). One important future direction for researchers in the area of psychophysiological response patterns to emotion regulation is to design large-scale, comprehensive studies that directly compare psychophysiological measures and emotion regulation strategies ideally using the same assessment and quantification of psychophysiological responses.

With psychophysiological recordings we cannot control which regulation strategies are really being applied by participants. The variability of autonomic responding across different emotion regulation contexts further complicates an accurate interpretation of effects and may be particularly problematic in studies focusing on just one psychophysiological outcome measure. Experiments using simultaneous recordings from multiple psychophysiological channels would be helpful to e.g. identify potential response patterns uniquely characterizing different emotion regulation strategies (e.g., pupil, heart rate, skin conductance, etc.). However, major progress is unlikely without coordinated effort across labs to systematically address these questions.

There is also a need for studies that carefully tease apart attention, arousal and other cognitive processes that may influence autonomic responses in order to gain a better understanding of the interpretation of autonomic responses during emotion regulation. Systematic variations in different experimental setups may help to dissociate the underlying cognitive and emotional processes that cause autonomic activity in order to draw clear inferences.

2.5.2 Conclusion

This meta-analysis represents the first attempt to determine the mean effects of different emotion regulation strategies on individual psychophysiological measures. Our results indicate that a) effects of reappraisal decreased heart rate and corrugator activity, whereas suppression increased sympathetic arousal but decreased respiration amplitude and mean arterial pressure, b) effects of autonomic measures, even if significant, were small and heterogeneous across studies, while electromyographic measures showed medium effect sizes and c) the study design,

control instruction and trial duration moderated some but not all effect sizes. As available studies were few, our findings remain preliminary. In order to use meta-analyses to compare effects of psychophysiological responses in different regulation contexts, more comparable.

2.6 Supplementary Material

2.6.1 Electrodermal Activity

Boucsein et al. (2012) define phasic electrodermal responses as short-lasting changes in electrodermal activity most commonly reported as a peak amplitude (including non-zero responses) or magnitude (including only responses above a defined minimum) after a specific stimulus. Phasic skin conductance responses are thus only a small fraction of the skin conductance level. Tonic skin conductance level refers to the slowly changing level of skin conductance "typically computed as a mean of several measurements taken during a specific time period." (Boucsein et al., 2012, p. 1026). In addition, non-specific skin conductance responses may be assessed during periods of nonstimulation and they are commonly computed as number of skin conductance responses per unit of time (nSCR). The inspection of included studies in the present review however showed that the distinction between short-lasting changes and slow-changing level in skin conductance are rather fluent in emotion regulation studies. In the studies included in our analyses, skin conductance levels (i.e. average values) were assessed over trials lasting between 4 and 590s seconds, whereas skin conductance responses were assessed over trial durations lasting between 5 and 1800 seconds. To overcome this problem, we created a taxonomy to divide skin conductance level and skin conductance response with definitions adapted to the emotion regulation literature. The taxonomy and a full list of studies with detailed explanation why we categorized them as either "skin conductance response" or "skin conductance level" can be found in Table 2.10. In brief, we distinguished between indices reflecting the maximum amplitude and indices reflecting the duration of the skin conductance signal (i.e. averages). In particular, we defined skin conductance responses as maximum amplitude, magnitude, or peak occurring in a particular time after stimulus onset. These amplitudes may be baseline corrected and averaged over time. Non-specific skin conductance responses occurring during longer periods of time were also regarded as "skin conductance responses", if they were reported as maximum amplitudes. We defined skin conductance level as the mean skin conductance over a specific period of time that can range between several seconds and several minutes. If skin conductance was calculated as the area under the curve or the integrated signal over a period of

time, we also defined these measures as skin conductance level, since they are also affected by the duration of the response.

2.6.2 Forest Plots

2.6.2.1 Distraction Strategies

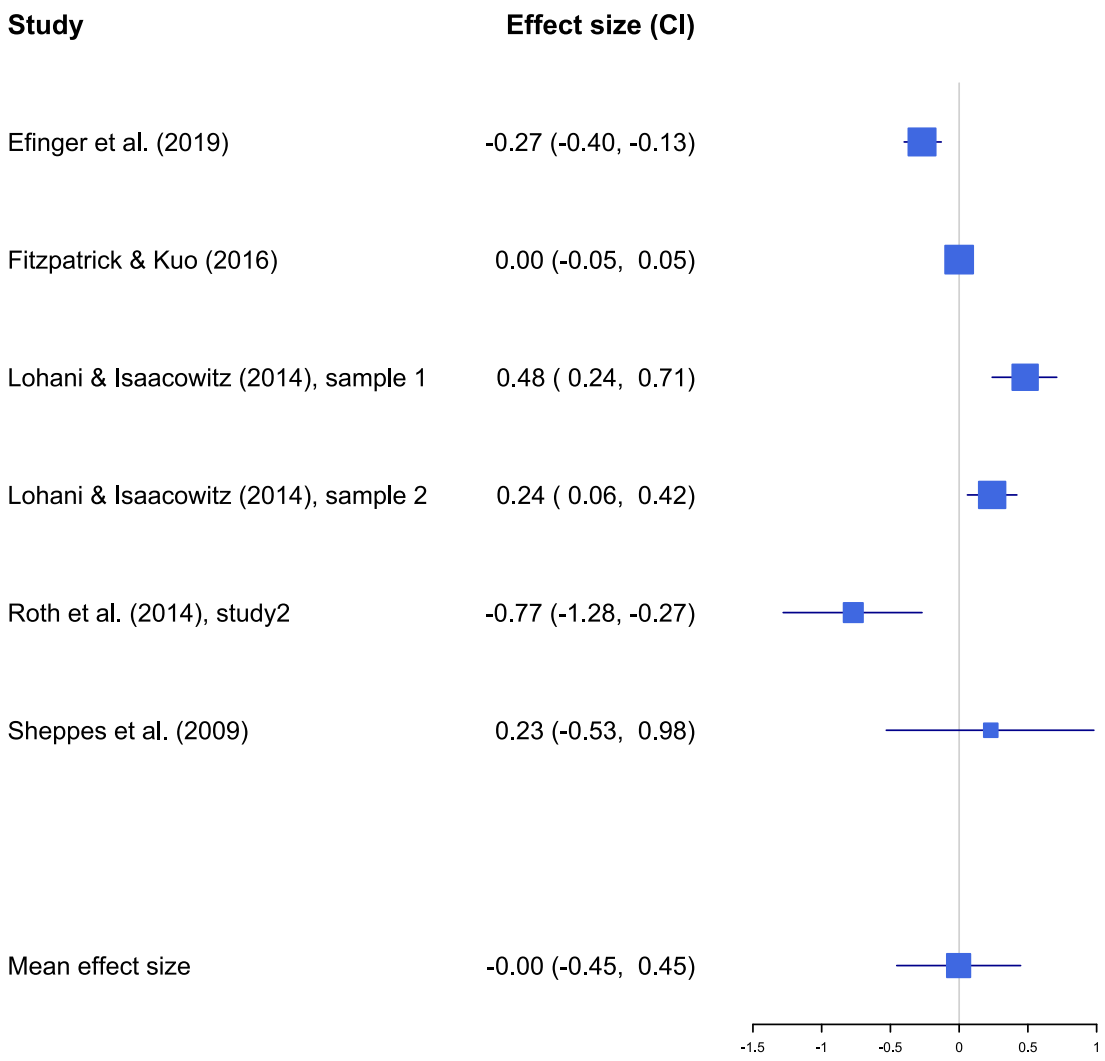

Figure 2.4. Statistics and results from the meta-analysis on skin conductance level (SCL) during distraction. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

2.6.2.2 Reappraisal Strategies

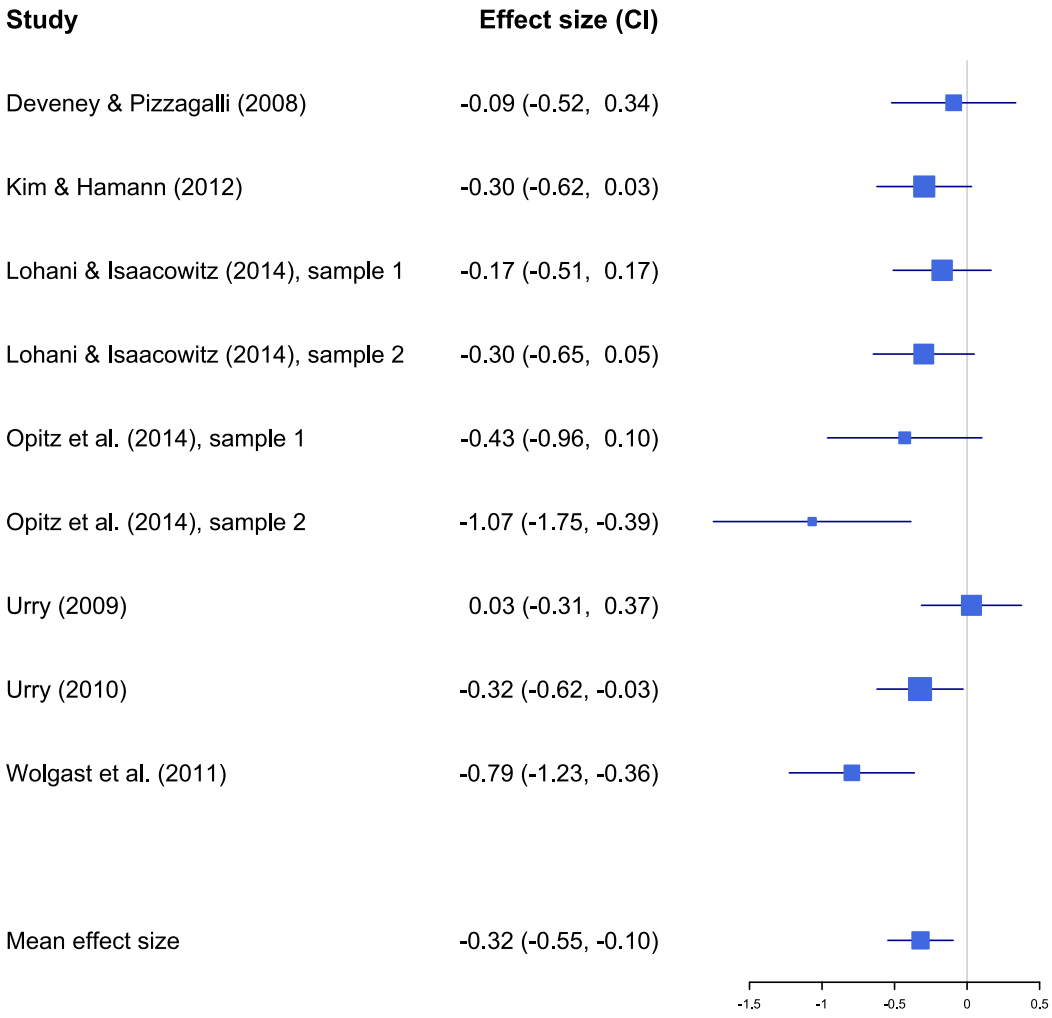

Figure 2.5. Statistics and results from the meta-analysis on corrugator activity (cEMG) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

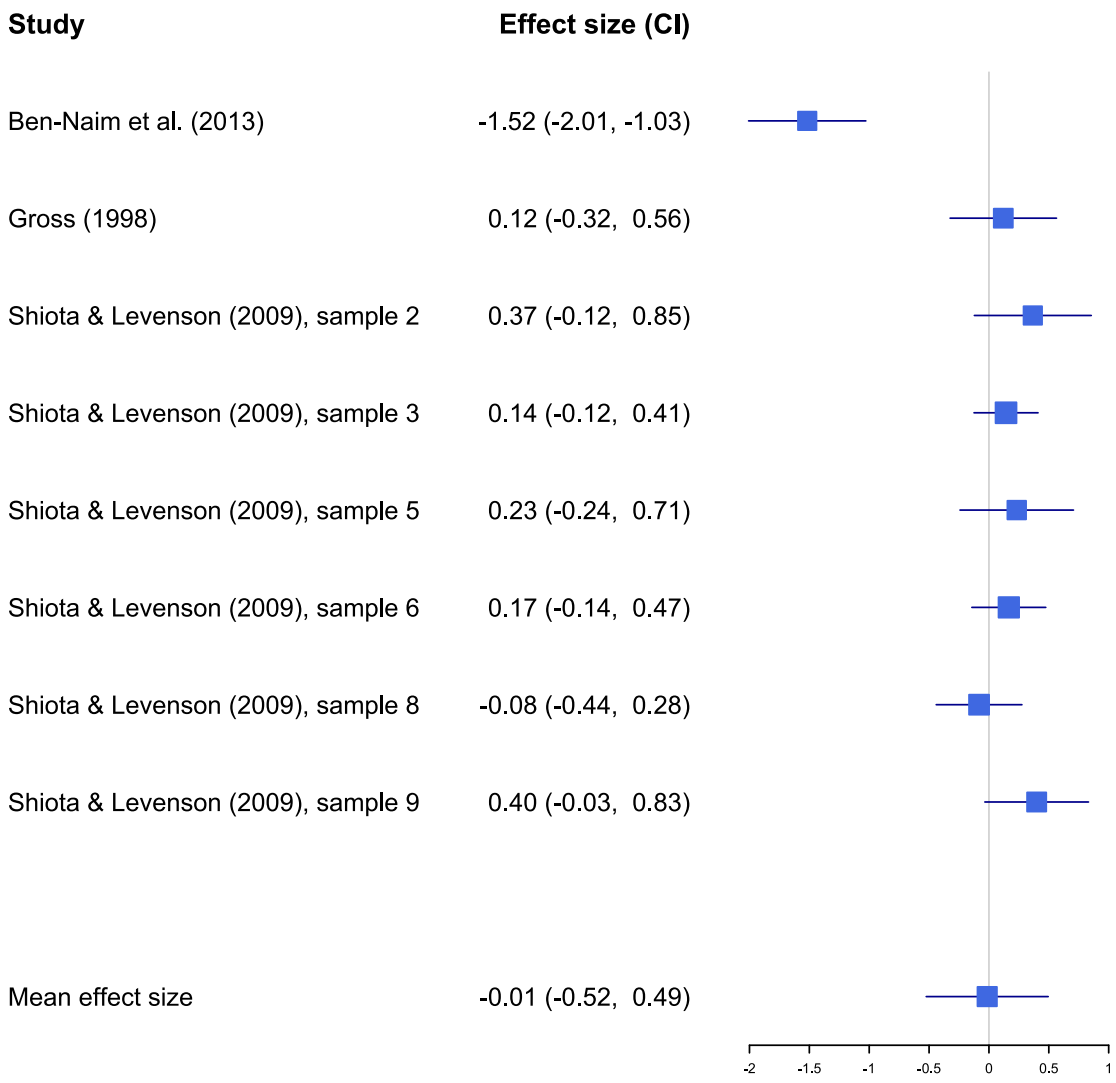

Figure 2.6. Statistics and results from the meta-analysis on finger pulse amplitude (FPA) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

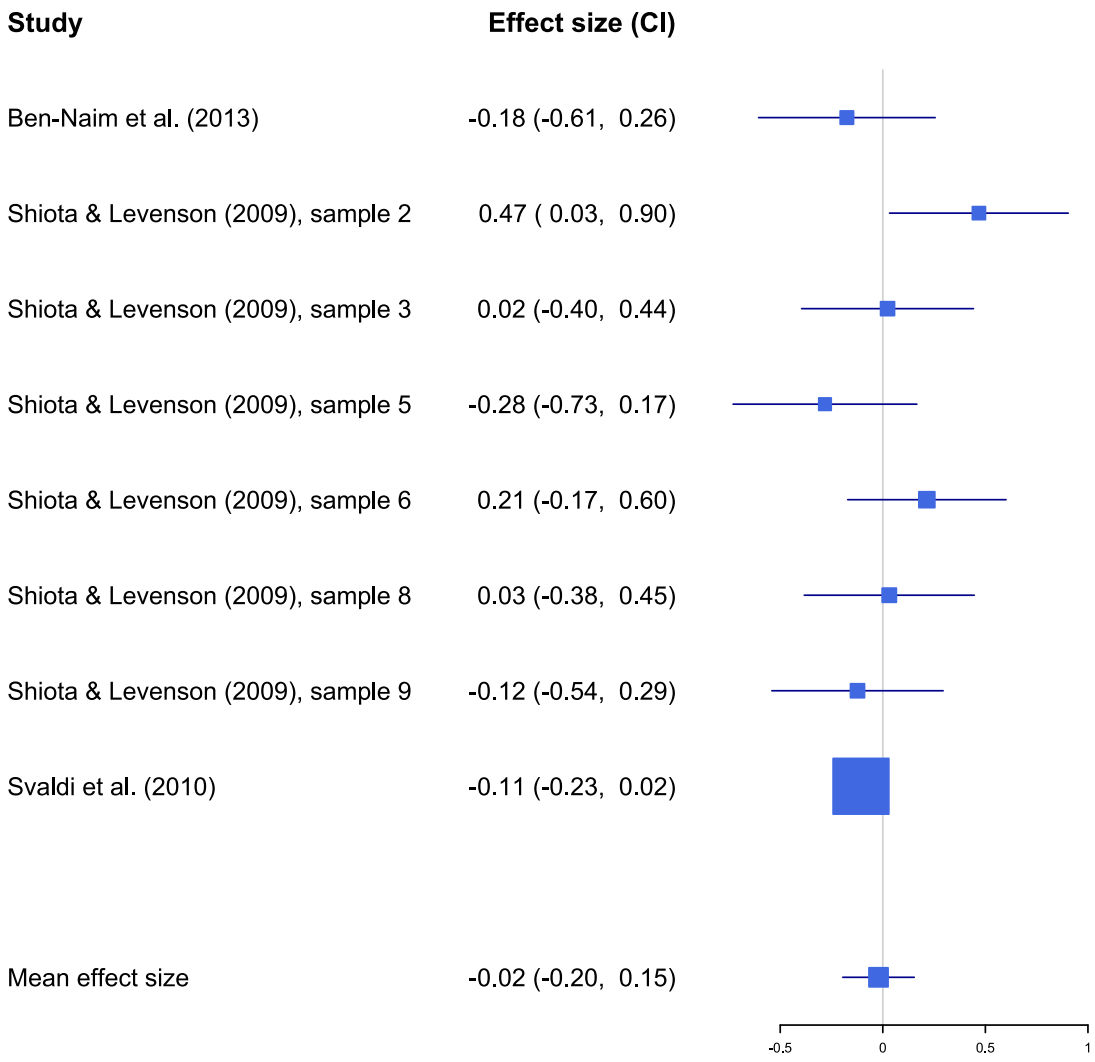

Figure 2.7. Statistics and results from the meta-analysis on finger pulse transit time (FPTT) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

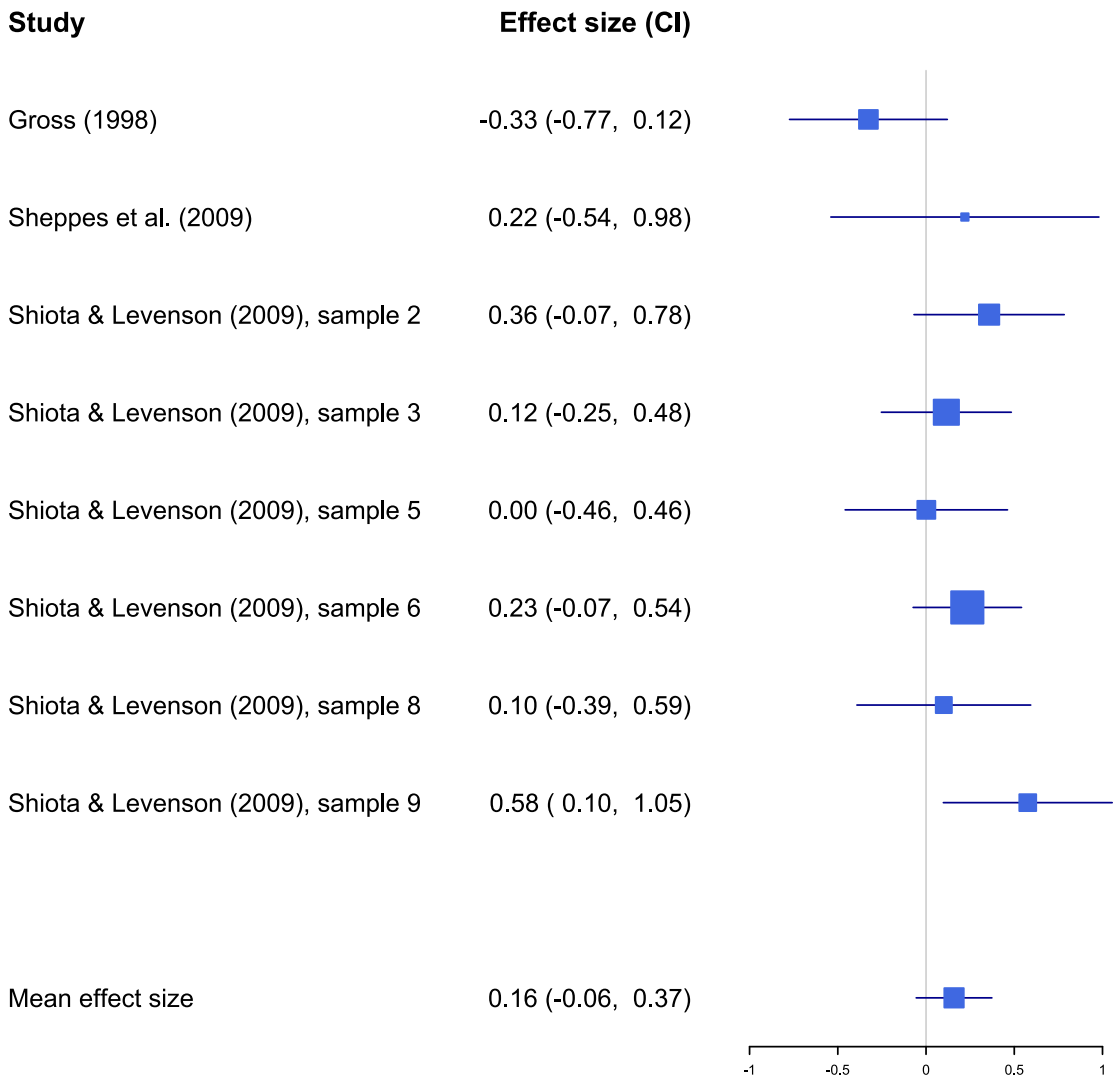

Figure 2.8. Statistics and results from the meta-analysis on finger temperature (FT) during re-appraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

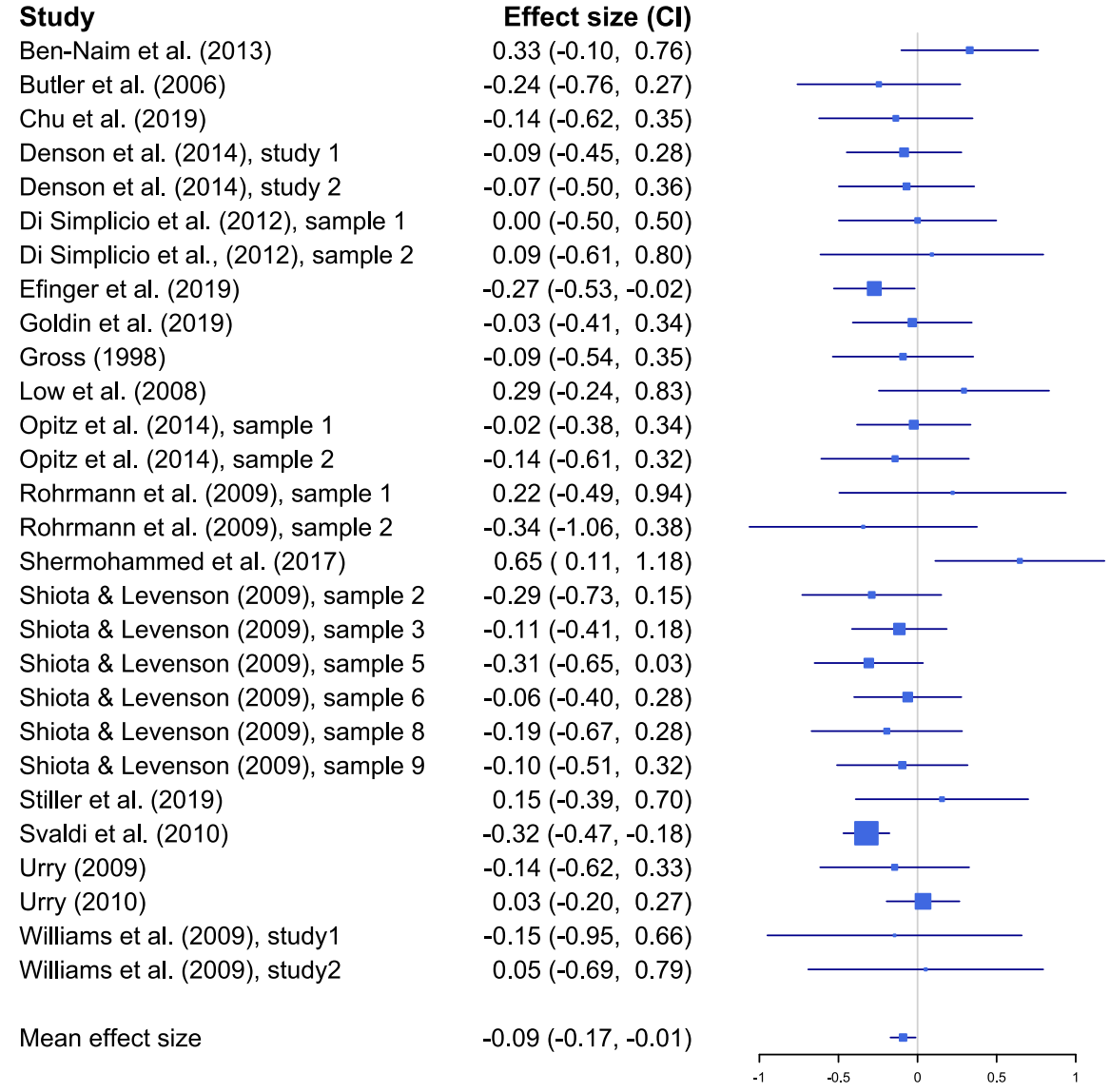

Figure 2.9. Statistics and results from the meta-analysis on heart rate (HR) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

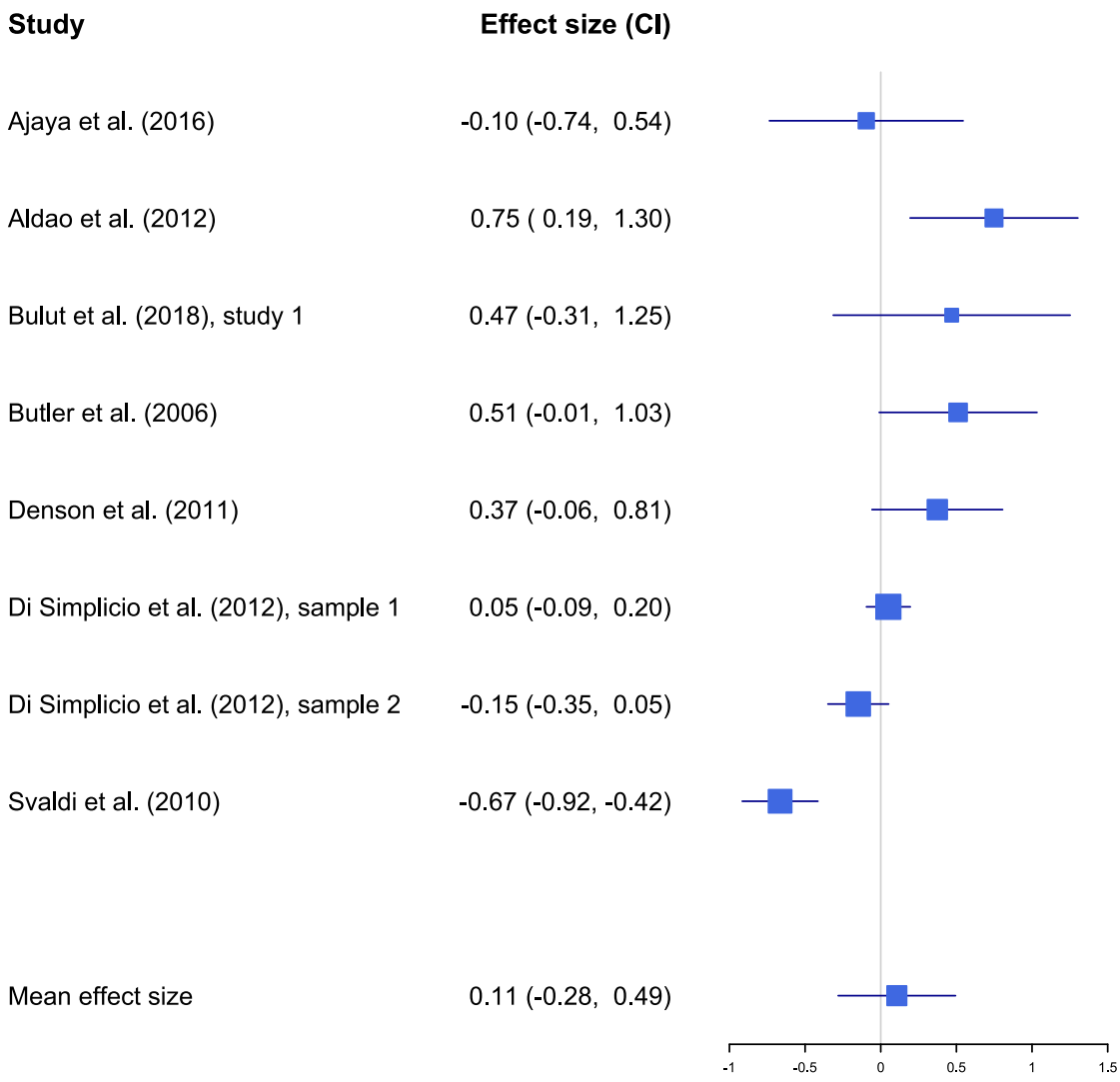

Figure 2.10. Statistics and results from the meta-analysis on heart rate variability (HRV) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

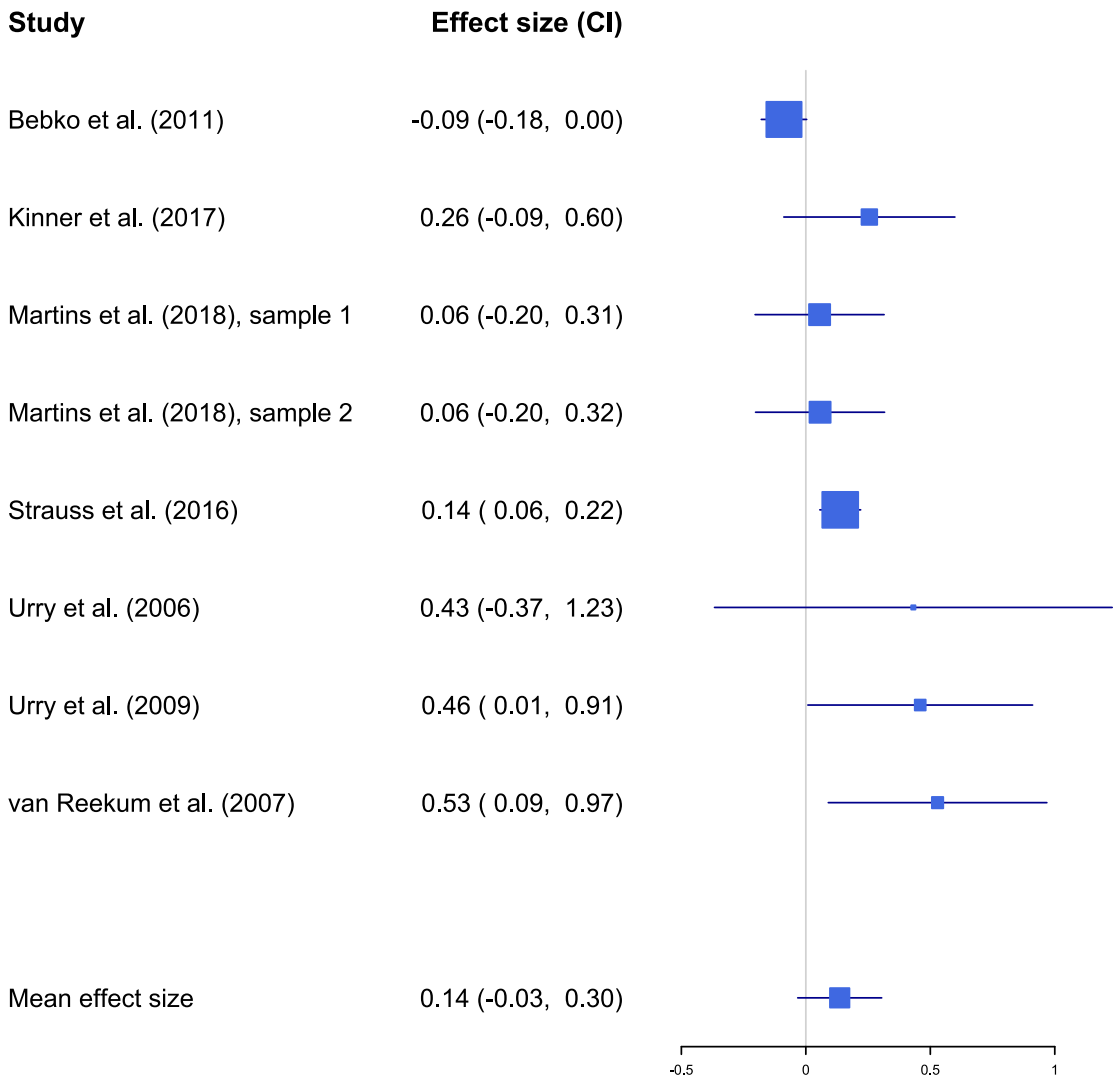

Figure 2.11. Statistics and results from the meta-analysis on pupil dilation (PD) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

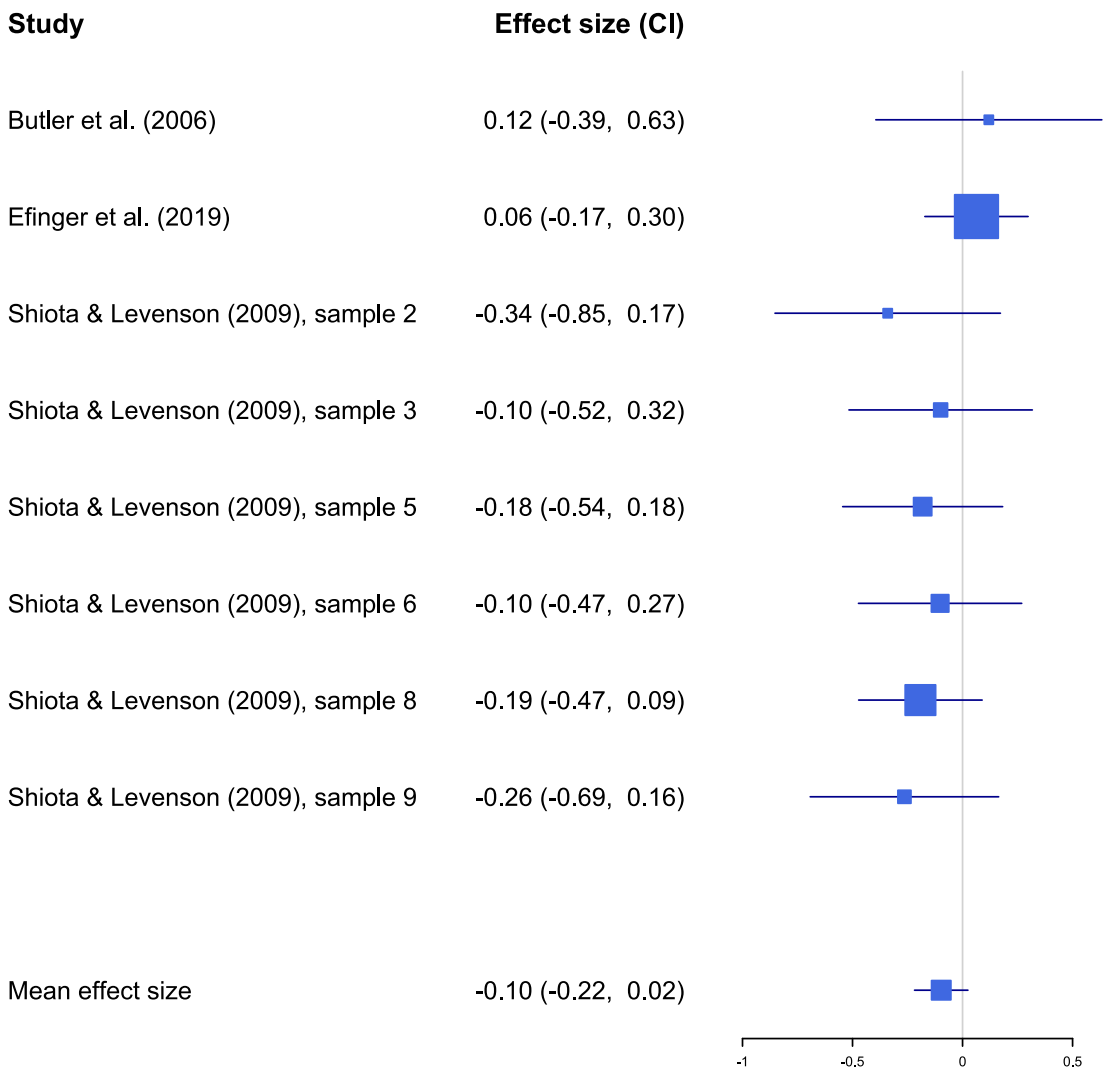

Figure 2.12. Statistics and results from the meta-analysis on respiration amplitude (RA) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

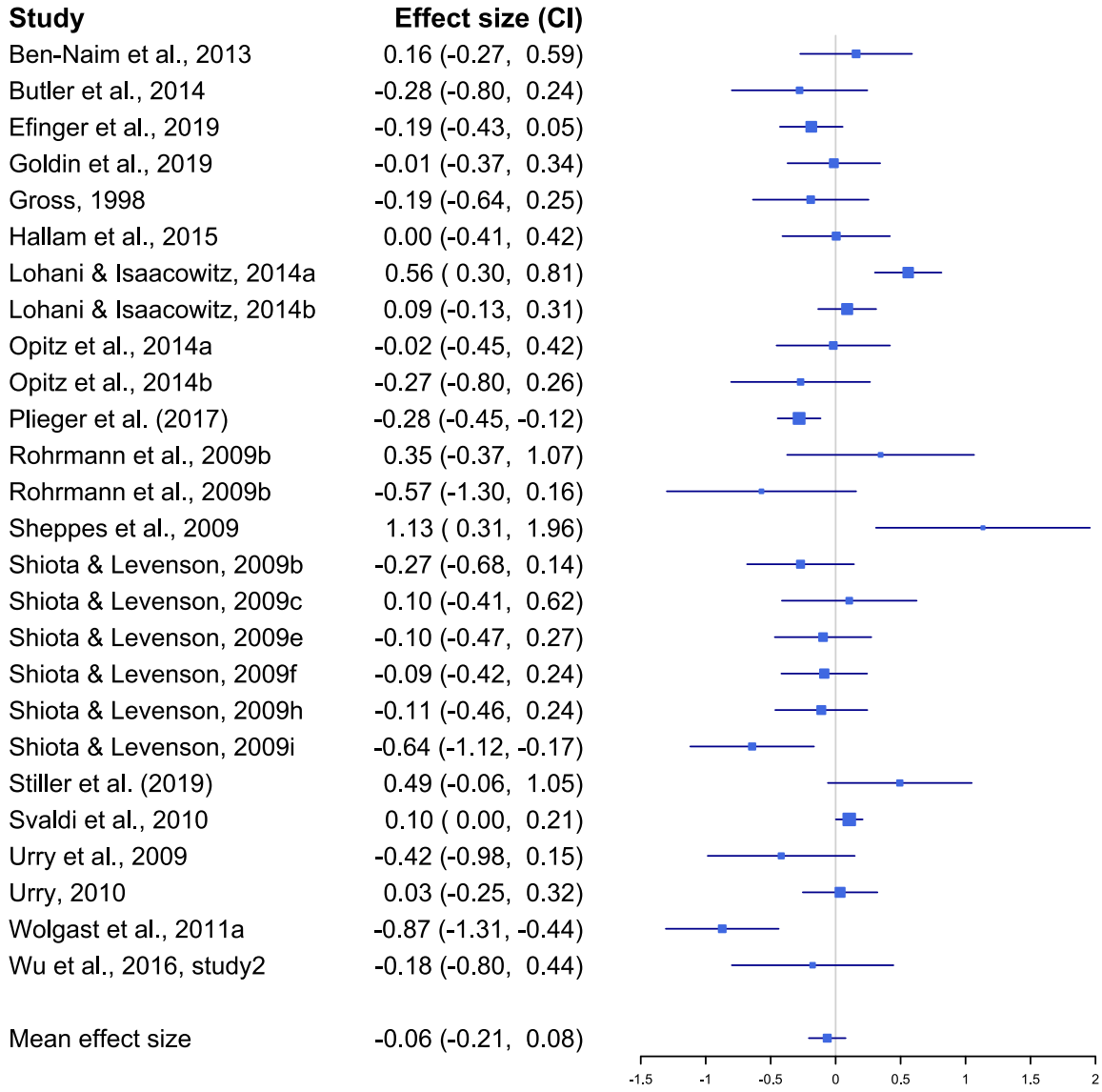

Figure 2.13. Statistics and results from the meta-analysis on skin conductance level (SCL) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

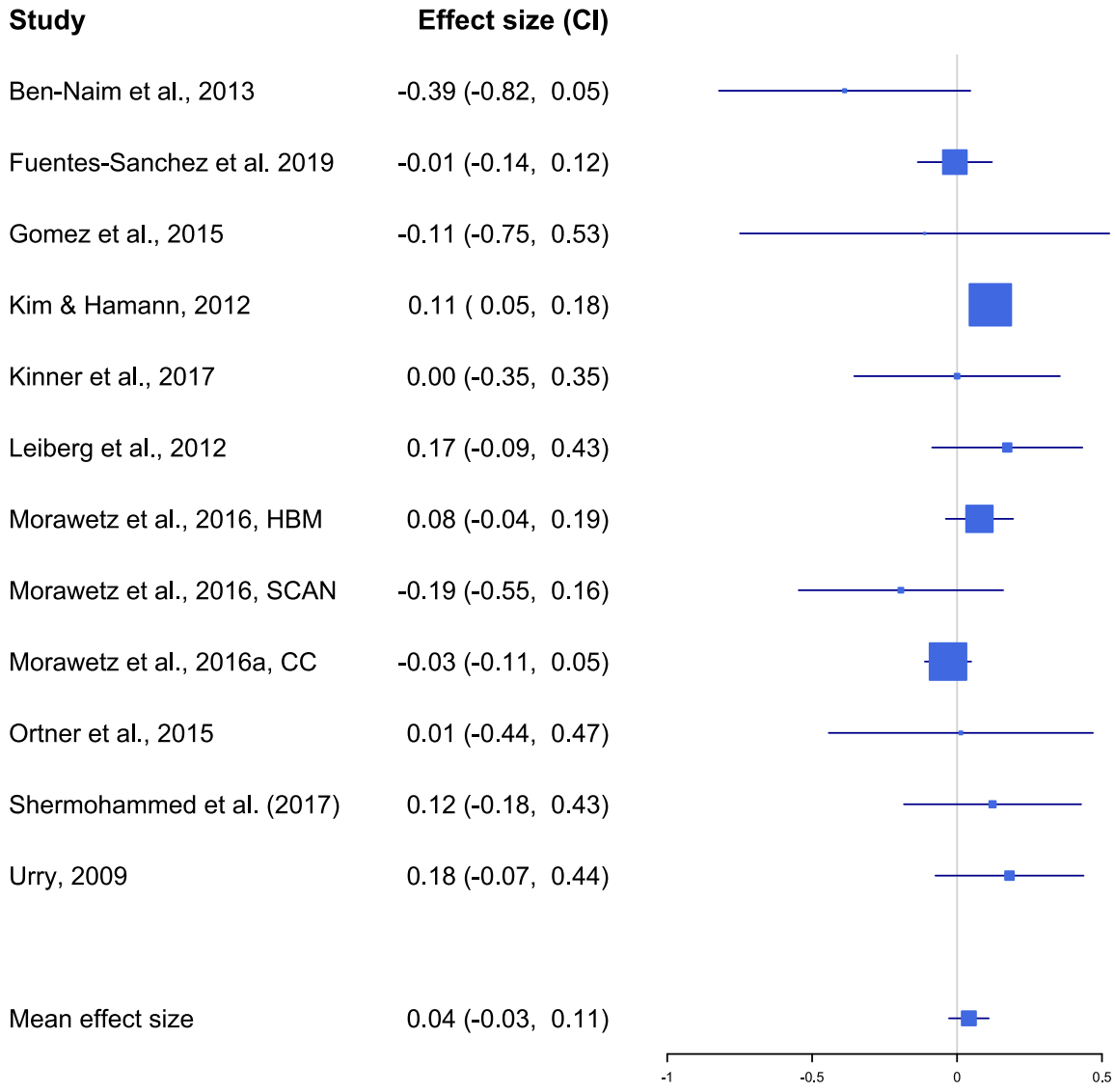

Figure 2.14. Statistics and results from the meta-analysis on skin conductance response (SCR) during reappraisal. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

2.6.2.3 Suppression Strategies

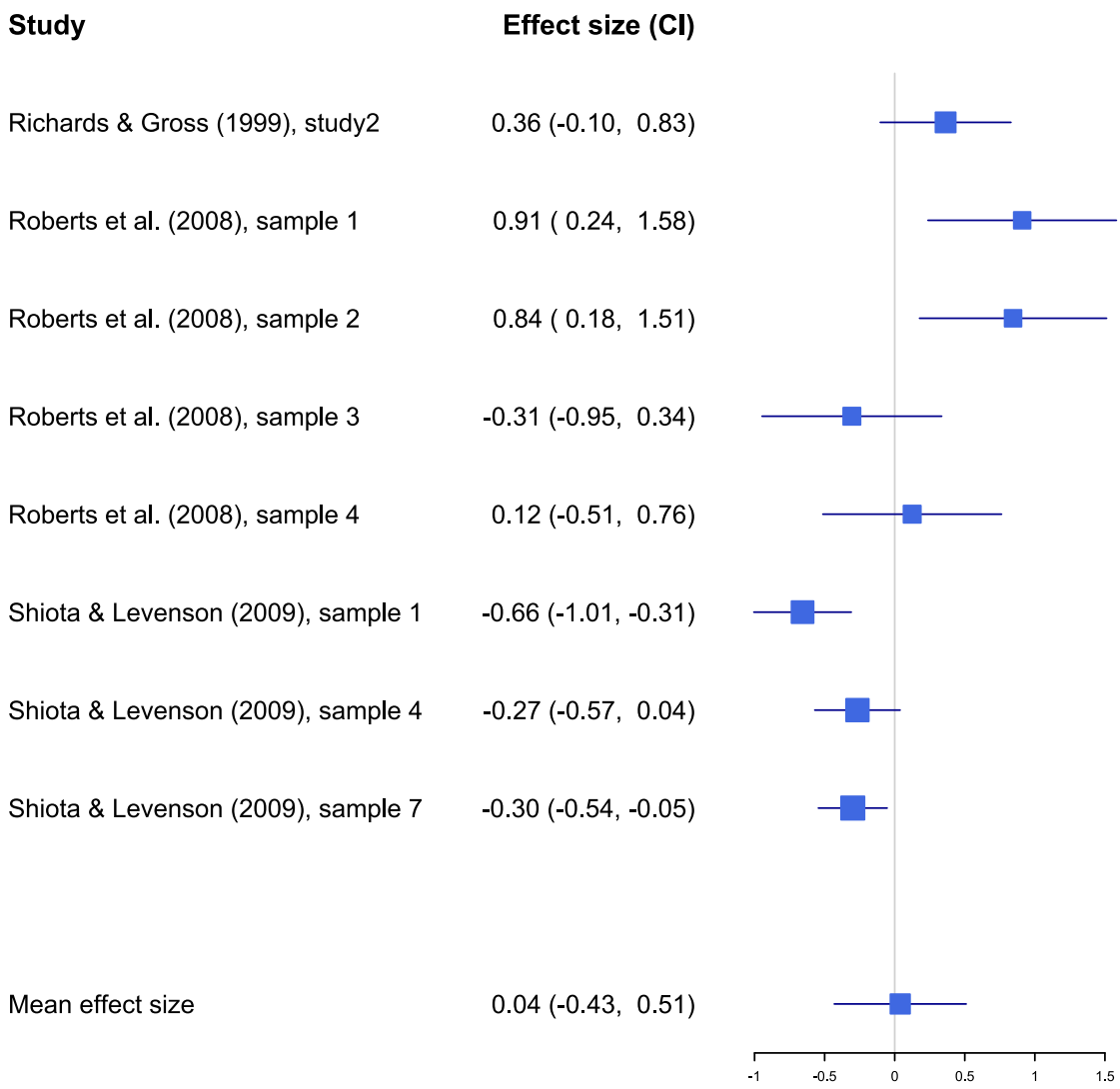

Figure 2.15. Statistics and results from the meta-analysis on diastolic blood pressure (DBP) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

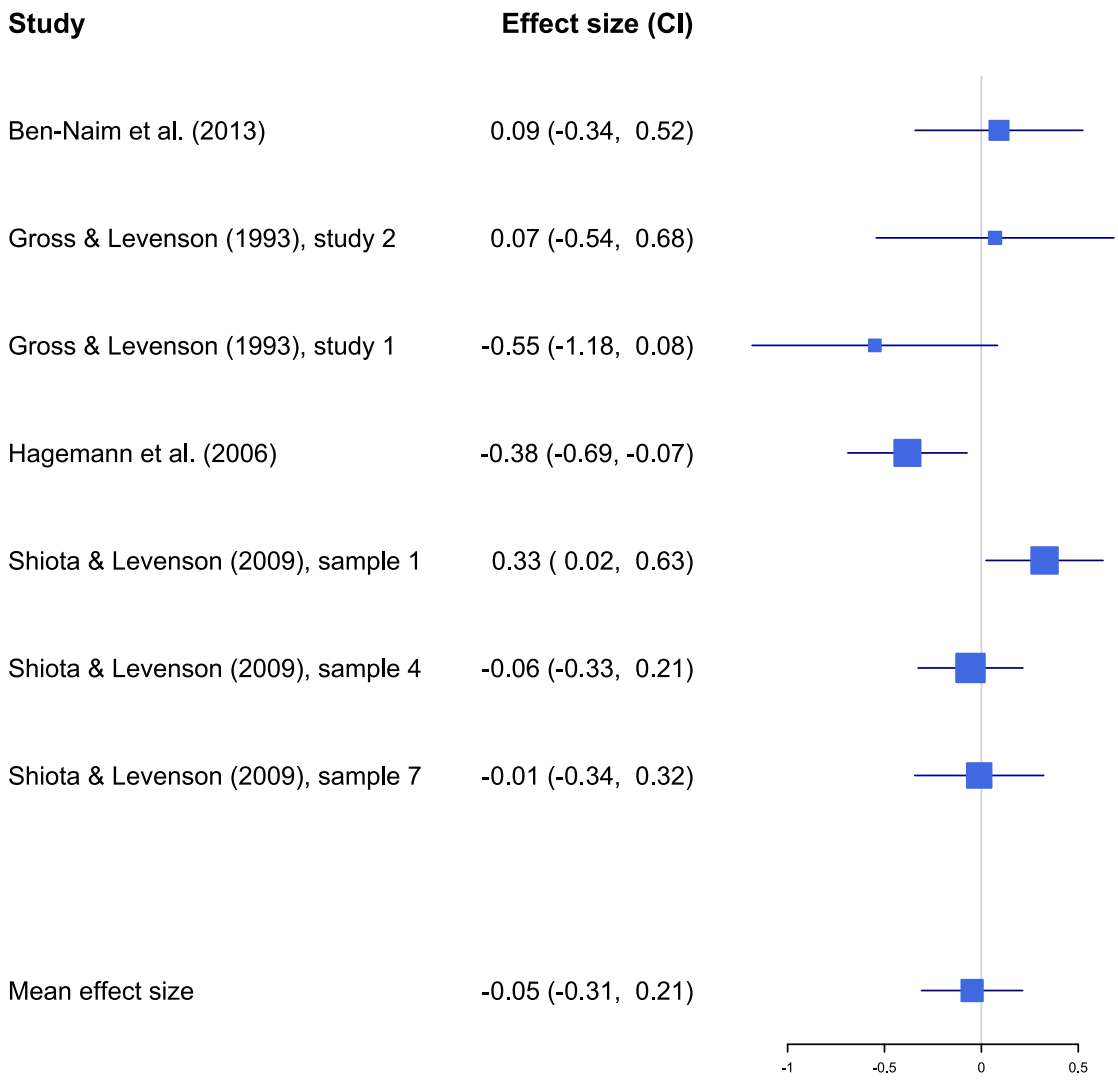

Figure 2.16. Statistics and results from the meta-analysis on ear pulse transit time (EPTT) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

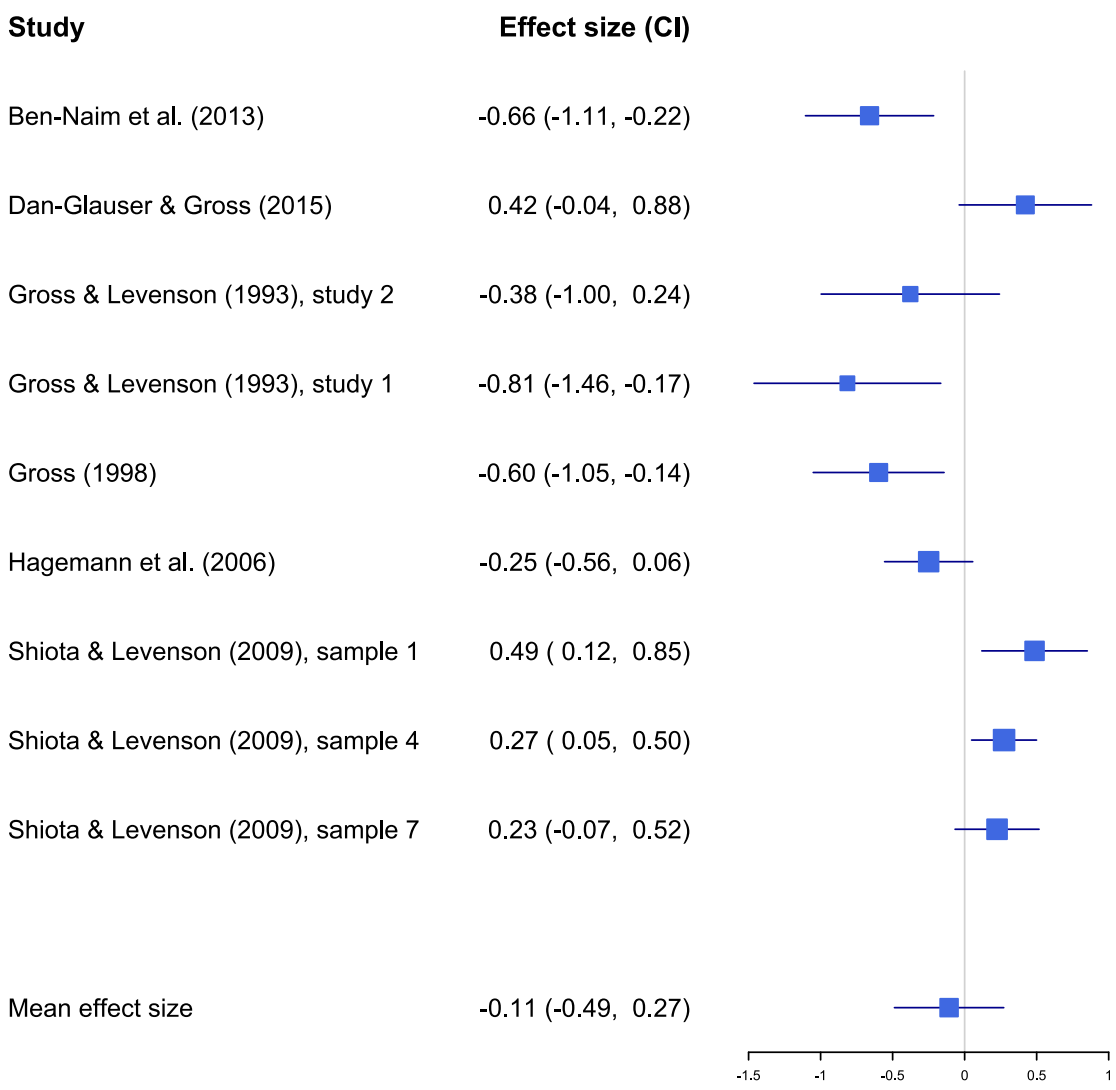

Figure 2.17. Statistics and results from the meta-analysis on finger pulse amplitude (FPA) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

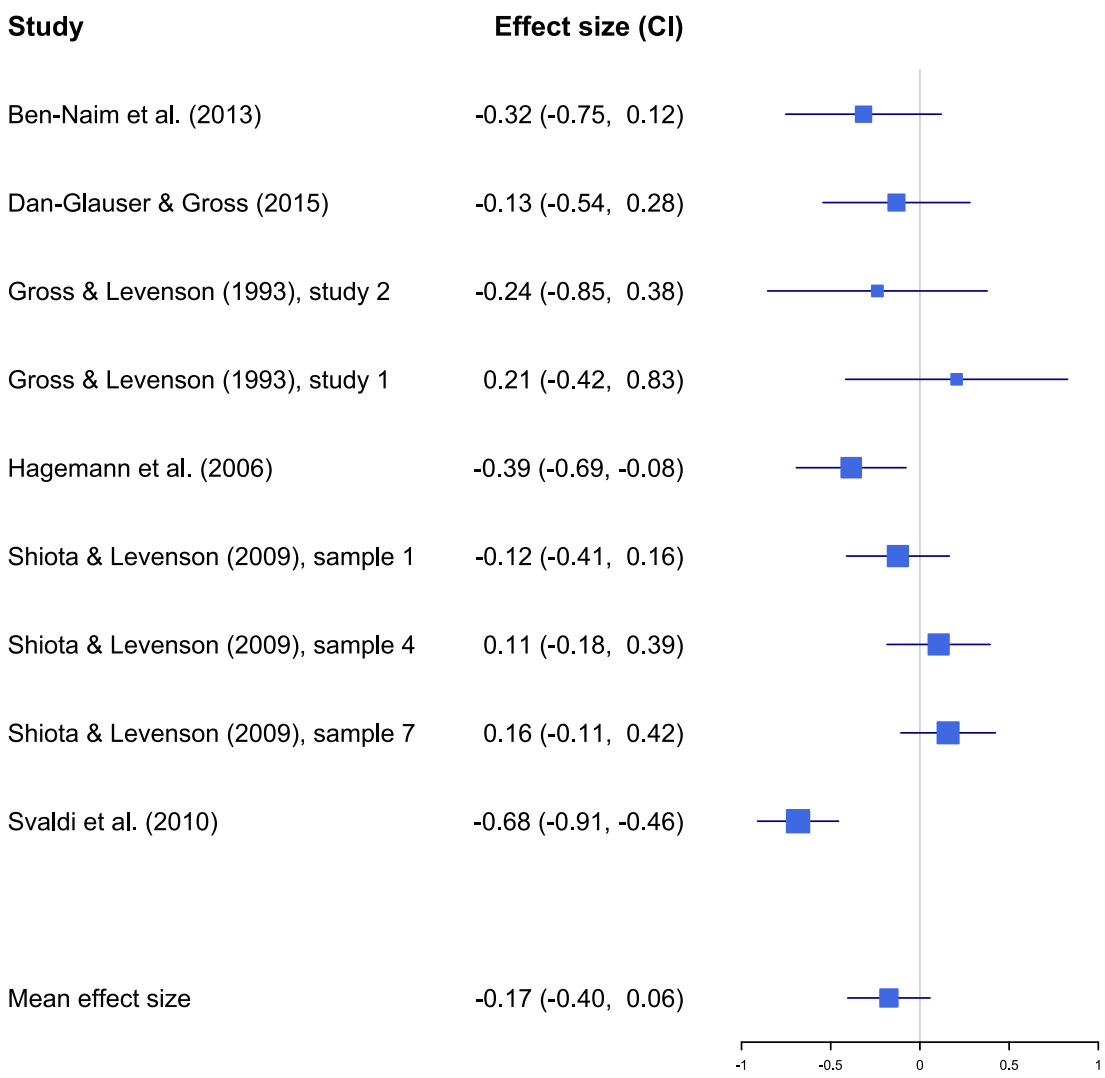

Figure 2.18. Statistics and results from the meta-analysis on finger pulse transit time (FPTT) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

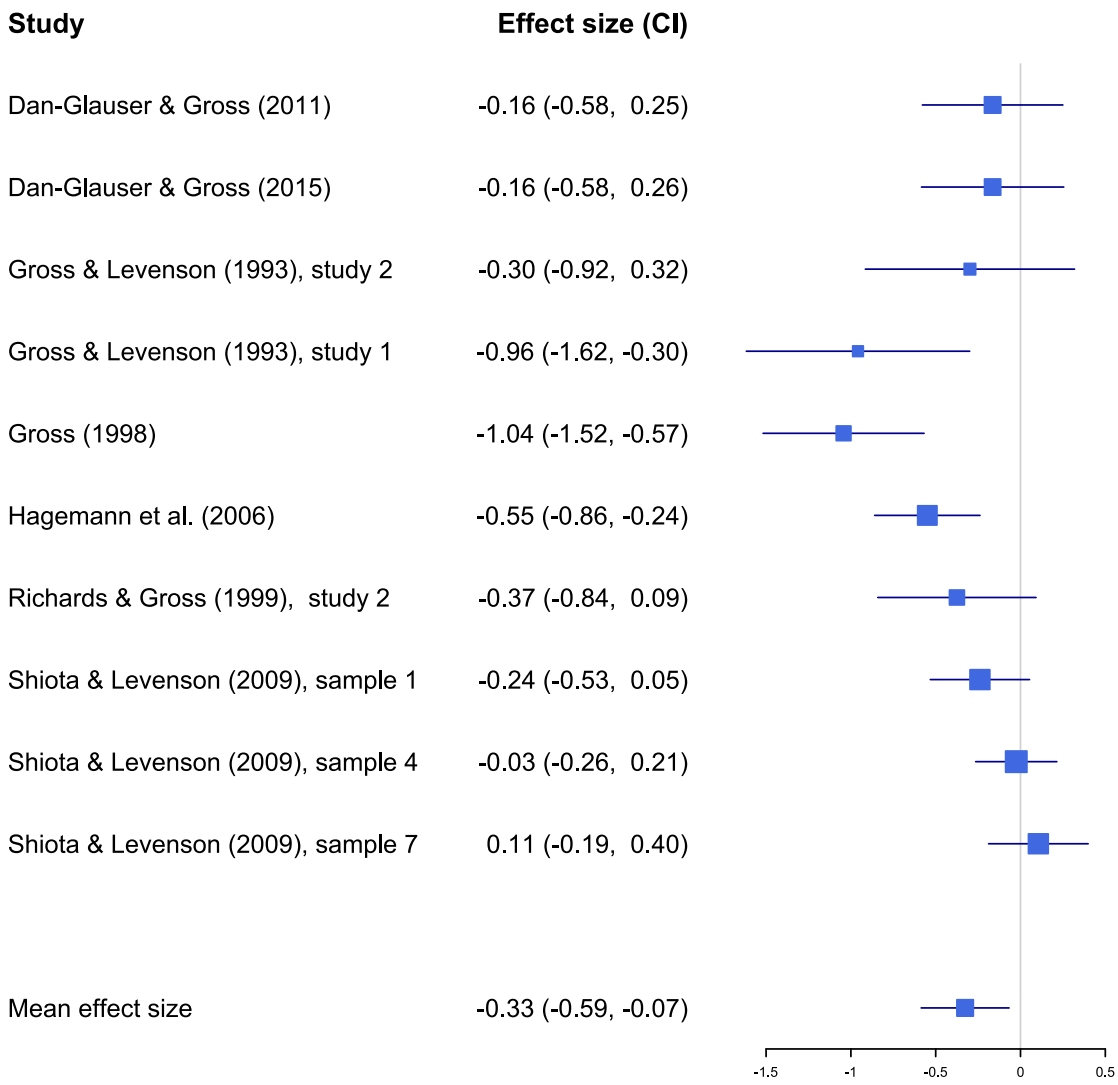

Figure 2.19. Statistics and results from the meta-analysis on finger temperature (FT) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

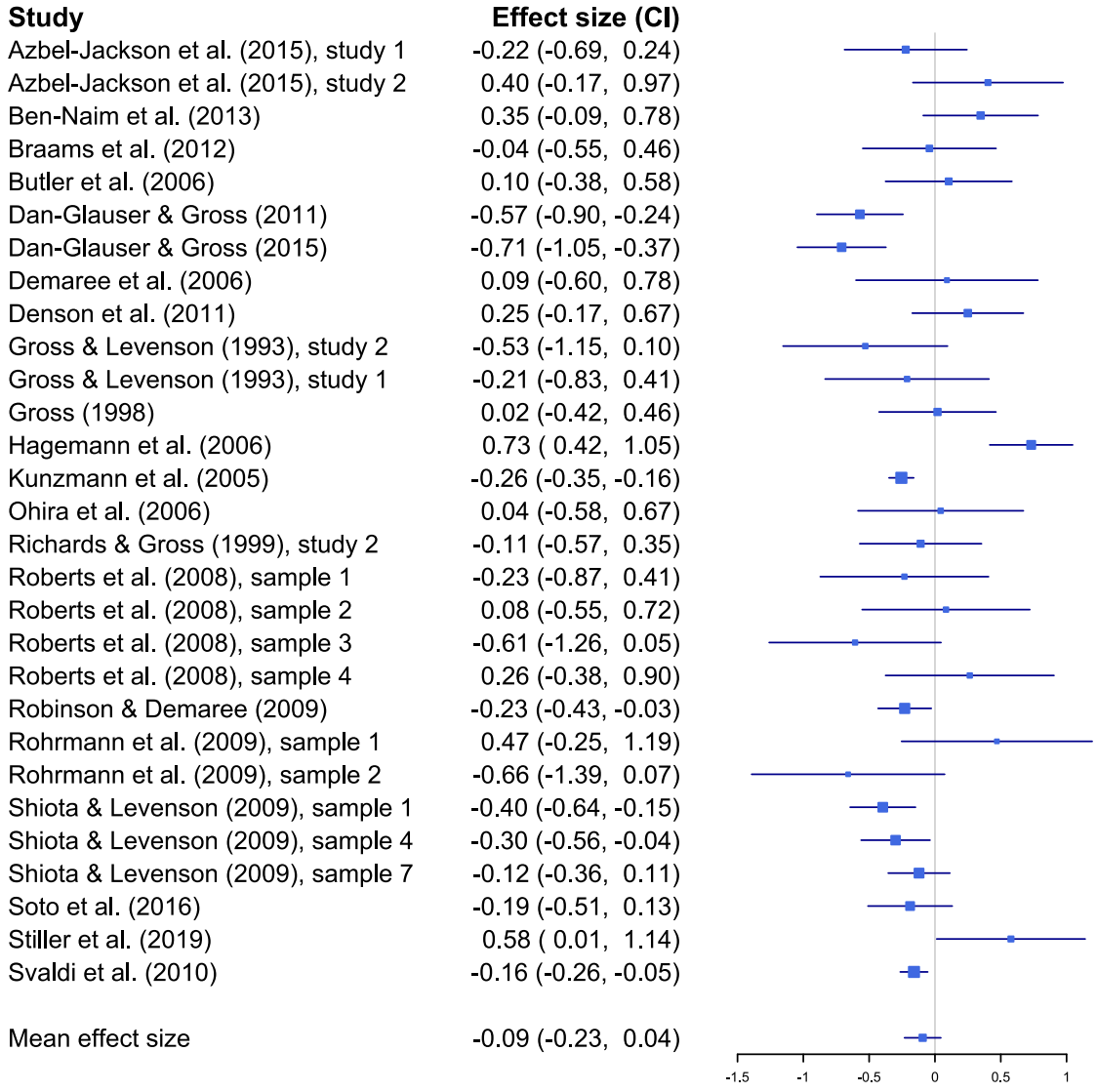

Figure 2.20. Statistics and results from the meta-analysis on heart rate (HR) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

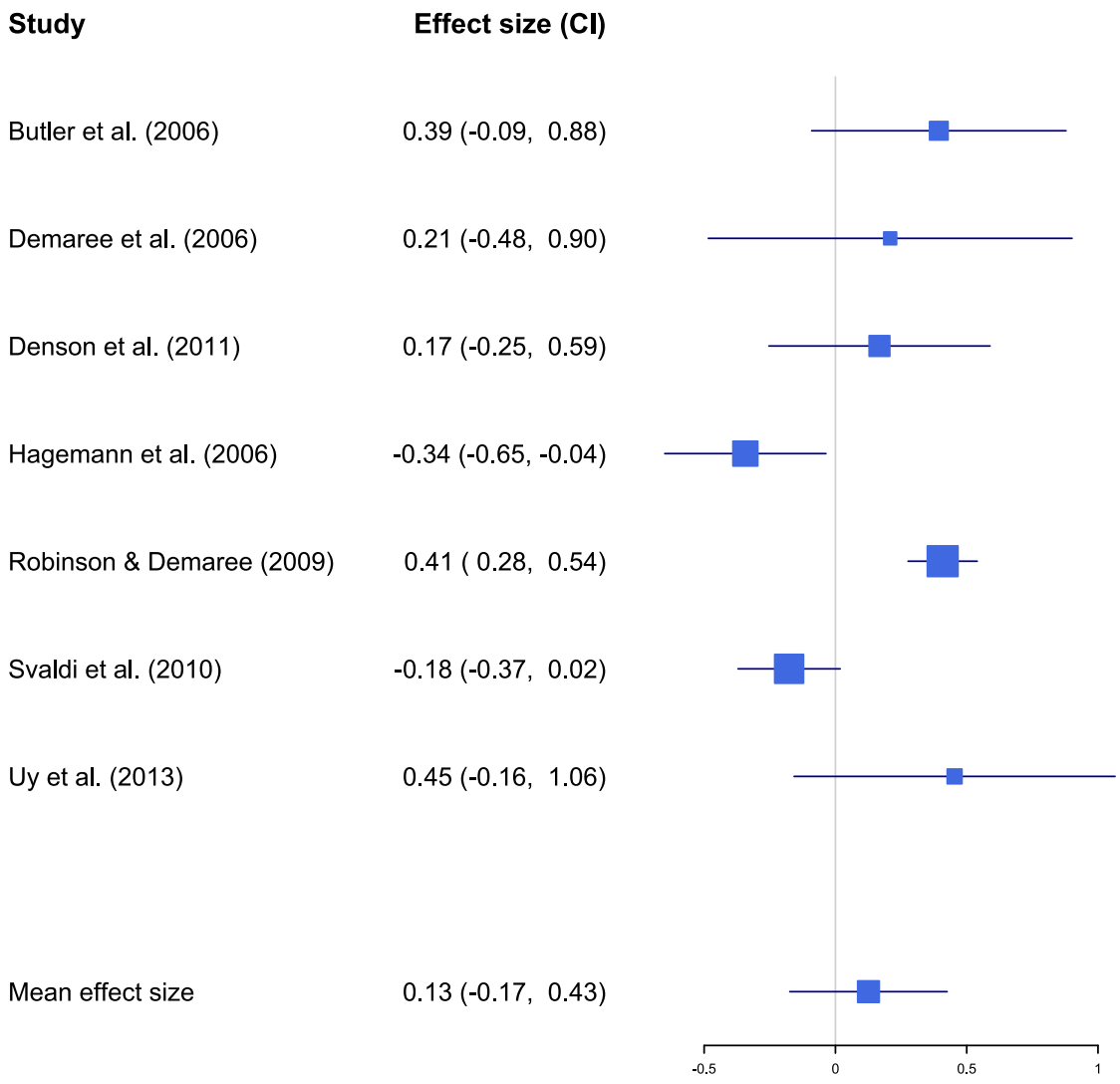

Figure 2.21. Statistics and results from the meta-analysis on heart rate variability (HRV) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

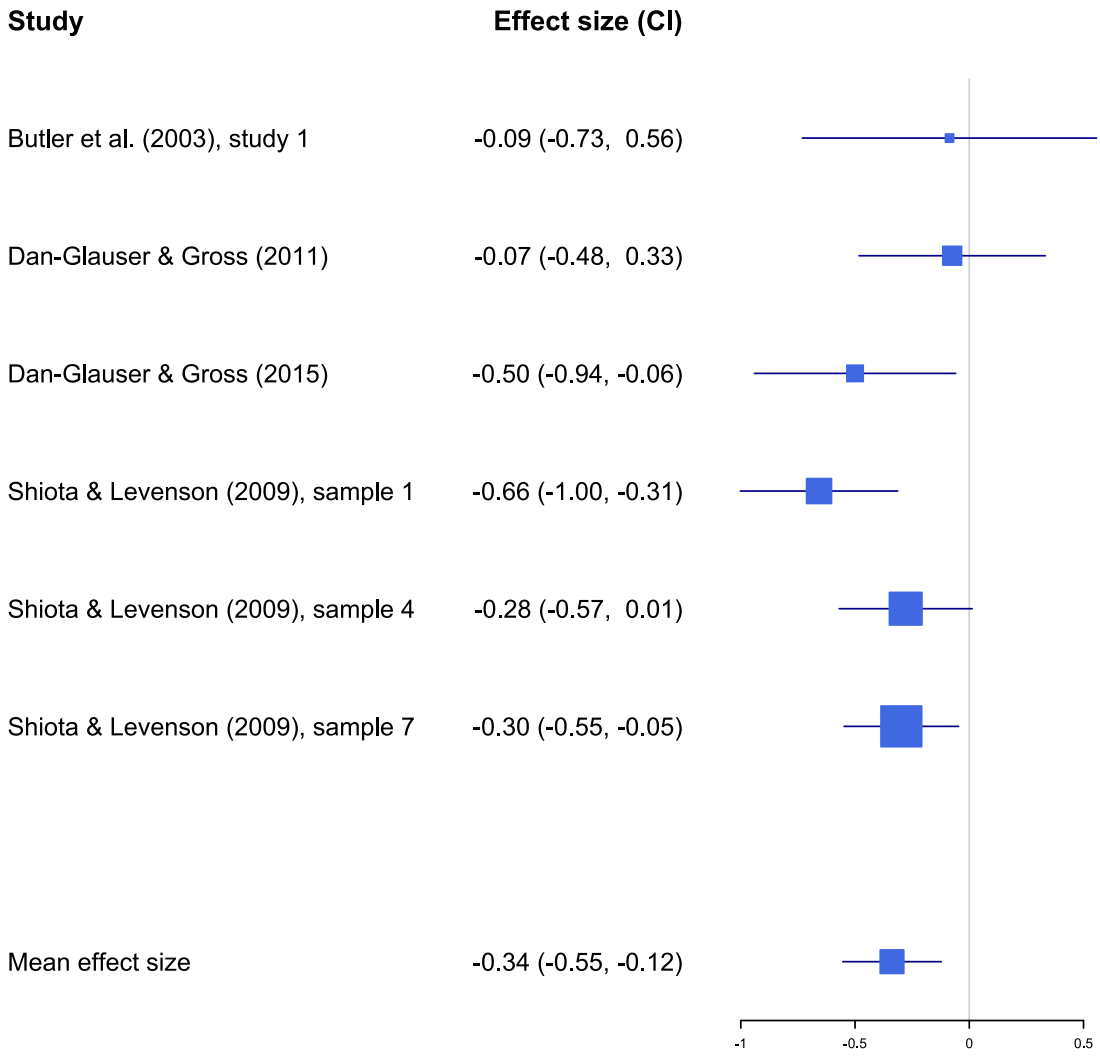

Figure 2.22. Statistics and results from the meta-analysis on mean arterial pressure (MAP) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

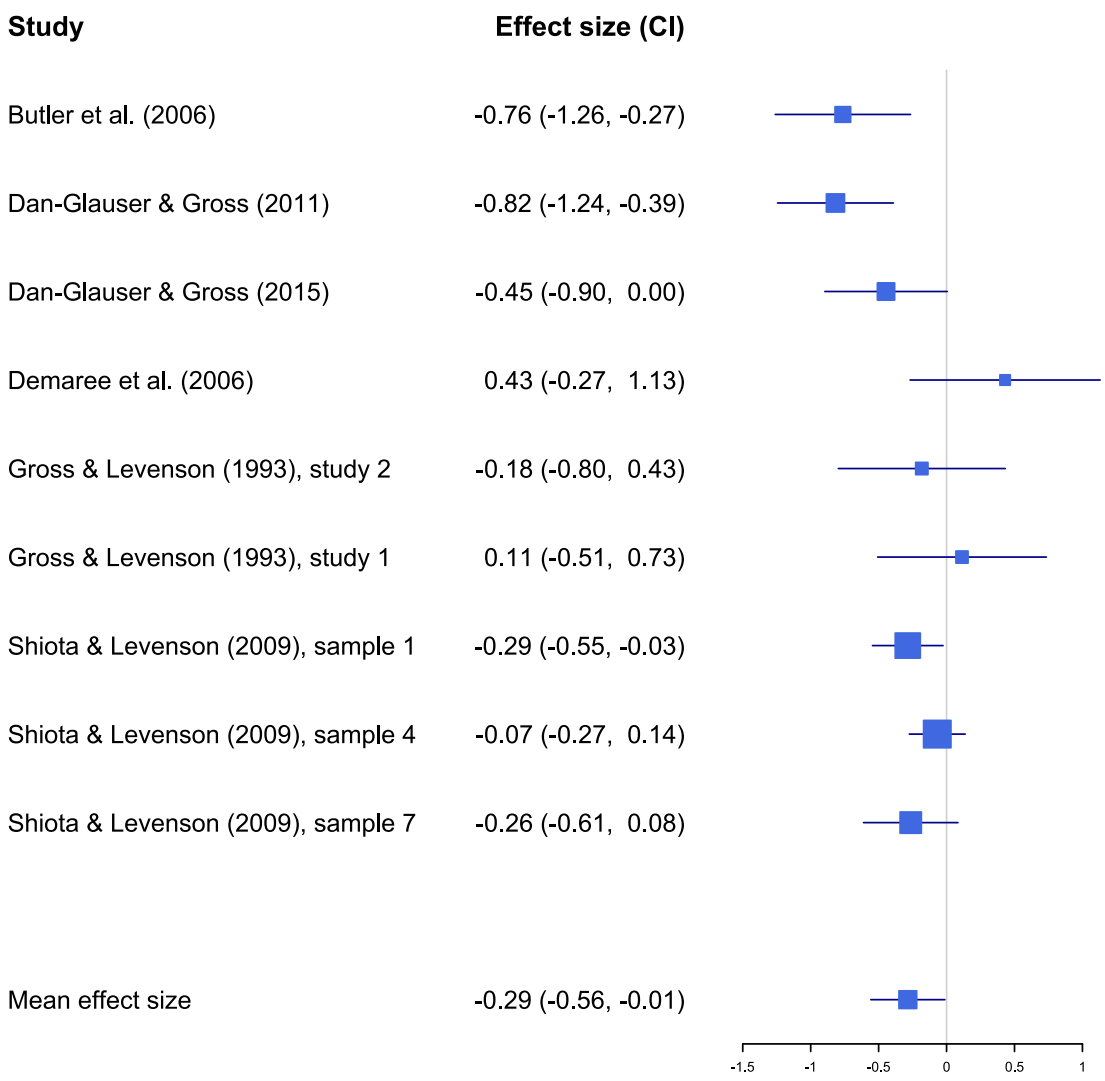

Figure 2.23. Statistics and results from the meta-analysis on respiration amplitude (RA) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

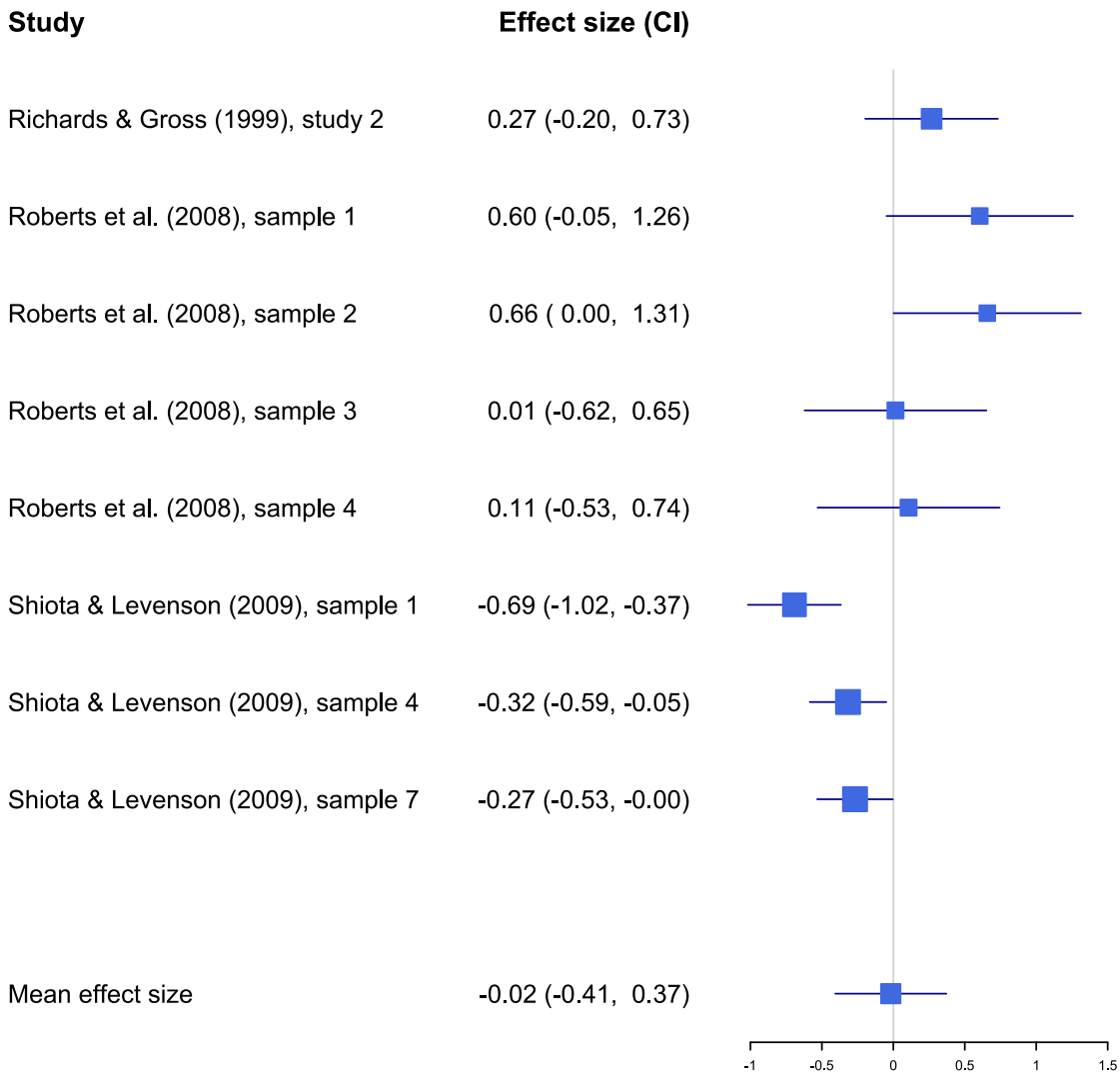

Figure 2.24. Statistics and results from the meta-analysis on systolic blood pressure (SBP) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

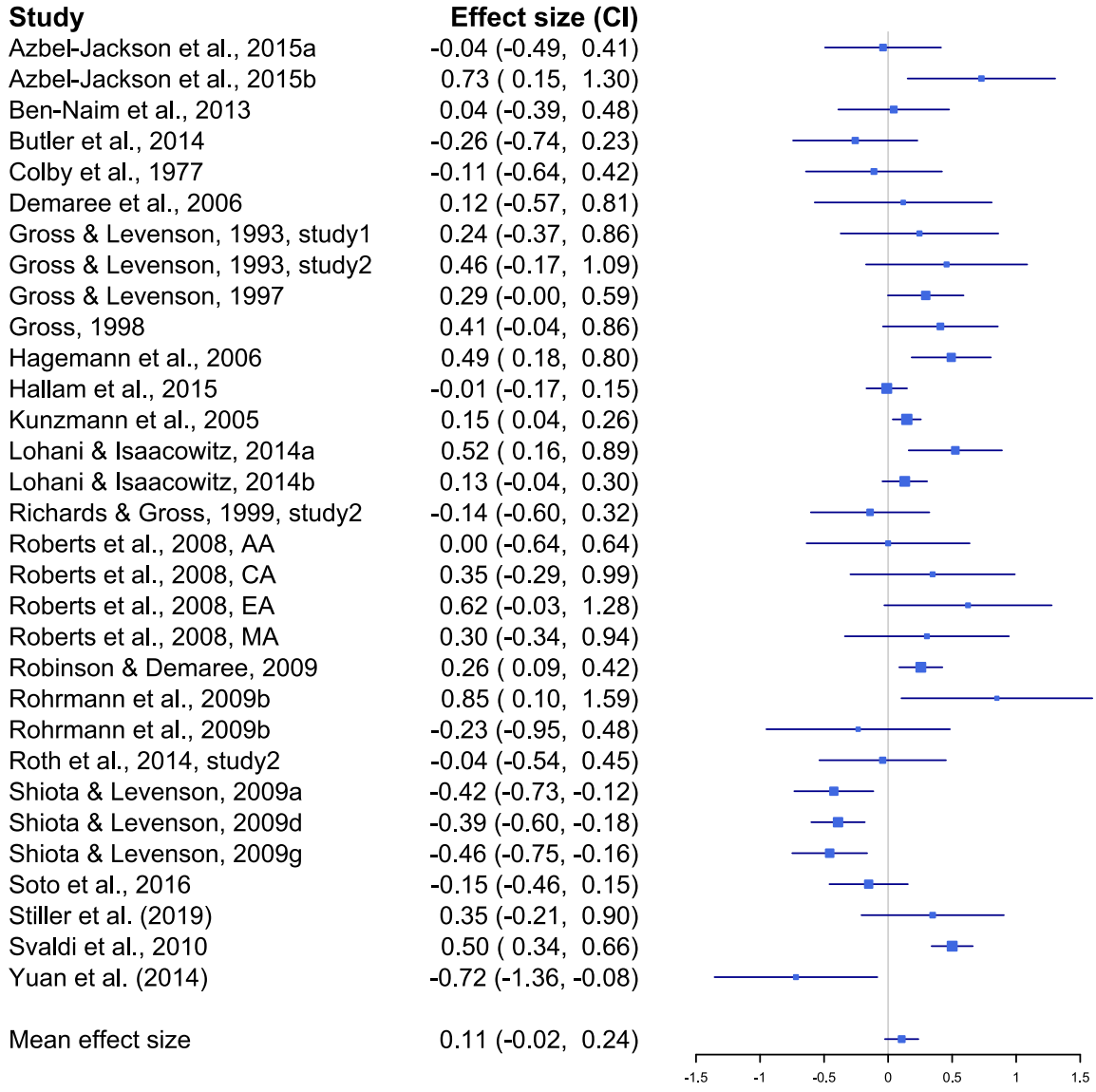

Figure 2.25. Statistics and results from the meta-analysis on skin conductance level (SCL) during suppression. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

2.6.2.4 Own Choice Strategies

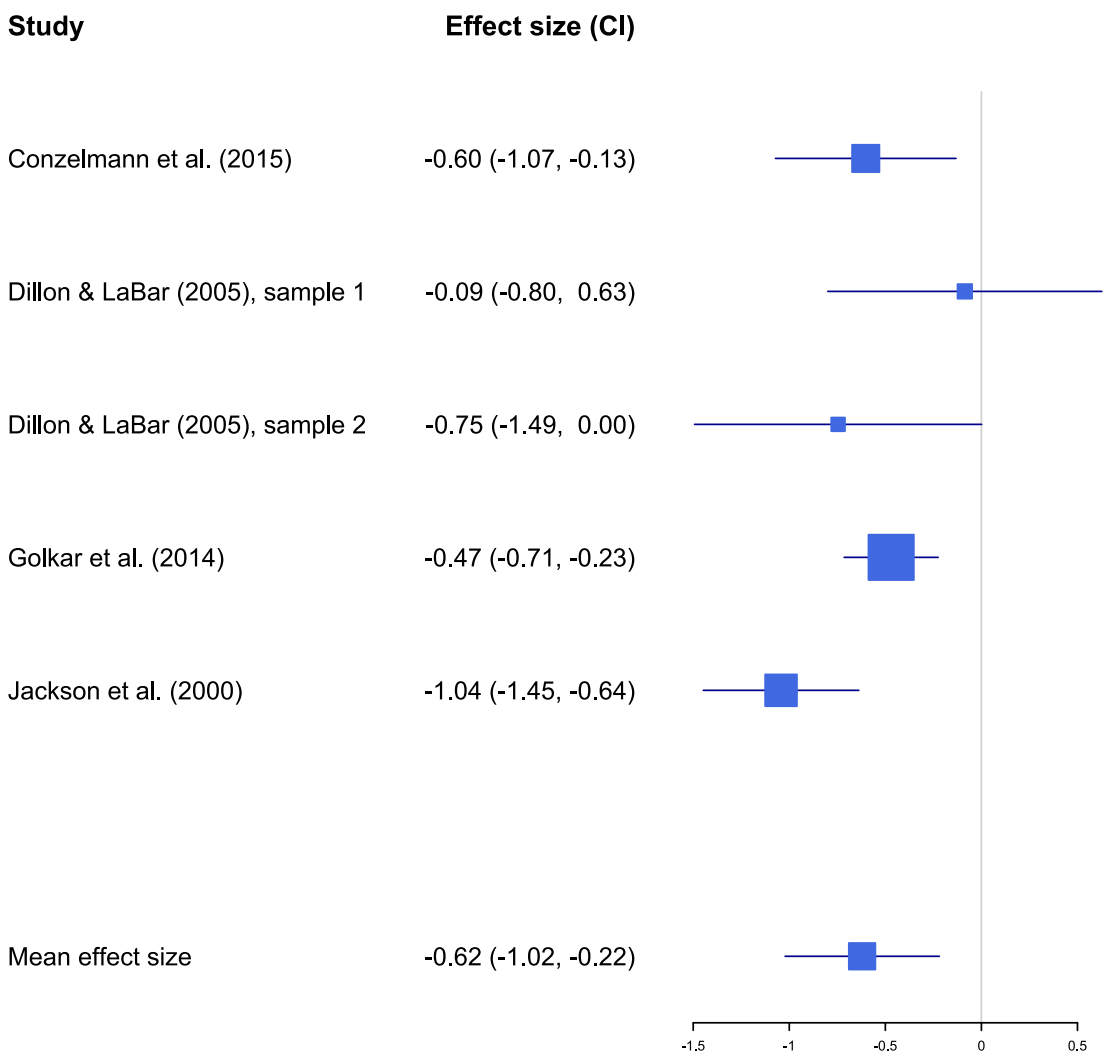

Figure 2.26. Statistics and results from the meta-analysis on emotion-modulated startle during downregulation with own choice. Each row represents one sample. Middle column: 'effect size' refers to effect size of the sample. 'CI' refers to the 95% confidence interval. The right column shows a forest plot with effect sizes and 95% confidence intervals (CI). The diamond size of the individual effect sizes of each sample refers to the relative weight of the effect.

2.6.3 Funnel Plots

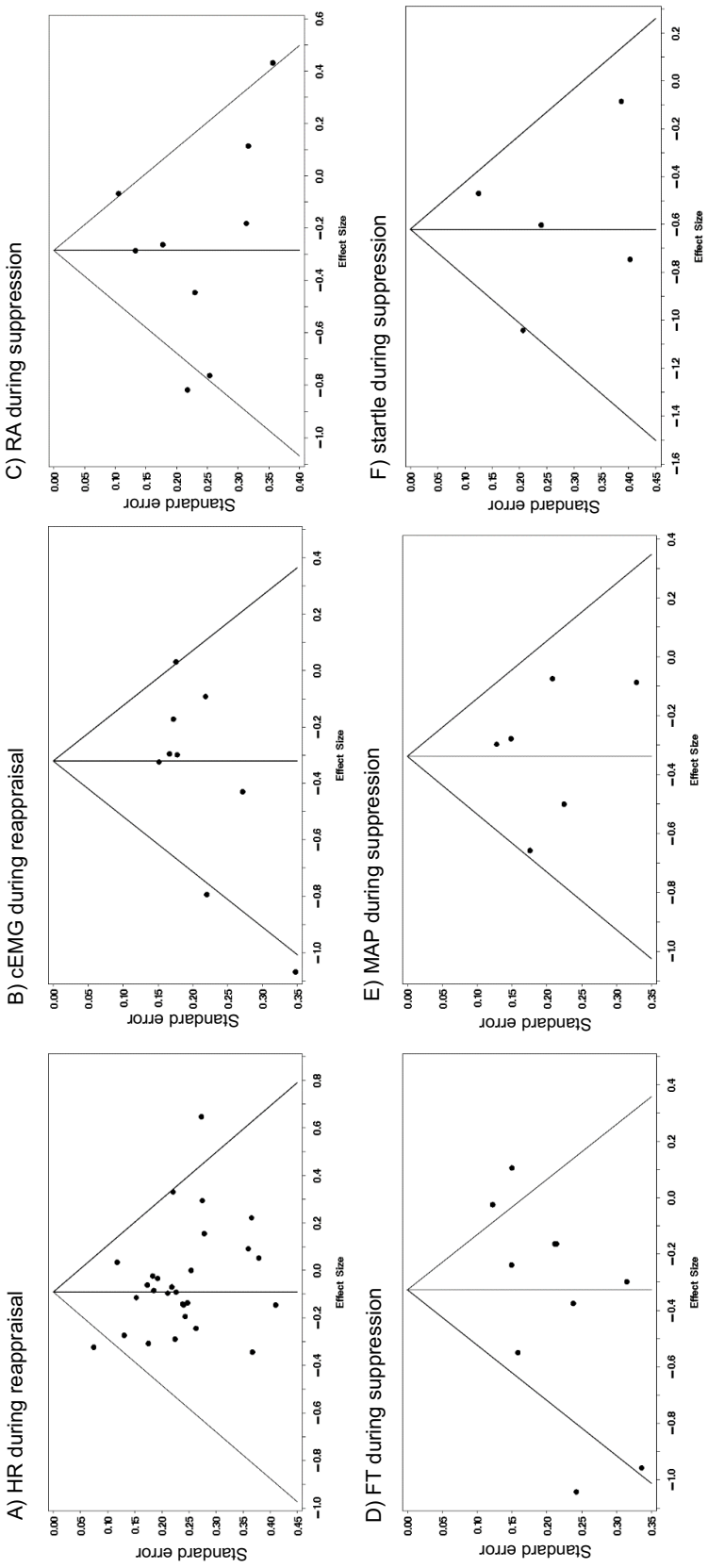

Figure 2.27. A)-F) Funnel plots with the effect sizes on the horizontal axis and their standard errors on the vertical axis for each significant meta-analysis.
Egger's test revealed that there was significant asymmetry only for the effect of reappraisal on heart rate (HR). cEMG = corrugator activity; FT = finger temperature; MAP = mean arterial pressure; RA = respiration amplitude; SCL = skin conductance level; EMG = startle.

Table 2.9

Study characteristics of identified emotion regulation studies inducing positive emotions.

Study name	D	Strategy	Measure	Emotion	CTL	TD (s)	EI	N_{total}	$\%_{women}$	Age_{mean}
Wu et al. (2016)	W	Reappraisal	cEMG	Positive	C4	6	I	75	49.33	
Giuliani et al. (2008)	W	Reappraisal	EPPT	Positive	C4	15	F	16	100	18.8
Giuliani et al. (2008)	W	Reappraisal	FPA	Positive	C4	15	F	16	100	18.8
Giuliani et al. (2008)	W	Reappraisal	FPTT	Positive	C4	15	F	16	100	18.8
Giuliani et al. (2008)	W	Reappraisal	FT	Positive	C4	15	F	16	100	18.8
Giuliani et al. (2008)	W	Reappraisal	HR	Positive	C4	15	F	16	100	18.8
Giuliani et al. (2008)	W	Reappraisal	MAP	Positive	C4	15	F	16	100	18.8
Gruber et al. (2014)	W	Reappraisal	nSCR	Positive	C1	200	F	23	52.17	35.2
Giuliani et al. (2008)	W	Reappraisal	RR	Positive	C4	15	F	16	100	18.8
Giuliani et al. (2008)	W	Reappraisal	SCR	Positive	C4	15	F	16	100	18.8
Gomez et al. (2015)	B	Reappraisal	SCR	Positive	C1	10	I	81	64.2	28.2
Wu et al. (2016)	W	Reappraisal	zEMG	Positive	C4	6	I	75	49.33	
Baur et al. (2015)	W	Own choice	cEMG	Positive	C2	4	I	41	78.05	21.1
Driscoll et al. (2009)	W	Own choice	HR	Positive	C1	8	I	10	70	35.2
Discroll et al. (2009)	W	Own choice	SCR	Positive	C1	8	I	10	70	35.2
Dillon and LaBar (2005), sample 1	W	Own choice	Startle	Positive	C3	12	I	48	77.08	22
Dillon & LaBar (2005), sample 2	W	Own choice	Startle	Positive	C3	12	I	48	77.08	22
Conzelmann et al. (2015)	W	Own choice	Startle	Positive	C3	8	I	31	48.39	22
Dan-Glauser and Gross (2011)	W	Suppression	cEMG	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	cEMG	Positive	C4	8	I	37	100	20.2

Table 2.9 *(continued)*

Study name	D	Strategy	Measure	Emotion	CTL	TD (s)	EI	N_{total}	$\%_{women}$	Age_{mean}
Dan-Glauser and Gross (2015)	W	Suppression	FPA	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	FPTT	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2011)	W	Suppression	FT	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	FT	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2011)	W	Suppression	HR	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	HR	Positive	C4	8	I	37	100	20.2
Ohira et al. (2006)	W	Suppression	HR	Positive	C4	60	I	10	100	24.2
Gross and Levenson (1997)	B	Suppression	HR	Amusement	C1	210	F	180	100	
Gross and Levenson (1997)	B	Suppression	HR	Amusement	C1	210	F	180	100	
Dan-Glauser and Gross (2011)	W	Suppression	MAP	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	MAP	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2011)	W	Suppression	RA	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	RA	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2011)	W	Suppression	RR	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	RR	Positive	C4	8	I	37	100	20.2
Gross and Levenson (1997)	B	Suppression	SCL	Amusement	C1	210	F	180	100	
Gross and Levenson (1997)	B	Suppression	SCL	Amusement	C1	210	F	180	100	
Ohira et al. (2006)	W	Suppression	SCR	Positive	C4	60	I	10	100	24.2
Kotwas et al. (2019)	W	Suppression	SCR	Positive	C4	45	F	34	70.59	31.7
Gomez et al. (2015)	B	Suppression	SCR	Positive	C1	10	I	81	64.2	28.2
Dan-Glauser and Gross (2011)	W	Suppression	zEMG	Positive	C4	8	I	37	100	20.2
Dan-Glauser and Gross (2015)	W	Suppression	zEMG	Positive	C4	8	I	37	100	20.2

Note. D = study design; B = between-subject design study; W = within-subject design study; cEMG = corrugator electromyography; DBP = diastolic blood pressure; EPTT = ear pulse transit time; FPA = finger pulse amplitude; FPTT = finger pulse transit time; FT = finger temperature; HR = heart rate; HRV = heart rate variability; MAP = mean arterial pressure; PD = pupil dilation; RA = respiration amplitude; RR = respiration rate; SBP = systolic blood pressure; SCL = skin conductance level; SCR = skin

conductance response; CTL = nature of control instruction; C1 = no instruction given (“view”); C2 = instruction not to regulate; C3 = instruction to maintain target emotion; C4 = instruction to respond naturally; C5 = a combination of C1-C4; TD = trial duration; EI = nature of emotion induction; F = film; I = images; N_{total} = number of participants in the study; $\%_{women}$ = Percent of women in the respective sample of the study. $N_{analyzed}$ = number of participants in the subsample.

Table 2.10

Taxonomy for coding electrodermal measures.

Study name	Strategy	Measure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Azbel-Jackson et al. (2015), study 1	Suppression	SCL	7	„Physiological data were baseline corrected by subtracting the signal recorded during the 1 s time window just before stimulus onset from all time points within a 7s time window immediately after picture onset. Data trials that fell more than 4 SDs from the within-subjects mean on a measure-by-measure basis for each participant were eliminated." (p. 1111)	Average
Azbel-Jackson et al. (2015), study 2	Suppression	SCL	7	„Physiological data were baseline corrected by subtracting the signal recorded during the 1 s time window just before stimulus onset from all time points within a 7s time window immediately after picture onset. Data trials that fell more than 4 SDs from the within-subjects mean on a measure-by-measure basis for each participant were eliminated." (p. 1111)	Average
Ben-Naim et al. (2013)	Suppression	SCR	900	Authors state that skin conductance response and skin conductance level recorded. No specific details about quantification provided.	Max. amp./peak
Ben-Naim et al. (2013)	Reappraisal	SCR	900	Authors state that skin conductance response and skin conductance level recorded. No specific details about quantification provided.	Max. amp./peak
Ben-Naim et al. (2013)	Suppression	SCL	900	Authors state that skin conductance response and skin conductance level recorded. No specific details about quantification provided.	Average
Ben-Naim et al. (2013)	Reappraisal	SCL	900	Authors state that skin conductance response and skin conductance level recorded. No specific details about quantification provided.	Average
Butler et al. (2014)	Suppression	SCL	570.6	Authors state that skin conductance level recorded. No specific details about quantification provided.	Author information
Butler et al. (2014)	Reappraisal	SCL	590.8	Authors state that skin conductance level recorded. No specific details about quantification provided.	Author information

Table 2.10 *(continued)*

Study name	Strategy	Mea-sure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Colby et al. (1977)	Suppression	SCL	6	"Skin resistance records were scored at the point of shock onset, 2 sec prior to shock onset and every sec after shock onset up to 6 sec after shock offset. The resistance scores were transformed to mhos of conductance. The phasic skin response for each trial was conducted by subtracting the average of the first two points from the average of the remainder for that particular trial." (p. 539)	Average
Demaree et al. (2006)	Suppression	SCL	120	Authors state that skin conductance level recorded. No specific details about quantification provided.	First author provided data on SCL via email.
Driscoll et al. (2009)	Own choice	SCR	8	„Skin conductance response magnitude was scored as the greatest change above 0.02 µS occurring in a 1–4 s time window following picture onset. A log transformation (log [SCR + 1]) was then performed to normalize the SCR data" (p. 63)	Max. amp./peak
Efinger et al. (2019)	Reappraisal	SCL	8	"Skin conductance level was exported as mean values for each trial." (p. 417)	Average
Efinger et al. (2019)	Distraction	SCL	8	"Skin conductance level was exported as mean values for each trial." (p. 417)	Average
Fitzpatrick and Kuo (2016)	Distraction	SCL	10	"The difference between average SCL during stimuli presentation and average SCL for two seconds prior to the stimuli presentation (during fixation cross presentation) was calculated for each experimental trial and used as the outcome variable." (p. 244)	Average
Fuentes-Sánchez et al. (2019)	Reappraisal	SCR	8	"For each trial, the peak response was scored as the maximum EDA value within a 1 to 6 s time window following picture onset, and amplitude was computed as the maximum electrodermal change score with respect to a baseline of 1 s prior to the picture onset." (p. 5)	Max. amp./peak

Table 2.10 *(continued)*

Study name	Strategy	Mea-sure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Goldin et al. (2019)	Reappraisal	SCL	12	"We computed mean values for […] skin conductance for each of the four conditions in the experimental task. We used a lag of 1.5 s, equivalent to one TR in the fMRI time series, to capture autonomic responses." (p. 5)	Average
Gomez et al. (2015)	Suppression	SCR	10	"To compute SCR change scores were calculated for each picture by subtracting the mean skin conductance of the interval between 1 and 4 sec after picture onset from the peak skin conductance of the interval between 1 and 4 sec after picture onset." (p. 14)	Max. amp./peak
Gomez et al. (2015)	Reappraisal	SCR	10	"To compute SCR change scores were calculated for each picture by subtracting the mean skin conductance of the interval between 1 and 4 sec after picture onset from the peak skin conductance of the interval between 1 and 4 sec after picture onset." (p. 14)	Max. amp./peak
Gross (1998a)	Suppression	SCL	64	"[…] laboratory software computed second-by-second averages for each of the five physiological measures throughout each baseline, instructional, film, and postfilm period. These second-by-second physiological values were later used to compute scores for each participant representing the averages of the physiological variables for the baseline, instructional, film, and postfilm periods. Change scores for the five measures were computed by subtracting baseline scores from instructional, film, and postfilm periods." (p. 228)	Average
Gross (1998a)	Reappraisal	SCL	64	"During the experimental sessions, laboratory software computed second-by-second averages for each of the five physiological measures throughout each baseline, instructional, film, and postfilm period. These seond-by-second physiological values were later used to compute scores for each participant representing the averages of the physiological variables for the baseline, instructional, film, and postfilm periods. Change scores for the five measures were computed by subtracting baseline scores from instructional, film, and postfilm periods." (p. 228)	Average

Table 2.10 *(continued)*

Study name	Strategy	Measure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Gross and Levenson (1993), study 1 and 2	Suppression	SCL	64	Authors state that skin conductance level was recorded. "Physiological measures were monitored continuously using an online data acquisition software package developed by our laboratory. This software computed second-by-second averages for each measure." (p. 974)	Average
Gross and Levenson (1997)	Suppression	SCL	210	Authors state that skin conductance level was recorded. "Change scores were created by subtracting prefilm period scores from film period scores for each variable." (p. 97)	Average
Hagemann et al. (2006)	Suppression	SCL	5	"Second-by-second values for each of the physiological measures, except RSA, were then reduced to mean values representing nonoverlapping time slices: (1) 2-min pretrial baseline, (2) 5-s pre-startle period, and (3) 6-s post-startle period (including the startle). Physiological reactivity scores were calculated for each measure by subtracting the pre- trial baseline from the pre- and post-startle periods." (p. 108)	Average
Hallam et al. (2015)	Suppression	SCL	10	"SCR traces were analysed in Ledalab v.3.2.9 using the Continuous Decomposition Analysis method to distinguish the phasic (driver) information from the underlying tonic sudomotor nerve activity. Raw SCR data were smoothed via convolution with a Hann window to reduce error noise and fitted to a bi-exponential Bateman function. Data were optimised by a conjugated gradient descent algorithm to reduce the error between them and the inbuilt SCR model. These processing steps allowed computation of a stimulus-locked 'integrated skin conductance response' (ISCR), a time-integration of the continuous phasic activity for each stimulus. This ISCR therefore represents an unbiased and time-sensitive measure of sympathetic activity in response to each stimulus. For investigating whether implementation intention and goal intention ER strategies may be associated with different skin conductance response, ISCRs from participants in both groups were averaged across epochs, within-subject." (p. 8)	Integrated signal

Table 2.10 *(continued)*

Study name	Strategy	Measure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Hallam et al. (2015)	Reappraisal	SCR	10	"SCR traces were analysed in Ledalab v.3.2.9 using the Continuous Decomposition Analysis method to distinguish the phasic (driver) information from the underlying tonic sudomotor nerve activity. Raw SCR data were smoothed via convolution with a Hann window to reduce error noise and fitted to a bi-exponential Bateman function. Data were optimised by a conjugated gradient descent algorithm to reduce the error between them and the inbuilt SCR model. These processing steps allowed computation of a stimulus-locked 'integrated skin conductance response' (ISCR), a time-integration of the continuous phasic activity for each stimulus. This ISCR therefore represents an unbiased and time-sensitive measure of sympathetic activity in response to each stimulus. For investigating whether implementation intention and goal intention ER strategies may be associated with different skin conductance response, ISCRs from participants in both groups were averaged across epochs, within-subject." (p. 8)	Integrated signal
Kim and Hamann (2012)	Reappraisal	SCR	24	"A skin conductance response was defined as the maximal positive deflection in skin conductance level with its onset occurring within a time window of .5 and 4 sec after each picture presentation. Peak-to-peak amplitude measurement was acquired for each picture. SCR amplitudes less than .03 MicroSiemens or continuously decreasing throughout the window were defined as 0 MicroSiemens." (p. 350)	Max. amp./peak
Kinner et al. (2017)	Reappraisal	SCR	5	"SCRs were defined as the maximum amplitude within a window of 1-8s after picture onset and calculated as baseline-to-peak amplitude differences of the largest deflection within a time window of 1-8s after picture onset. The baseline was the skin conductance level immediately preceding the inflection point. For each condition individual SCRs were averaged across the 10 trials." (p. 512)	Max. amp./peak

Table 2.10 *(continued)*

Study name	Strategy	Measure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Kotwas et al. (2019)	Suppression	SCR	45	Only phasic responses of skin conductance were recorded with the amplifier in the AC position. […] The mean amplitude of peaks of SCR were measured only during the 45 seconds duration of each film. SCR was obtained by averaging peaks' amplitude for each film. SCR below .01 microSiemens was not considered." (p. 8)	Max. amp./peak
Kunzmann et al. (2005)	Suppression	SCL	117	Authors states that skin conductance level was recorded. No specific details about quantification provided.	Author information
Leiberg et al. (2012)	Reappraisal	SCR	6	"EDA responses were determined as the difference between the maximum in the regulation phase (4,000 - 10,000 ms after picture onset) and the maximum in the pre-regulation phase (1,000 - 3,500 ms after picture onset). All trials were included in the analysis, regardless of the size of the response (i.e. reported values are a measure of EDA magnitude)." (p. 2467)	Max. amp./peak
Lohani and Isaacowitz (2014), sample 1 and sample 2	Suppression	SCL	300	"For each participant, difference scores for mood, SCL and EMG data were calculated by subtracting the mean activity during baseline period from the respective film period (no-regulation, attentional deployment, positive reappraisal and suppression).For SCL and corrugator, mean values of the respective measures while participants had watched a neutral video were used as baseline activity." (p. 685)	Average
Lohani & Isaacowitz (2014), sample 1 and sample 2	Reappraisal	SCL	300	"For each participant, difference scores for mood, SCL and EMG data were calculated by subtracting the mean activity during baseline period from the respective film period (no-regulation, attentional deployment, positive reappraisal and suppression).For SCL and corrugator, mean values of the respective measures while participants had watched a neutral video were used as baseline activity." (p. 685)	Average

Table 2.10 *(continued)*

Study name	Strategy	Mea-sure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Lohani & Isaacowitz (2014), sample 1 and sample 2	Distraction	SCL	300	"For each participant, difference scores for mood, SCL and EMG data were calculated by subtracting the mean activity during baseline period from the respective film period (no-regulation, attentional deployment, positive reappraisal and suppression)." (p. 685)	Average
Morawetz et al. (2016a)	Reappraisal	SCR	8	"Values for phasic SCRs were extracted as the difference between a local minimum and the succeeding local maximum within the response window." (p. 603)	Max. amp./peak
Morawetz et al. (2016b)	Reappraisal	SCR	8	"Skin conductance responses were defined as a deflection of at least .01 microSiemens occurring 1-8 s after stimulus onset. Only runs including more than 10% SCRs exceeding the above criterion were used for analysis. Values for phasic SCRs were extracted as the difference between a local minimum and the succeeding local maximum within the response window." (p. 572)	Max. amp./peak
(Morawetz et al., 2016c)	Reappraisal	SCR	8	"Values for phasic SCRs were extracted as the difference between a local minimum and the succeeding local maximum within the re- sponse window." (p. 1925)	Max. amp./peak
Ohira et al. (2006)	Suppression	SCR	60	"The amplitude of the maximum peak of SCR was measured in every 10s-time window during the presentation of each stimulus. Mean scores of SCRs during each block were analyzed statistically." (p. 723)	Max. amp./peak
Opitz et al. (2014), sample 1 and sample 2	Reappraisal	SCL	8	"For the continuous peripheral physiological measures, we summarized raw activity for two periods of interest in each trial, pre- instruction activity (reactivity period; mean of activity occurring 4 s after picture onset) and post-instruction activity (regulation period; mean of activity occurring 8 s after both ER manipulations)." (p. 6)	Average
Ortner (2015)	Reappraisal	SCR	8	"Skin conductance levels were collected throughout the task, and SCRs to each picture were calculated by subtracting the mean skin conductance response for the 1,000 ms prior to picture onset from the maximum during the 8,000 ms of picture presentation." (p. 566)	Max. amp./peak

Table 2.10 *(continued)*

Study name	Strategy	Measure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Plieger et al. (2017)	Reappraisal	SCL	4.5	"After smoothing the data, we conducted a continuous decomposition analysis (CDA) to separate the phasic parts from the tonic parts. Hence, the skin conductance response (SCR) reported here is adjusted by tonic activity. We measured phasic activity for 5 s starting with the onset of each stimulus. Phasic activity was averaged across all blocks of negative stimuli and across all blocks of neutral stimuli." (p. 33)	Average
Richards and Gross (1999)	Suppression	SCL	84	"Responses were digitized using custom software, which also computed second-by-second period averages for each of the five measures. In addition, change scores were calculated for each physiological measure by subtracting the baseline average from each slide-viewing period average." (p. 1038)	Average
Roberts et al. (2008), sample 1, sample 2, sample 3 and sample 4	Suppression	SCL	62	"For the remaining six measures (systolic and diastolic blood pressure, cardiac interbeat interval, RSA, skin conductance level, and general somatic activity), separate change scores were computed by subtracting mean response during the two-minute pre-film baseline from mean response during the film period." (p. 5)	Average
Robinson and Demaree (2009)	Suppression	SCL	120	Author state that skin conductance level was recorded. "Skin conductance data were also collected via the use of Biopac TSD203 transducers with Biopac Skin Conductance Electrode Paste placed at the non-dominant middle and fourth fingers. Data were amplified using Biopac's GSR100C amplifier using a gain of 10 lmhos and a low-pass filter of 10 Hz. Mindware's EDA 2.1 computer program identified all galvanic skin responses (SCR), as defined as a .05 lS increase in skin conductance." (p. 15). Report Galvanic skin conductance level in results section. No further specification about skin conductance responses.	Author information

Table 2.10 *(continued)*

Study name	Strategy	Mea-sure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Rohrmann et al. (2009), sample 1 and sample 2	Suppression	SCL	60	"Because of the high interindividual variability of the skin conductance level a range correction according to Schandry (1998) was performed [corrected EDA-value = EDA-value minus minimal EDA-value/maximum EDA- value minus minimal EDA-value] in order to obtain values independent from the individual absolute level. Heart rate (bpm) as well as skin conductance (mS) values were averaged across the period of the neutral and the amputation film." (p. 219)	Average
Rohrmann et al. (2009), sample 1 and sample 2	Reappraisal	SCL	60	"Because of the high interindividual variability of the skin conductance level a range correction according to Schandry (1998) was performed [corrected EDA-value = EDA-value minus minimal EDA-value/maximum EDA- value minus minimal EDA-value] in order to obtain values independent from the individual absolute level. Heart rate (bpm) as well as skin conductance (mS) values were averaged across the period of the neutral and the amputation film." (p. 2019)	Average
Roth et al. (2014)	Suppression	SCL	197	"The second-by-second SCL values were averaged for the two epochs of the baseline and the fear-eliciting film in each session. A difference score was calculated by subtracting the baseline mean SCL from the value obtained during the fear-eliciting film." (p. 6)	Average
Roth et al., (2014)	Distraction	SCL	197	"The second-by-second SCL values were averaged for the two epochs of the baseline and the fear-eliciting film in each session. A difference score was calculated by subtracting the baseline mean SCL from the value obtained during the fear-eliciting film." (p. 6)	Average
Sheppes et al. (2009)	Reappraisal	SCL	190	Authors state that skin conductance level was recorded. "For all measures, we applied the mean change from the pre-instruction baseline score, using the AcqKnowledge software (Biopac Systems, Goleta, CA)." (p. 93)	Average

Table 2.10 *(continued)*

Study name	Strategy	Measure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Sheppes et al. (2009)	Distraction	SCL	190	"For all measures, we applied the mean change from the pre-instruction baseline score, using the AcqKnowledge software (Biopac Systems, Goleta, CA)." (p. 93)	Average
Shermohammed et al. (2017)	Reappraisal	SCR	8	"The amplitude of the skin conductance response for each trial was calculated using the maximum change from base to peak in the .5-4.5 sec after picture onset. Amplitudes below .02 Microsec were scored as 0." (p. 1807)	Max. amp./peak
Shiota and Levenson (2009), sample 1, sample 4, sample 7	Suppression	SCL	180	Author state that skin conductance level was recorded. No specific details about quantification provided.	Author information
Shiota and Levenson (2009), sample 2, sample 3, sample 5, sample 6, sample 8, sample 9	Reappraisal	SCL	180	Author state that skin conductance level was recorded. No specific details about quantification provided.	Author information
Soto et al. (2016)	Suppression	SCL	58	"Mean reactivity levels were calculated for both IBI and SCL by subtracting the mean of the 1-minute prefilm baseline from mean responses during each film." (p. 45)	Average
Stiller et al. (2019)	Suppression	SCL	165	No specific information. Average skin conductance reported.	Average
Stiller et al. (2019)	Reappraisal	SCL	165	No specific information. Average skin conductance reported.	
Svaldi et al. (2010)	Suppression	SCL	211	Authors state that skin conductance level was recorded. No specific details about quantification provided.	Author information
Svaldi et al. (2010)	Reappraisal	SCL	125	Authors state that skin conductance level was recorded. No specific details about quantification provided.	Author information

Table 2.10 *(continued)*

Study name	Strategy	Mea-sure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Urry (2009)	Reappraisal	SCR	8	"Reappraisal- related change was thus captured by subtracting the baseline signal recorded during the 100-ms (EDA, HR) or 250-ms (EMG) time bin just before instruction delivery from all time points within an 8-s window immediately after this baseline period. During anticipatory trials the instruction was delivered 2 s before picture onset, and during online trials the instruction was delivered 4 s after picture onset. (…) The data were then aggregated across trials within each cell (described by reappraisal goal, timing, and period) for each participant. The summary statistic (mean for corrugator EMG and HR, maximum for EDA) was then tested with a multivariate GLM to assess the effects of reappraisal timing (anticipatory vs. online), reappraisal goal (increase, maintain, decrease), and time period (early, late)." (p. 788)	Max. amp./peak
Urry (2010)	Reappraisal	SCL	4	"Reappraisal-related change was captured by subtracting baseline signal recorded during the 100-ms (EDA, HR) or 250-ms (EMG) time bin just prior to instruction delivery from all subsequent time bins within a 6-second window immediately following this baseline period. […] Following baseline correction, the data were then averaged across time bins and trials within each cell (described by reappraisal goal and gaze direction) for each participant." (p. 128)	Average
Urry et al. (2009)	Reappraisal	SCL	8	"The data were downsampled to 20 Hz offline, averaged across trials within each condition in 1-s time bins […]. Regulation-related change in EDA was captured by subtracting EDA during the 4-s picture period from each of eight 1-s time bins following delivery of the reappraisal instruction. An area-under-the-curve index of EDA was computed by summing across all time points within each of three time intervals, early, middle, and late, as described below for the pupil diameter measure." (p. 6)	Area under the curve

Table 2.10 *(continued)*

Study name	Strategy	Measure	TD (s)	Quantification of electrodermal response (Description in paper)	Reason
Wolgast et al. (2011)	Reappraisal	SCL	153	"Average values for SCL were calculated for each film-clip using the software EDFBrowser and SPSS." (p. 861)	Average
Wu et al. (2016), study 2	Reappraisal	SCL	180	"Biotrace+ software (Mind Media B.V., Netherlands) supplied with the NeXus-10 was applied to data reduction, artifact control, and computation of average SCL scores for each participant for each 3-min relaxation period and film clip." (p. 5)	
Yuan et al. (2014)	Suppression	SCL	1800	No specific information about quantification provided. "Average values of SCR were computed for each phase of interest (e.g., rest, task, recovery) using the BioTrace+ software." (p. 4)	Average

Note. SCL = skin conductance level; SCR = skin conductance response; TD = Trial duration; Text in quotations marks are passages directly quoted from the paper

3 STUDY II: EMOTION-MODULATED STARTLE REFLEX DURING REAPPRAISAL: PROBE TIMING AND BEHAVIORAL CORRELATES

An adapted version of this chapter has been published as ‘Zaehringer, J., Schmahl, C., Ende, G., & Paret, C. (2018). Emotion-modulated startle reflex during reappraisal: Probe timing and behavioral correlates. *Behavioral Neuroscience, 132*(6), 573-579. doi:10.1037/bne0000271‘

3.1 Abstract

Down-regulation of negative emotions has been shown to reliably inhibit the emotion-modulated startle reflex, but it remains unclear whether the timing of the startle probe influences the meaningful quantification of emotion regulation. Moreover, it is not known, whether the degree of startle inhibition corresponds to the subjective attenuation of negative emotions. Therefore, the two main goals of the study were first to systematically analyze the effect of probe time on startle inhibition. Second, we aimed to explore the association between subjectively perceived down-regulation of arousal and valence and the degree of startle inhibition. We presented negative and neutral pictures to $N = 47$ participants. Pictures were paired with the instruction to reappraise or to maintain the emotions elicited by these pictures. Probes were delivered at three different times during a 12.5 s regulation phase and the startle response was measured with electromyography. Valence and arousal ratings were assessed after each trial. Results revealed no significant impact of probe time on startle inhibition during reappraisal. Startle inhibition and perceived down-regulation of arousal were significantly and positively correlated, whereas perceived down-regulation of valence was not. The results provide important implications for future studies in terms of startle probe timing and shed light onto the interpretation of startle inhibition as an indicator of subjective attenuation of negative emotions.

3.2 Introduction

Emotions add flavor to life and substantially shape the adaptive and survival responses to emotional stimuli (Ekman & Davidson, 1994). Emotions are organized in an appetitive and a defensive motivational system based on valence and arousal dimensions (Lang, 1995; Lang, Simons, & Balaban, 1997). Viewing of negative pictures activates the aversive motivational system and defensive reflexes, which vary with the intensity (i.e. arousal) of the stimulus (Bradley, Codispoti, Cuthbert, et al., 2001). The startle response is a defensive reflex and is typically measured as the contraction of the *orbicularis oculi* muscle in response to auditory

probes in human studies (Lang et al., 1990). A substantial body of work (Davis et al., 1995; Lang et al., 1990) has demonstrated that startle amplitudes are increased and decreased in negative and positive emotional states, respectively (i.e. emotion-modulated startle), with responses being more potentiated and inhibited during viewing of highly arousing stimuli (Bradley, Codispoti, Cuthbert, et al., 2001; Vrana et al., 1988).

Humans can purposefully control the type, intensity and occurrence of their emotions in a context-dependent manner, which is known as emotion regulation (Gross, 1998b, 2002). The most prominent and well-studied emotion regulation strategy is reappraisal, which is implemented before the behavioral response has fully unfolded (Gross, 1998a). The emotion-modulated startle can be used as a specific measure of reappraisal because it is inhibited when down-regulating and potentiated when up-regulating negative emotions (Adolph & Pause, 2012; Bernat et al., 2011; Conzelmann et al., 2015; Dillon & LaBar, 2005; Driscoll et al., 2009; Grillon et al., 2015; Jackson et al., 2000; Lee et al., 2009; Lissek et al., 2007). It has been demonstrated that the downregulation of negative emotions through reappraisal involves a significant reduction of amygdala activity (Buhle et al., 2014). Since projections from the amygdala modulate the startle reflex (Davis, 2000), down-regulation of the amygdala likely mediates startle inhibition with reappraisal.

Several questions however remain unanswered. As reappraisal processes are dynamic, the startle response may change as a function of probe timing. Previous studies demonstrated greater startle potentiation for probes delivered later compared to those delivered earlier during a 6-12s period of emotional picture viewing (Bradley, Codispoti, Cuthbert, et al., 2001; Sutton et al., 1997). When it comes to reappraisal, it is unclear whether probe timing has an impact on the reliability of the startle reflex to quantify emotion regulation. According to the implementation and maintenance model (IMMO; Kalisch, 2009; Paret et al., 2011) reappraisal is divided into an early and late phase. In the early phase, participants choose and implement a strategy, whereas in the late phase they maintain the strategy in working memory and monitor its success. In light of this, reappraisal might need several seconds until it effectively reduces negative emotions. Thus, startle modulation may become more pronounced as soon as the maintenance of reappraisal predominates. In line with this, two studies have observed small decreases of the startle amplitudes when probes were delivered 3 seconds into the regulation phase, whereas large decreases were observed when probes were delivered 8-11 seconds into the emotion regulation phase (Dillon & LaBar, 2005; Jackson et al., 2000). Another study delivered the startle probe 2 seconds into the reappraisal phase and reported non-significant startle inhibition (Eippert et al., 2007). Moreover, most pronounced amygdala down-regulation was observed after

probe presentation, suggesting that the probe might have been given too early to reliably detect reappraisal effects. The main goal of our study was thus to determine whether the timing of the startle probe has an influence on the reliability of the startle reflex to quantify reappraisal. This is a question worth exploring, as reappraisal is the emotion regulation strategy most often assessed in emotion regulation studies. Our study will thus help avoiding misinterpretation of non-significant effects in future studies.

A second issue concerns the question whether the emotion-modulated startle reflex as an objective measure of emotion down-regulation is correlated with subjective attenuation of negative emotions. One way to measure subjective attenuation of negative emotions is to assess the perceived valence of one's emotions and one's level of arousal after each regulation trial. The difference of arousal as well as valence between regulation and control condition then serves as a measure of subjective attenuation of negative emotions. Startle modulation by emotional stimuli has been shown to be sensitive to changes in both perceived stimulus valence and emotional arousal during picture viewing (VanOyen Witvliet & Vrana, 1995; Vrana et al., 1988). However, studies have found that the instruction to decrease emotions attenuated startle responses for both negatively and positively rated pictures (Bernat et al., 2011; Conzelmann et al., 2015; Dillon & LaBar, 2005), suggesting that startle reflex in response to emotion regulation may vary with perceived arousal, but not with valence. Few studies have assessed valence and arousal ratings in addition to the emotion-modulated startle (Bernat et al., 2011; Conzelmann et al., 2015; Dillon & LaBar, 2005) and only one of these studies (Bernat et al., 2011) assessed them during the experimental task, but the latter did not provide an analysis of covariance. We are not aware of any study that directly explored the relationship between subjective attenuation of valence and arousal of negative emotions and the degree of startle inhibition during reappraisal.

To address these questions, we conducted a reappraisal experiment with early, middle and late probe presentation (2 seconds, 7 seconds and 12 seconds into the regulation phase). Furthermore, we assessed how negative and how aroused participants felt after each trial. At the end of the experiment, we assessed the habitual use of cognitive emotion regulation strategies. On the basis of both the implementation and maintenance model and previous findings assessing both early and late startle probes (Dillon & LaBar, 2005; Jackson et al., 2000), we hypothesized that the difference of startle amplitudes in the reappraisal versus the control condition significantly increases with startle probe time (i.e. 2 seconds <7 seconds <12 seconds). We also hypothesized that down-regulation of arousal significantly correlates with startle inhibition through reappraisal. Another goal of our study was to explore whether the frequency to which

participants engage in reappraisal and other cognitive emotion regulation strategies in daily life would be associated with startle inhibition, as previous studies have shown that both difficulties with emotion regulation and the frequent use of reappraisal are associated with alterations in psychophysiological responding (Mauss et al., 2007; Williams et al., 2015).

3.3 Methods

3.3.1 Participants

47 healthy right-handed females (age: M = 23.9, SD = 5.5) participated in this study. Only females were studied to avoid confounding effects of gender differences (Bradley, Codispoti, Sabatinelli, & Lang, 2001). Exclusion criteria checked beforehand in a telephone interview comprised any mental disorder according to the DSM-V criteria (American Psychological Association, 2013), any history of psychiatric or neurological treatment, drug use, regular intake of medicine, current pregnancy and a BMI below 16.5. All participants had normal vision and hearing and were fluent in German. The study was approved by the Ethics Committee of the Medical Faculty Mannheim/ Heidelberg University and was conducted according to the Declaration of Helsinki. After full explanation of the study, all subjects gave written informed consent prior to participation and were paid 30 Euro for participation.

3.3.2 Apparatus and Procedure

To increase motivation, subjects were told at the beginning of the experiment that an additional amount of 15 Euro would be paid to the top 25% regulators, as determined by their "physiological responses" to the stimuli. At the end of the experiment, all participants received the additional 15 Euro and were debriefed (cf. Jackson et al., 2000). Participants were instructed either to view negative and neutral pictures without modifying their emotions or to down-regulate their feelings toward negative pictures by means of reappraisal strategies. Reappraisal strategies comprised distancing oneself from the depicted content or to reinterpret the depicted content in a positive or neutral manner. Oral practice trials were conducted before the computer task until the participant was able to apply reappraisal properly. Furthermore, participants were instructed not to turn away their gaze or to close their eyes, nor to focus exclusively on non-emotional parts of the picture. For a detailed description of the experimental procedure see Figure 3.1. The participants' eyes were tracked by a camera system (SMI BeGaze, Teltow, Germany) to encourage subjects to stay with the task but data were not analyzed. Trial order was pseudorandomized and counterbalanced with no more than two consecutive conditions of

the same type. In total, the paradigm consisted of 63 trials (7 trials per condition) and lasted 40 min. After trial 21 and trial 42 the task paused for a couple of minutes to prevent fatigue.

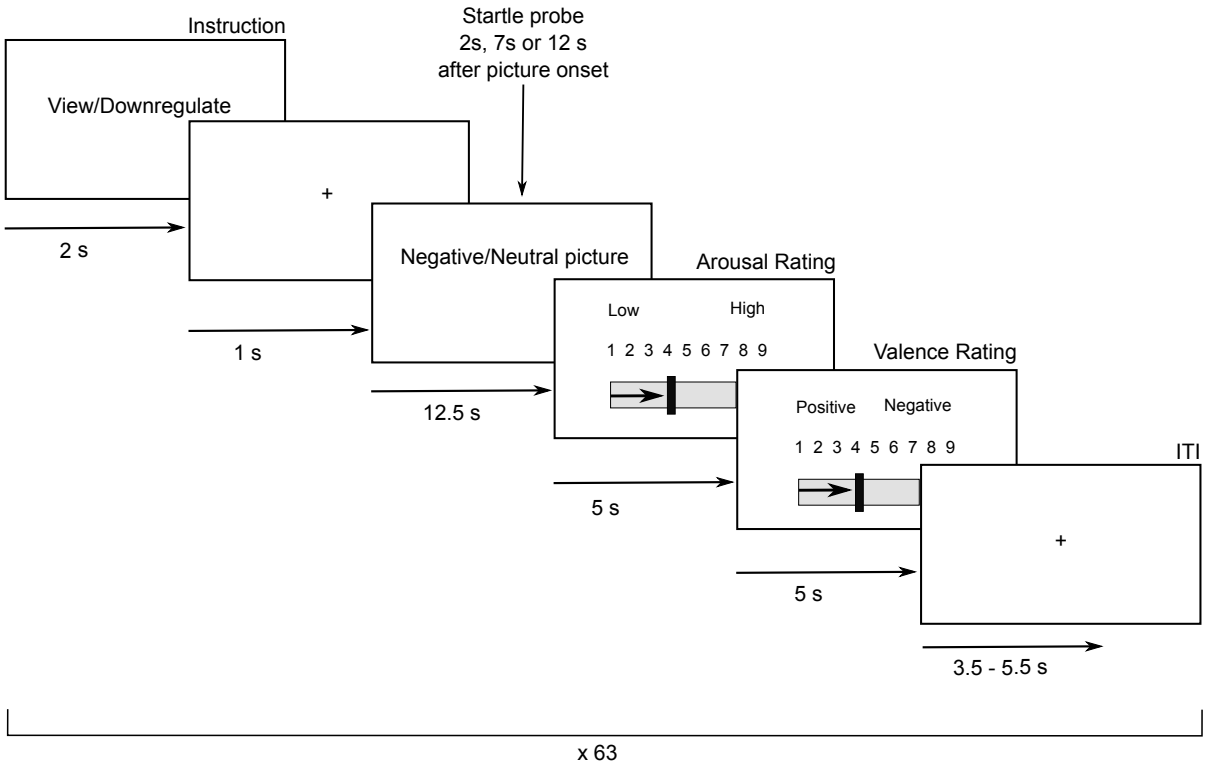

Figure 3.1. Experimental procedure. Each trial began with a 2,000 milliseconds presentation of an instructional cue (view, downregulate), followed by a fixation cross displayed for 1,000 milliseconds. Next, a neutral or negative picture was presented for 12,500 milliseconds. A startle probe (50 milliseconds, 95dB white noise burst) was presented through headphones either at 2 seconds (probe A), 7 seconds (probe B) or 12 seconds (probe C) into the regulation phase. Probes were balanced across conditions, and no more than two trials of same probe type were presented in consecutive trials. Self-assessment Manikins (SAM Ratings; Bradley & Lang, 1994) were presented after picture offset. Participants rated on a 1–9 Likert scale how positive/negative and aroused/calm they felt at that moment. Lower scores on the valence scale indicate that they felt more positive; lower scores on the arousal scale indicate that they felt calmer. By pressing buttons on a keyboard, subjects moved a bar from left to right to select SAMs corresponding to their subjective valence and arousal. The initial bar position was random, and the final position of the bar at the end of the rating was logged. Valence and arousal rating scales were displayed consecutively for 5,000 milliseconds each, with a 1,500 milliseconds time lag between the ratings. Intertrial intervals were jittered between 3,500 and 5,500 milliseconds.

3.3.3 Picture Stimuli

Stimuli were taken from the standardized picture series (Lang et al., 2008; Marchewka, Żurawski, Jednoróg, & Grabowska, 2014) and were presented with the Presentation software (Neurobehavioral Systems, Berkeley, CA) in semi-randomized order with restriction of no more than two consecutive trials from the same condition, and no more than three consecutive trials with negative pictures. Sets of 42 negative pictures (valence: $M = 2.36$, $SD = .68$; arousal:

M = 6.86, *SD* = .23) and 21 neutral pictures (valence: *M* = 5.21, *SD* = .59; arousal: *M* = 2.57, *SD* = .26) were created (normative ratings based on representative samples (Lang, Bradley, & Cuthbert, 2009; Marchewka, Zurawski, Jednorog & Grabowska, 2014). Details about the stimuli we used in our experiment can be derived from the supplement. Arousal and valence ratings differed significantly between the sets (both *p*s <.001). The pictures were further divided into two sets (balanced for content, valence, and arousal), which resulted in three conditions depending on instruction and picture type: View neutral pictures (LookNeu), view negative pictures (LookNeg) and down-regulate emotions while viewing negative pictures with reappraisal (RegNeg). Assignment of negative picture sets to LookNeg and RegNeg condition was alternated between subjects.

3.3.4 Measures.

3.3.4.1 Emotion-modulated Startle

The eye blink was measured by electromyogram (EMG). Two Ag-AgCl electrodes were placed on the orbicularis oculi muscle below the left eye, and a ground electrocardiogram electrode was attached on the lower rip bow[1]. The raw EMG signal was sampled at 1000 Hz, and the gain was amplified by 2000. High-pass (50 Hz) and low-pass (500 Hz) filters were applied to the data with AcqKnowledge software (BIOPAC Systems; Goleta, CA). EMG data were integrated over 10 samples and analyzed offline with Clip, a C++based, semi-automated program (Kinzig, Schulz, Curio, & Schächinger, 2008). Startle response was defined as the difference between peak (20–120 milliseconds after stimulus onset) and baseline (50 milliseconds prior to stimulus onset) signal. Trials including movement artifacts, excessive baseline activity (exceeding 2 standard deviations [SD] above baseline mean), or non-responses (peak < four SD below baseline mean) were excluded (mean % = 9.56 [*SD* = 6.52]) of all trials across participants). Startle data from six participants were excluded because of excessive noise (more than 30% missing). Finally, amplitudes were z-standardized within participants and transformed to *T*-scores with mean = 50 and SD = 10. Responses were averaged across participants for each condition.

3.3.4.2 Cognitive Emotion Regulation Questionnaire (CERQ)

To measure how frequently participants apply cognitive emotion regulation strategies in daily life, we administered the German adaptation of the Cognitive Emotion Regulation Questionnaire (CERQ; Garnefski, Kraaij, & Spinhoven, 2002) after the emotion regulation task. The

CERQ is a 36-item questionnaire, comprising the following nine conceptually different subscales: *Self-blame, Acceptance, Planning, Positive Refocusing, Rumination, Positive Reappraisal, Putting into Perspective, Catastrophizing and Other-blame.*

3.3.4.3 Postexperimental Success Ratings

After the emotion regulation task, participants indicated the subjectively experienced regulation success on a 9-point Likert scale ranging from 1 = not at all to 9 = very much. Questions were adapted from Gallo, Keil, McCulloch, Rockstroh, and Gollwitzer (2009): 1. *How much have you tried to reduce your negative feelings?*; 2. *How difficult was it to reduce your negative feelings?*; 3. *How well did you succeed in realizing the goal expressed in the reappraisal instruction?*

3.3.5 Data Analysis

Data analysis was performed using SPSS version 24 (SPSS Inc., Chicago, IL, USA). Before interaction analysis of emotion regulation and startle probe timing, we subtracted the mean amplitudes of the LookNeu conditions from the RegNeg and from the LookNeg conditions. We then conducted a repeated-measures ANOVA with 2 *condition* (RegNeg, LookNeg) x 3 *probe timing* (2 seconds, 7 seconds, 12 seconds) levels. The Greenhouse-Geisser correction was used to correct sphericity, and corrected *p* values are reported. Additionally, to analyze the correlation between startle inhibition and subjective attenuation of negative emotions, we calculated difference scores (LookNeg minus RegNeg) for startle amplitudes (*T*-scores) and valence and arousal ratings. Furthermore, CERQ subscales were correlated with startle inhibition using Pearson's correlation coefficient and assessed significance at the $p < .05$ level.

3.4 Results

3.4.1 Ratings (Manipulation Check)

All participants adhered to the reappraisal task as evidenced by the post-hoc success ratings (see Table 3.1 for means and SDs for each question). Induction of negative emotions was successful, evidenced by paired *t*-tests showing that arousal and valence ratings were significantly higher in the LookNeg (arousal: $M = 4.31$, $SD = .23$; valence: $M = 5.96$, $SD = .22$) than in the LookNeu (arousal: $M = 1.57$, $SD = .10$; valence: $M = 2.89$, $SD = .18$) condition [arousal: $t(46) = 12.25$, $p < .001$; valence: $t(46) = 14.51$, $p < .001$]. Moreover, all participants were able to down-regulate their level of arousal and negative emotional state with reappraisal as evidenced

by lower arousal (M = 3.18, SD = .18) and less negative/more positive valence (M = 4.85, SD = .20) in the RegNeg condition compared to the LookNeg condition [arousal: $t(46) = -6.23$, $p < .001$; valence: $t(46) = -7.71$, $p < .001$].

3.4.2 Emotion-modulated Startle

The 2 (condition) x 3 (startle probe) repeated measure ANOVA revealed a significant main effect of *condition*, $F(1,40) = 6.105$, $p = .02$, $eta^2 = .13$. The main effect of *probe timing* was not significant, $F(2,80) = .67$, $p = .60$, $eta^2 = .01$. Contrary to our hypothesis, the interaction *condition* x *probe timing* was also not significant, $F(2,80) = .51$, $p = .60$, $eta^2 = .01$ (Figure 3.2A).

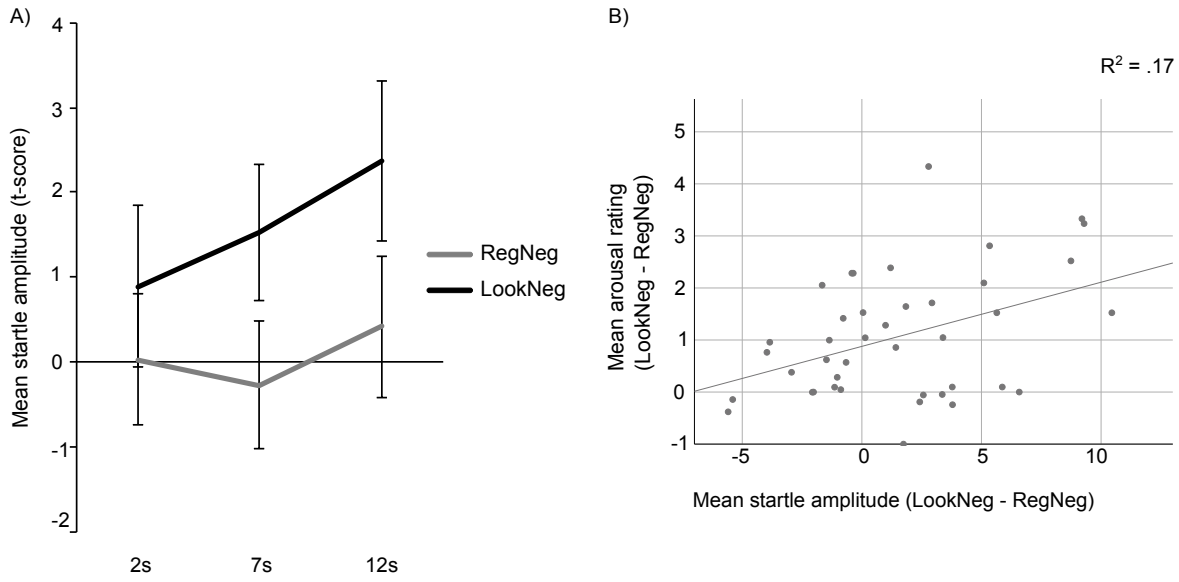

Figure 3.2. A) Mean startle amplitudes in the RegNeg and LookNeg condition across the three startle probe times (A: 2s, B: 7s, C: 12s). Mean amplitudes represent T score converted difference scores (RegNeg minus LookNeu and LookNeg minus LookNeu). Error bars represent standard errors of means (SEM). B) Significant and positive Pearson's correlation between arousal down-regulation and startle inhibition indicating that individuals who performed better at down-regulating arousal also performed better at inhibiting their startle response during reappraisal.

3.4.3 Correlations

3.4.3.1 Arousal and Valence Ratings

Pearson's correlational analysis showed that the difference scores (LookNeg minus RegNeg) of startle amplitudes were significantly and positively associated with difference scores (LookNeg minus RegNeg) of arousal ratings (r =.41, p = .01, n = 41), indicating that participants who showed a stronger subjective attenuation of arousal also showed a stronger inhibition of the startle amplitude during reappraisal (Figure 3.2B). In contrast, difference scores (LookNeg minus RegNeg) of startle amplitudes did not significantly correlate with difference scores of valence ratings (r = .26, p = .11, n = 41), suggesting that the inhibition of startle during reappraisal is not predictive for the subjective attenuation of valence.

3.4.3.2 CERQ

For descriptive statistics on CERQ subscales see Table 3.2. Internal consistency of the CERQ was good, as evidenced by Cronbach's alpha = 0.77. Correlations between difference scores (LookNeg minus RegNeg) of startle amplitudes and CERQ subscales and success ratings were not significant (all ps > .05).

3.5 Discussion

The primary goal of this study was to examine effects of startle probe timing on the meaningful quantification of emotion regulation through the startle eye-blink. In addition, we analyzed whether the degree of startle inhibition during reappraisal was predictive for the subjective attenuation of negative emotions. Results demonstrated that subjects successfully reduced negative emotions evidenced by startle inhibition, arousal, valence and post-hoc success ratings. Contrary to our expectations, startle inhibition was independent of probe timing. In other words, whether probes were delivered at 2, 7 or 12 seconds into the reappraisal phase did not significantly affect the assessment of emotion regulation. In accordance with our second hypothesis, startle inhibition was significantly and positively correlated with subjective attenuation of arousal but not with subjective reduction of valence. We were able to replicate the finding that down-regulating negative emotions with cognitive emotion regulation strategies inhibits emotion-modulated startle (Conzelmann et al., 2015; Grillon et al., 2015; Jackson et al., 2000; Lee et al., 2009; Lissek et al., 2007; Piper & Curtin, 2006). Although temporal differences were not significant, a visual inspection of results shows that startle probes delivered at ≥7 seconds are useful to quantify reappraisal effects (see Figure 3.2A). Descriptively, startle probes delivered

at 2 seconds produced smaller effects and might be less sensitive than later probes (Dillon & LaBar, 2005; Eippert et al., 2007; Jackson et al., 2000), though these differences are not significant.

Startle inhibition during reappraisal correlated with the perceived downregulation of arousal. This finding is in line with our second hypothesis, which was based on previous literature suggesting that the regulation effect on the emotion-modulated startle follows the pattern of variations in arousal of pictures (Bernat et al., 2011; Dillon & LaBar, 2005). The present results provide important new information as they demonstrate that changes in the defensive tendency measured with emotion-modulated startle also reflect changes in perceived levels of arousal. Efferent pathways from the amygdala, found to modulate startle (Davis, 1992), are involved in regulating arousal-related responses to aversive stimuli (LeDoux, Iwata, Cicchetti, & Reis, 1988; Reyes, Carvalho, Vakharia, & Van Bockstaele, 2011). Hence, subjective arousal as well as defensive tendency may be down-regulated due to amygdala inhibition, and the shared neural mechanism may account for the finding. Alternatively, it could be that subjects may have rated arousal based on self-observation of startle intensity. In contrast, we found no significant correlation between the perceived downregulation of valence and startle inhibition. The result raises the possibility that in the context of emotion regulation the affective modulation of the startle response might be a particularly strong indicator of one's level of arousal rather than a measure of the valence of one's emotional state. This assumption is complemented by prior research on emotional picture viewing showing that both the degree of affective modulation of the startle response and amygdala activity become more pronounced as the level of subjective arousal of pictures increases (Cuthbert, Bradley, & Lang, 1996; Phan et al., 2004; Sabatinelli, Bradley, Fitzsimmons, & Lang, 2005).

No correlations were found between startle inhibition and CERQ subscales, suggesting that startle inhibition during reappraisal might not be indicative for a frequent use of reappraisal and other forms of cognitive emotion regulation strategies. This is contrary to previous studies showing that people who frequently engage in reappraisal show different psychophysiological responding in aversive situations than people who use reappraisal less frequently (Mauss et al., 2007; Memedovic, Grisham, Denson, & Moulds, 2010).

The study is not without limitations. To limit attentional demands and fatigue due to long task duration, ITI was short and startle probes were presented in every trial. As a result, late probe types B and C, in particular, were highly predictable. However, subtraction of LookNeu condition probes controlled for probe anticipation effects. Moreover, the study may have been underpowered to show the effect of probe timing on startle inhibition. A post-hoc power analysis

based on our results using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that 262 participants would be necessary to achieve reasonable power (1-β>.80) in order to prove significance, given a true interaction effect. In our view, the cost to assess more data would outweigh the benefit to detect a potential but marginal effect. Finally, generalization of our results is limited because all participants were female college students.

3.6 Conclusion

Startle probes delivered at ≥7 seconds are useful to quantify reappraisal effects, though earlier probes did not yield significantly worse effects. Moreover, the successful downregulation of perceived arousal is reflected by a decline in the defensive tendency, measured with the emotion-modulated startle response. In contrast, down-regulation of emotional valence is not correlated with a reduction of the startle response.

3.7 Supplementary Material

3.7.1 Supplementary Methods

3.7.1.1 Picture Stimuli

Stimuli were taken from the standardized picture series (Lang et al., 2008; Marchewka et al., 2014) and were presented with the Presentation software (Neurobehavioral Systems, Berkeley, CA) in semi-randomized order with restriction of no more than two consecutive trials from the same condition, and no more than three consecutive trials with negative pictures. Sets of 42 negative pictures[8] (valence: *M* = 2.36, *SD* = .68; arousal: *M* = 6.86, *SD* = .23) and 21 neutral pictures[9] (valence: *M* = 5.21, *SD* = .59; arousal: *M* = 2.57, *SD* = .26) were created (normative ratings based on representative samples (Lang, Bradley, & Cuthbert, 2008; Marchewka et al., 2014). Arousal and valence ratings differed significantly between the sets (both *p*s <.001). The pictures were further divided into two sets (balanced for content, valence, and arousal), which

[8] IAPS pictures: 1050, 1120, 1201, 9635, 3051, 3261, 3266, 3400, 3550, 6021, 6231, 6250, 6263, 6370, 6550, 6838, 9163, 9490, 9908, 9940, 1202

Pictures derived from Marchewka et al. (2014): Animals_052, Faces_009, Faces_016, Faces_018, Faces_149, Faces_159, Faces_364, Objects_121, Objects_006, Objects_011, Objects_125, People_088, People_127, People_073, People_201, People_211, People_212, People_214, People_222, People_226, People_231

[9] IAPS pictures: 7004, 7000, 7490, 7950, 7187, 7026, 2381, 7041, 7217, 7052, 7006, 7080, 7185, 7100, 7035, 5471, 5740, 2038, 2580, 5510, 5530

resulted in three conditions depending on instruction and picture type: View neutral pictures (LookNeu), view negative pictures (LookNeg) and down-regulate emotions while viewing negative pictures with reappraisal (RegNeg). Assignment of negative picture sets to LookNeg and RegNeg condition was alternated between subjects.

3.7.1.2 Emotion Regulation Questionnaire

Participants rated the frequency of the use of the following emotion regulation strategies (in %) directly after they completed the emotion regulation task: 1. Reappraisal, 2. Distancing, 3. Distraction, 4. Relaxation, 5. Mindfulness, 6. Attentional deployment, 7. Other strategies. Percentages were then ranked within each participant from 1 (mostly used) to 7 (least used).

3.7.2 Supplementary Results

3.7.2.1 Emotion Regulation Questionnaire

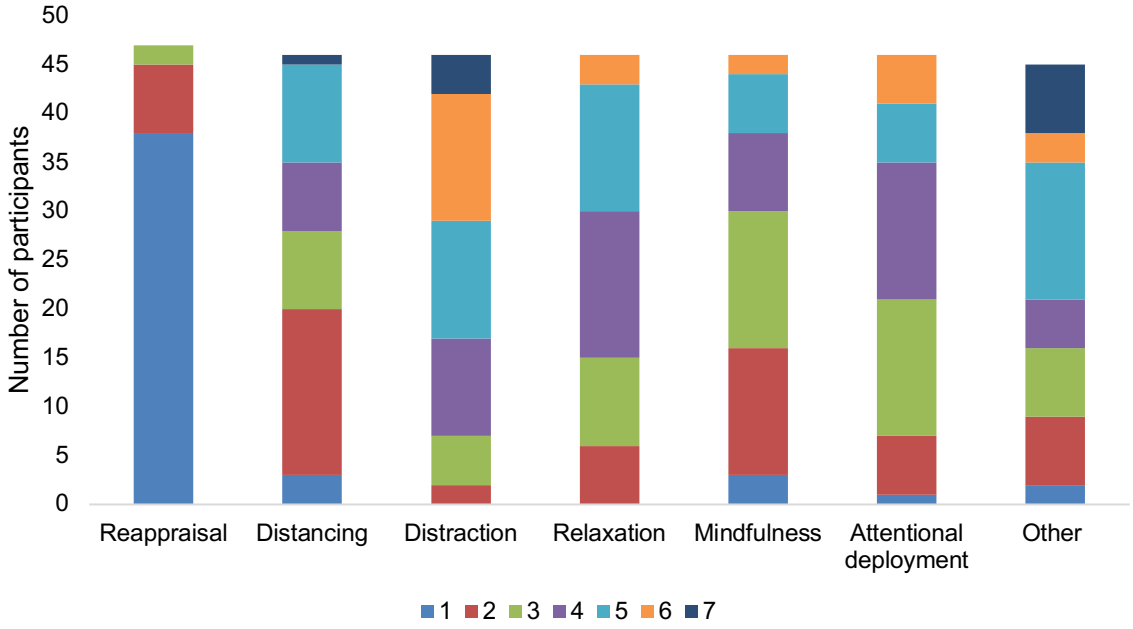

Figure 3.3. Number of participants reporting to use a strategy most often (1) to least often (7) during the regulation task.

In line with instructions, subjects reported that they had dominantly used reappraisal to regulate emotions during the experiment (see Figure 3.3).

3.7.2.2 Post-hoc Success Ratings

Table 3.1

Descriptive statistics of post-hoc success ratings.

Success ratings	*N*	Min.	Max.	*M*	*SD*	Skewness		Kurtosis	
						M	*SE*	*M*	*SE*
How much have you tried to reduce your negative feelings?	47	1.00	9.00	7.45	1.73	-2.05	0.35	5.03	0.68
How difficult was it to reduce your negative feelings?	47	1.00	9.00	4.02	1.74	0.64	0.35	0.00	0.68
How well did you succeed in realizing the goal expressed in the reappraisal instruction?	47	3.00	9.00	6.30	1.59	-0.45	0.35	-5.55	0.68

Note. SD = standard deviation; *SE* = standard error.

3.7.2.3 CERQ Subscales

Table 3.2

Descriptive statistics of CERQ subscales.

CERQ subscale	*N*	Min.	Max.	*M*	*SD*	Skewness		Kurtosis	
						M	*SE*	*M*	*SE*
Self-blame	47	5.00	17.00	12.19	2.46	-0.09	0.35	0.55	0.68
Acceptance	47	6.00	19.00	13.60	2.86	-0.51	0.35	0.03	0.68
Rumination	47	4.00	17.00	10.26	2.89	0.22	0.35	-0.00	0.68
Refocusing	47	7.00	18.00	12.96	2.65	-0.35	0.35	-0.53	0.68
Planning	47	7.00	20.00	13.23	3.02	-0.23	0.35	-0.20	0.68
Reappraisal	47	4.00	20.00	11.89	2.72	-0.07	0.35	1.56	0.68
Putting into perspective	47	4.00	14.00	9.15	2.17	-0.31	0.35	0.60	0.68
Catastrophizing	47	4.00	13.00	7.30	2.20	0.36	0.35	-0.35	0.68
Other blame	47	5.00	15.00	9.55	2.39	0.15	0.35	-0.56	0.68

Note. *SD* = standard deviation; *SE* = standard error.

Table 3.3

Pearson Correlations between startle inhibition (LookNeg – RegNeg) and CERQ subscales.

		Rumination	Self-blame	Acceptance	Refocusing	Planning	Reappraisal	Putting into perspective	Catastrophi-zing	Other blame
Startle inhibition (LookNeg minus RegNeg)	*r*	-0.06	0.15	0.04	0.20	0.07	.02	-0.24	0.04	0.02
	p	0.72	0.34	0.81	0.22	0.67	0.91	0.14	0.83	0.91
	N	41	41	41	41	41	41	41	41	41

Note. *r* = Pearson's correlation; *N* = number of participants

4 STUDY III: IMPROVED EMOTION REGULATION AFTER NEUROFEEDBACK: A SINGLE-ARM TRIAL IN PATIENTS WITH BORDERLINE PERSONALITY DISORDER

An adapted version of this chapter has been published as 'Zaehringer, J., Ende, G., Santangelo, P., Kleindienst, N., Ruf, M., Bertsch, K., Bohus, M., Schmahl, C. & Paret, C. (2019). Improved emotion regulation after neurofeedback: A single-arm trial in patients with borderline personality disorder. *NeuroImage Clinical*, 24:102032. doi:10.1016/j.nicl.2019.102032'

4.1 Abstract

Real-time functional magnetic resonance imaging (fMRI) neurofeedback training of amygdala hemodynamic activity directly targets a neurobiological mechanism, which contributes to emotion regulation problems in borderline personality disorder (BPD). However, it remains unknown which outcome measures can assess changes in emotion regulation and affective instability, associated with amygdala downregulation in a clinical trial. The current study directly addresses this question. Twenty-four female patients with a DSM-IV BPD diagnosis underwent four runs of amygdala neurofeedback. Before and after the training, as well as at a six-weeks follow-up assessment, participants completed measures of emotion dysregulation and affective instability at diverse levels of analysis (verbal report, clinical interview, ecological momentary assessment, emotion-modulated startle, heart rate variability, and fMRI). Participants were able to downregulate their amygdala blood oxygen-dependent (BOLD) response with neurofeedback. There was a decrease of BPD symptoms as assessed with the Zanarini rating scale for BPD (ZAN-BPD) and a decrease in emotion-modulated startle to negative pictures after training. Further explorative analyses suggest that patients indicated less affective instability, as seen by lower hour-to-hour variability in negative affect and inner tension in daily life. If replicated by an independent study, our results imply changes in emotion regulation and affective instability for several systems levels, including behavior and verbal report. Conclusions are limited due to the lack of a control group. A randomized controlled trial (RCT) will be needed to confirm effectiveness of the training.

4.2 Introduction

Emotion dysregulation is considered a hallmark of borderline personality disorder (BPD) (Glenn & Klonsky, 2009; Sanislow et al., 2002; Schmahl et al., 2014), characterized by heightened reactivity to negative stimuli, with impairments in the implementation and maintenance of adaptive and appropriate emotion regulation strategies, as well as heightened experience of negative affect (Carpenter & Trull, 2013). On a neural level, a key feature of BPD is hyperactivation of the amygdala in response to negative and neutral stimuli (Schulze et al., 2019), likely reflecting the emotion dysregulation observed in BPD patients (Schmahl et al., 2014).

Current emotion regulation models implicate downregulation of the amygdala as a mechanism to control emotions in clinical contexts (Buhle et al., 2014; Etkin et al., 2015). A normalization of amygdala activation and improved emotion regulation were found during Dialectical Behavior Therapy (DBT) in BPD patients, suggesting that amygdala response is an important indicator of BPD remission (Goodman et al., 2014; Schmitt, Winter, Niedtfeld, Herpertz, & Schmahl, 2016). However, it is not clear whether decreased amygdala response mediates BPD remission. Until recently, probing this has been virtually impossible, as techniques to tackle subcortical activation were limited to highly invasive deep-brain stimulation.

With the emergence of real-time functional magnetic resonance imaging (fMRI), modulation of emotion brain circuitry became feasible (Linhartová et al., 2019). With feedback from brain activation in real-time, dubbed neurofeedback, healthy subjects (Brühl et al., 2014; Herwig et al., 2019; Keynan et al., 2016; Paret et al., 2014) and patients (Nicholson et al., 2017; Paret et al., 2016) were able to reduce their amygdala activation during real-time fMRI. The benefits of this new technique are two-fold: first, assessing behavioral sequels of neuromodulation provide a better understanding of mechanisms that contribute to reduced amygdala activation in BPD. Second, the potential to address dysregulated neurobiological mechanisms with neurofeedback could be used for BPD treatment. However, before addressing these goals, primary outcome measures for clinical trials must be identified.

Emotion dysregulation in BPD has been studied with a plethora of measures, such as emotional picture-viewing tasks (Krause-Utz et al., 2012), clinical interviews (Zanarini et al., 2003), retrospective questionnaires (Glenn & Klonsky, 2009; Gratz, Rosenthal, Tull, Lejuez, & Gunderson, 2006; Salsman & Linehan, 2012) and affective variability in ecological momentary assessment (EMA; Ebner-Priemer et al., 2007; Nica & Links, 2009; Santangelo et al., 2017). In addition, psychophysiological indices such as resting heart rate variability (HRV) and startle modulation have been used to study emotion dysregulation in BPD (Ebner-Priemer et al., 2005; Thompson et al., 2018). BPD patients show lower resting HRV than controls (Koenig, Kemp,

Feeling, Thayer, & Kaess, 2016), which is indicative of less regulation ability (Appelhans & Luecken, 2006). Cognitive emotion regulation diminishes emotion-modulated startle in healthy individuals (Jackson et al., 2000; Zaehringer et al., 2018) and BPD patients (Thompson et al., 2018), as this downregulation correlates with downregulation of affective states (Zaehringer et al., 2018). Similarly, studies report associations of amygdala hyperactivation and BPD diagnosis (Schulze et al., 2019), outlining a pathway of amygdala regulation via self-injury (Reitz et al., 2015), and reporting a coincidence of amygdala normalization with response to psychotherapy (Goodman et al., 2014). Yet, little action has been shown to map amygdala hyperactivation with behavioral correlates of emotion dysregulation and affective instability. It is unknown what aspects of BPD symptomatology improve with normalization of amygdala activation. Thus, evidence is very limited, impeding informed selection of a primary outcome measure for clinical trials that assess amygdala neuromodulation. The present study addressed exactly this question, i.e., what aspects of emotion dysregulation improve following amygdala neurofeedback? Moreover, because dysfunction of emotion neurocircuitry manifests through dysregulated behavior, including the verbal report of symptoms collected in standard psychometric assessments (LeDoux & Pine, 2016), we used a multimodal assessment of psychopathology, explained below. BPD patients underwent four sessions of neurofeedback training and received a test battery directly before training, both after training and at 6-weeks follow-up. The test battery included a multimodal assessment of emotion regulation of self-report, EMA, behavioral, and fMRI measures. We hypothesized that BPD patients would downregulate the amygdala with neurofeedback. In addition, we hypothesized that BPD patients would show significant changes at several system levels, i.e., verbal report, everyday experience, and behavior and brain responses. Specifically, we hypothesized reductions in emotion dysregulation and improved clinical symptoms, enhanced emotion regulation as shown by increased resting HRV, improved emotion regulation in an established laboratory task (Jackson et al., 2000), and decreased amygdala response to emotional pictures. In addition, we explored changes in everyday experience as well as changes in a number of aspects of emotion regulation and BPD psychopathology.

4.3 Methods

4.3.1 Participants

Twenty-six female patients with at least 5 BPD criteria, according to the DSM-V (American Psychiatric Association, 2013) participated in the present study. All participants were on stable medication (see Table 4.1 for details on medication) during the course of the study. In case

participants receiving psychotherapeutic treatment, they were required to maintain it throughout the study. Two patients were excluded after completion (one patient reported amphetamine consumption during participation, and one patient fell asleep during the neurofeedback runs).
The diagnostic assessment comprised the Structured Interview for DSM-IV Axis-I (SKID-I; First, Gibbon, & Spitzer, 1997) and the International Personality Disorder Examination (IPDE; Loranger, 1999). Patients were excluded from our study in cases of severe somatic illness and if exclusion criteria related to MRI were fulfilled (metal implants, left-handedness, claustrophobia, and pregnancy). Further exclusion criteria were alcohol or substance abuse within the last 6 months, lifetime psychotic disorder, bipolar affective disorder, or mental retardation.
A total of n = 108 individuals were initially screened for our study. N = 77 had to be excluded because they did not fulfill our inclusion criteria, were not interested in the first place or were interested but ultimately did not participate. Thus n = 31 participants were allocated to our study and n = 26 of them received the full neurofeedback training. A detailed flow chart of the study is shown in Figure 4.9 in the supplement.
Descriptive statistics of demographic variables are reported in Table 4.1. The study was approved by the Ethics Committee of the Medical Faculty Mannheim / Heidelberg University and was conducted according to the Declaration of Helsinki. All subjects gave written informed consent prior to participation and received financial compensation (120 Euros). The research protocol was registered on ClinicalTrials.gov (NCT02866110) and the *Deutsches Register für Klinische Studien* (drks.de; DRKS00009363).

4.3.2 Procedure

Participants took part in four runs of amygdala neurofeedback training. Runs were administered on 3 different days, with run 2 and 3 being administered consecutively on the second training day. Training days were scheduled 2-7 days apart from each other. At baseline (T0) and after completion of amygdala neurofeedback training (T1), the test battery was administered. All measures except EMA were assessed again at 6-weeks follow-up (T2). For details of the procedure, see Figure 4.1. The consensus on the reporting and experimental design of clinical and cognitive-behavioral neurofeedback studies (CRED-nf checklist; Ros et al., 2019) can be found in the supplement on page 10.

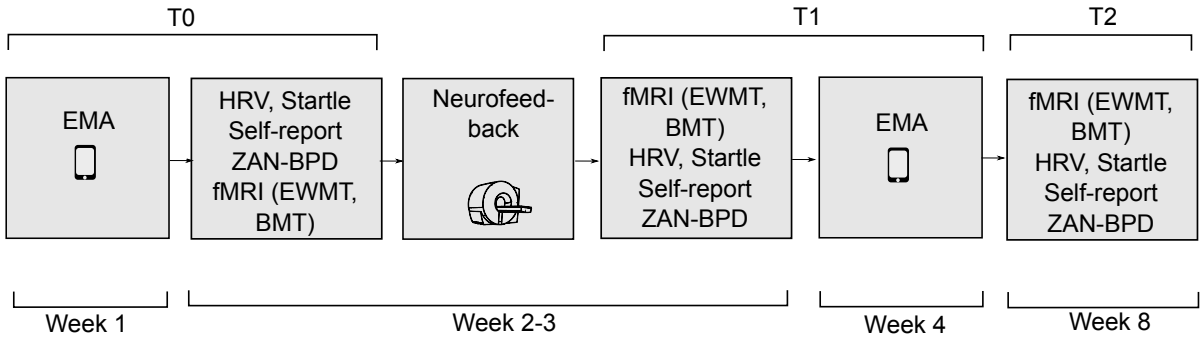

Figure 4.1. Participants received a total of 4 runs of amygdala neurofeedback training (weeks 2-3). Runs were administered on 3 different days, with run 2 and 3 being administered consecutively on the second training day. At baseline (T0), participants completed an emotion test battery: ecological momentary assessment (EMA) was assessed on 4 consecutive weekdays before neurofeedback started (week 1). At the beginning of week 2, the Zanarini Rating Interview (ZAN-BPD), self-report questionnaires, heart rate variability (HRV), and an emotion regulation task with emotion-modulated startle (Startle) was administered. Participants also answered an Emotional Working Memory Task (EWMT) and a Backward Masking Task (BMT) during fMRI, immediately before completion of the first neurofeedback session. From 2-7 days later, participants completed the next neurofeedback session (visit 2), followed by 2-7 days for their third and final session (visit 3). During these sessions, participants were instructed to downregulate a thermometer, with activity of the right amygdala, while watching aversive pictures. Details of the neurofeedback procedure can be found in the supplement. Immediately after the last neurofeedback session (end of week 3), the test battery was administered a second time (T1). The follow-up visit (T2) was completed 6 weeks after visit 3, and was identical to T1, excluding EMA (week 8).

4.3.3 Neurofeedback

4.3.3.1 Procedure

Subjects were instructed to look at negative pictures (without feedback, 'view' condition), or downregulate a colored thermometer bar, representing brain activation while watching negative pictures ('down' condition), respectively. Participants were not given a particular strategy to downregulate. Rather, they were told to assess what strategy worked best for them. In the 'view' condition, a picture with negative emotional content was presented for 18 seconds, followed by a fixation cross on a grey background ('rest,' 12 seconds). In the 'down' condition, pictures were presented with feedback. After each neurofeedback session, participants were asked which strategies they used to downregulate (s. supplement for details). For details, see Figure 4.2. Three participants had to be excluded from the statistical analysis due to technical problems in session 2 and 4 (logfiles were not available).

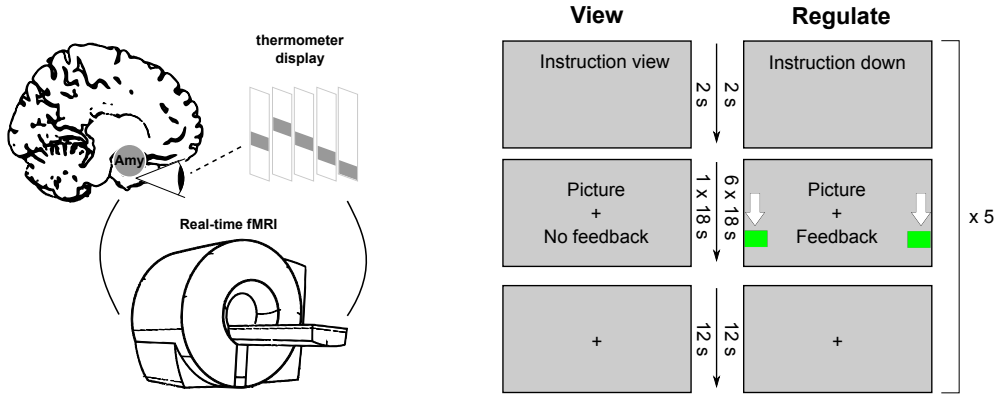

Figure 4.2. Experimental procedure of a neurofeedback training run. Participants viewed aversive pictures, with a feedback signal from their amygdala BOLD response, which is depicted as a thermometer. They were instructed to try to downregulate the temperature, representing their brain activation. They were also told to consider the temporal delay of the BOLD response, resulting in a time lag of the thermometer response (2-5 seconds). Furthermore, they should not close their eyes or shift their gaze from the screen and avoid focusing exclusively on the thermometer or borders of the picture. They should not hold their breath or move their heads. After participants entered the scanner, anatomical and fieldmap scans were acquired. Before the first neurofeedback run, a demonstration trial was presented without fMRI scanning. Subjects were instructed beforehand to look at the picture (without feedback), or to downregulate the thermometer signal. The neurofeedback consisted of 'down' and 'view' conditions, respectively. In the 'view' condition, a picture with negative emotional content was shown for 18 s, followed by a fixation cross on a grey background ('rest,' 12 seconds). In the 'down' condition, pictures were presented with feedback. Six pictures were presented in a 'down' block, each for 18 seconds (108 seconds total). The order of conditions was fixed, with alternating 'view' and 'down' blocks. In total, there were 5 'down' blocks and five 'view' blocks. After the last block, participants were instructed to rate their perceived regulation success ('Were you able to regulate the display?') on a 10-level visual analogue scale (results can be found in the supplement).

4.3.3.2 Online fMRI Data Analysis

The neurofeedback signal was computed as the fMRI percent of signal change, relative to the global mean, and updated every second and displayed as a colored bar. The BOLD signal data were calculated online from voxels within a right amygdala mask, produced with the Harvard-Oxford brain atlas with a probability threshold of 25%. Details of fMRI acquisition, real-time fMRI analysis, and feedback presentation are in the supplement, and were published by Paret et al. (2018).

4.3.4 Assessments

4.3.4.1 Verbal Report: Interviewer- and Self-Assessment

Self-assessment included several questionnaires on different aspects of BPD psychopathology and emotion regulation.

To test our hypothesis on changes in emotion dysregulation, we used the difficulties with the Emotion Regulation Scale (DERS; Gratz & Roemer, 2004), a 36-item questionnaire that assesses levels of emotion regulation problems. The DERS is comprised of six subscales (nonacceptance of emotional response difficulties in engaging in goal-directed behavior, impulse control difficulties, lack of emotional awareness, limited access to emotion regulation strategies, and lack of emotional clarity). The DERS was found to have adequate construct and predictive validity and good test-retest reliability over a period of 4-8 weeks (ρ_I = .88; Gratz & Roemer, 2004).

To test our hypothesis on changes in clinical status, we assessed the Zanarini rating scale for BPD (ZAN-BPD; Zanarini et al., 2003). The ZAN-BPD is a semi-structured interview and reflects a 1-week time frame. The nine criteria for BPD were rated on a five-point anchored rating scale of 0-4 by trained psychologists (JZ, CP, SM), yielding a total score between 0-36. The ZAN-BPD demonstrates good reliability (Cronbach's α= 0.85), with convergent and discriminant validity (Zanarini et al., 2003). One participant was excluded for the statistical analysis of the ZAN-BPD, because she did not do the interview at T2.

In addition to these two measures, we assessed several other questionnaires which were used for explorative analyses: The German version of the Emotion Regulation Skills Questionnaire (SEK-27; Berking & Znoj, 2008; Grant, Salsman, & Berking, 2018) was used to assess emotion regulation skills. The SEK-27 is a 27-item self-report instrument that utilizes a 5-point Likert-type scale (0 = not at all to 4 = almost always) to assess the respondent's adaptive emotion regulation skills the previous week. The SEK-27 comprises six subscales: (1) awareness, (2) clarity, (3) understanding, (4) modification, (5) acceptance, and (6) tolerance. In addition to the subscales, the SEK-27 provides a total score, computed as the average of all items. The SEK-27 showed adequate internal consistency (Cronbach's α= .90 for the total score, and .68-.81 for subscales), as well as adequate test-retest reliability (r = .75 for total score). The Affective Lability Scale (ALS; Harvey, Greenberg, & Serper, 1989), a 54-item self-report scale, was used to measure changeable affect. ALS items assessed subjects' perception of their tendency to vary between what they considered a normal mood versus those of anger (ANG), depression (DEP), elation (ELA), and anxiety (ANX), with a tendency to oscillate between depression and elation

(BIP), or between states of anxiety and depression (ANXDEP). Each item was rated on a four-point scale (scored 0–3 inclusive) from ‘‘very undescriptive’’ to ‘‘very descriptive’’ of themselves. The ALS total is the mean of six subscales for individual affect shifts, and showed good internal consistency (among subscales, alpha range = .76–.86). The Toronto Alexithymia Scale (TAS-26; Bagby, Parker, & Taylor, 1994), a 26-item scale, was used to measure alexithymia in three dimensions: *difficulty identifying feelings*, *difficulty-describing feelings*, and *externally-oriented thinking*. The TAS-26 displays adequate reliability, ranging from r = .67 to r = .84. We further used the Dissociation Tension Scale (DSS; Stiglmayr, Schmahl, Bremner, Bohus, & Ebner-Priemer, 2009) to assess dissociative symptoms, with the short version of the UPPS Impulsive Behavior Scale (Cyders, Littlefield, Coffey, & Karyadi, 2014) to control baseline impulsivity.

4.3.4.2 Everyday Experience: EMA

To measure affective instability and emotion regulation during participants’ everyday lives, we used a smartphone programmed with the movisensXS app (Movisens GmbH, Karlsruhe, Germany) as an electronic diary. The e-diary emitted a prompting signal according to a stratified random schedule, with 12 assessments per day between 9:00 a.m. and 10:00 p.m. on four consecutive workdays. Thus, the 13-hour assessment period of each day was divided into 12 intervals, with assessments scheduled at random within each one. At each prompt, we assessed participants’ current affective state using five questions about positive affect (PA) and five questions about negative affect (NA), based on the affective circumplex model (Russell, 1980). To assess participants’ current dissociative state, we used the DSS-4, including an item asking about aversive tension (5 items; Stiglmayr et al., 2009). We also assessed participants’ perceived control over their emotions with two items (“When the phone rang, I felt like I could control my feelings” and **“**When the phone rang, I felt overwhelmed by my feelings”**)**. The wording of all items can be found in the supplement. We determined the person-mean of the repeated assessments, as well as the mean squared for successive differences (MSSD) as an established instability index for each person and for both assessment periods (i.e., for both the pre- (T0) and post-(T1) neurofeedback training EMA assessment).

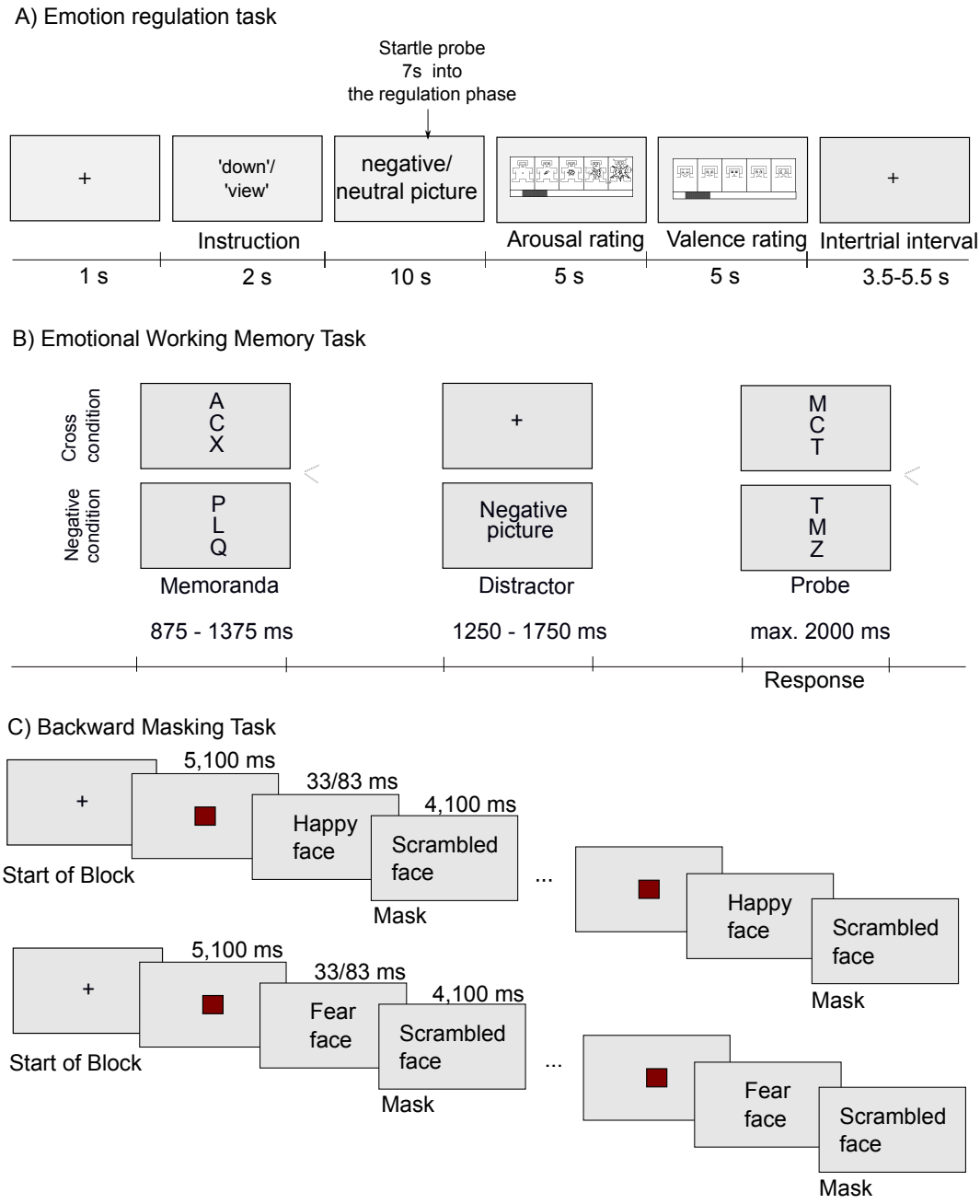

Figure 4.3. A) Experimental procedure of the emotion regulation task. Participants were instructed either to view negative and neutral pictures without modifying their emotions ('view'; 'neutral' condition, respectively) or to downregulate their feelings toward negative pictures ('down' condition). Each trial began with a 2,000 milliseconds presentation of an instructional cue ('view', 'down'), followed by a fixation cross displayed for 1,000 milliseconds. Next, a neutral or negative picture was presented for 10,000 milliseconds. A startle probe (50 milliseconds s, 95 dB white noise burst) was presented through headphones at 6,500 milliseconds – 9,500 milliseconds into the regulation phase). Self-assessment Manikins (SAM Ratings; Bradley & Lang, 1994) were presented after presentation of each picture. Participants rated a 1-9 Likert scale on how positive / negative and aroused / calm they felt at that moment. Lower scores on the valence scale indicated that they felt more positive; lower scores on the arousal scale indicated that they felt calmer. Intertrial intervals were jittered between 3,500 and 5,500 milliseconds. Details about stimuli and the procedure can be found in the supplement. B) Experimental procedure of the Emotional Working Memory Task (EWMT), which is an adapted

Sternberg item recognition task (Sternberg, 1966) modified by Oei and colleagues (Krause-Utz et al., 2012; Oei et al., 2012). Each trial started with the presentation of a set of three letters (memoranda, 875–1,375 milliseconds). After a delay phase of 1,250 –1,750 milliseconds, another set of three letters appeared on the screen (probe, 2,000 milliseconds). Next, a blank screen appeared (intertrial interval, 550–1,050 milliseconds). Participants had to press the left or right button to indicate whether they recognized one of the memoranda-letters in the probe. In half of the trials, one of three memoranda was present. During the delay interval, no distractor (i.e., a fixation cross; 'cross' condition) or a distractor (i.e., an aversive picture; 'negative' condition) was presented. Details of stimuli and the procedure can be found in the supplement. C) Experimental Procedure of the Backward Masking Task (BMT). Participants were instructed to identify photographs of faces expressing happy or fearful facial expressions. The BMT had a total of four conditions: happy or fearful facial expressions presented for 33 milliseconds or 83 milliseconds. A total of 4 blocks per condition were presented. Each block consisted of 8 consecutive pictures. Each block began with a fixation cross. Next, eight faces were shown for 33 milliseconds or 83 milliseconds, preceded by a red rectangle on a grey background for 5,100 milliseconds and followed by a mask (scrambled face) for 4,100 milliseconds. Details of the stimuli, procedure, and behavioral results are in the supplement.

4.3.4.3 Behavior and Peripheral Physiology: Emotion Regulation Test, Resting HRV

To test changes in emotion regulation, we assessed the emotion-modulated startle during an emotion regulation paradigm, modified from Jackson et al. (2000). For details of the procedure, see Figure 4.3A. In brief, participants were instructed to view negative or neutral pictures ('view,' 'neutral' condition) or to downregulate emotions in response to negative pictures ('down' condition). Seven seconds into the regulation phase, a burst of white noise was presented for 50 milliseconds at 104 dB[10] (startle probe). The eye blink was measured by electromyogram (EMG). The raw EMG signal was sampled at 1,000 Hz, and the gain was amplified by 2,000. High-pass (50 Hz) and low-pass (500 Hz) online filters were applied to the data with AcqKnowledge software (version 3.4; BIOPAC Systems; Goleta, CA, USA). EMG data were integrated over 10 samples and analyzed offline with Clip, a C++ based, semi-automated program (Kinzig et al., 2008). Emotion-modulated startle response was defined as the difference between peak (20–120 milliseconds after stimulus onset) and baseline (20 milliseconds prior to stimulus onset) signal. Amplitudes were transformed to *T*-scores with mean = 50 and *SD* = 10. Responses were averaged in participants for each condition. Emotion-modulated startle amplitudes in the 'neutral' condition were subtracted from the 'down' and the 'view' conditions for statistical comparison. Four participants (n = 4) were excluded for statistical analysis of the

[10] 104 dB startle probes elicited severe discomfort in two participants. Therefore, we turned down the volume to 95 dB for them. We reran startle analyses without the two participants, without the results changing. Thus, we report data from all participants.

startle data due to technical problems at T0 or T1 (n = 2). One participant did not show a startle response at all and one participant did not complete the psychophysiological tests at T2.

To test changes in resting HRV, we recorded the electrocardiogram (ECG) for 6 minutes at 1,000 Hz, with a gain of 2,000. High-pass (50 Hz) and low-pass (500 Hz) online filters were applied to the data with AcqKnowledge software. Offline, ECG waveforms were transformed into the heart rate (beats per minute) and analyzed with Kubios software (Amsterdam, The Netherlands) (Tarvainen, Niskanen, Lipponen, Ranta-Aho, & Karjalainen, 2014). Resting HRV was calculated as the ratio of low frequency (LF) power in the .04-0.15 Hz range and high frequency (HF) power in the 0.15 - 0.40 Hz range, indicative of sympathetic to parasympathetic autonomic balance (HF/LF). Three participants (n = 3) were excluded: two (n = 2) were excluded due to technical problems at T0 or T1, and one (n = 1) did not complete the psychophysiological tests at T2.

4.3.4.4 Brain Responses

To test changes in amygdala response during shortly presented pictures, we conducted the EWMT and the BMT in the fMRI. Details of these procedures can be found in Figure 4.3B and C. Three participants (n = 3) were excluded for the statistical analysis of the EWMT, due to missing button presses all three times (n = 2), and technical problems at T1 (n = 1; logfile not available). One participant was excluded for the statistical analysis of the BMT due to technical problems at T1 (logfile not available).

4.3.5 Offline Data Analysis

4.3.5.1 Neurofeedback Data

4.3.5.1.1 Preprocessing

Data analysis was performed with Matlab (vR2012a)-based SPM12 package (v6225, Wellcome Trust Center for Neuroimaging, London, UK). Preprocessing included slice timing, which was corrected with reference to the middle slice of a volume, realignment of the scans to the first scan of the series, with rigid body transformation and correction of geometric distortions using a voxel displacement map (VDM); this was produced based on fieldmap scans. The functional scans were not warped, given the VDM parameters and corrected for susceptibility-by-movement artifacts (Andersson, Hutton, Ashburner, Turner, & Friston, 2001). A mean image of the functional scans was next computed and coregistered to the anatomical scan of the subject; this scan was segmented with six standard SPM tissue probability maps and normalized to MNI

space. These parameters were used for normalization of functional images. Images were resampled to 2 mm isometric voxels. Functional data were smoothed using an 8 mm kernel (full width at half maximum, FWHM) to account for between-subject variation in anatomical localization. Finally, a high-pass filter (256 seconds cut-off) was added to the general linear model (GLM) to remove slow signal drifts. An autoregressive model was used to account for serial correlations.

4.3.5.1.2 Amygdala Region-of-Interest (ROI) Analysis

We estimated HRFs using the inverse logit model by Lindquist, Loh, Atlas, and Wager (2009) to investigate the hemodynamic amygdala response. First, the eigenvariate was extracted from voxels corresponding to the right amygdala, with the same mask being used for neurofeedback. The eigenvariate was also adjusted for condition effects ('down' and 'view'). HRFs were fitted to each picture presentation interval. The HRF amplitude represents the magnitude of the event-related BOLD response. In addition, we analyzed the area under the curve (AUC). Amplitude estimates and AUC values were compared with SPSS statistics software (v23, IBM Corp. Armonk, NY, USA).

4.3.5.1.3 Amygdala Downregulation Success

We quantified down-regulation success by creating two different indices: First deltas of amygdala amplitudes/AUCs ('view' minus 'down') between the first and last neurofeedback run were created. However, as this index assumes linear improvement and may misrepresent actual learning slopes, we complemented this by calculating the best performance (Paret et al., 2019) of each participant. That is, we determined the largest delta between the 'view' and 'down' condition for each participant across all four neurofeedback runs.

4.3.5.2 EWMT and BMT

4.3.5.2.1 fMRI Acquisition and Analysis

For fMRI acquisition, a 3 Tesla MRI Scanner (Trio, Siemens Medical Solutions, Erlangen, Germany) with a 32-channel head coil was used. T1-weighted anatomical images were acquired with a Magnetization Prepared Rapid Acquisition Gradient Echo sequence (TE =3.03 milliseconds, TR=2.3 seconds, 192 slices and FOV= 256 x 256 mm). Functional images of both EWMT and BMT tasks were acquired with a gradient echo T2* weighted echo-planar-imaging (EPI) sequence with a field of view = 210 mm x 210 mm, voxel size = 3 mm x 3 mm x 3 mm, echo time = 30 milliseconds, TR = 2000 milliseconds with 40 contiguous 3 mm sagittal slices in a 64 x 64 matrix. Head movement artifacts and scanning noise were reduced with head cushions

and headphones in the scanner coil. Preprocessing was comprised of adjusting for variable acquisition time over slices (slice-timing), head motion correction (realignment), normalization of images into a standard three-dimensional space defined by the Montreal Neurological Institute (MNI), and spatial smoothing using an 8 mm Gaussian kernel to increase signal-to-noise ratio.

4.3.5.2.2 First-Level Analysis

For the EWMT, we modeled regressors for the memoranda, probe and response phase, respectively. In addition, each condition was modeled (negative, cross). Parameter estimates from the contrast of interest (negative > cross) were entered into group-level *t*-tests. For the BMT, we modeled regressors for each condition (happy faces 33 milliseconds, happy faces 83 milliseconds, fearful faces 33 milliseconds, and fearful faces 83 milliseconds). All regressors were convolved with the HRF implemented in SPM12. Parameter estimates from the contrast of interest (all conditions versus implicit baseline) were entered into group-level *t*-tests. To test our hypotheses, voxel-wise *t*-tests of parameter estimates for the EWMT contrast negative > cross, and the BMT contrast (all conditions *versus* implicit baseline) were conducted on the first level. The mean contrast value was then extracted from all voxels of the right amygdala, based on the neurofeedback mask.

4.3.6 Statistical Analysis of Assessments

Validated statistical software (SPSS v25; IBM Inc., Armonk, NY, USA) was used for analyses. Missing variables were estimated from available items, based on a Stochastic Regression Imputation (SRI) approach, which improves deterministic regression imputation by imputing a value which includes a random error (van Ginkel, Sijtsma, van der Ark, & Vermunt, 2010), hereby avoiding both bias and overfitting (Enders, 2006). For missing self-report items, the regression model underlying SRI was based on all other items from the questionnaire (within the same assessment). For missing neurofeedback, ZAN-BPD, psychophysiological information, EWMT, and BMT variables, the regression model underlying SRI was based on all other conditions available across assessments. We used the stochastic regression imputation SPSS syntax provided by van Ginkel et al. (2010). All variables (including original and imputed data) were entered into repeated-measures: ANOVA with time (T0, T1, and T2) and condition (if available) as within-subject factors (*$p < .05$). If Mauchly's sphericity test was significant, Greenhouse-Geisser correction was applied to the degrees of freedom.

To limit the risk of false positive results, results from original data are reported in case they differed from results with imputed data. If results from original data do not differ from those

with imputed data, the original data without imputation are not reported. We repeated analyses on measures with a-priori hypotheses (i.e. our primary endpoints: ZAN-BPD and DERS total score, emotion-modulated startle, resting HRV and amygdala reactivity to BMT and EWMT) with a conservative correction for multiple tests (i.e. Bonferroni-correction).

In addition, we ran correlation analyses between amygdala down-regulation (deltas of amygdala amplitudes/AUCs subtracting the 'down' from the 'view' condition) and primary endpoints (i.e. emotion-modulated startle ['view' minus 'down'], resting HRV, amygdala activity to the EWMT [negative > cross contrast] and BMT [all conditions *versus* implicit baseline], ZAN-BPD total score and DERS total score at T0 and T1, respectively). We also ran correlations between down-regulation success indices (deltas of amygdala down-regulation [run 1 minus run 4], best performance) and changes in emotion-modulated startle, resting HRV, amygdala activity to the EWMT and BMT, ZAN-BPD total score and DERS total score using Pearson's r correlation coefficient. Changes in amygdala reactivity during the EWMT (negative > cross contrast) and BMT (all conditions *versus* implicit baseline), emotion-modulated startle ('view' minus 'down'), resting HRV, ZAN-BPD total score and DERS total scores were calculated by subtracting means at T1 from T0. Correlation analyses were limited to our primary endpoints. We did not run correlations with the remaining outcome measures to avoid an increase in chances of false discovery due to multiple testing.

4.4 Results

4.4.1 Amygdala Downregulation Success

Participants downregulated the amygdala BOLD amplitude, $F(1,23) = 9.40$, $p = .01$, $eta^2 = .30$. This effect was driven by a significant difference between 'down' and 'view' at the fourth training run, $t(23) = -2.51$, $p = .02$, $d = -0.51$ - whereas at the first, $t(23) = -.51$, $p = .61$, second, $t(23) = -.77$, $p = .45$, and third, $t(23) = -1.71$, $p = .10$, training run, amygdala BOLD amplitude did not significantly differ between 'down' and 'view' conditions (see Figure 4.4.A). Interaction between condition and run was not significant, $F(3,69) = .43$, $p = .73$, $eta^2 = .02$, showing that the observed improvement of amygdala downregulation over time did not pass the significance level.

Statistical analysis did not support the trend for improvement via training. Similar results were seen for the amygdala AUC: participants could downregulate the amygdala AUC, $F(1,23) = 13.30$, $p < .01$, $eta^2 = .37$, yet this effect was driven by a significant difference between 'down' and 'view' at the second, $t(23) = 2.48$, $p = .02$, $d = .50$, and fourth training run, $t(23) = -2.76$, $p = .01$, $d = -0.56$. In the first, $t(23) = -.97$, $p = .34$ and third, $t(23) = -1.38$, $p = .18$ training run,

amygdala AUC did not significantly differ between 'down' and 'view' conditions (see Figure 4.4.B). Interaction between condition and run was not significant, $F(3,69) = .67$, $p = .57$, $eta^2 = .03$.

From the amygdala AUC and amplitudes (delta between 'view' and 'down') we determined the best run (i.e. largest delta) for each participant. When considering the AUC, 6-8 participants each showed best performance during run 2-4, respectively, whereas only two participants showed best performance at the first run (see Figure 4.11). When considering the amygdala amplitude, best performance was more equally distributed across runs.

4.4.2 Verbal Report

The main effect of time of the ZAN-BPD total score revealed that overall BPD symptoms lessened over time, $F(2, 46) = 5.13$, $p = .010$ (uncorrected[11]), $eta^2 = .18$. Post hoc paired *t*-tests showed a significant reduction from T0 to T1, $t(23) = 3.17$, $p = .004$, $d = .65$, no significant change from T1 to T2, $t(23) = -.62$, $p = .54$, $d = -.13$, and a significant reduction from T0 to T2, $t(23) = 2.22$, $p = .036$, $d = .45$ (see Figure 4.5A). A main effect of time of the DERS total score indicated how difficulties with emotion regulation did change over time, $F(2,46) = 3.78$, $p = .03$ (uncorrected[2]), $eta^2 = .14$ (see Figure 4.5B). Post hoc paired *t*-tests showed a significant reduction from T1 to T2, $t(23) = 2.42$, $p = .025$, $d = .49$ and from T0 to T2, $t(23) = -2.40$, $p = .024$, $d = .49$. Original data without imputation revealed nonsignificant main effect of time of the DERS total score, $F(2,40) = 2.48$, $p = .10$, $eta^2 = .11$.

Explorative analyses showed a significant main effect of time for the SEK total score, $F(2,46) = 5.90$, $p = .01$, $eta^2 = .20$, indicating change of emotional competence over time (see Figure 4.5C). Post hoc paired *t*-tests revealed that emotion regulation skills and their efficacy significantly increased from T1 to T2, $t(23) = -2.71$, $p = .01$, $d = -.55$, and significantly increased from T0 to T2, $t(23) = -2.73$, $p = .01$, $d = -.56$. Overall alexithymia symptoms did not significantly change over time, as indicated by the TAS-26 total score, $F(1.36, 31.22) = 2.76$, $p = .10$, $eta^2 = .10$ (see Figure 4.5D). The subscales 'Difficulty describing feelings' and 'External thinking' did not significantly change, whereas 'Identification of one's feelings' did significantly change over time, $F(2,46) = 6.25$, $p < .01$, $eta^2 = .21$ (see Table 4.1). No significant main effect of time was found for the total score of the ALS and the DSS-21 (see Table 4.1).

[11] no correction for multiple testing was applied.

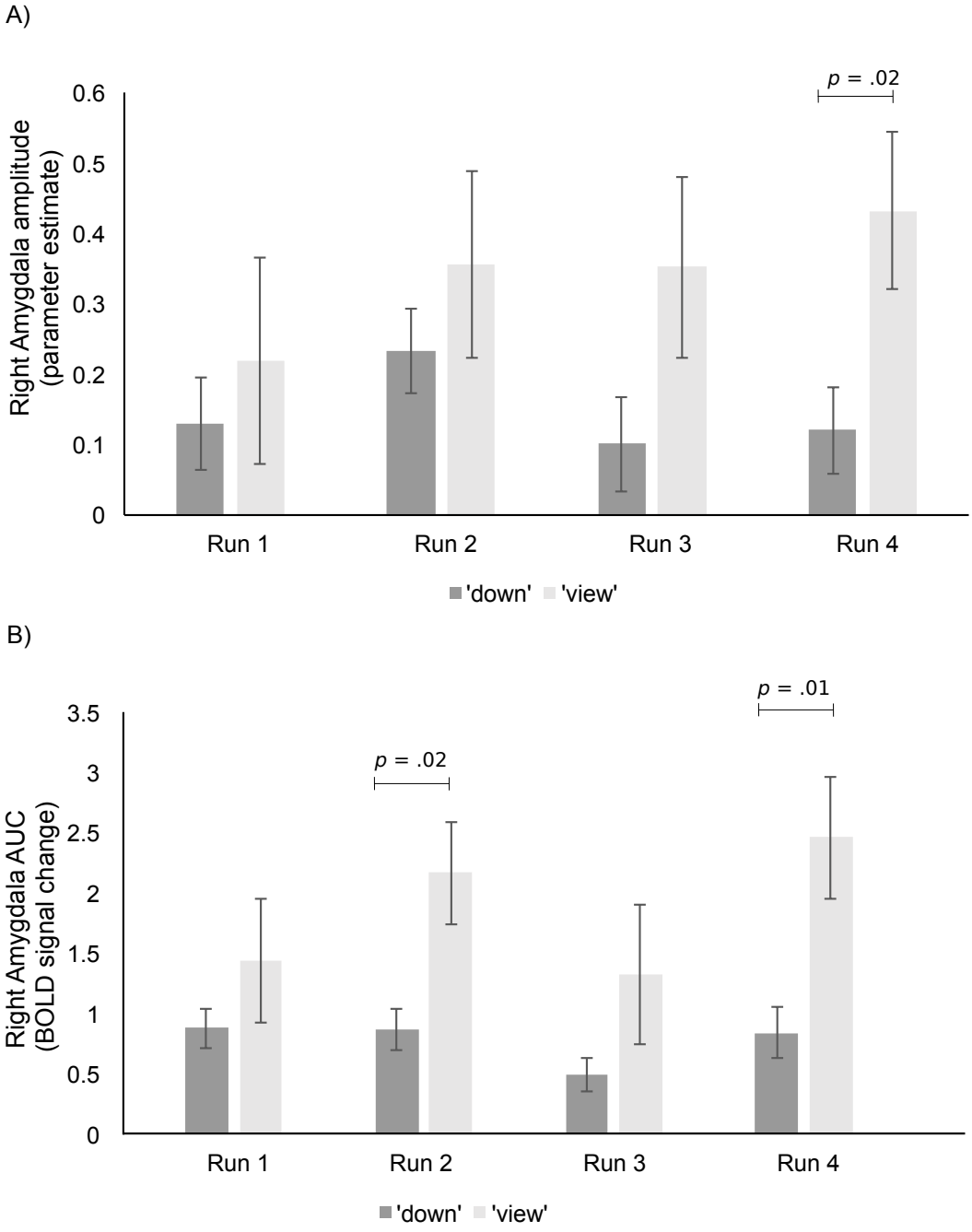

Figure 4.4. Amygdala amplitude and AUC in the 'down' and 'view' conditions at each neurofeedback session. A) Participants significantly downregulated the amygdala amplitude with neurofeedback comparing the 'down' with the 'view' condition at run 4. B) participants significantly downregulated the amygdala AUC at run 2 and run 4. Error bars represent standard error of the mean (*SEM*). AUC = area under the curve.

Original data without imputation revealed a nonsignificant main effect of time for the SEK total score, $F(1.59, 27.07) = 3.34$, $p = .06$, $eta^2 = .16$, but a significant main effect of time for the

TAS total score, $F(2,36) = 5.33$, $p = .01$, $eta^2 = .27$. Post hoc paired *t*-test of original TAS total scores revealed a significant reduction from T1 to T2, $t(18) = 2.78$, $p = .01$ and from T1 to T2, $t(18) = 2.56$, $p = .02$. Results of the original TAS subscales can be found in the supplement. To follow a conservative approach, we further discuss and interpret the original instead of the imputed data in case they differ from the imputed data. A detailed perspective of interviewer- and self-assessment results at T0, T1, and T2 can be found in Table 4.1.

4.4.3 Everyday Experience: EMA

Explorative paired *t*-tests contrasting T0 and T1 revealed a significant reduction of mean negative affect (NA), $t(23) = 3.46$, $p < .01$, $d = .70$, a significant reduction of mean inner tension, $t(23) = 3.27$, $p < .01$, $d = .67$, a nonsignificant reduction of mean dissociative symptoms, $t(23) = 1.85$, $p = .08$, $d = .38$, and a significant increase of mean emotion regulation control, $t(23) = -2.07$, $p = .05$, $d = -.42$ (see Figure 4.6.A). No significant effects were found for mean positive affect (PA), $t(23) = 1.28$, $p = .21$, $d = .26$. Paired *t*-tests of the MSSDs revealed a significant reduction of instability in PA, $t(23) = 2.30$, $p = .03$, $d = .47$, NA, $t(23) = 2.73$, $p = .01$, $d = .56$, and inner tension, $t(23) = 3.41$, $p < .01$, $d = .18$ (see Figure 4.6.B). No significant effects were found for the instability of dissociative symptomatology, $t(23) = 1.71$, $p = .10$, $d = .35$. Adherence to prompts was 69.21% (*SD* =18.18) at T0 and 63.87% (*SD* = 17.34) at T1, which is satisfactory. There was no significant difference in adherence between T0 and T1, $t(23) = .16$, $p = .12$.

4.4.4 Behavior: Emotion Regulation Test and Resting HRV

As hypothesized, patients could downregulate negative emotions more effectively after training, indexed by a significant decrease of the emotion-modulated startle in the ‘down’ compared to the ‘view’ condition after training, $F(2,46) = 4.23$, $p = .02$, $eta^2 = .16$ (uncorrected). There was no significant main effect of time, $F(2,46) = .90$, $p = .42$ and condition, $F(1,23) = .39$, $p = .54$. The interaction was due to a significant difference between the ‘down’-‘neutral’ and the ‘view’-‘neutral’ condition at T1, $t(23) = -2.15$, $p = .04$, $d = -.44$. In T0 and T2, in contrast, patients did not significantly decrease startle in the ‘down’-‘neutral’ vs ‘view’- ‘neutral’ comparison (see Figure 4.7A). Results from original data revealed similar results, except for the post hoc paired *t*-tests of original emotion-modulated startle data: Emotion-modulated startle was lower in the ‘down’ than the ‘view’ condition at T1, but this effect was only at the trend-level, $t(17) = -2.01$, $p = .06$, $d = -.47$.

Table 4.1

Means (SD) of demographics, clinical characteristics and questionnaires.

	T0		T1		T2		*Test-Statistics*			
							F	*df*	*p*	
Demographics										
Age mean (SD)	33.42	(11.10)								
Clinical Characteristics										
Number of BPD criteria fulfilled (DSM-IV)	6.61	(1.03)								
Borderline Symptoms ZAN-BPD										
Total (SD)	7.48	(3.88)	5.09	(3.99)	5.55	(3.33)	5.13	2, 46	.01	T0>T1
Affect (SD)	2.9	(1.36)	2.09	(1.41)	2.30	(1.25)	3.43	2, 46	.04	T0>T1
Cognition (SD)	2.17	(1.74)	1.35	(1.46)	1.25	(1.17)	4.52	2, 46	.02	T1>T2
Impulsivity (SD)	0.96	(1.46)	0.96	(1.43)	0.85	(0.80)	0.11	1.57, 36.14	.85	
Interpersonal Relationships (SD)	1.48	(1.13)	0.96	(1.30)	1.15	(1.51)	3.82	2, 46	.03	T0>T1
Difficulties with Emotion Regulation Scale (DERS)										
Total (SD)	123.29	(17.49)	121.71	(18.73)	113.20	(20.88)	3.80	2, 46	.03	
Nonaccepatence (SD)	20.79	(5.32)	21.62	(5.56)	19.45	(4.33)	2.08	2, 46	.14	
Goals (SD)	20.08	(3.37)	19.25	(3.76)	18.44	(3.72)	3.70	2, 46	.03	T0>T2
Impulse (SD)	18.21	(4.01)	17.63	(3.77)	16.96	(4.18)	1.02	2, 46	.37	
Awareness (SD)	20.21	(5.30)	19.75	(5.75)	18.79	(5.38)	3.18	2, 46	.05	
Strategies (SD)	28.79	(4.73)	28.58	(.,52)	25.86	(5.98)	4.52	2, 46	.02	T0>T2,T1>T2
Clarity (SD)	15.21	(4.18)	14.88	(3.95)	13.69	(3.76)	1.99	2, 46	.25	
Affect Lability Scale (ALS)										
Total (SD)	89.42	(16.86)	92.89	(17.56)	88.98	(18.66)	0,45	2, 46	.64	
Depression	17.33	(4.53)	18.76	(4.31)	19.18	(3.88)	1.37	2, 46	.26	
Elation	20.04	(5.69)	18.98	(6.11)	18.65	(6.68)	.48	2, 46	.62	
Depression Elation	13.88	(4.06)	14.51	(4.51)	14.89	(4.34)	.45	2, 46	.67	
Anxiety	12.17	(2.53)	13.90	(3.51)	12.25	(3.62)	2.41	2, 46	.10	
Anger	12.63	(2.60)	12.10	(3.26)	10.89	(3.43)	2.28	2, 46	.11	
Anxiety Depression	13.38	(3.66)	14.59	(4.08)	13.12	(4.40)	0.90	2, 46	.40	
Toronto Alexithymia Scale (TAS-26)										
Total (SD)	54.17	(10.29)	53.50	(9.56)	50.54	(10.78)	2.76	1.36, 31.22	.096	
Identification of one's feelings	22.67	(4.05)	21.29	(3.98)	19.82	(4.84)	6.25	2, 46	.004	T0>T2
Difficulty Describing Feelings	15.96	(3.61)	16.58	(3.93)	15.39	(3.83)	1.39	2, 46	.26	
External thinking	15.54	(4.81)	15.63	(4.99)	15.33	(4.97)	0.05	1.47, 33.72	.95	

Table 4.1 *(continued)*

	T0		T1		T2		*Test-Statistics*			
							F	*df*	*p*	
Emotion Regulation Skills Questionnaire (SEK-27)										
Total (SD)	78.61	(21.31)	77.87	(20.27)	86.83	(18.26)	5.90	2, 46	.01	T0<T2
Awareness	9	(3.24)	9.71	(3.45)	10.27	(2.90)	3.85	2, 46	.03	T1<T2
Clarity	9	(2.47)	9.00	(2.96)	9.63	(2.49)	4.65	10.31, 2.22	.03	T0<T2
Sensations	8.58	(2.98)	8.25	(2.94)	10.33	(2.63)	4.08	2, 46	.02	T0<T2
Understanding	9.33	(3.23)	8.71	(3.16)	9.73	(2.43)	4.64	2, 46	.02	T1<T2
Accepatnce	8.5	(2.87)	8.37	(2.39)	8.59	(2.76)	5.96	2, 46	.01	T1<T2
Tolerance	8.06	(2.86)	7.83	(2.55)	9.77	(2.86)	0.90	2, 46	.41	T1<T2
Readiness to confront distressing situations	9.58	(3.12)	9.25	(3.30)	9.45	(2.71)	0.17	2, 46	.84	
Self-support	8.75	(2.79)	8.79	(2.70)	9.77	(2.85)	3.62	2, 46	.04	
Modification	7.79	(2.77)	7.96	(2.74)	8.91	(2.26)	2.88	2, 46	.07	
UPPS										
Urgency (SD)	2.9	(0.46)	-	-	-	-				
Pre (SD)	2.37	(0.45)	-	-	-	-				
Pers (SD)	2.54	(0.51)	-	-	-	-				
SS (SD)	2.76	(0.54)	-	-	-	-				
DSS-21										
Intensity (SD)	18.49	(11.65)	14.80	(9.42)	14.48	(10.43)	2.16	1.36, 31.32	.12	
Duration (SD)	18.49	(1.28)	2.24	(1.42)	2.18	(1.36)	1.47	1.23, 35.41	.24	
Current comorbidities N (%)										
Major Depression	6	(24%)								
Major Depression lifetime	22	(88%)								
Dysthymia	4	(16%)								
Double Depression	3	(12%)								
Panic Disorder	3	(12%)								
Social Phobia Disorder	4	(16%)								
Specific Phobia	5	(20%)								
PTBS	6	(24%)								
Anorexia Nervosa	1	(4%)								
Bulimia Nervosa	1	(4%)								
Binge Eating Disorder	2	(8%)								

Table 4.1 *(continued)*

	T0		T1	T2	*Test-Statistics*		
					F	*df*	*p*
Psychotropic Medication N (%)							
SSRI	3.00	(12.50)					
SNRI	4.00	(16.70)					
Tricyclica	3.00	(12.50)					
Other Antidepressants	3.00	(12.50)					
Neuroleptics	5.00	(20.80)					
Anticonvulsants	2.00	(8.30)					
Unmedicated	10.00	(41.70)					

Note. SD =standard deviation; UPPS = Impulsive Behavior Scale; DSS-21 = Dissociation Tension Scale; PTBS = posttraumatic stress disorder; SSRI = Selective serotonin reuptake inhibitor; SNRI = Serotonin–norepinephrine reuptake inhibitor.

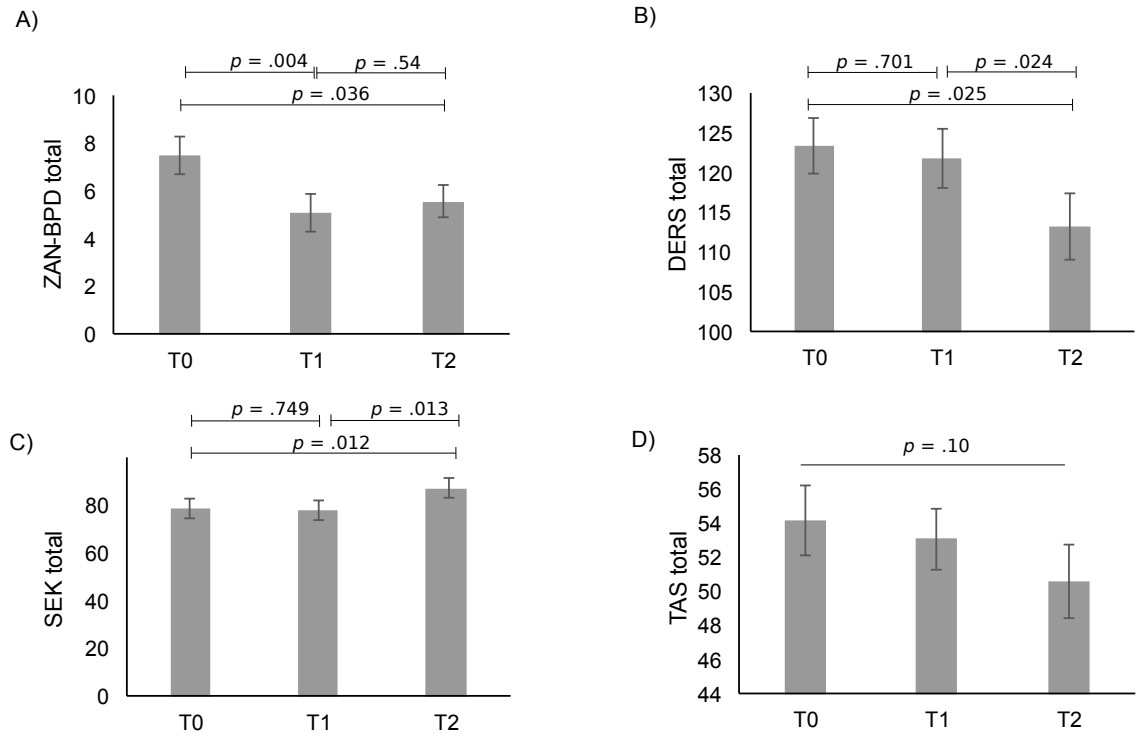

Figure 4.5. Diagnostic interview and self-assessment results at T0, T1, and T2. A) BPD psychopathology significantly ($p < .05$, uncorrected) improved from T0-T1 and from T0-T2. B) Difficulties with the Emotion regulation Scale (DERS) total score significantly[3] reduced from T1-T2 and from T0-T2, indicating that difficulties in emotion regulation decreased over time. C) Self-assessment of the emotional competencies (Selbsteinschaetzung Emotionaler Kompetenzen, SEK) total score significantly increased from T1-T2 and from T0-T2, showing an increase in emotional competence over time. D) A trend in reduction of alexithymia was observed but did not pass the significance test. Error bars represent standard error of the mean (SEM).

Arousal ratings of the emotion regulation test ('down'-'neutral'; 'view'-'neutral' significantly *changed* over time, corroborated by a significant main effect of time, $F(2,46) = 18.64$, $p < .01$, $eta^2 = .51$. A significant main effect of condition indicated that overall arousal ratings were significantly lower in the 'down'-'neutral' than in the 'view'-'neutral' condition, $F(1,23) = 3.33$, $p = .03$, $eta^2 = .23$. Post hoc *t*-tests between the 'down' and 'view' condition at T0, T1, and T2, respectively, revealed no significant effects, all *p*s > .10. Valence ratings of the emotion regulation test ('down'-'neutral;' 'view'-'neutral') substantially changed over time, corroborated by the main effect of time, $F(2,46) = 3.22$, $p = .05$, $eta^2 = .16$. Interaction of time and condition was not significant.

Resting HRV did not change over time, $F(1.33, 23.870) = 1.27$, $p = .23$, $eta^2 = .07$ (see Figure 4.7B).

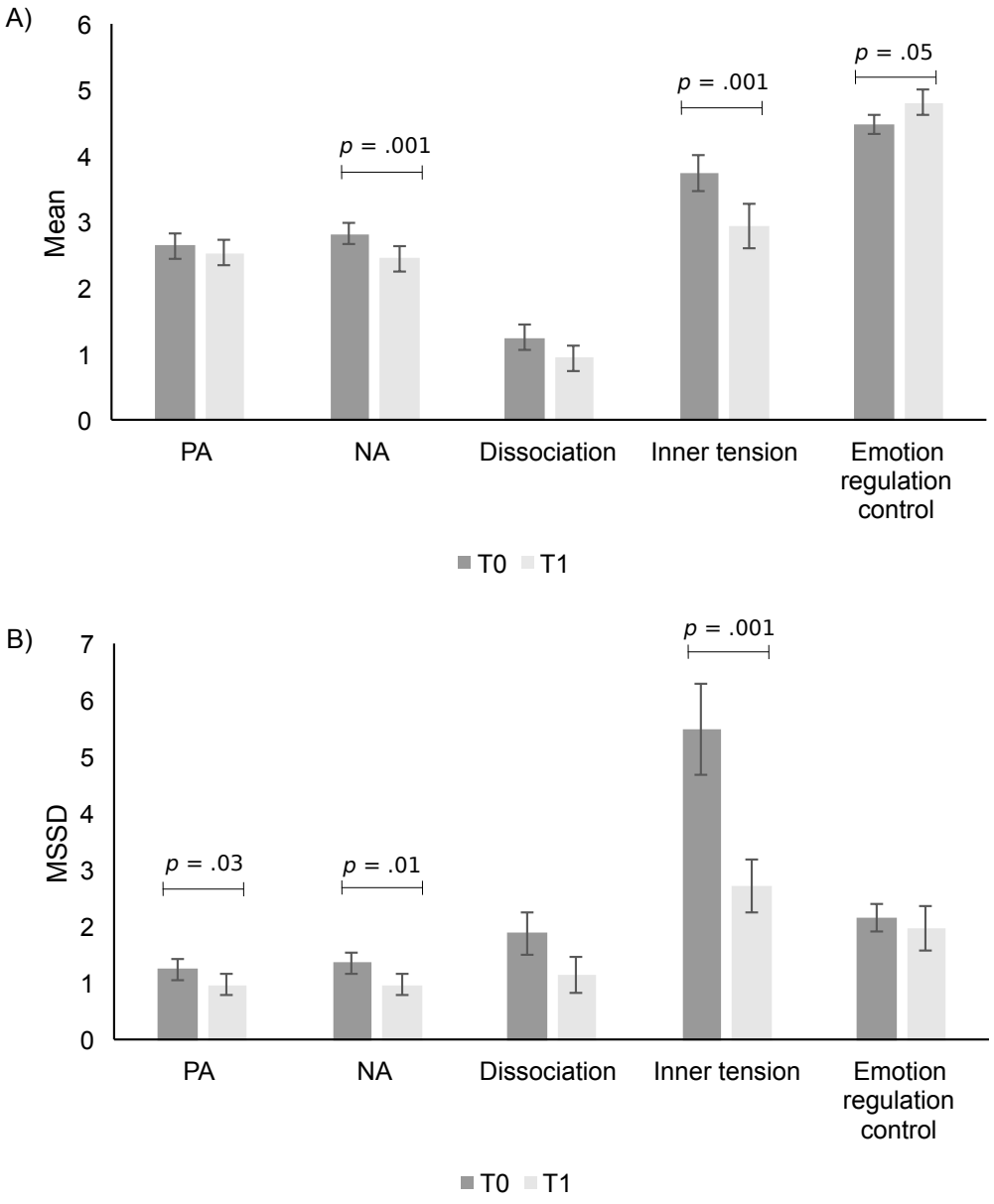

Figure 4.6. Ecological Momentary Assessment (EMA) data were assessed before (T0) and after (T1) neurofeedback training. A) Mean negative affect (NA) and inner tension significantly decreased, and perceived control over one's own emotions increased from T0-T1. Perceived control over one's own emotions was assessed with two items: asking how much participants felt they can control / cannot control their emotions now (see supplement for exact wording). Mean dissociation and positive affect (PA) did not significantly change from T0-T1. B) Mean squared successive differences (MSSD; i.e., hour-to-hour variability) of PA, NA, and inner tension significantly decreased from T0-T1, while the hour-to-hour variability of dissociation and perceived control over one's own emotions did not significantly change from T0-T1. Error bars represent standard error of the mean (SEM).

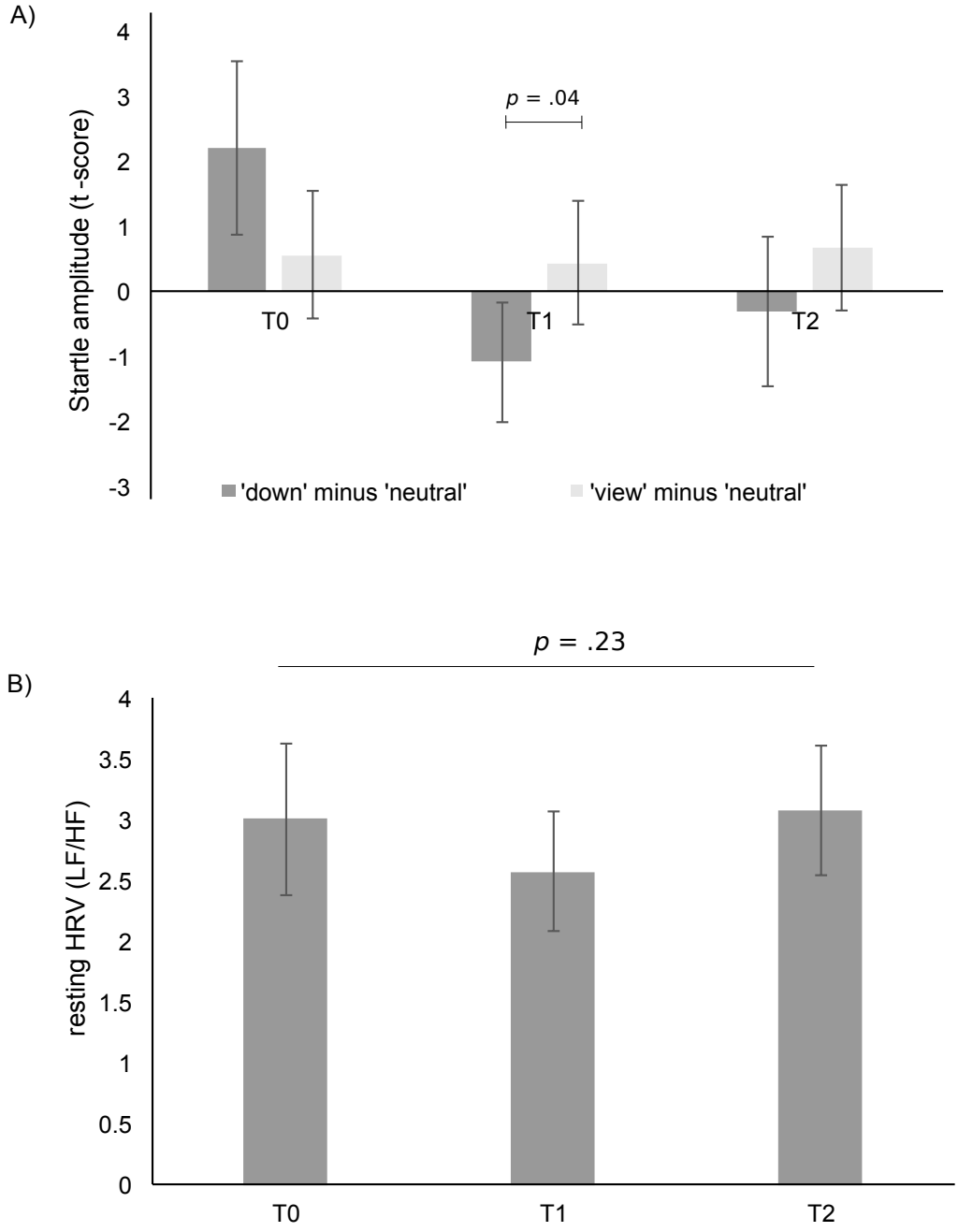

Figure 4.7. A) Mean startle amplitudes in the 'down' and 'view' condition at each assessment (T0, T1, and T2). Mean amplitudes represent the T-score converted to difference scores ('down' minus 'neutral' and 'view' minus 'neutral'). Results indicate a significant reduction of emotion-modulated startle amplitude in the 'down' versus 'view' condition at T1, but not at T0 or T2. B) Resting HRV did not significantly change over the course of the study. Error bars represent standard error of the mean (*SEM*).

4.4.5 Brain Responses: EWMT and BMT

EWMT accuracy was not significantly different between conditions and did not improve, all *p*s > .10. EWMT reaction times were significantly increased in the 'negative' versus 'cross' condition, $F(2,46) = 3.66$, $p = .03$, $eta^2 = .14$, but did not change over time, as corroborated by a nonsignificant interaction of time and condition, $p > .10$. Contrary to our hypothesis, amygdala reactivity did not change over EWMT sessions, indicated by a nonsignificant interaction effect, $F(2,46) = .87$, $p = .43$. $eta^2 = .04$ (see Figure 4.8A).

Regarding the BMT, amygdala activity to quickly-presented happy and fearful faces did not significantly change after neurofeedback, although a decreasing trend could be observed, $F(2,46) = 2.74$, $p = .08$, $eta^2 = 0.11$ (see Figure 4.8B).

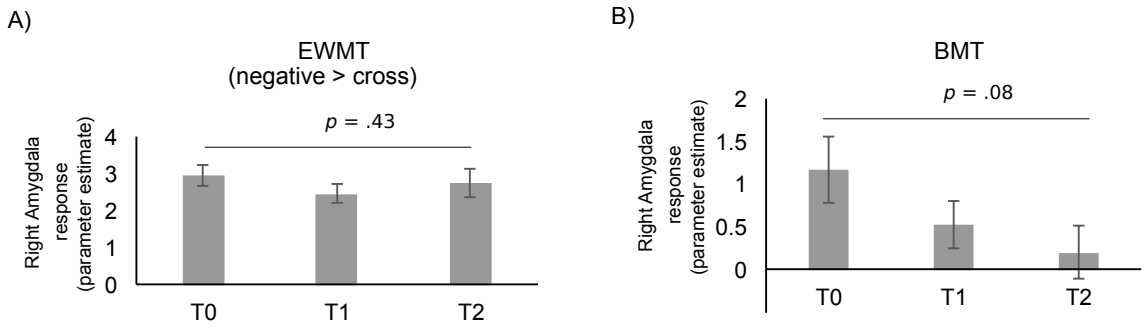

Figure 4.8. A) Right amygdala hemodynamic response during the Emotional Working Memory Task (EWMT) for each visit (T0, T1, T2). Amygdala hemodynamic response was assessed using fMRI during exposure to negative pictures, versus pictures depicting a fixation cross (negative > cross). B) Right amygdala hemodynamic response during the Backward Masking Task (BMT) for each visit (T0, T1, and T2). Amygdala hemodynamic response was assessed using fMRI during exposure to fearful and happy faces. Error bars indicate standard error of the mean (*SEM*).

4.4.6 Correction for Multiple Testing

We repeated analyses on primary endpoints (ZAN-BP and DERS total score, emotion-modulated startle, resting HRV, amygdala reactivity to EWMT and BMT) using a Bonferroni-corrected alpha-level of $p = .008$. None of the main effects remain statistically significant with Bonferroni-correction.

4.4.7 Correlations between Neurofeedback Success and Outcome Measures

There were no significant correlations between amygdala down-regulation at run 4 and any of the primary endpoints (see Table 4.5). Amygdala amplitude down-regulation at run 1 correlated significantly and positively with resting HRV at T0 (r = .51, p = .01, N = 24; not significant with Bonferroni-correction), and significantly and negatively with the ZAN-BPD total score at T0 (r = -.45, p = .03, N = 24; not significant with Bonferroni-correction). Changes in downregulation of amygdala amplitude and AUC during neurofeedback and changes in primary endpoints did not significantly correlate (see Table 4.5 for results). Similarly, there were no significant correlations between participants' best performance and changes in measures with a-priori hypotheses.

4.5 Discussion

This is the first study assessing alterations in a variety of emotion processing and emotion regulation measures after amygdala neurofeedback training for BPD. This is an important step towards advancing neurobiological models and treatment for BPD, using endogenous neuromodulation with neurofeedback. Our results show that BPD patients were able to downregulate amygdala activation with neurofeedback. BPD psychopathology, emotion dysregulation, and affective instability improved at several levels of analysis, including self-report, startle modulation, and experience in everyday life. With regard to our primary endpoints, effects failed to pass significance level when applying a conservative correction for multiple tests. Therefore, our results need to be treated as preliminary and should be replicated by an independent study. In line with our hypotheses, we observed changes in the ZAN-BPD interview, suggesting that subjectively-experienced BPD symptoms improved over the course of the study. These results are in accordance with other studies reporting associations of amygdala normalization and reductions in BPD psychopathology, and are in harmony with the notion that amygdala response is a critical mechanism of remission with BPD (Goodman et al., 2014; Schnell & Herpertz, 2007).

In addition, the present results on our EMA analyses indicate that negative affect and affective instability experienced in daily life reduced over time as well. Affective instability in BPD is supposed to arise from high sensitivity of neural systems involved in the generation of an emotional state, in combination with a severe emotion regulation deficit (Koenigsberg, 2010; Putnam & Silk, 2005). Increased amygdala activation has been interpreted as impairment in top-down control of the prefrontal cortex and may therefore contribute to affective instability (Dillon & Pizzagalli, 2007; Herpertz, Schneider, Schmahl, & Bertsch, 2018; Schulze et al.,

2016). Amygdala neurofeedback training might be specifically suited to target the neural mechanisms of affective instability in precision psychiatry, although more research is needed for corroboration.

On the physiological level, we found an improvement in emotion regulation after training, evidenced by reduced startle-response, which suggests that participants improved their ability to regulate negative emotions. The neural pathway of the emotion-modulated startle involves midbrain neurons, mainly controlled by the central nucleus of the amygdala (Rosen, Hitchcock, Sananes, Miserendino, & Davis, 1991). Enhanced amygdala activation leads to enhanced startle response (Davis, Walker, & Lee, 1999; Rosen & Davis, 1988). Given the strong relation between emotional dysregulation, enhanced amygdala activation, and enhanced startle response, our results suggest that emotion-modulated startle is a sensitive measure for investigating therapeutic effects of amygdala neuromodulation. Improvements in emotion regulation, assessed with the emotion-modulated startle, however, faded to the follow-up test; that is, some training effects did not persist for 6 weeks. Future studies must gain more stable effects, such as adding booster sessions or homework between sessions (Paret et al., 2019).

Contrary to our hypotheses, we did not find significant changes at the brain level. That is, amygdala response to negative pictures and facial expressions did not significantly lessen after neurofeedback training. In addition, no significant changes in resting HRV were observed. A possible explanation could be that these tasks simply do not measure the mechanisms that are trained with neurofeedback. During the EWMT and BMT, participants viewed emotional pictures, but were not explicitly told to regulate their emotions. Rather, these tasks measure the spontaneous response to negative stimuli. Likewise, resting HRV is a measure of autonomic flexibility representing the capacity for spontaneously regulated emotional responses (Appelhans & Luecken, 2006). In contrast, participants showed improvements in the emotion regulation test after training. The emotion regulation test explicitly instructed participants to downregulate negative emotions. In other words, our treatment might not alter the spontaneous response to negative emotions. Rather, participants might have acquired new or already-strengthened existing emotion regulation skills.

With respect to alexithymia, i.e. the difficulty to cognitively process emotions, our results suggest a reduction in these symptoms after training. However, we highlight the explorative fashion of this finding and we stress that only the original data showed significant reductions. Nonetheless, our results are in line with a recent study showing that amygdala electrical fingerprint neurofeedback resulted in a larger reduction of alexithymia scores compared to a control group (Keynan et al., 2019). Conversely, neurofeedback studies to increase the amygdala response

showed that the ability to identify or describe one's own emotions (as indicated by a subscale of the Toronto alexithymia scale; TAS), was correlated with the effectivity to increase amygdala activity (Young, Misaki, et al., 2017; Young et al., 2014; Zotev et al., 2011), which suggests that individuals with less symptoms of alexithymia might have better prerequisites to learn increasing their amygdala activity with neurofeedback. Together with our results, these studies indicate that the ability to identify and describe one's own feelings is directly related to the ability to gain control over the amygdala, however further studies are needed to fully understand the relation between alexithymia and amygdala neurofeedback.

Overall, patients were able to downregulate the amygdala BOLD response with feedback, which is in line with our prior study (Paret et al., 2016). However, when looking at each run individually, we could not observe a significant downregulation effect in all four runs. Rather, the difference between the 'down' and 'view' condition descriptively seemed to increase over time (although the interaction of run and condition did not pass the significance level). In particular, significant downregulation of the amygdala amplitude was achieved at the fourth training run and downregulation of the AUC was achieved at the second and fourth run. This implies that in BPD patients multiple training runs are necessary to observe amygdala downregulation with neurofeedback.

In addition, we determined participants' best performance (i.e. the run with the largest delta between 'view' and 'down'). Both downregulation of the amygdala BOLD response and best performance did not correlate significantly with any of our primary endpoints. Thus, evidence for a mechanistic relationship between amygdala regulation and emotion dysregulation is still missing. The lack of significance may be a function of several causes, including lack of power and technical issues. For example, the neurofeedback training was optimized to increase absolute training time but was less optimal in terms of quantifying downregulation of the amygdala, as the view condition of each session was comprised of only five pictures, while the 'down' condition was comprised of 25 pictures. Additionally, shifts of behavior, physiology, and cognition during an emotional response are often loosely coupled (Bonanno & Keltner, 2004), and as such, a significant correlation is not necessarily observable, particularly in small sample sizes. Placebo-controlled trials are necessary to corroborate that neurofeedback training is indeed causal for improvement in emotion regulation.

4.5.1 Limitations

Several limitations merit comment. Most importantly, the present study lacks a control group, so that our results do not allow conclusions about the specificity and efficacy of neurofeedback

training. It is possible that factors other than the neurofeedback training itself account for the results. For example, it could be that the motivation to try a new treatment approach, psychosocial factors or effects of repeated exposure of tasks (i.e. practice effects) led to the observed changes. Therefore, replication in a RCT is necessary. In addition, we assessed a large number of different outcome measures. Testing many different outcome measures in a single patient cohort is the only way to identify potential behavioral targets for a new, technically-demanding, and cost-intensive technique (such as neurofeedback), given the current database and limited financial resources. Multiple comparisons however bare the risk of false discovery. To overcome this issue, we repeated statistical tests of primary endpoints with a conservative correction for multiple tests (i.e. Bonferroni-correction). No statistical tests survived significance testing with correction. Notwithstanding such disenchanting outcome, several comparisons (e.g., ZAN-BPD, startle response) achieved medium effect sizes. With appropriate sample size, future studies might replicate this finding and achieve significant outcomes.

Finally, the fixed order of the EWMT and BMT in the experiment might induce bias. Both tasks were performed prior to the first scanning session and immediately after the last neurofeedback training. At the end of the last scanning session, participants might have been fatigued and less capable or motivated to concentrate. Similarly, results from the ZAN-BPD should be interpreted with caution, as EMA assessment was conducted one week before the ZAN-BPD interview and may have biased the effect, as interviewers were not blinded to treatment.

4.6 Conclusion

Until now, it has been unclear which aspects of psychopathology and emotion regulation may change with neurofeedback-aided amygdala downregulation. The present study provides the first preliminary empirical basis for informed decision-making in primary outcome measures of larger clinical trials of amygdala neurofeedback training. We show that general BPD psychopathology, as well as different aspects of emotion dysregulation, improve after training, although these effects do not remain statistically significant after a conservative correction for multiple tests. If confirmed by an independent study, our results suggest that the ZAN-BPD, emotion regulation (assessed with emotion-modulated startle), and EMA are appropriate measures to quantify these improvements. Future placebo-controlled trials must confirm that neurofeedback treatment is effective in improving emotion regulation in those with BPD. Future trials will allow for the development of new therapy concepts, including neurofeedback that can be incorporated into clinical practice. In addition, the causal pathway through amygdala hyperactivation, regarding symptoms of emotion dysregulation, can also be tested.

4.7 Supplementary Material

4.7.1 Supplementary Methods

4.7.1.1 Flow chart of the study

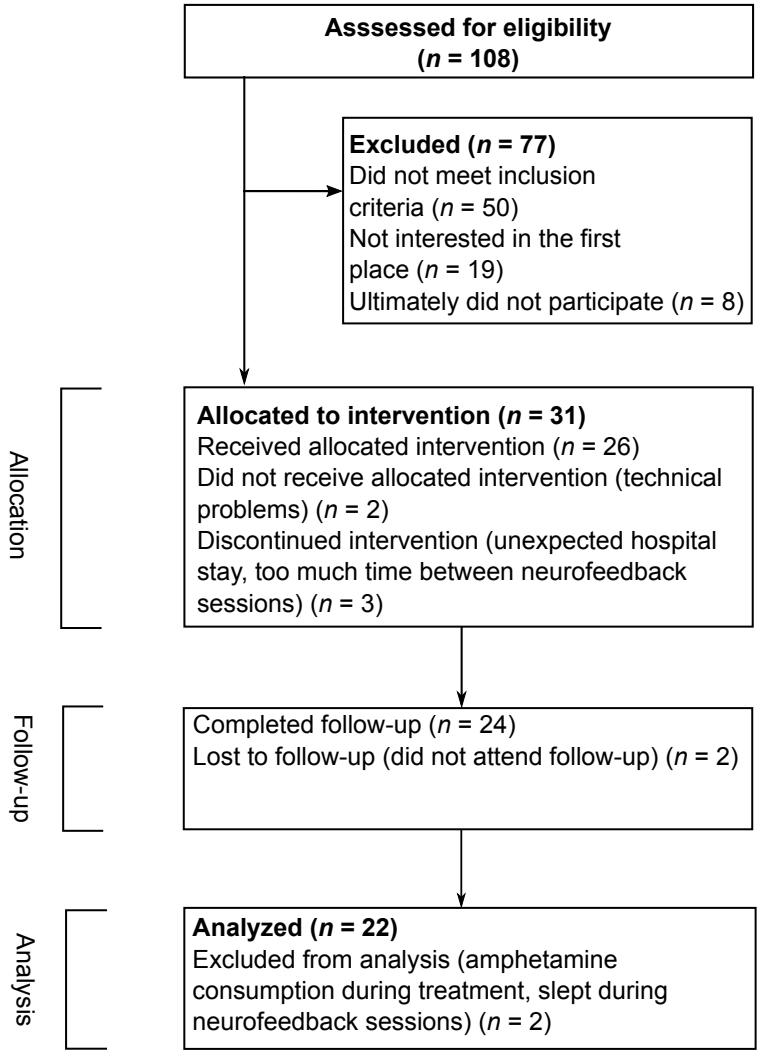

Figure 4.9. CONSORT flow chart for the study.

4.7.1.2 Recruitment

Recruitment was done by the central project of the KFO-256, a Clinical Research Unit funded by the German Research Foundation (DFG; KFO-256). Interested patients were told about the aims of the study as follows:

"fMRI neurofeedback is an innovative technique with which brain regulation can be improved. Using fMRI, real-time brain activation is measured and fed back to the patient in real-time. In

this vein, self-regulation and emotion processing may be improved. In this study we would like to investigate the therapeutic effects of this technique for borderline personality disorder."

4.7.1.3 Real-time fMRI neurofeedback

4.7.1.3.1 Procedure

BPD patients participated in a total of three neurofeedback training days with an interval of 2–7 days between subsequent sessions. On the second training day, participants underwent two consecutive training runs, whereas on the first and third training day they underwent one training run. A neurofeedback training run lasted 15 minutes.

Before the neurofeedback training started, participants were told that they would see negative pictures and a feedback signal from their brain depicted as a colored thermometer bar on each side of the picture. They were further instructed to down-regulate the thermometer bar and that there would be a temporal delay of the BOLD response, which caused a time lag of the thermometer response (2-5 seconds). In addition, they were instructed to keep their eyes open, not shift their gaze away from the screen, not to focus exclusively on the thermometer and the edges of the picture, not to control their breath and to keep their heads still throughout the experiment. Participants then entered the scanner and the experiment started. After the anatomical scans were acquired, a demo-feedback trial was presented without fMRI scanning to get participants accustomed to the training. Subjects were instructed to either look at the picture (without receiving feedback), or to down-regulate the thermometer signal, respectively. The neurofeedback run thus consisted of 'down' and 'view' conditions. In the 'down' condition, the pictures were presented with feedback. In the 'view' condition, a picture with aversive content was shown. A neurofeedback run comprised ten blocks, with five 'down' blocks consisting of six pictures presented for 18 seconds (108 seconds total) and with five 'view' blocks consisting of one picture presented for 18 seconds. The order of conditions was fixed with alternating view and down blocks. After the last block, participants were instructed to rate their perceived regulation success ('Were you able to regulate the display?') on a 10-level visual analogue scale.

4.7.1.3.2 Stimuli Used for Neurofeedback Runs

For the neurofeedback training runs, 132 negative stimuli were taken from standardized picture series (Dan-Glauser & Scherer, 2011; Lang et al., 2008; Marchewka et al., 2014; Wessa et al., 2010) and were completed with 8 pictures from the internet. Stimuli were presented with the Presentation software (Neurobehavioral Systems, Berkeley, CA). For each patient, pictures were randomly assigned to runs.

Stimuli with these original codes were used:

EmoPics: 208, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 241, 243, 244, 245, 246, 251, 252, 253, 254.
IAPS: 2053, 2301, 2345, 2375, 2456, 2457, 2458, 2691, 2694, 2700, 2703, 2718, 2795, 2799, 2800, 2900, 2981, 3180, 3215, 3220, 3500, 3530, 6220, 6231, 6311, 6313, 6520, 6550, 6563, 6831, 6940, 8485, 9050, 9163, 9183, 9185, 9230, 9250, 9332, 9342, 9414, 9415, 9419, 9423, 9424, 9425, 9426, 9428, 9429, 9435, 9480, 9901, 9902, 9905, 9920.
EmoMadrid: EM0001, EM0123, EM0319, EM0321, EM0322, EM0326, EM0331, EM0350, EM0356, EM0363, EM0392, EM0393, EM0410, EM0420, EM0447, EM0484, EM0488, EM0568, EM0570, EM0572, EM0577, EM0586, EM0597, EM0605, EM0633, EM0692
NAPS: Faces_012_v, Faces_021_h, Faces_148_h, Faces_150_h, Faces_303_h, Landscapes_002_h, Objects_006, Objects_011_h, Objects_012_h, Objects_125_h, Objects_143_h, Objects_275_v, People_071_h, People_079_h, People_118_h, People_139_h, People_210_h, People_243_h.

4.7.1.3.3 Neurofeedback Strategies Used

After each neurofeedback session, participants indicated which strategies they used to down-regulate the feedback signal and to what extent. We used an inhouse questionnaire that was created based on our experience of previous neurofeedback studies. In the questionnaire, participants were provided with different cognitive strategies such as reappraisal, distraction and mindfulness strategies and were instructed to indicate on a visual analogue scale to what extent they have used the strategy during the neurofeedback session.
The following strategies were provided:

1. I tried to reappraise the depicted situation (e.g. I told myself that this is just a film).
2. I tried to distance myself from the depicted situation (e.g. I told myself that this doesn't pertain to me).
3. I tried to distract myself with different thoughts or activities (e.g. mental arithmetic, counting numbers).
4. I tried to relax (e.g. by focusing on my breath or by doing relaxation exercises).
5. I tried to look at the picture without evaluating it (e.g. by describing what I see on the picture without evaluating it in a positive or negative manner).
6. I tried to shift my attention away from aversive details of the picture and to e.g. relocate my attention towards neutral details of the picture or towards the thermometer display.
7. I tried to remember positive emotional memories e.g. with my family or friends.
8. I tried to think of a positive story of the people depicted on the picture (e.g. the person will get help, or the person is doing better).

9. I tried to think of a reason or a story behind the depicted situation.
10. I tried to integrate the picture into a political or historical context.

Participants were instructed to mark on a continuous analogue scale, how often they used the strategies provided. The manually set marks were digitally transformed into a percent maximum possible ranging from 0 to 100. Later, question 1, 2, 8, 9 and 10 were combined to reappraisal strategies, question 3 and 7 were combined to distraction strategies, by adding the mean values of the respective strategies and subdividing it by the number of strategies. Thus, we analyzed 5 strategies: reappraisal, distraction, relaxation, mindful viewing, paying attention to neutral aspects.

4.7.1.3.4 Real-time fMRI Analysis and Feedback Presentation

Scan volumes were immediately transferred from the scanner site to a computer in order to preprocess and analyze the data using SPM8 (Wellcome Department of Cognitive Neurology, London, UK). For that, the T1-weighted scan was first segmented and normalized to Montreal Neurological Imaging (MNI) space. Anatomical masks of the regions of interest (ROI) for feedback calculation were then moved to subject-native space. The BOLD signal data from voxels within a right amygdala mask was taken for further processing, The mask was created with the Harvard-Oxford brain atlas with a probability threshold of 25%. We also recorded BOLD signal data from a rectangular ROI (3x30x30 mm in AC-PC orientation, center of mass = [0,-16,-5], MNI coordinates). This mask served as a control for global signal fluctuation unrelated to functional brain activation. The perpendicular distance from the control ROI to the right amygdala mask comprised 7 mm in sagittal direction. Functional images were realigned to the first volume and BOLD signal data from all voxels within each ROI were averaged. The average time course was processed with a modified Kalman filter (Koush, Zvyagintsev, Dyck, Mathiak, & Mathiak, 2012) and detrended using Matlab's (R2014b) detrend function. Detrending began with the 35th volume (i.e. before feedback started) so that the filter and detrend functions were stabilized. Next, percent signal change from the global mean was calculated. On a second computer, stimulus presentation software received the data via TCP/IP. The feedback was displayed as a colored rectangle moving up and down, altering from dark red at the maximum over light green to dark green at the minimum. Resolution of the display consisted of six units and ranged between two percent signal change above and below baseline. The display was refreshed with a frequency of 1 Hz. Details of fMRI acquisition, real-time fMRI analysis and feedback presentation can be found in the supplement and has been published by Paret et al. (2018).

4.7.1.4 Emotion Regulation Test

4.7.1.4.1 Procedure

Participants completed the task directly before (T0) and after (T1) the neurofeedback training as well as after 6 weeks (T2). They were instructed either to view negative and neutral pictures without modifying their emotions ('view'; 'neutral' condition respectively) or to down-regulate their feelings toward negative pictures ('down' condition). Furthermore, participants were instructed not to turn away their gaze or to close their eyes, nor to focus exclusively on non-emotional parts of the picture. Before the emotion regulation task started, ten startle probes were presented consecutively, to control for habituation effects. In total, the paradigm consisted of 36 trials (12 trials per condition) and lasted 20 min. Each trial began with a 2,000 milliseconds presentation of an instructional cue ('view', 'down'), followed by a fixation cross displayed for 1,000 milliseconds. Next, a neutral or negative picture was presented for 10,000 milliseconds. A startle probe (50 milliseconds, 95dB white noise burst) was presented through headphones at 6,500 milliseconds – 9,500 milliseconds into the regulation phase). Self-assessment Manikins (SAM Ratings; Bradley & Lang, 1994) were presented after presentation of each picture. Participants rated on a 1-9 Likert scale how positive/negative and aroused/calm they felt at that moment. Lower scores on the valence scale indicate that they felt more positive; lower scores on the arousal scale indicate that they felt calmer. By pressing buttons on a keyboard, subjects selected SAMs corresponding to their subjective valence and arousal. The initial rating position was random and the current selection after 5 s was logged. Intertrial intervals were jittered between 3,500 and 5,500 milliseconds Picture stimuli were presented in semi-randomized order with restriction of no more than two consecutive trials from the same condition, and no more than three consecutive trials with negative pictures. The participants' eyes were tracked by a camera system (SMI BeGaze, Teltow, Germany) to encourage subjects to comply with instructions, data were not analyzed. Ten percent of trials did not include a startle probe (2 trials per condition).

4.7.1.4.2 Stimuli Used for the Emotion Regulation Test

Stimuli were taken from standardized picture series (Lang et al., 2008; Marchewka et al., 2014) and were presented with the Presentation software (Neurobehavioral Systems, Berkeley, CA). Arousal and valence were calculated from normative ratings provided with the published data sets. Pictures were matched between T-assessments according to arousal and valence. Stimuli with the following original codes were used:

T0: Negative: 247, 1033, 1304, 2205, 3185, 6212, 6510, 9413, 9433, 9921, People_034, Animals_033, Animals_052, People_147, Faces_003, Faces_028, Faces_041, Faces_272, People_002, People_007, People_039, People_072, People_136, People_203. Neutral: 2381, 5510, 5740, 7003, 7025, 7052, 7150, 7161, 7175, 7185, 7236, 7950.
T1: Negative: 234, 239, 250, 252, 3181, 6244, 9042, 9412, Animals_001, Animals_014, Animals_034, Faces_014, Faces_146, Faces_174, Faces_274, Faces_290, Faces_369, Objects_139, Objects_283, People_019, People_084, People_124, People_143, Animals_076- Neutral: 2038, 6150, 7000, 7004, 7006, 7010, 7035, 7041, 7045, 7187, 7205, 7490.
T2: Negative: 209, 249, 2683, 235, 325, 6231, 9904, 2811, Animals_003, Animals_006, Faces_035, People_145, Faces_151, People_033, People_231, Animals_085, People_004, People_073, Faces_007, Faces_034, Faces_019, People_001, Faces_293, People_038. Neutral: 7001, 7026, 7217, 7080, 5471, 2580, 7090, 7224, 5635, 7100, 7009, 7002.

4.7.1.5 Emotional Working Memory Task (EWMT)

4.7.1.5.1 Procedure

The EWMT is an adapted Sternberg item recognition task (Sternberg, 1966), modified by Oei and colleagues (Krause-Utz et al., 2014; Krause-Utz et al., 2012; Oei et al., 2012). The present version comprised 40 trials with jittered durations: each trial started with the presentation of a set of three letters (memoranda, 875 – 1,375 milliseconds). After a delay phase of 1,250 – 1,750 milliseconds another set of three letters appeared on the screen (probe, 2,000ms). Next, a blank screen appeared (inter trial interval, 550 – 1,050 milliseconds). Participants had to press the left or right button to indicate whether they recognized one of the memoranda-letters in the probe. In half of the trials, one of the three memoranda was present in the probe. During the delay interval, either no distractor (i.e. a fixation cross; 'cross' condition) or a distractor (i.e. an aversive picture; 'negative' condition) was presented. Target-present and target-absent trials were equal across the 'negative' and the 'cross' condition. The presentation of the two conditions (write conditions) was balanced in a pseudo-random manner with no more than two consecutive conditions of the same type. Stimuli were presented in semi-randomized order with restriction of no more than two consecutive trials from the same condition (write conditions), and no more than three consecutive trials with negative pictures. Order of condition and assignment of pictures to assessments (i.e. T0, T1, T2) was alternated between subjects.

4.7.1.5.2 Stimuli Used for the EWMT

Stimuli were taken from standardized picture series (Lang et al., 2008; Marchewka et al., 2014) and were presented with the Presentation software (Neurobehavioral Systems, Berkeley, CA).

Pictures were matched between T-assessments according to arousal and valence calculated from normative ratings provided in original publication of the data sets. The following stimuli were used:
T0: 284, 242, 1202, 3550, 6021, 9620, 3195, 9810, 6370, 1300, 9635, 1202, 3212, 6350. People_022_h, Faces_363_h, People_209_h, Faces_284_h, People_031_h, People_128_h.
T1: 233, 3300, 8230, 9800, 9254, 6570, 9670, 9325, 3103, 9600, 9321, 2730, 3030, 9410, People_214, People_058, People_140, People_088, Objects_001, People_246.
T2: 237, 3350, 9181, 9300, 9400, 9910, 6415, 9075, 6834, 1930, 6312, 9252, 1120, 6230, Animals_008, People_225, People_037, Faces_362, Objects_149, Faces_010.

4.7.1.6 Backward Masking Task (BMT)

4.7.1.6.1 Procedure

Participants were instructed to identify whether faces expressing either happy or fearful facial expressions are male or female. To keep them attentive throughout the task, participants were told to identify whether the presented faces were male or female via button press. Faces were presented for 33 milliseconds or 83 milliseconds. Thus, the BMT had a total of four conditions: Happy or fearful facial expressions either presented for 33 milliseconds or 83 milliseconds. A total of 4 blocks per condition were presented. Each block consisted of 8 pictures shown consecutively. Each block began with a fixation cross. Next, eight faces were shown for 33ms or 83ms, each preceded by a red rectangle on a grey background for 5,100 milliseconds and followed by a mask (scrambled face) for 4,100 milliseconds. Stimuli were taken from the Karolinska Directed Emotional Faces set (KDEF; Lundqvist, Flykt, & Öhman, 1998).

4.7.1.6.2 Behavioral Data of the BMT

Participants were instructed to identify the sex of the faces to keep them attentive throughout the task by pressing either the left or right button, respectively. Button presses were analyzed by summing up right, wrong and missed button presses per session across all conditions.

4.7.1.7 EMA

The following questions were used for EMA assessments:
Positive affect:
At the moment I feel happy
At the moment I feel lucky
At the moment I feel relaxed
At the moment I feel content

At the moment I feel enthusiastic

Negative affect:

At the moment I feel sad

At the moment I feel irritated

At the moment I feel angry

At the moment I feel depressed

At the moment I feel fearful

Subjective control over emotions:

When the phone rang I felt like I could control my feeling

When the phone rang I felt overwhelmed by my feelings

Inner tension:

At the moment I feel an aversive inner tension

Dissociation:

When the phone rang I felt like my body did not belong to me

When the phone rang I felt like people or things or the world were not real

When the phone rang I had problems to hear right, e.g. noise around me appeared as if it came from far away

When the phone rang I felt like my body or some body parts were insensitive to pain

4.7.1.8 Consensus on the Reporting and Experimental Design of clinical and cognitive-behavioural Neurofeedback studies (CRED-nf) best practices checklist 2019

Table 4.2

Cred-nf best practice checklist 2019.

Domain	Item #	Checklist item	Reported on page #
Pre-experiment			
	1a	Pre-register experimental protocol and planned analyses	129
	1b	Justify sample size	In our opinion, N = 25 subjects were sufficient to reach our aims
Control groups			
	2a	Employ control group(s) or control condition(s)	Control condition:131, 140; control group: 127, 153
	2b	When leveraging experimental designs where a double-blind is possible, use a double-blind	No blinding possible
	2c	Blind those who rate the outcomes, and when possible, the statisticians involved	Not applicable
	2d	Examine to what extent participants and experimenters remain blinded	Not applicable
	2e	In clinical efficacy studies, employ a standard-of-care intervention group as a benchmark for improvement	Not applicable
Control measures			
	3a	Collect data on psychosocial factors	157, 165, 166
	3b	Report whether participants were provided with a strategy	131
	3c	Report the strategies participants used	165
	3d	Report methods used for online-data processing and artifact correction	132, 158
	3e	Report condition and group effects for artifacts	-

Table 4.2 *(continued)*

Domain	Item #	Checklist item	Reported on page #
Feedback specifications			
	4a	Report how the online-feature extraction was defined	132, 158
	4b	Report and justify the reinforcement schedule	132
	4c	Report the feedback modality and content	132
	4d	Collect and report all brain activity variable(s) and/or contrasts used for feedback, as displayed to experimental participants	158, 138, 140
	4e	Report the hardware and software used	131, 137, 158
Outcome measures			
Brain	5a	Report neurofeedback regulation success based on the feedback signal	-
	5b	Plot within-session and between-session regulation blocks of feedback variable(s), as well as pre-to-post resting baselines or contrasts	140
	5c	Statistically compare the experimental condition/group to the control condition(s)/group(s) (not only each group to baseline measures)	Not applicable
Behaviour	6a	Include measures of clinical or behavioural significance, defined a priori, and describe whether they were reached	133 ff, 141 ff
	6b	Run correlational analyses between regulation success and behavioural outcomes	151, 168
Data storage			
	7a	Upload all materials, analysis scripts, code, and raw data used for analyses, as well as final values, to an open access data repository, when feasible	-

4.7.2 Supplementary Results

4.7.2.1 Data without Imputation

4.7.2.1.1 Neurofeedback Downregulation

Original data of Amygdala AUCs revealed a significant main effect of condition, $F(1,18) = 7.57$, $p = .01$, $eta^2 = .30$, and non-significant main effect of time, $F(3,54) = 2.53$, $p = 0.08$, $eta2 = .12$, but no significant interaction of time and condition, $F(3,54) = .96$, $p = .42$, $eta^2 = .05$. Post-hoc paired *t*-tests of original data revealed a significant effect between regulate and view at session 4, $t(18) = -2.52$, $p = .02$. Original data of Amygdala amplitudes revealed a main effect of condition, $F(1,18) = 7.97$, $p = .01$, $eta^2 = .31$. There were no significant main effect of time and no significant interaction of time and condition. Post-hoc paired *t*-tests of original data revealed a significant effect between regulate and view at session 4, $t(18) = -2.84$, $p = .01$.

4.7.2.1.2 Neurofeedback Downregulation Strategies

Descriptively, participants used mindful viewing the most and distraction strategies the least across all sessions (see Figure 4.10).

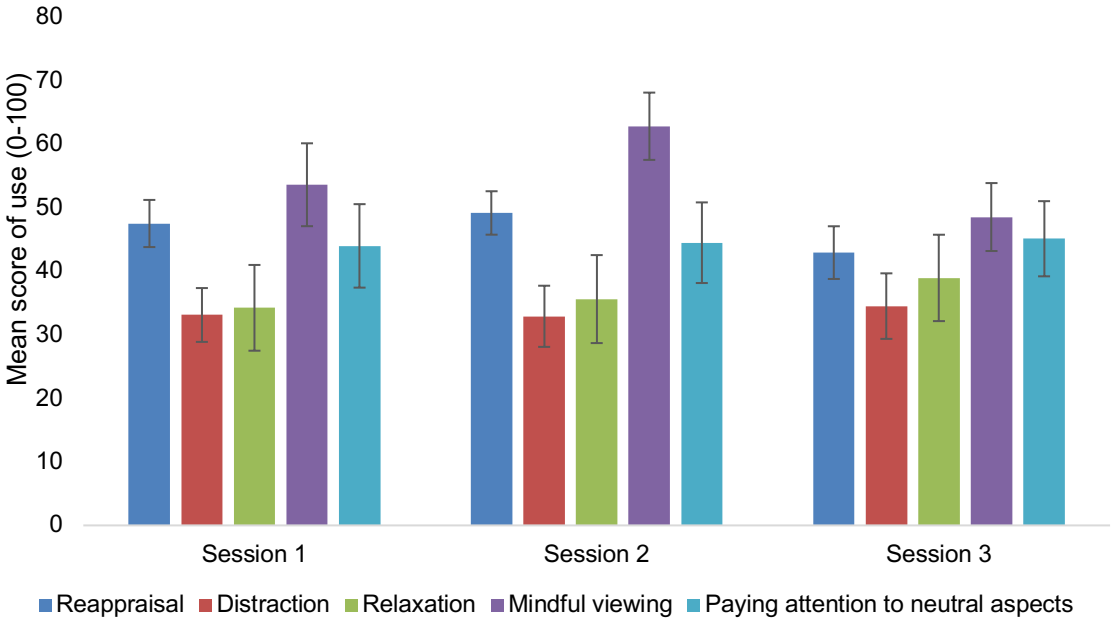

Figure 4.10. Mean extent of use of different strategies during the neurofeedback sessions. Error bars represent standard error of mean (*SEM*).

4.7.2.1.3 TAS-26

Original data of the TAS-26 revealed a significant main effect of time of the TAS total score, $F(1.47, 27.91) = 6.85$, $p = .01$, $eta^2 = .27$, and of the subscale 'Identification of one's feelings', $F(2, 40) = 6.63$, $p < .01$, $eta^2 = .25$. Post-hoc paired *t*-tests revealed a significant decrease of 'Identification of one's feelings' from T1 to T2, $t(20) = 2.06$, $p = .05$, and from T1 to T2, $t(20) = 3.05$, $p = .01$, and a significant decrease of TAS total score from T1 to T2, $t(18) = 2.78$, $p = .01$, and from T0 to T2, $t(18) = 2.55$, $p = .02$). No significant main effect of time was found for the subscales, Difficulty describing feelings', $F(1.56, 29.64) = .67$, $p = .52$, $eta^2 = .03$), and 'external thinking', $F(2, 38) = 2.59$, $p = .09$, $eta^2 = .12$.

4.7.2.1.4 Emotion Regulation Test

As hypothesized, patients down-regulated negative emotions more effectively after training, indexed by a significant decrease of the emotion-modulated startle in the 'down' compared to the 'view' condition after training, $F(2,34) = 3.27$, $p = .05$, $eta^2 = .16$. There was no significant main effect of time, $F(2,34) = 1.09$, $p = .35$, and condition, $F(1,17) = .28$, $p = .60$. Post-hoc paired *t*-tests between the 'down'-'neutral' and the 'view'-'neutral' condition revealed a trend-level effect at T1, $t(17) = -2.01$, $p = .06$, $d = -.47$. At T0 and T2, in contrast, patients did not significantly decrease startle in the 'down'-'neutral' vs 'view'- 'neutral'.

4.7.2.2 Data with Imputation

4.7.2.2.1 Rating of Subjective Regulation Success

Participants were instructed to rate their perceived regulation success on a 10-level visual analogue scale (0 = not at all; 10 = very much). Mean success ratings ranged between 4.00 and 4.63 across sessions, suggesting that participants estimated their regulation success in the medium range. Success ratings did not significantly change over time, $F(3,69) = 1.13$, $p = .34$. Descriptive data of success ratings can be derived from Table 4.3.

Table 4.3

Mean ratings of subjective downregulation success during individual neurofeedback sessions.

	Session 1	Session 2	Session 3	Session 4
Mean success (*SD*)	4.00 (1.50)	4.63 (1.86)	4.56 (1.67)	4.33 (1.34)

Note. SD = standard deviation.

4.7.2.2.2 Best Performance

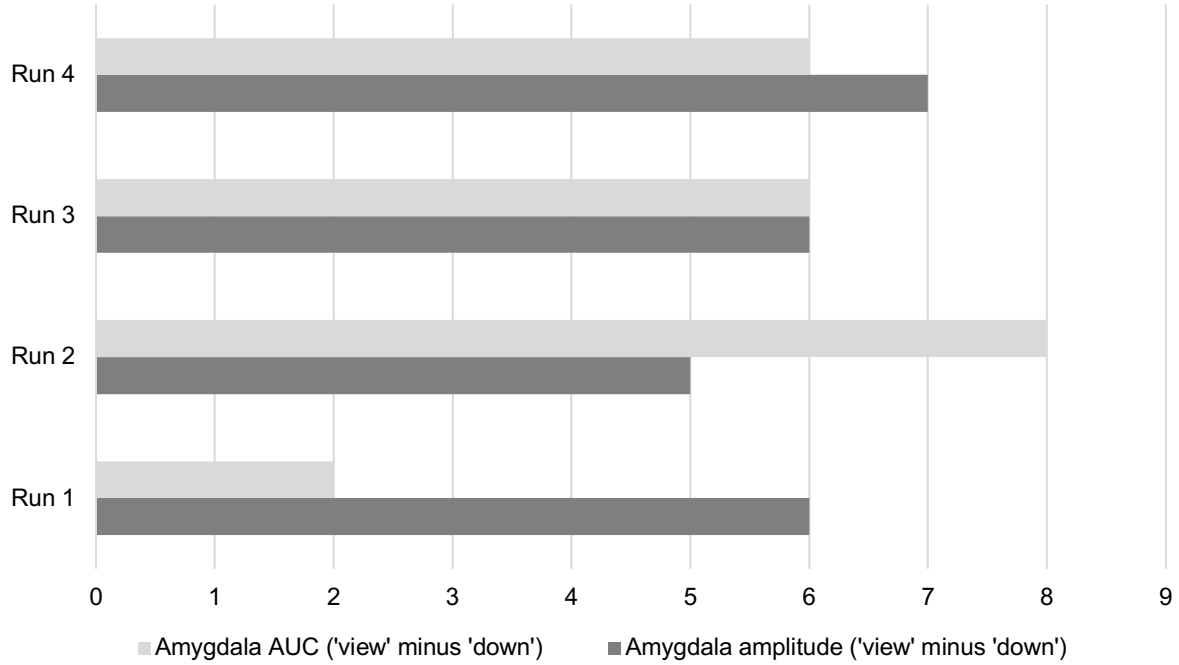

Figure 4.11. Best performance (i.e. largest delta between 'down' and 'view') of amygdala downregulation. Bars represent for each run the number of participants that reached their best performance. AUC = area under the curve.

4.7.2.2.3 Behavioral Results of the BMT

Mean number of right, $F(2,46) = .64$, $p = .54$, wrong, $F(2,46) = .59$, $p = .56$, and missed trials, $F(2,46) = 1.33$, $p = .28$, did not significantly change over time. Descriptive data can be derived from Table 4.4.

Table 4.4

Mean (SD) of number of correct, wrong and missed trials during the BMT at T0, T1 and T2.

	Right answer	Wrong answer	Missed trials
T0	109.56 (14.87)	14.51 (11.00)	3.97 (5.45)
T1	110.49 (10.52)	13.36 (8.51)	4.10 (4.33)
T2	107.53 (22.39)	13.31 (10.32)	7.15 (15.29)

Note. SD = standard deviation.

4.7.2.2.4 Correlations between Downregulation Success and Primary Endpoints

Table 4.5

Correlations between Amygdala downregulation and primary endpoints.

		Startle		EWMT		BMT		HRV		ZAN total		DERS total	
		T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1
AUC session 1	*r*	0.10	-0.23	-0.11	0.00	0.03	0.11	0.03	-0.10	-0.300	-0.30	-0.01	0.16
	p	0.65	0.27	0.61	1.00	0.90	0.62	0.88	0.64	0.154	0.16	0.96	0.46
	N	24	24	24	24	24	24	24	24	24	24	24	24
AUC session 4	*r*	0.17	0.04	-0.02	0.33	-0.03	0.39	0.15	-0.12	0.143	0.18	-0.29	-0.11
	p	0.44	0.86	0.93	0.11	0.89	0.06	0.47	0.56	0.504	0.40	0.17	0.62
	N	24	24	24	24	24	24	24	24	24	24	24	24
Amplitude session 1	*r*	0.01	0.01	-0.01	-0.06	-0.05	0.02	0.51*	-0.07	-0.45*	-0.31	-0.05	0.13
	p	0.95	0.96	0.98	0.79	0.80	0.92	0.01	0.74	0.029	0.14	0.80	0.53
	N	24	24	24	24	24	24	24	24	24	24	24	24
Amplitude session 4	*r*	-0.02	0.13	-0.21	0.35	-0.14	0.29	0.21	-0.13	-0.005	0.17	-0.33	-0.16
	p	0.93	0.54	0.32	0.09	0.52	0.18	0.32	0.53	0.983	0.43	0.12	0.45
	N	24	24	24	24	24	24	24	24	24	24	24	24

Note. AUC = area under the curve; Startle = startle amplitude 'down' minus 'view'; *r* = Pearson's Correlation; *p* Sig. (2-tailed); EWMT = Emotional Working Memory Task; BMT = Backward Masking Task; HRV = heart rate variability; DERS = Difficulties with Emotion Regulation Scale.

* Correlation is significant at the 0.05 level (2-tailed).

Table 4.6

Correlations between downregulation success indices and changes in primary endpoints.

T0 minus T1		Startle	EWMT	BMT	HRV	DERS total	ZAN total
Best performance (AUC)	*r*	-0.265	0.079	0.090	0.127	-0.137	-0.081
	p	0.210	0.712	0.677	0.554	0.525	0.706
	N	24	24	24	24	24	24
Best performance (Amplitude)	*r*	-0.097	-0.178	0.272	-0.336	0.043	0.266
	p	0.653	0.406	0.198	0.108	0.844	0.209
	N	24	24	24	24	24	24
AUC (run 1 minus run 4)	*r*	0.070	-0.092	0.180	-0.090	-0.004	0.151
	p	0.747	0.669	0.401	0.677	0.985	0.482
	N	24	24	24	24	24	24
Amplitude (run 1 minus run 4)	*r*	0.052	0.240	0.127	-0.098	-0.059	0.303
	p	0.808	0.259	0.555	0.647	0.783	0.149
	N	24	24	24	24	24	24

Note. AUC= area under the curve; *r* = Pearson's Correlation; *p* Sig. (2-tailed); EWMT= Emotional Working Memory Task; BMT =Backward Masking Task; HRV = heart rate variability DERS = Difficulties with Emotion Regulation Scale.

5 GENERAL DISCUSSION

5.1 Overall Rationale and Review

The feasibility to treat emotion regulation problems with fMRI neurofeedback has been demonstrated by a number of studies (Linhartová et al., 2019). Because BPD patients show amygdala hyperreactivity which is assumed to underlie the severe emotion regulation problems they suffer from (Schulze et al., 2011), amygdala neurofeedback training might be a candidate training to reduce symptoms in BPD and to specifically target emotion dysregulation.

However, the field has yet to prove that amygdala fMRI neurofeedback has an effect above and beyond placebo. One major problem is the lack of controlled trials and a lack of critical outcome variables in fMRI neurofeedback experiments. In a recent systematic review including 99 neurofeedback experiments (Thibault, MacPherson, Lifshitz, Roth, & Raz, 2018), 38 studies did not assess a control group. 39 studies included a control group and 11 of them showed behavioral improvements that were significantly greater in the experimental than in the control group. The authors conclude that many fMRI neurofeedback experiments would lack the key variables required to demonstrate the behavioral and clinical benefits of this new technique.

With regard to amygdala neurofeedback, assessed outcome measures have been quite diverse (Barreiros, Almeida, Baia, & Castelo-Branco, 2019). In addition, we have not fully understood how amygdala normalization is mapped onto behavioral correlates in BPD. There is no "gold-standard" to assess emotion (dys)regulation, but rather available tools differ largely between each other and psychophysiological measures have produced mixed results. This in turn impeded the selection of primary outcome measures for future randomized controlled trials (RCT) of amygdala neurofeedback in BPD patients.

To fill this gap, three empirical studies were conducted in the present thesis, that were devoted to identifying emotion regulation effects of psychophysiological responses that may also be used in clinical trials, and to inform about potential outcome measures of amygdala neurofeedback in BPD.

In particular, in **Study I** we screened and selected relevant emotion regulation studies and identified psychophysiological effects as well as important moderating variables using a meta-analytic strategy. **Study II** focused on the emotion-modulated startle to measure emotion regulation and investigated whether the startle probe timing has an impact on the effects. **Study III** tested emotion regulation on several system levels before and after amygdala neurofeedback in BPD, including the paradigm piloted in Study II.

With that, the present thesis provides the groundwork for amygdala neurofeedback RCTs which are urgently needed to demonstrate that amygdala neurofeedback has a therapeutic value in BPD.

5.2 Summary and Integration into Previous Research

In the following section, results of all three studies will be summarized and discussed. First, the present findings on the psychophysiological effects of emotion regulation (Study I) will be integrated into existing literature. The following paragraph addresses potential influencing factors of emotion regulation effects on autonomic measures and the emotion-modulated startle, in particular the timing of the startle probe (Study II) as well as the duration of the emotion regulation period (Study I). The last paragraph then concentrates on potential outcome measures of amygdala neurofeedback (Study III). At some passages data of the present studies are related and compared to each other.

5.2.1 What are the Effects of Emotion Regulation on Psychophysiology?

In Study I, we conducted meta-analyses of psychophysiological effects of emotion regulation strategies. Results demonstrate that autonomic responses of reappraisal and suppression are heterogenous with small to non-significant effect sizes ranging between d = -.09 and d = -.34 (see Figure 2.3). The effects are compatible with effect sizes reported in a previous meta-analysis, which did not differentiate between individual psychophysiological measures (Webb et al., 2012). Compared to the autonomic effects, we found a significant and consistent decrease of the emotion-modulated startle across studies. In addition, corrugator EMG showed consistent and medium effects, too. This raises the question whether emotion regulation impacts indices of facial expressivity and eye-blink to a different degree than physiological measures of arousal (e.g., skin conductance). Given the inconsistent and small effects of autonomic responses, it may be even asked whether emotion regulation can be successfully assessed with autonomic measures at all. Possible reasons for the non-significant effects of autonomic measures will be discussed in section 5.2.2.

With regard to the emotion-modulated startle, the conducted meta-analysis of Study I included only five studies in total. However, startle studies that could not be considered in Study I support the findings of the meta-analysis: Except for one study (Eippert et al., 2007), they also report a significant inhibition of the emotion-modulated startle during emotion downregulation of negative emotions (Fuentes‐Sánchez et al., 2019; Grillon et al., 2015; Lee et al., 2009; Leiberg et al., 2012; Lissek et al., 2007; Piper & Curtin, 2006).

In Study II, we replicated the finding of Study I in that healthy subjects successfully reduced negative emotions evidenced by a significant startle inhibition during the reappraisal of negative emotions. Descriptively, the effect became larger with increasing probe timing, indicating that later startle probes are useful to quantify emotion regulation with the emotion-modulated startle.

In Study III, we assessed the emotion-modulated startle during explicit emotion downregulation before and after neurofeedback training in BPD patients. We found increased startle inhibition during downregulation compared to the no-regulation control condition after neurofeedback training, which suggests that participants improved their ability to regulate negative emotions. BPD patients did not exhibit a significant inhibition of emotion-modulated startle during emotion downregulation (versus no-regulation) before the neurofeedback training, whereas a significant inhibition was observed after the training, indicating a normalization of the emotion-modulated startle after neurofeedback.

In sum, results of Study I, II and III demonstrate that the emotion-modulated startle is a robust measure of emotion regulation which can be used as a potential variable to assess training-related effects of amygdala neurofeedback.

Our findings encourage possible reasons that might have accounted for the consistent effects of the emotion-modulated startle (for a detailed review see section 2.5). The methodology and quantification of the startle response is relatively straightforward (Grillon & Baas, 2003). In Study I, there was great concordance across studies with regard to the quantification of the startle response, a procedure that has been described in detail in Blumenthal et al. (2005). Second, animal and human subject studies have shown that the amygdala directly modulates the auditory startle reflex via modulation of midbrain neurons (Davis, 1992; Kuhn et al., 2020; Rosen & Davis, 1988). In other words, emotion downregulation of negative emotions may affect the startle response by modulating the amygdala. Emotional priming as indexed by the startle reflex modulation is assumed to provide an unequivocal index of amygdala activation that is independent of different task demands (Grillon & Baas, 2003). The direct modulation via the brain's motivational system and the consistent methodology thus might have contributed to successful replications of the effect in our studies.

The startle can track the valence dimension of emotion, which, according to Grillon & Baas (2003), could also be a valuable methodological advantage for the study of emotion. The authors reason that other psychophysiological indices such as electrodermal and cardiovascular measures do not robustly differentiate between positive and negative emotions. More specifically, arousing negative emotional scenes potentiate startle amplitudes, while arousing positive

scenes cause inhibition. Researcher have assumed that the startle is a measure of general affective valence, while skin conductance reflects arousal (Lang et al., 1990).

Yet, other research suggests that the interpretation of the startle as a measure of valence might be oversimplified: When it comes to the *regulation* of an emotion for example, previous literature suggests that the modulatory effect of the emotion-modulated startle during explicit emotion regulation follows the pattern of variations in arousal of pictures rather than valence (Bernat et al., 2011; Dillon & LaBar, 2005). In line with these findings, Study II of the present dissertation found that startle inhibition during reappraisal correlated with the perceived downregulation of arousal. In contrast, we found no significant correlation between the perceived downregulation of valence and startle inhibition. Our results demonstrate that changes in the defensive tendency measured with emotion-modulated startle also reflect changes in perceived levels of arousal. Further studies are needed to understand the relationship between startle and the arousal and valence dimensions.

To date, we can conclude that our results together with the strong relation between emotional dysregulation, enhanced amygdala activation, and enhanced startle response (Davis et al., 1999; Rosen & Davis, 1988) suggest that emotion-modulated startle is a sensitive measure for investigating therapeutic effects of amygdala neuromodulation. The association between startle and both valence and arousal dimension in the context of emotion regulation remains to be studied in future research.

5.2.2 What Influences the Effects of Emotion Regulation on Psychophysiology?

Small to non-significant effects in our meta-analysis (Study I) were partly due to the fact that the direction of effect sizes across individual studies were heterogenous and thus cancelled out each other. In addition, confidence intervals around effect sizes were very large in some studies, suggesting a high variability of the autonomic responses. The inconsistencies implicate that there are factors not yet understood, which drive autonomic effects in one or the other direction. In Study I, we identified a number of moderator variables that influenced effect sizes of the autonomic measures, such as study design, trial duration and the nature of the control condition (see section 2.4 for a detailed discussion). The moderator analyses suggest that the autonomic effects of suppression and reappraisal are in parts influenced by important aspects of the study set-up.

In addition to these moderators, we also observed tremendous variation in the assessment and quantification of the autonomic measures, which was especially obvious in electrodermal responses, where the selection of baseline activity and the quantification of the response varied

largely between the individual studies (for an overview see Table 2.10). Therefore, it could be that the non-significant effects on autonomic measures reported in Study I may also be due to the variable assessment methods across studies. More research is needed to carefully investigate the influence of assessment and quantification methods on psychophysiological effects of emotion regulation, especially in electrodermal responses. In addition, we need better standardization of psychophysiological assessment and quantification across studies.

Besides this, the present findings from both Study I and II indicate that the duration of the regulation period seemed to be a somewhat important factor for autonomic effects and startle inhibition during emotion regulation. Both the extended process model (Gross, 2015) and the implementation and maintenance model (IMMO; Kalisch, 2009) agree on the fact that the extended period of emotion regulation can be separated into different stages. Hence, from a theoretical standpoint it was hypothesized that emotion regulation might need some time until it effectively reduces negative emotions as emotion regulation strategies need to be activated and implemented first. The present psychophysiological effects are in line with that.

Study I revealed that the duration of the emotion regulation period significantly moderated effects of reappraisal on skin conductance response and effects of suppression on skin conductance level, diastolic and systolic blood pressure, in that the effects became more negative (i.e. the mean levels decreased) with increasing trial duration. Although sample size was very small and thus should be treated with caution, the findings suggest that both suppression and reappraisal might need some time until they effectively decrease sympathetic arousal (e.g., decreased skin conductance, decreased blood pressure).

Study II examined whether effects of reappraisal assessed with the emotion-modulated startle increase with startle probe time. Contrary to our expectations, startle inhibition was independent of probe timing. That is, whether probes were delivered at 2, 7 or 12 seconds into the reappraisal phase did not significantly affect the assessment of emotion regulation. Descriptively however , startle probes delivered at 2 seconds into the regulation phase produced smaller effects and thus might be less sensitive than later probes (Dillon & LaBar, 2005; Eippert et al., 2007; Jackson et al., 2000).

As such we may speculate that the duration of the regulation period (and in startle experiments the timing of the startle probe), is a variable that should still be considered carefully in experimental studies. Further research is needed to fully understand the temporal dynamics of psychophysiological effects of emotion regulation. For example, Study II could be replicated with longer trial durations (and probe times beyond 12 seconds into the regulation phase) to analyze the full-blown emotion regulation response.

5.2.3 What are the Effects of Amygdala Neurofeedback on Emotion Dysregulation?

Based on the results of Study I and II**,** Study III assessed emotion dysregulation before and after an amygdala neurofeedback training using a range of outcome modalities such as psychophysiology, self-report and neural assessments. In line with our hypotheses, we observed changes in BPD psychopathology after neurofeedback training, suggesting that subjectively experienced BPD symptoms improved over the course of the intervention. This result corresponds with studies reporting associations of amygdala normalization with reductions in BPD psychopathology, and are in harmony with the notion that amygdala response is a critical mechanism of remission with BPD (Goodman et al., 2014; Schmitt et al., 2016; Schnell & Herpertz, 2007). The improvement in BPD psychopathology also dovetails with other clinical neurofeedback studies targeting the emotion circuit in the brain. Young, Siegle, et al. (2017), for example, found a 40% reduction in depressive symptoms in an amygdala upregulation neurofeedback intervention compared to a placebo neurofeedback group in unmedicated depressive patients. Another neurofeedback study in patients with depression found that depressive symptoms decreased by 43%, and 38% of patients showed a remission of depressive symptoms, although the intervention and control groups did not differ significantly (Mehler et al., 2018). Together with improvements in general BPD psychopathology as reported in study III, these studies imply that a neurofeedback intervention specifically designed to target the emotion brain circuit may also improve more general psychopathological symptoms.

In addition, the present results of Study III indicate that negative affect and affective instability experienced in daily life were reduced after neurofeedback training. These effects were particularly large compared to the other outcome measures of Study III, with a Cohen's d ranging between $d = .56$ and $d = .70$. Negative affect and affective instability were assessed with ecological momentary assessment (EMA) over the course of four days before and after neurofeedback training. The results are particularly encouraging as they indicate a possible transfer of the neurofeedback training to everyday life situations. However, as we did not assess EMA at the 6-week follow-up assessment, it remains to be shown if these changes last in the long-term.

Participants also showed improvements in the emotion regulation test after training, indicated by a greater inhibition of the startle response. The emotion regulation test explicitly instructed participants to downregulate negative emotions. In other words, amygdala neurofeedback treatment might particularly support new or strengthen already existing emotion regulation skills, which is in line with the skill acquisition hypothesis outlined in Gevensleben et al. (2014).

At the same time, we also found significant improvements in alexithymia, i.e. the difficulty to cognitively process emotions, after neurofeedback training. Results are compatible with a recent study showing that amygdala electrical fingerprint neurofeedback resulted in a larger reduction of alexithymia scores compared to a control group (Keynan et al., 2019). Neurofeedback studies to increase the amygdala response showed that the ability to identify or describe one's own emotions was correlated with the effectivity to increase amygdala activity (Young, Misaki, et al., 2017; Young et al., 2014; Zotev et al., 2011), suggesting that patients with less alexithymia might have better prerequisites to learn increasing their amygdala activity with neurofeedback. Together with our results, these studies indicate that the ability to identify and describe one's own feelings is related to the ability to gain control over the amygdala. However, it needs to be further studied whether the ability to identify and describe one's feelings predicts amygdala neurofeedback success or rather improves with the training.

In terms of the extended process model of emotion regulation (Gross, 2015), the effects found in alexithymia and the emotion-modulated startle indicate that amygdala neurofeedback might target different stages of emotion regulation: the selection and implementation of specific emotion regulation skills might be strengthened, and the identification of intrinsic emotional states might be increased. Further research is needed, which will be discussed in detail in section 5.4.2.

5.3 Important Limitations

Several methodological aspects of the studies presented in this thesis have to be reviewed critically regarding sample characteristics and the applied paradigms and designs.

5.3.1 Sample Characteristics

In Study II and III only female subjects participated. Thus, the majority of results presented in this work are limited to women. We chose to focus on female participants, due to gender differences in the involvement of prefrontal and limbic regions during emotion regulation (Domes et al., 2010; McRae, Ochsner, Mauss, Gabrieli, & Gross, 2008) and due to gender differences in emotional responsivity (Bradley, Codispoti, Sabatinelli, et al., 2001). As men and women are equally affected by BPD (Bayes & Parker, 2017; Grant et al., 2008), it will be important to replicate our findings in male patients.

Moreover, patients in Study III were allowed to take psychotropic medication with the restriction that the medication needed to be constant over the course of the study. Studies however indicate that psychotropic medication can mediate amygdala hyperreactivity in BPD patients

(Schulze et al., 2016). Specifically, amygdala hyperactivity was found only in unmedicated BPD patients. Medication thus might have already dampened amygdala activity and thus participants might have had difficulties to further downregulate the amygdala.

5.3.2 Study Design Characteristics

With regard to the assessment of psychophysiological effects of emotion regulation, Study I highlights the importance of corresponding assessment and quantification methods of autonomic measures across literature. We observed tremendous variations in quantification methods across the literature, in particular in studies assessing electrodermal responses. Due to the small number of studies, we were not able to account for such variation using moderator variables. We encourage future researchers to use similar research methodology and terminology of electrodermal responses and other autonomic measures as suggested by the state of the art literature (e.g., Boucsein et al., 2012) to make studies more comparable.

Moreover, Study II might have been underpowered to show evidence for the expected effect of probe timing on startle inhibition. A post hoc power analysis based on our results using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that 262 participants would be necessary to achieve reasonable power ($1 - \beta > .80$) in order to prove significance, given a true interaction effect.

Regarding Study III, a major limitation is the lack of a control group. Experimental studies that lack a no-feedback or sham feedback control group cannot differentiate whether the change in the target region's activity is caused by the feedback or by other elements such as mental strategies, attention, and motivation. Well-controlled study protocols are urgently needed for neurofeedback studies (for a review, see Sorger, Scharnowski, Linden, Hampson, & Young, 2019). Besides this, both downregulation of the amygdala BOLD response and best performance did not correlate significantly with any of our primary outcome measures. Thus, evidence for a mechanistic relationship between amygdala regulation and emotion dysregulation is still missing. The lack of significance may be a function of several causes, including lack of power and methodological issues. For example, the neurofeedback training was optimized to increase absolute training time but was less optimal in terms of quantifying downregulation of the amygdala. Additionally, shifts of behavior, physiology, and cognition during an emotional response are often loosely coupled (Bonanno & Keltner, 2004), and as such, a significant correlation is not necessarily observable, particularly in small sample sizes. To overcome the problem of multiple comparisons, we repeated statistical tests of primary endpoints in Study III with a

conservative correction for multiple tests (i.e. Bonferroni-correction). No statistical tests survived significance testing with correction. It should however be mentioned that several comparisons (e.g., ZAN-BPD, startle response) achieved medium effect sizes. Finally, improvements in emotion regulation, assessed with the emotion-modulated startle, faded to the follow-up test; that is, some training effects did not persist for 6 weeks. Overall, these results limit the potential effectiveness of the amygdala neurofeedback training in its current version as a intervention to target emotion dysregulation.

In Study III we moreover identified the ZAN-BPD interview as a suitable measure that may track changes in BPD psychopathology with amygdala neurofeedback training. However, the interview in Study III was conducted by the Study investigators, who were not blind to the timing of the assessment. It will be important for future neurofeedback studies to assign independent clinicians, who are blind to the assessment, and to replicate our results.

In addition, there are also some important methodological limitations to the neural tasks in Study III. First, participants in the EWMT performed very well already at the first training session, with a rate of correct responses of about 80%. This raises the question whether the EWMT is sensitive enough to distinguish between the intervention and control condition. It may be that longer periods of distraction would have increased the difficultly of the task. Moreover, a potentially significant reduction in amygdala response after the neurofeedback training could have been achieved with a simple confrontation with emotional stimuli, too (without working memory task)[12].Future studies using the EWMT should therefore include a control condition presenting emotional pictures but without a distractor. A limitation of the Backward Masking Task (BMT) is the lack of a proper control condition, which did not allow to identify changes in amygdala activation specific to the emotional faces. Future studies should include a neutral control condition with neutral faces or scrambled pictures.

5.4 Research Implications

5.4.1 Implication for Emotion Regulation Research

Findings from the present thesis, in particular Study I, underscore the benefits of using multiple levels of assessment when investigating the effects of emotion regulation to obtain a richer picture towards the success and limitations associated with emotion regulation. Research using

[12] Although we did not find significant reductions in amygdala response during the EWMT, this limitation is important to mention as it may also apply to other EWMT studies.

several indices of emotion regulation ability can have a substantial impact into how emotion regulation success is understood in the research literature. Findings from Study I highlight the variability of autonomic responding across different emotion regulation contexts. Studies focusing on just one autonomic outcome measure may be prone to misinterpretations, as a one-to-one correspondence between psychophysiological and psychological processes rarely exists (Cacioppo et al., 2007). Psychophysiological changes are often difficult to interpret because they are prone to several context factors as outlined in Study I and are also influenced by other psychological factors such as fatigue, drowsiness or stress in general (McRae & Shiota, 2017). Therefore, to obtain a comprehensive understanding of how emotion regulation is expressed, it is necessary for ongoing research to complete a thorough assessment using a range of subjective, expressive, physiological, and cognitive outcomes measures. These measures may also be assessed in a follow-up meta-analysis to see how the effects of individual psychophysiological effects, subjective experience and behavior are related to each other.

Beyond extending the assessment of emotion regulation, future studies should also incorporate more ecologically valid instruments of emotion regulation. Much of the emotion regulation research to date has contrasted one (or more) specific types of emotion regulation with a no-regulation control condition and the present thesis is no exception to that. Viewing emotional pictures and receiving relatively "pure" forms of emotion regulation every few seconds is important for both theoretical and statistical power reasons. While such a research strategy will continue to be a valuable approach in the future, it may not be the most reflective of every-day emotion regulation.

One avenue to follow could be to use emotional stimuli that are closer to the rich and complex emotional experiences in daily life such as emotional films (Gross & Levenson, 1995). Although (Morawetz et al., 2016a) did not find significant differences in emotion regulation when using films versus pictures, emotional films are usually presented over longer periods of time than pictures and thus allow the emotional response to unfold over time. With regard to emotion regulation paradigms assessed in patient populations such as BPD, personally more relevant emotional stimuli such as (autobiographical) scripts (Barnow et al., 2012; Zeier, Sandner, & Wessa, 2019) may also be used to induce strong negative emotions.

Beyond these ecological considerations, it might also be helpful to rethink how to operationalize emotion regulation success. The majority of emotion regulation studies that were included in Study I focused on intrapersonal emotion regulation. From this perspective, emotion regulation success is viewed as an intrapersonal process to cause a desirable internal emotional state (Gross, 1998a, 2015). The way emotion regulation success has been conceptualized, assessed

and discussed in Study II and III is no exception to this. From an evolutionary viewpoint however, emotions also serve to coordinate social interactions (Keltner, Haidt, & Shiota, 2006). In this way, emotions may also be seen as an interpersonal process with which individuals generate emotions in each other. If emotion regulation is studied as an interpersonal process, success can also be defined as the ability to effectively manage social interactions. One of the core features of BPD is disturbed social interactions (American Psychiatric Association, 2013) and it is assumed that emotion dysregulation leads to these interpersonal problems (Linehan, 1993). Hence, it may also be interesting to study emotion regulation in BPD during dyadic interactions in the laboratory, as has been successfully done with healthy individuals (Butler et al., 2003; Peters et al., 2014).

The extended process model offers many ways to operationalize emotion regulation success, e.g. implementing a given strategy, the ability to perceive the target emotion and to flexibly switch a strategy if needed in a context-appropriate manner. Each of these aspects are also prone to dysfunctions in psychopathologies (Gross, 2015; Sheppes et al., 2015). A limitation to the classical explicit emotion regulation paradigm of the present thesis is that it rather focuses on the implementation of an (explicitly instructed) emotion regulation strategy (Sheppes et al., 2015), whereas the valuation and selection stage is not considered because a strategy instruction (e.g., downregulate, reappraise) is provided by the researchers. Future studies should shift the focus from one element of one regulatory stage (e.g., implementation) to examining various elements of different regulatory stages. Recently, emotion regulation studies have tried to focus on other stages of the extended process model, e.g. using emotion regulation choice paradigms (Murphy & Young, 2020; Sauer et al., 2016; Shafir, Guarino, Lee, & Sheppes, 2017; Sheppes et al., 2014). Sauer et al. (2016) for example contrasted distraction and reappraisal in healthy and BPD patients and found that both groups preferred distraction over reappraisal under high-intensity stimuli. More research is needed to understand situations in which both healthy individuals and patients choose less fortunate emotion regulation strategies. Such paradigms may also be used in the future to implement psychophysiological responses, such as the emotion-modulated startle.

5.4.2 Implications for Future Amygdala Neurofeedback Studies

Despite promising effects on several system levels obtained in Study III, effects did not surpass the significance level when correcting for multiple comparisons. Some of the effects even faded to the follow-up sessions. Future clinical studies on amygdala neurofeedback should a) include larger sample sizes to achieve significant effects that surpass the correction for multiple testing,

b) should add booster sessions or mental training between sessions (Paret et al., 2019; Sulzer et al., 2013) to gain more stable effects and c) should improve paradigms to measure changes in emotion regulation. In section 5.4.1, suggestions for general improvements in emotion regulation assessments were already outlined. With regard to amygdala neurofeedback studies, pre-post data of transfer runs may also improve the understanding of whether and to what extent amygdala neurofeedback training can change the activation of neural circuitries of emotion regulation without feedback. A recent meta-analysis on amygdala neurofeedback highlighted the inclusion of a practice and a transfer run before and after the neurofeedback training as a relevant factor for the interpretation of the causality of the effects (Barreiros et al., 2019). In addition to this, emotion regulation may be assessed in every-day life to increase ecological validity. The results of Study III indicate promising effects of negative affect and affective instability assessed with ecological momentary assessment (EMA). In light of these results, we could speculate that the neurofeedback effects on emotion regulation might be captured equally well in a more ecologically valid environment. To date, no study has tested emotion regulation in daily life with EMA (Colombo et al., 2019), but it is already possible to assess psychophysiological measures with EMA (Raugh, Chapman, Bartolomeo, Gonzalez, & Strauss, 2019). The development of EMA assessment of emotion regulation paired with psychophysiological recordings might be a new avenue to understand what triggers emotion regulation choices in daily life and how this might be improved via amygdala neurofeedback.

Future studies could also study associations between amygdala neurofeedback and the different stages of emotion regulation proposed in the extended process model (Gross, 2015; Paret & Hendler, 2020). At present, there is relatively modest empirical support for the extended process model of emotion regulation (Sheppes et al., 2015) and the learning mechanisms of amygdala neurofeedback have not been fully understood yet (Paret & Hendler, 2020). In general, the empirical evaluation of basic processes of the extended process model requires assessing the three basic elements (perception, valuation, and action) as they process different types of information in the three regulatory stages of emotion regulation (identification, selection and implementation). As mentioned above, classical emotion regulation paradigms considered in the present thesis captured only a fraction of these stages, e.g. action stage of implementation. The study of multiple stages of the extended process model might allow to understand the mechanisms of amygdala neurofeedback and vice versa (Paret & Hendler, 2020). During neurofeedback, the contingent reinforcement of behavior may help to select and implement mental strategies that are successful for neurofeedback control via operant learning (Caria, 2016; Strehl, 2014). In addition, participants may also refine the perception of intrinsic processes that are

correlated with the feedback via associative learning (Kotchoubey, Kubler, Strehl, Flor, & Birbaumer, 2002; Schwartz, Collura, Kamiya, & Schwartz, 2016). A necessary requirement for operant and associative learning is to monitor the feedback signal (Paret et al., 2018). As such, amygdala neurofeedback training may refine the ability to perceive, implement and monitor an internal emotional state that corresponds to the feedback but also helps to select and implement a certain strategy that is associated to successful control of the feedback signal. In other words, neurofeedback may tackle and improve abilities that are required during different stages of the extended emotion regulation process model (i.e. identification, selection and implementation stage with perception, valuation and action as sub-stages). A paradigm that captures these stages may be tested before and after amygdala neurofeedback training to understand what aspect of neurofeedback learning may transfer to emotion regulation abilities: Does amygdala neurofeedback help to perceive an emotion, select appropriate emotion regulation strategies or implement and switch the strategy if needed? In addition, these emotion regulation success metrices may be used to predict amygdala neurofeedback success.

Finally, Study III implies that multiple neurofeedback training runs are necessary to observe amygdala downregulation with neurofeedback in BPD. As real-time fMRI neurofeedback is cost-intensive and not all clinical institutions have access to MRI, the question arises how real-time fMRI neurofeedback can be realized in a clinical setting in the future. Recently, promising results have been achieved with amygdala-electrical fingerprint, a novel imaging approach using fMRI-inspired electroencephalography (EEG). Keynan et al. (2016) demonstrated that learned downregulation of the amygdala-electrical fingerprint, which was shown to reliably predict amygdala BOLD activity in the first place, facilitated volitional downregulation of amygdala BOLD activity via real-time fMRI, which manifested as reduced amygdala reactivity to visual stimuli. More recently, healthy individuals undergoing a stressful military training program underwent amygdala-electrical fingerprint neurofeedback sessions and showed reduced alexithymia and better stress coping abilities following the training relative to controls (Keynan et al., 2019). Amygdala electrical-fingerprint neurofeedback has the advantage to be more cost-efficient than fMRI amygdala neurofeedback. It remains to be tested in the future whether amygdala-electrical fingerprint neurofeedback is an effective training to target emotion regulation problems in BPD.

5.5 Conclusion and Clinical Outlook

The present thesis answered three important questions: What psychophysiological responses are suited to measure emotion regulation in general and what aspects of the study design moderates these effects? Does the timing of the startle probe influence effects of emotion regulation? And finally, how can improvements in emotion dysregulation after amygdala neurofeedback training be measured in BPD? Based on our results, important implications for future emotion regulation and amygdala neurofeedback studies were outlined above. However, the present thesis may also provide practical implications for clinical research and therapy.

There has been a long tradition to consider suppressing affect a poorly effective emotion regulation strategy because it increases subjective and physiological arousal (e.g., Gross & Levenson, 1993; Gross, 1998b). With respect to physiological effects, our results on suppression are less clear. Suppression strategies decreased finger temperature, indicative of increased sympathetic arousal, however effects were rather small and stemmed from a handful of studies. In addition, we did not find that suppression significantly increased skin conductance level, which is a pure measure of sympathetic arousal. With respect to skin conductance level, effects were highly contradictory across studies and were moderated by important study characteristics. This implies that the effectivity of a strategy to decrease sympathetic arousal might depend on the context in which it is being applied. More research is needed to understand both the short- and long-term effects of certain emotion regulation strategies in different contexts and how this is related to psychophysiological responding.

With regard to Study III, it should be emphasized that replication of the robustness of results is a necessary precondition before drawing any conclusion for clinical practice. However, potential clinical implications still may spark future research. First of all, results imply that amygdala neurofeedback may address the acquisition and refinement of emotion regulation skills. At the same time, results also indicate that amygdala neurofeedback training might increase the ability to perceive internal emotional states as seen by improvements in alexithymia. Dialectical Behavior Therapy (DBT; Linehan, 1993) already has a lot of excellent suggestions for BPD patients how to reduce stress (e.g., distress tolerance skills) and strong negative emotions (e.g., “Emotionssurfing” in the German version) and suggestions how to increase the ability to perceive internal emotion states via mindfulness strategies. If amygdala neurofeedback has been shown to be effective in BPD, the implementation of this treatment into DBT skills training might therefore offer valuable possibilities for patients to further implement and practice emotion regulation skills.

As a concluding remark, during our training sessions in Study III, we often received the feedback from participants that for the first time they were “able to control at least something in their lives” – that is, the feedback thermometer. This gives rise to the hope that amygdala neurofeedback training indeed has a positive impact to patients with BPD. But only a randomized controlled trial will tell for sure whether this impact is beyond placebo.

6 SUMMARY

The way we regulate emotions is a powerful determinant of behavior and directly impacts affect and physiology. Many mental disorders, such as borderline personality disorder, are in large part disorders of emotion dysregulation. Because of its important role in mental health, research has endeavored to understand the mechanisms and biological underpinnings of emotion regulation and to create trainings and specific clinical programs that aim to augment the ability to regulate emotions. The assessment of psychophysiological responses represents an important complementary method to quantify emotion regulation in both studies on healthy individuals and studies assessing clinical emotion regulation trainings. However, psychophysiological effects have been inconsistent across literature, which impedes informed decisions about suitable psychophysiological variables of emotion regulation experiments and clinical trainings. A new technique assumed to improve emotion regulation is amygdala neurofeedback training. Because patients with borderline personality disorder show hyperreactivity of the amygdala likely underlying the severe emotion regulation problems they suffer from, amygdala neurofeedback training may be a candidate training to improve emotion regulation in these patients. Until now, it has been unclear which aspects of psychopathology and emotion regulation may change with neurofeedback-aided amygdala downregulation in borderline personality disorder, which is urgently needed to determine a primary outcome measure for future randomized controlled trails. To fill these gaps, the present doctoral thesis identified the effects of psychophysiological responses of emotion regulation as well as important moderators and identified primary outcome measures of emotion dysregulation after neurofeedback training in patients with borderline personality disorder.

In total, three studies were conducted. In Study I, a total of 1353 studies on psychophysiological responses of emotion regulation were screened through a systematic search of articles and meta-analyses were used to evaluate effect sizes of instructed downregulation strategies on common autonomic and electromyographic measures. Following this, Study II systematically tested effects of the startle probe timing on startle responses during emotion regulation in 47 healthy individuals. Study II aimed at optimizing emotion regulation assessment with the emotion-modulated startle that was then used in Study III. In Study III, a four-session amygdala neurofeedback training was tested in 24 female patients with borderline personality disorder. Before and after the neurofeedback training, as well as at a 6-week follow-up assessment, measures of emotion dysregulation and borderline personality disorder psychopathology were tested at diverse levels of analysis.

Results from Study I demonstrate that effects of emotion regulation on autonomic measures, even if significant, were small and heterogeneous across studies, while electromyographic measures were more homogeneous and revealed medium effect sizes. Important study characteristics such as the study design, control instruction and trial duration moderated some autonomic effects of suppression and reappraisal. Study II demonstrated a significant inhibition of the startle response with emotion downregulation. Startle probes delivered at >7 seconds into the regulation phase were useful to quantify reappraisal effects, although earlier probes did not yield significantly smaller effects. Finally, Study III demonstrated that the inhibition of the startle with emotion downregulation increased after the training, suggesting improved emotion regulation abilities. In addition, we could demonstrate that general BPD psychopathology as well as affective instability and negative affect in daily life improved after training. However, after correction for multiple comparisons, observed effect sizes did not surpass the significance level and some effects (e.g., startle) faded to the 6-week follow-up assessment.

In sum, the present thesis provides the groundwork for future randomized controlled trials of amygdala neurofeedback training and enables future laboratory and clinical studies to gain more stable effects in psychophysiological measurements of emotion regulation. In particular, the findings implicate that with regard to emotion regulation research, autonomic measures appear to be highly variable and thus should be selected carefully. In addition, we need more comparable psychophysiological set-ups in the empirical study of emotion regulation. The emotion-modulated startle not only proved to be a robust measure to quantify emotion regulation effects in general, but also appeared to be suitable to track improvements in emotion regulation in the context of a neurofeedback training targeting emotion dysregulation. With respect to emotion regulation outcome measures for future amygdala neurofeedback studies, further improvement of the specific paradigms is needed. In addition, the neurofeedback training itself should be optimized in terms of e.g. training time and booster sessions. Future placebo-controlled trials must then confirm that the treatment is effective in improving emotion regulation in those with borderline personality disorder.

7 REFERENCES

Adolph, D., & Pause, B. M. (2012). Different time course of emotion regulation towards odors and pictures: Are odors more potent than pictures? *Biological Psychology, 91*(1), 65-73. doi:10.1016/j.biopsycho.2012.05.003

Ajaya, Y., Peckham, A. D., & Johnson, S. L. (2016). Emotion regulation and mania risk: Differential responses to implicit and explicit cues to regulate. *Journal of Behavior Therapy and Experimental Psychiatry, 50*, 283-288. doi:10.1016/j.jbtep.2015.10.004

Aldao, A., & Mennin, D. S. (2012). Paradoxical cardiovascular effects of implementing adaptive emotion regulation strategies in generalized anxiety disorder. *Behaviour Research and Therapy, 50*(2), 122-130. doi:10.1016/j.brat.2011.12.004

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review, 30*(2), 217-237. doi:10.1016/j.cpr.2009.11.004

Alpers, G. W., Adolph, D., & Pauli, P. (2011). Emotional scenes and facial expressions elicit different psychophysiological responses. *International Journal of Psychophysiology, 80*(3), 173-181. doi:10.1016/j.ijpsycho.2011.01.010

American Psychiatric Association. (2013). *Diagnostic and Statical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association.

Andersson, J. L., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling geometric deformations in EPI time series. *NeuroImage, 13*(5), 903-919. doi:0.1006/nimg.2001.0746

Ansell, E. B., Sanislow, C. A., McGlashan, T. H., & Grilo, C. M. (2007). Psychosocial impairment and treatment utilization by patients with borderline personality disorder, other personality disorders, mood and anxiety disorders, and a healthy comparison group. *Comprehensive Psychiatry, 48*(4), 329-336. doi:10.1016/j.comppsych.2007.02.001

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology, 10*(3), 229-240. doi:10.1037/1089-2680.10.3.229

Attwell, D., & Iadecola, C. (2002). The neural basis of functional brain imaging signals. *Trends in Neuroscience, 25*(12), 621-625. doi:10.1016/s0166-2236(02)02264-6

Azbel-Jackson, L., Butler, L. T., Ellis, J. A., & van Reekum, C. M. (2015). Stay calm! Regulating emotional responses by implementation intentions: Assessing the impact on physiological and subjective arousal. *Cognition & Emotion*, 1-15. doi:10.1080/02699931.2015.1049515

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*(1), 23-32. doi:10.1016/0022-3999(94)90005-1

Banks, S. J., Eddy, K. T., Angstadt, M., Nathan, P. J., & Phan, K. L. (2007). Amygdala-frontal connectivity during emotion regulation. *Social Cognitive and Affective Neuroscience, 2*(4), 303-312. doi:10.1093/scan/nsm029

Barnow, S., Limberg, A., Stopsack, M., Spitzer, C., Grabe, H. J., Freyberger, H. J., & Hamm, A. (2012). Dissociation and emotion regulation in borderline personality disorder. *Psychological Medicine, 42*(4), 783-794. doi:10.1017/S0033291711001917

Barreiros, A. R., Almeida, I., Baia, B. C., & Castelo-Branco, M. (2019). Amygdala Modulation During Emotion Regulation Training With fMRI-Based Neurofeedback. *Frontiers in Human Neuroscience, 13*, 89. doi:10.3389/fnhum.2019.00089

Baskin-Sommers, A. R., Curtin, J. J., Larson, C. L., Stout, D., Kiehl, K. A., & Newman, J. P. (2012). Characterizing the anomalous cognition-emotion interactions in externalizing. *Biological Psychology, 91*(1), 48-58. doi:10.1016/j.biopsycho.2012.05.001

Baskin-Sommers, A. R., Vitale, J. E., Maccoon, D., & Newman, J. P. (2012). Assessing emotion sensitivity in female offenders with borderline personality symptoms: results from a fear-potentiated startle paradigm. *Journal of Abnormal Psychology, 121*(2), 477-483. doi:10.1037/a0026753

Bateman, A. W., & Fonagy, P. (2007). *Mentalization-based treatment for borderline persoanlity disorder: A practical guide.* Oxford: Oxford University Press.

Baur, R., Conzelmann, A., Wieser, M. J., & Pauli, P. (2015). Spontaneous emotion regulation: Differential effects on evoked brain potentials and facial muscle activity. *International Journal of Psychophysiology, 96*(1), 38-48. doi:10.1016/j.ijpsycho.2015.02.022

Bayes, A., & Parker, G. (2017). Borderline personality disorder in men: A literature review and illustrative case vignettes. *Psychiatry Reseach, 257*, 197-202. doi:10.1016/j.psychres.2017.07.047

Bebko, G. M., Franconeri, S. L., Ochsner, K. N., & Chiao, J. Y. (2011). Look before you regulate: differential perceptual strategies underlying expressive suppression and cognitive reappraisal. *Emotion, 11*(4), 732-742. doi:10.1037/a0024009

Ben-Naim, S., Hirschberger, G., Ein-Dor, T., & Mikulincer, M. (2013). An experimental study of emotion regulation during relationship conflict interactions: The moderating role of attachment orientations. *Emotion, 13*(3), 506-519. doi:10.1037/a0031473

Berking, M., & Znoj, H. (2008). Entwicklung und Validierung eines Fragebogens zur standardisierten Selbsteinschätzung emotionaler Kompetenzen (SEK-27). *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie, 56*(2), 141-153. doi:10.1024/1661-4747.56.2.141

Bernat, E. M., Cadwallader, M., Seo, D., Vizueta, N., & Patrick, C. J. (2011). Effects of Instructed Emotion Regulation on Valence, Arousal, and Attentional Measures of Affective Processing. *Developmental Neuropsychology, 36*(4), 493-518. doi:10.1080/87565641.2010.549881

Berntson, G. G., Quigley, K. S., Norman, G. J., & Lozano, D. L. (2016). Cardiovascular Psychophysiology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (4 ed., pp. 183–216). Cambridge: Cambridge University Press.

Birbaumer, N., Ruiz, S., & Sitaram, R. (2013). Learned regulation of brain metabolism. *Trends in Cognitive Sciences, 17*(6), 295-302. doi:10.1016/j.tics.2013.04.009

Black, D. W., Blum, N., Pfohl, B., & Hale, N. (2004). Suicidal behavior in borderline personality disorder: prevalence, risk factors, prediction, and prevention. *Journal of Personality Disorders, 18*(3: Special issue), 226-239. doi:10.1521/pedi.18.3.226.35445

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology, 42*(1), 1-15. doi:10.1111/j.1469-8986.2005.00271.x

Bohus, M., & Schmahl, C. (2007). Psychopathologie und Therapie der Borderline-Persönlichkeitsstörung. *Deutsches Ärzteblatt, 103*(49), 3345-3352. doi:10.1007/s00115-007-2341-x

Bohus, M., Schmahl, C., Herpertz, S. C., Lieb, K., Berger, M., Roepke, S., . . . Lyssenko, L. (2016). Leitliniengerechte stationare psychiatrisch-psychotherapeutische Behandlung der Borderline-Personlichkeitsstorung: Normative Personalbedarfsermittlung. *Der Nervenarzt, 87*(7), 793-745. doi:10.1007/s00115-016-0132-y

Bonanno, G., & Keltner, D. (2004). Brief Report The coherence of emotion systems: Comparing “on‐line” measures of appraisal and facial expressions, and self‐report. *Cognition & Emotion, 18*(3), 431-444. doi:10.1080/02699930341000149

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Reseach Synthesis Methods, 1*(2), 97-111. doi:10.1002/jrsm.12

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology, 49*(8), 1017-1034. doi:10.1111/j.1469-8986.2012.01384.x

Braams, B. R., Blechert, J., Boden, M. T., & Gross, J. J. (2012). The effects of acceptance and suppression on anticipation and receipt of painful stimulation. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(4), 1014-1018. doi:10.1016/j.jbtep.2012.04.001

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276-298. doi:10.1037/1528-3542.1.3.276

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion, 1*(3), 300-319. doi:10.1037/1528-3542.1.3.300

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1993). Pictures as prepulse: attention and emotion in startle modification. *Psychophysiology, 30*(5), 541-545. doi:10.1111/j.1469-8986.1993.tb02079.x

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. doi:10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane, L. Nadel, R. D. Lane, & L. Nadel (Eds.), *Cognitive neuroscience of emotion.* (pp. 242-276). New York, NY, US: Oxford University Press.

Bradley, M. M., Lang, P. J., & Cuthbert, B. N. (1993). Emotion, novelty, and the startle reflex: habituation in humans. *Behavioral Neuroscience, 107*(6), 970-980. doi:10.1037/0735-7044.107.6.970

Brühl, A. B., Scherpiet, S., Sulzer, J., Stämpfli, P., Seifritz, E., & Herwig, U. (2014). Real-time neurofeedback using functional MRI could improve down-regulation of amygdala activity during emotional stimulation: a proof-of-concept study. *Brain Topography, 27*(1), 138-148. doi:10.1007/s10548-013-0331-9

Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., . . . Ochsner, K. N. (2014). Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies. *Cerberal Cortex, 24*, 2981–2990. doi:10.1093/cercor/bht154

Bulut, N. S., Würz, A., Küpeli, N. Y., Bulut, G. Ç., & Sungur, M. Z. (2018). Heart rate variability response to affective pictures processed in and outside of conscious awareness: three consecutive studies on emotional regulation. *International Journal of Psychophysiology, 129*, 18-30. doi:10.1016/j.ijpsycho.2018.05.006

Butler, E. A., Egloff, B., Wlhelm, F. H., Smith, N. C., Erickson, E. A., & Gross, J. J. (2003). The social consequences of expressive suppression. *Emotion, 3*(1), 48-67. doi:10.1037/1528-3542.3.1.48

Butler, E. A., Gross, J. J., & Barnard, K. (2014). Testing the effects of suppression and reappraisal on emotional concordance using a multivariate multilevel model. *Biological Psychology, 98*, 6-18. doi:10.1016/j.biopsycho.2013.09.003

Butler, E. A., Wilhelm, F. H., & Gross, J. J. (2006). Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. *Psychophysiology, 43*(6), 612-622. doi:10.1111/j.1469-8986.2006.00467.x

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (Vol. 2, pp. 173-191). New York: Guilford Press.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45*(1), 16-28.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2007). *Handbook of Psychophysiology* (Vol. 2). New York: Cambridge University Press.

Campbell-Sills, L., & Barlow, D. H. (2007). Incorporating Emotion Regulation into Conceptualizations and Treatments of Anxiety and Mood Disorders. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 542-559). New York: Guilford Press.

Caria, A. (2016). Self-Regulation of Blood Oxygenation Level Dependent Response: Primary Effect or Epiphenomenon? *Frontiers in Neuroscience, 10*, 117. doi:10.3389/fnins.2016.00117

Caria, A., Veit, R., Sitaram, R., Lotze, M., Weiskopf, N., Grodd, W., & Birbaumer, N. (2007). Regulation of anterior insular cortex activity using real-time fMRI. *NeuroImage, 35*(3), 1238-1246. doi:10.1016/j.neuroimage.2007.01.018

Carpenter, R. W., & Trull, T. J. (2013). Components of emotion dysregulation in borderline personality disorder: a review. *Current Psychiatry Reports, 15*(1), 335. doi:10.1007/s11920-012-0335-2.

Chu, V. C., Lucas, G. M., Lei, S., Mozgai, S., Khooshabeh, P., & Gratch, J. (2019). Emotion regulation in the Prisoner's Dilemma: Effects of reappraisal on behavioral measures and cardiovascular measures of challenge and threat. *Frontiers in Human Neuroscience, 13*, 50. doi:10.3389/fnhum.2019.00050

Cohen, J. E. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Colby, C. Z., Lanzetta, J. T., & Kleck, R. E. (1977). Effects of the Expression of Pain on Autonomic and Pain Tolerance Responses to Subject‐Controlled Pain. *Psychophysiology, 14*(6), 537-540. doi:10.1111/j.1469-8986.1977.tb01194.x

Colombo, D., Fernandez-Alvarez, J., Garcia Palacios, A., Cipresso, P., Botella, C., & Riva, G. (2019). New Technologies for the Understanding, Assessment, and Intervention of Emotion Regulation. *Frontiers in psychology, 10*, 1261. doi:10.3389/fpsyg.2019.01261

Conzelmann, A., McGregor, V., & Pauli, P. (2015). Emotion regulation of the affect-modulated startle reflex during different picture categories. *Psychophysiology, 52*(9), 1257-1262. doi:10.1111/psyp.12450

Côté, S., Gyurak, A., & Levenson, R. W. (2010). The ability to regulate emotion is associated with greater well-being, income, and socioeconomic status. *Emotion, 10*(6), 923-933. doi:10.1037/a0021156

Cox, R. W., Jesmanowicz, A., & Hyde, J. S. (1995). Real-time functional magnetic resonance imaging. *Magnetic Resonance in Medicine, 33*(2), 230-236. doi:10.1002/mrm.1910330213

Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbui, C., & Cuijpers, P. (2017). Efficacy of Psychotherapies for Borderline Personality Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry, 74*(4), 319-328. doi:10.1001/jamapsychiatry.2016.4287

Cuthbert, B. N., Bradley, M. M., & Lang, P. J. (1996). Probing picture perception: Activation and emotion. *Psychophysiology, 33*(2), 103-111. doi:10.1111/j.1469-8986.1996.tb02114.x

Cyders, M. A., Littlefield, A. K., Coffey, S., & Karyadi, K. A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. *Addictive Behavior, 39*(9), 1372-1376. doi:10.1016/j.addbeh.2014.02.013

D'Agostino, A., Covanti, S., Rossi Monti, M., & Starcevic, V. (2017). Reconsidering Emotion Dysregulation. *Psychiatric Quarterly, 88*(4), 807-825. doi:10.1007/s11126-017-9499-6

Dan-Glauser, E. S., & Gross, J. J. (2011). The temporal dynamics of two response-focused forms of emotion regulation: experiential, expressive, and autonomic consequences. *Psychophysiology, 48*(9), 1309-1322. doi:10.1111/j.1469-8986.2011.01191.x

Dan-Glauser, E. S., & Gross, J. J. (2015). The temporal dynamics of emotional acceptance: Experience, expression, and physiology. *Biological Psychology, 108*, 1-12. doi:10.1016/j.biopsycho.2015.03.005

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior Reseach Methods, 43*(2), 468-477. doi:10.3758/s13428-011-0064-1

Daros, A. R., & Williams, G. E. (2019). A Meta-analysis and Systematic Review of Emotion-Regulation Strategies in Borderline Personality Disorder. *Harvard Review of Psychiatry, 27*(4), 217-232. doi:10.1097/hrp.0000000000000212

Davis, M. (1992). The role of the amygdala in fear-potentiated startle: implications for animal models of anxiety. *Trends in Pharmacological Sciences, 13*, 35-41. doi:10.1016/0165-6147(92)90014-W

Davis, M. (2000). The role of the amygdala in conditioned and unconditioned fear and anxiety. In J. P. Aggleton (Ed.), *The amygdala* (Vol. 2, pp. 213-287). Oxford, England: Oxford University Press.

Davis, M., Campeau, S., Kim, M., & Falls, W. A. (1995). Neural systems of emotion: The amygdala's role in fear and anxiety. In J. L. McGaugh, N. M. Weinberger, & G. Lynch (Eds.), *Brain and memory: Modulation and mediation of neuroplasticity* (pp. 3-40). New York, NY: Oxford University Press.

Davis, M., Walker, D. L., & Lee, Y. S. (1999). Neurophysiology and neuropharmacology of startle and its affective modulation. In M. E. Dawson, A. M. Schell, & A. H. Böhmelt (Eds.), *Startle modification: implications for neuroscience, cognitive science, and clinical science* (pp. 95–114). Cambridge: University Press.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2016). The Electrodermal System. In G. G. Berntson, J. T. Cacioppo, & L. G. Tassinary (Eds.), *Handbook of psychophysiology* (4 ed., pp. 217-243). Cambridge: Cambridge University Press.

De Witte, N. A., Sutterlin, S., Braet, C., & Mueller, S. C. (2017). Psychophysiological correlates of emotion regulation training in adolescent anxiety: Evidence from the novel PIER task. *Journal of Affective Disorders, 214*, 89-96. doi:10.1016/j.jad.2017.03.012

Delgado, M. R., Nearing, K. I., LeDoux, J. E., & Phelps, E. A. (2008). Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. *Neuron, 59*(5), 829-838. doi:10.1016/j.neuron.2008.06.029

Demaree, H. A., Schmeichel, B. J., Robinson, J. L., & Everhart, D. E. (2004). Behavioural, affective, and physiological effects of negative and positive emotional exaggeration. *Cognition & Emotion, 18*(8), 1079-1097. doi:10.1080/02699930441000085

Demaree, H. A., Schmeichel, B. J., Robinson, J. L., Pu, J., Everhart, D. E., & Berntson, G. G. (2006). Up- and down-regulating facial disgust: affective, vagal, sympathetic, and respiratory consequences. *Biological Psychology, 71*(1), 90-99. doi:10.1016/j.biopsycho.2005.02.006

Denson, T. F., Creswell, J. D., Terides, M. D., & Blundell, K. (2014). Cognitive reappraisal increases neuroendocrine reactivity to acute social stress and physical pain. *Psychoneuroendocrinology, 49*, 69-78. doi:10.1016/j.psyneuen.2014.07.003

Denson, T. F., Grisham, J. R., & Moulds, M. L. (2011). Cognitive reappraisal increases heart rate variability in response to an anger provocation. *Motivation and Emotion, 35*(1), 14-22. doi:10.1007/s11031-011-9201-5

Detenber, B. H., Simons, R. F., & Bennett Jr, G. G. (1998). Roll 'em!: The effects of picture motion on emotional responses. *Journal of Broadcasting & Electronic Media, 42*(1), 113-127. doi:10.1080/08838159809364437

Deveney, C. M., & Pizzagalli, D. A. (2008). The cognitive consequences of emotion regulation: An ERP investigation. *Psychophysiology, 45*(3), 435-444. doi:10.1111/j.1469-8986.2007.00641.x

Di Simplicio, M., Costoloni, G., Western, D., Hanson, B., Taggart, P., & Harmer, C. J. (2012). Decreased heart rate variability during emotion regulation in subjects at risk for psychopathology. *Psychological Medicine, 42*(8), 1775-1783. doi:10.1017/S0033291711002479

Diekhof, E. K., Geier, K., Falkai, P., & Gruber, O. (2011). Fear is only as deep as the mind allows: a coordinate-based meta-analysis of neuroimaging studies on the regulation of negative affect. *NeuroImage, 58*(1), 275-285. doi:10.1016/j.neuroimage.2011.05.073

Diers, K., Weber, F., Brocke, B., Strobel, A., & Schönfeld, S. (2014). Instructions matter: a comparison of baseline conditions for cognitive emotion regulation paradigms. *Frontiers in Psychology, 5*, 347. doi:10.3389/fpsyg.2014.00347

Dillon, D. G., & LaBar, K. S. (2005). Startle modulation during conscious emotion regulation is arousal-dependent. *Behavioral Neuroscience, 119*(4), 1118-1124. doi:10.1037/0735-7044.119.4.1118

Dillon, D. G., & Pizzagalli, D. A. (2007). Inhibition of action, thought, and emotion: a selective neurobiological review. *Applied and Preventive Psychology, 12*(3), 99-114. doi:10.1016/j.appsy.2007.09.004

Dixon-Gordon, K. L., Aldao, A., & De Los Reyes, A. (2015). Emotion regulation in context: Examining the spontaneous use of strategies across emotional intensity and type of emotion. *Personality and Individual Differences, 86*, 271-276. doi:10.1016/j.paid.2015.06.011

Domes, G., Schulze, L., Bottger, M., Grossmann, A., Hauenstein, K., Wirtz, P. H., . . . Herpertz, S. C. (2010). The neural correlates of sex differences in emotional reactivity and emotion regulation. *Human Brain Mapping, 31*(5), 758-769. doi:10.1002/hbm.20903

Driscoll, D., Tranel, D., & Anderson, S. W. (2009). The effects of voluntary regulation of positive and negative emotion on psychophysiological responsiveness. *International Journal of Psychophysiology, 72*(1), 61-66. doi:10.1016/j.ijpsycho.2008.03.012

Ebner-Priemer, U. W., Badeck, S., Beckmann, C., Wagner, A., Feige, B., Weiss, I., . . . Bohus, M. (2005). Affective dysregulation and dissociative experience in female patients with borderline personality disorder: a startle response study. *Journal Psychiatry Research, 39*(1), 85-92. doi:10.1016/j.jpsychires.2004.05.001

Ebner-Priemer, U. W., Eid, M., Kleindienst, N., Stabenow, S., & Trull, T. J. (2009). Analytic strategies for understanding affective (in)stability and other dynamic processes in psychopathology. *Journal of Abnormal Psychology, 118*(1), 195-202. doi:10.1037/a0014868

Ebner-Priemer, U. W., Kuo, J., Kleindienst, N., Welch, S. S., Reisch, T., Reinhard, I., . . . Bohus, M. (2007). State affective instability in borderline personality disorder assessed by ambulatory monitoring. *Psychological Medicine, 37*(7), 961-970. doi:10.1017/S0033291706009706

Edelmann, R. J., & Baker, S. R. (2002). Self‐reported and actual physiological responses in social phobia. *British Journal of Clinical Psychology, 41*(1), 1-14. doi:10.1348/014466502163732

Efinger, L., Thuillard, S., & Dan-Glauser, E. (2019). Distraction and reappraisal efficiency on immediate negative emotional responses: role of trait anxiety. *Anxiety, Stress, & Coping, 32*(4), 412-427. doi:10.1080/10615806.2019.1597859

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal,* (315), 629-634. doi:10.1136/bmj.315.7109.629

Ehring, T., Tuschen-Caffier, B., Schnulle, J., Fischer, S., & Gross, J. J. (2010). Emotion regulation and vulnerability to depression: spontaneous versus instructed use of emotion suppression and reappraisal. *Emotion, 10*(4), 563-572. doi:10.1037/a0019010

Eippert, F., Veit, R., Weiskopf, N., Erb, M., Birbaumer, N., & Anders, S. (2007). Regulation of emotional responses elicited by threat-related stimuli. *Human Brain Mapping, 28*(5), 409-423. doi:10.1002/hbm.20291

Ekman, P., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions.* New York, NY: Oxford University Press.

Emmert, K., Kopel, R., Sulzer, J., Bruhl, A. B., Berman, B. D., Linden, D. E. J., . . . Haller, S. (2016). Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: How is brain regulation mediated? *NeuroImage, 124*(Pt A), 806-812. doi:10.1016/j.neuroimage.2015.09.042

Enders, C. K. (2006). A primer on the use of modern missing-data methods in psychosomatic medicine research. *Psychosomatic Medicine, 68*(3), 427-436. doi:10.1097/01.psy.0000221275.75056.d8

Enriquez-Geppert, S., Smit, D., Pimenta, M. G., & Arns, M. (2019). Neurofeedback as a Treatment Intervention in ADHD: Current Evidence and Practice. *Current Psychiatry Reports, 21*(6), 46. doi:10.1007/s11920-019-1021-4

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience, 16*(11), 693. doi:10.1038/nrn4044

Fassbinder, E., Rudolf, S., Bussiek, A., Kroger, C., Arnold, R., Greggersen, W., . . . Schweiger, U. (2007). Effektivität der dialektischen Verhaltenstherapie bei Patienten mit Borderline-Persönlichkeitsstörung im Langzeitverlauf. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 57*(3-4), 161-169. doi:10.1055/s-2006-952015

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Reseach Methods, 39*(2), 175-191. doi:10.3758/bf03193146

Fechir, M., Schlereth, T., Purat, T., Kritzmann, S., Geber, C., Eberle, T., . . . Birklein, F. (2008). Patterns of sympathetic responses induced by different stress tasks. *The Open Neurology Journal, 2*, 25-31. doi:10.2174/1874205x00802010025

First, M. B., Gibbon, M., & Spitzer, R. L. (1997). *User's guide for the structured clinical interview for DSM-IV axis II personality disorders: SCID-II*. Washington, DC: American Psychiatric Press.

Fitzpatrick, S., & Kuo, J. R. (2016). The impact of stimulus arousal level on emotion regulation effectiveness in borderline personality disorder. *Psychiatry Research, 241*, 242-248. doi:10.1016/j.psychres.2016.05.004

Frank, D., Dewitt, M., Hudgens-Haney, M., Schaeffer, D., Ball, B., Schwarz, N., . . . Sabatinelli, D. (2014). Emotion regulation: Quantitative meta-analysis of functional activation and deactivation. *Neuroscience and Biobehavioral Reviews, 45*, 202-211. doi:/10.1016/j.neubiorev.2014.06.010

Fuentes‐Sánchez, N., Jaén, I., Escrig, M. A., Lucas, I., & Pastor, M. C. (2019). Cognitive reappraisal during unpleasant picture processing: Subjective self‐report and peripheral physiology. *Psychophysiology*, e13372. doi:10.1111/psyp.13372

Gallo, I. S., Keil, A., McCulloch, K. C., Rockstroh, B., & Gollwitzer, P. M. (2009). Strategic automation of emotion regulation. *Journal of Personality and Social Psychology, 96*(1), 11-31. doi:10.1037/a0013460

Garnefski, N., Kraaij, V., & Spinhoven, P. (2002). *CERQ: Manual for the use of the Cognitive Emotion Regulation Questionnaire*. Leiderdorp, the Netherlands: DATEC.

Gerin, M. I., Fichtenholtz, H., Roy, A., Walsh, C. J., Krystal, J. H., Southwick, S., & Hampson, M. (2016). Real-Time fMRI Neurofeedback with War Veterans with Chronic PTSD: A Feasibility Study. *Frontiers in Psychiatry, 7*, 111. doi:10.3389/fpsyt.2016.00111

Gevensleben, H., Moll, G. H., Rothenberger, A., & Heinrich, H. (2014). Neurofeedback in attention-deficit/hyperactivity disorder - different models, different ways of application. *Frontiers in Human Neuroscience, 8*, 846. doi:10.3389/fnhum.2014.00846

Gevensleben, H., Rothenberger, A., Moll, G. H., & Heinrich, H. (2012). Neurofeedback in children with ADHD: validation and challenges. *Expert Review of Neurotherapeutics, 12*(4), 447-460. doi:10.1586/ern.12.22

Giuliani, N. R., McRae, K., & Gross, J. J. (2008). The up- and down-regulation of amusement: experiential, behavioral, and autonomic consequences. *Emotion, 8*(5), 714-719. doi:10.1037/a0013236

Glenn, C. R., & Klonsky, E. D. (2009). Emotion dysregulation as a core feature of borderline personality disorder. *Journal of Personality Disorders, 23*(1), 20-28. doi:10.1521/pedi.2009.23.1.20.

Goldin, P. R., Moodie, C. A., & Gross, J. J. (2019). Acceptance versus reappraisal: Behavioral, autonomic, and neural effects. *Cognitive Affective and Behavioral Neuroscience, 19*, 927-944. doi:10.3758/s13415-019-00690-7

Golkar, A., Johansson, E., Kasahara, M., Osika, W., Perski, A., & Savic, I. (2014). The influence of work-related chronic stress on the regulation of emotion and on functional connectivity in the brain. *PLoS ONE, 9*(9), e104550. doi:10.1371/journal.pone.0104550

Gomez, P., Scholz, U., & Danuser, B. (2015). The Down-Regulation of Disgust by Implementation Intentions: Experiential and Physiological Concomitants. *Applied Psychophysiology and Biofeedback, 40*(2), 95-106. doi:10.1007/s10484-015-9280-2

Goodman, M., Carpenter, D., Tang, C. Y., Goldstein, K. E., Avedon, J., Fernandez, N., . . . Triebwasser, J. (2014). Dialectical behavior therapy alters emotion regulation and amygdala activity in patients with borderline personality disorder. *Journal of Psychiatyr Research, 57*, 108-116. doi:10.1016/j.jpsychires.2014.06.020

Goodman, M., Tomas, I. A., Temes, C. M., Fitzmaurice, G. M., Aguirre, B. A., & Zanarini, M. C. (2017). Suicide attempts and self‐injurious behaviours in adolescent and adult patients with borderline personality disorder. *Personality and Mental Health, 11*(3), 157-163. doi:10.1002/pmh.1375

Grant, B. F., Chou, S. P., Goldstein, R. B., Huang, B., Stinson, F. S., Saha, T. D., . . . Pickering, R. P. (2008). Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of Clinical Psychiatry, 69*(4), 533-545. doi:doi:10.4088/jcp.v69n0404

Grant, M., Salsman, N. L., & Berking, M. (2018). The assessment of successful emotion regulation skills use: Development and validation of an English version of the Emotion Regulation Skills Questionnaire. *PLoS ONE, 13*(10), e0205095. doi:doi.org/10.1371/journal.pone.0205095

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41-54. doi:10.1023/B:JOBA.0000007455.08539.94

Gratz, K. L., Rosenthal, M. Z., Tull, M. T., Lejuez, C., & Gunderson, J. G. (2006). An experimental investigation of emotion dysregulation in borderline personality disorder. *Journal of Abnormal Psychology, 115*(4), 850-855. doi:10.1037/0021-843X.115.4.850.

Greenwald, M. K., Cook, E. W., & Lang, P. J. (1989). Affective judgment and psychophysiological response: dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology, 3*(1), 51-64.

Grillon, C., & Baas, J. (2003). A review of the modulation of the startle reflex by affective states and its application in psychiatry. *Clinical Neurophysiology, 114*(9), 1557-1579. doi:10.1016/s1388-2457(03)00202-5

Grillon, C., Quispe-Escudero, D., Mathur, A., & Ernst, M. (2015). Mental fatigue impairs emotion regulation. *Emotion, 15*(3), 383-389. doi:10.1037/emo0000058

Gross, J. J. (1998a). Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74*(1), 224-237. doi:10.1037/0022-3514.74.1.224

Gross, J. J. (1998b). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271-299. doi:10.1037/1089-2680.2.3.271

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*(3), 281-291. doi:10.1017.S0048577201393198

Gross, J. J. (2014). Emotion Regulation: Conceptual and Empirical Foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (2nd ed., pp. 3-20). New York: Guilford Press.

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*(1), 1-26. doi:10.1080/1047840X.2014.940781

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348-362. doi:10.1037/0022-3514.85.2.348

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology, 64*(6), 970-986. doi:10.1037//0022-3514.64.6.970

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion, 9*(1), 87-108. doi:10.1080/02699939508408966

Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology, 106*(1), 95-103. doi:10.1037/0021-843x.106.1.95

Gross, J. J., & Thompson, R. A. (2007). Emotion Regulation: Conceptual Foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation.* (pp. 3-24). New York, NY, US: Guilford Press.

Gruber, J., Hay, A. C., & Gross, J. J. (2014). Rethinking emotion: Cognitive reappraisal is an effective positive and negative emotion regulation strategy in bipolar disorder. *Emotion, 14*(2), 388-396. doi:10.1037/a0035249

Gunderson, J. G. (2009). Borderline personality disorder: ontogeny of a diagnosis. *American Journal of Psychiatry, 166*(5), 530-539. doi:10.1176/appi.ajp.2009.08121825

Gunderson, J. G., Stout, R. L., McGlashan, T. H., Shea, M. T., Morey, L. C., Grilo, C. M., . . . Skodol, A. E. (2011). Ten-year course of borderline personality disorder: psychopathology and function from the Collaborative Longitudinal Personality Disorders study. *Archives of General Psychiatry, 68*(8), 827-837. doi:10.1001/archgenpsychiatry.2011.37

Gyurak, A., & Etkin, A. (2014). A Neurobiological Model of Implicit and Explicit Emotion Regulation. In J. J. Gross (Ed.), *Handbook of emotion regulation* (2nd ed.). New York: The Guilford Press.

Gyurak, A., Gross, J. J., & Etkin, A. (2011). Explicit and implicit emotion regulation: a dual-process framework. *Cognition & Emotion, 25*(3), 400-412.

Hagemann, T., Levenson, R. W., & Gross, J. J. (2006). Expressive suppression during an acoustic startle. *Psychophysiology, 43*(1), 104-112. doi:10.1111/j.1469-8986.2006.00382.x

Hallam, G. P., Webb, T. L., Sheeran, P., Miles, E., Wilkinson, I. D., Hunter, M. D., . . . Farrow, T. F. D. (2015). The Neural Correlates of Emotion Regulation by Implementation Intentions. *PLoS ONE, 10*(3), e0119500. doi:10.1371/journal.pone.0119500

Haller, S., Birbaumer, N., & Veit, R. (2010). Real-time fMRI feedback training may improve chronic tinnitus. *European Radiology, 20*(3), 696-703. doi:10.1007/s00330-009-1595-z

Hamilton, J. P., Glover, G. H., Bagarinao, E., Chang, C., Mackey, S., Sacchet, M. D., & Gotlib, I. H. (2016). Effects of salience-network-node neurofeedback training on affective biases in major depressive disorder. *Psychiatry Research: Neuroimaging, 249*, 91-96. doi: 10.1016/j.pscychresns.2016.01.016

Hamilton, J. P., Glover, G. H., Hsu, J. J., Johnson, R. F., & Gotlib, I. H. (2011). Modulation of subgenual anterior cingulate cortex activity with real-time neurofeedback. *Human Brain Mapping, 32*(1), 22-31. doi:10.1002/hbm.20997

Hamm, A. O., Cuthbert, B. N., Globisch, J., & Vaitl, D. (1997). Fear and the startle reflex: blink modulation and autonomic response patterns in animal and mutilation fearful subjects. *Psychophysiology, 34*(1), 97-107. doi:10.1111/j.1469-8986.1997.tb02420.x

Harvey, P. D., Greenberg, B. R., & Serper, M. R. (1989). The affective lability scales: development, reliability, and validity. *Journal of Clinical Psychology, 45*(5), 786-793. doi:10.1002/1097-4679(198909)45:5<786::AID-JCLP2270450515>3.0.CO;2-P

Hawk, L. W., & Cook, E. W. (2000). Independence of valence modulation and prepulse inhibition of startle. *Psychophysiology, 37*(1), 5-12. doi:10.1111/1469-8986.3710005

Hazlett, E. A., Speiser, L. J., Goodman, M., Roy, M., Carrizal, M., Wynn, J. K., . . . New, A. S. (2007). Exaggerated affect-modulated startle during unpleasant stimuli in borderline personality disorder. *Biological Psychiatry, 62*(3), 250-255. doi:10.1016/j.biopsych.2006.10.028

Hazlett, E. A., Zhang, J., New, A. S., Zelmanova, Y., Goldstein, K. E., Haznedar, M. M., . . . Chu, K. W. (2012). Potentiated amygdala response to repeated emotional pictures in borderline personality disorder. *Biological Psychiatry, 72*(6), 448-456. doi:10.1016/j.biopsych.2012.03.027

Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *journal of Educational Statistics, 7*(2), 119-137. doi:10.3102/10769986007002119

Hellrung, L., Dietrich, A., Hollmann, M., Pleger, B., Kalberlah, C., Roggenhofer, E., . . . Horstmann, A. (2018). Intermittent compared to continuous real-time fMRI neurofeedback boosts control over amygdala activation. *NeuroImage, 166*, 198-208. doi:10.1016/j.neuroimage.2017.10.031

Herpertz, S. C., Dietrich, T. M., Wenning, B., Krings, T., Erberich, S. G., Willmes, K., . . . Sass, H. (2001). Evidence of abnormal amygdala functioning in borderline personality disorder: a functional MRI study. *Biological Psychiatry, 50*(4), 292-298. doi:10.1016/S0006-3223(01)01075-7

Herpertz, S. C., & Koetting, K. (2005). Startle response in inpatients with borderline personality disorder vs. healthy controls. *Journal of Neural Transmission 112*(8), 1097-1106. doi:10.1007/s00702-004-0249-1

Herpertz, S. C., Kunert, H. J., Schwenger, U. B., & Sass, H. (1999). Affective responsiveness in borderline personality disorder: a psychophysiological approach. *American Journal of Psychiatry, 156*(10), 1550-1556. doi:10.1176/ajp.156.10.1550

Herpertz, S. C., Schneider, I., Schmahl, C., & Bertsch, K. (2018). Neurobiological mechanisms mediating emotion dysregulation as targets of change in borderline personality disorder. *Psychopathology, 51*(2), 96-104. doi:10.1159/000488357

Herwig, U., Lutz, J., Scherpiet, S., Scheerer, H., Kohlberg, J., Opialla, S., . . . Weidt, S. (2019). Training emotion regulation through real-time fMRI neurofeedback of amygdala activity. *NeuroImage, 184*, 687-696. doi:10.1016/j.neuroimage.2018.09.068.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta‐analysis. *Statistics in Medicine, 21*(11), 1539-1558. doi:10.1002/sim.1186

Hofmann, S. G., & Asmundson, G. J. (2008). Acceptance and mindfulness-based therapy: New wave or old hat? *Clinical Psychology Review, 28*(1), 1-16. doi:10.1016/j.cpr.2007.09.003

Hofmann, S. G., Heering, S., Sawyer, A. T., & Asnaani, A. (2009). How to handle anxiety: The effects of reappraisal, acceptance, and suppression strategies on anxious arousal. *Behaviour Research and Therapy, 47*(5), 389-394. doi:10.1016/j.brat.2009.02.010

Jackson, D. C., Malmstadt, J. R., Larson, C. L., & Davidson, R. J. (2000). Suppression and enhancement of emotional responses to unpleasant pictures. *Psychophysiology, 37*(4), 515-522. doi:10.1017/S0048577200990401

Jamieson, J. P., Mendes, W. B., & Nock, M. K. (2013). Improving acute stress responses: The power of reappraisal. *Current Directions in Psychological Science, 22*(1), 51-56. doi:10.1177/0963721412461500

Jamieson, J. P., Nock, M. K., & Mendes, W. B. (2012). Mind Over Matter: Reappraising Arousal Improves Cardiovascular and Cognitive Responses to Stress. *Journal of Experimental Psychology-General, 141*(3), 417-422. doi:10.1037/a0025719

Joormann, J., & Gotlib, I. H. (2010). Emotion regulation in depression: relation to cognitive inhibition. *Cognition & Emotion, 24*(2), 281-298. doi:10.1080/02699930903407948

Joormann, J., & Vanderlind, W. M. (2014). Emotion Regulation in Depression The Role of Biased Cognition and Reduced Cognitive Control. *Clinical Psychological Science, 2*(4), 402-421. doi:10.1177/2167702614536163

Kalisch, R. (2009). The functional neuroanatomy of reappraisal: Time matters. *Neuroscience and Biobehavioral Reviews, 33*(8), 1215-1226. doi:10.1016/j.neubiorev.2009.06.003

Kalisch, R., Wiech, K., Critchley, H. D., Seymour, B., O'doherty, J. P., Oakley, D. A., . . . Dolan, R. J. (2005). Anxiety reduction through detachment: subjective, physiological, and neural effects. *Journal of Cognitive Neuroscience, 17*(6), 874-883. doi:10.1162/0898929054021184

Kalokerinos, E. K., Greenaway, K. H., & Denson, T. F. (2015). Reappraisal but not suppression downregulates the experience of positive and negative emotion. *Emotion, 15*(3), 271-275. doi:10.1037/emo0000025

Keltner, D., Haidt, J., & Shiota, M. N. (2006). Social Functionalism and the Evolution of Emotions. In M. Schaller, D. Kenrick, & J. Simpson (Eds.), *Evolution and Social Psychology* (pp. 115-142). New York: Psychology Press.

Keynan, J. N., Cohen, A., Jackont, G., Green, N., Goldway, N., Davidov, A., . . . Hendler, T. (2019). Electrical fingerprint of the amygdala guides neurofeedback training for stress resilience. *Nature Human Behavior, 3*(1), 63-73. doi:10.1038/s41562-018-0484-3

Keynan, J. N., Meir-Hasson, Y., Gilam, G., Cohen, A., Jackont, G., Kinreich, S., . . . Gyurak, A. (2016). Limbic Activity Modulation Guided by Functional Magnetic Resonance Imaging–Inspired Electroencephalography Improves Implicit Emotion Regulation. *Biological Psychiatry, 80*(6), 490-496. doi:0.1016/j.biopsych.2015.12.024

Kim, S. H., & Hamann, S. (2012). The effect of cognitive reappraisal on physiological reactivity and emotional memory. *International Journal of Psychophysiology, 83*(3), 348-356. doi:0.1016/j.ijpsycho.2011.12.001

Kinner, V. L., Kuchinke, L., Dierolf, A. M., Merz, C. J., Otto, T., & Wolf, O. T. (2017). What our eyes tell us about feelings: Tracking pupillary responses during emotion regulation processes. *Psychophysiology, 54*(4), 508-518. doi:10.1111/psyp.12816

Kinzig, C., Schulz, A., Curio, C., & Schächinger, H. (2008). Clip-Software for the analysis of psychophysiological data. Trier: Univeristy of Trier.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1-2), 76-81. doi:10.1159/000119004

Koenig, J., Kemp, A. H., Feeling, N. R., Thayer, J. F., & Kaess, M. (2016). Resting state vagal tone in borderline personality disorder: A meta-analysis. *Progress in Neuropsychopharmacology and Biological Psychiatry, 64*, 18-26. doi:10.1016/j.pnpbp.2015.07.002

Koenigsberg, H. W. (2010). Affective instability: toward an integration of neuroscience and psychological perspectives. *Journal of Personality Disorders, 24*(1), 60-82. doi:10.1521/pedi.2010.24.1.60.

Koenigsberg, H. W., Fan, J., Ochsner, K. N., Liu, X., Guise, K. G., Pizzarello, S., . . . Goodman, M. (2009). Neural correlates of the use of psychological distancing to regulate responses to negative social cues: a study of patients with borderline personality disorder. *Biological Psychiatry, 66*(9), 854-863. doi:10.1016/j.biopsych.2009.06.010

Koenigsberg, H. W., Harvey, P. D., Mitropoulou, V., Schmeidler, J., New, A. S., Goodman, M., . . . Siever, L. J. (2002). Characterizing affective instability in borderline personality

disorder. *American Journal of Psychiatry, 159*(5), 784-788. doi:10.1176/appi.ajp.159.5.784

Koenigsberg, H. W., Siever, L. J., Lee, H., Pizzarello, S., New, A. S., Goodman, M., . . . Prohovnik, I. (2009). Neural correlates of emotion processing in borderline personality disorder. *Psychiatry Research: Neuroimaging, 172*(3), 192-199. doi:10.1016/j.pscychresns.2008.07.010

Kohl, S. H., Veit, R., Spetter, M. S., Gunther, A., Rina, A., Luhrs, M., . . . Hallschmid, M. (2019). Real-time fMRI neurofeedback training to improve eating behavior by self-regulation of the dorsolateral prefrontal cortex: A randomized controlled trial in overweight and obese subjects. *NeuroImage, 191*, 596-609. doi:10.1016/j.neuroimage.2019.02.033

Kohn, N., Eickhoff, S., Scheller, M., Laird, A., Fox, P., & Habel, U. (2014). Neural network of cognitive emotion regulation—An ALE meta-analysis and MACM analysis. *NeuroImage, 87*, 345-355. doi:10.1016/j.neuroimage.2013.11.001

Koole, S. L. (2009). The psychology of emotion regulation: An integrative review. *Cognition & Emotion, 23*(1), 4-41. doi:10.1080/02699930802619031

Kotchoubey, B., Kubler, A., Strehl, U., Flor, H., & Birbaumer, N. (2002). Can humans perceive their brain states? *Consciousness and Cognition, 11*(1), 98-113. doi:10.1006/ccog.2001.0535

Kotwas, I., Micoulaud-Franchi, J. A., Khalfa, S., McGonigal, A., Bastien-Toniazzo, M., & Bartolomei, F. (2019). Subjective and physiological response to emotions in temporal lobe epilepsy and psychogenic non-epileptic seizures. *Journal of Affective Disorders, 244*, 46-53. doi:10.1016/j.jad.2018.10.004

Koush, Y., Rosa, M. J., Robineau, F., Heinen, K., Rieger, S. W., Weiskopf, N., . . . Scharnowski, F. (2013). Connectivity-based neurofeedback: Dynamic causal modeling for real-time fMRI. *NeuroImage, 81*, 422-430. doi:10.1016/j.neuroimage.2013.05.010

Koush, Y., Zvyagintsev, M., Dyck, M., Mathiak, K. A., & Mathiak, K. (2012). Signal quality and Bayesian signal processing in neurofeedback based on real-time fMRI. *NeuroImage, 59*(1), 478-489. doi:10.1016/j.neuroimage.2011.07.076

Krause-Utz, A., Elzinga, B. M., Oei, N. Y., Paret, C., Niedtfeld, I., Spinhoven, P., . . . Schmahl, C. (2014). Amygdala and dorsal anterior cingulate connectivity during an emotional working memory task in borderline personality disorder patients with interpersonal trauma history. *Frontiers in Human Neuroscience, 8*, 848. doi:10.3389/fnhum.2014.00848

Krause-Utz, A., Oei, N., Niedtfeld, I., Bohus, M., Spinhoven, P., Schmahl, C., & Elzinga, B. (2012). Influence of emotional distraction on working memory performance in borderline personality disorder. *Psychological Medicine, 42*(10), 2181-2192. doi:10.1017/S0033291712000153

Krause-Utz, A., Winter, D., Niedtfeld, I., & Schmahl, C. (2014). The latest neuroimaging findings in borderline personality disorder. *Current Psychiatry Reports, 16*(3), 438. doi:10.1007/s11920-014-0438-z

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84*(3), 394-421. doi:10.1016/j.biopsycho.2010.03.010

Kuhn, M., Wendt, J., Sjouwerman, R., Büchel, C., Hamm, A., & Lonsdorf, T. B. (2020). The neurofunctional basis of affective startle modulation in humans: evidence from combined facial electromyography and functional magnetic resonance imaging. *Biological Psychiatry, 87*(6), 548-558. doi:10.1016/j.biopsych.2019.07.028

Kunzmann, U., Kupperbusch, C. S., & Levenson, R. W. (2005). Behavioral inhibition and amplification during emotional arousal: a comparison of two age groups. *Psychology of Aging, 20*(1), 144-158. doi:10.1037/0882-7974.20.1.144

Kuo, J. R., Fitzpatrick, S., Metcalfe, R. K., & McMain, S. (2015). A multi-method laboratory investigation of emotional reactivity and emotion regulation abilities in borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 50*, 52-60. doi:10.1016/j.jbtep.2015.05.002

Kuo, J. R., & Linehan, M. A. (2009). Disentangling Emotion Processes in Borderline Personality Disorder: Physiological and Self-Reported Assessment of Biological Vulnerability, Baseline Intensity, and Reactivity to Emotionally Evocative Stimuli. *Journal of Abnormal Psychology, 118*(3), 531-544. doi:10.1037/a0016392

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, 50*(5), 372-385. doi:10.1037/0003-066X.50.5.372

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97*(3), 377-395. doi: 0.1037/0033-295x.97.3.377

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report. A-8*. Gainsville, FL: University of Florida.

Lang, P. J., Simons, R. F., & Balaban, M. T. (1997). *Attention and orienting: Sensory and motivational processes.* Mahwah, NJ: Psychology Press.

Lang, S., Kotchoubey, B., Frick, C., Spitzer, C., Grabe, H. J., & Barnow, S. (2012). Cognitive reappraisal in trauma-exposed women with borderline personality disorder. *NeuroImage, 59*(2), 1727-1734. doi:10.1016/j.neuroimage.2011.08.061

Lapate, R. C., Rokers, B., Li, T., & Davidson, R. J. (2014). Nonconscious emotional activation colors first impressions: A regulatory role for conscious awareness. *Psychological Science, 25*(2), 349-357. doi:10.1177/0956797613503175

Larsen, R. J. (2000). Toward a science of mood regulation. *Psychological Inquiry, 11*(3), 129-141. doi:10.1207/S15327965PLI1103_01

LeDoux, J. (2007). The amygdala. *Current Biology, 17*(20), R868-R874. doi:10.1016/j.cub.2007.08.005

LeDoux, J., Iwata, J., Cicchetti, P., & Reis, D. J. (1988). Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *Journal of Neuroscience, 8*(7), 2517-2529. doi:10.1523/JNEUROSCI.08-07-02517.1988

LeDoux, J., & Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: a two-system framework. *American Journal of Psychiatry, 173*(11), 1083-1093. doi:10.1176/appi.ajp.2016.16030353

Lee, H., Shackman, A. J., Jackson, D. C., & Davidson, R. J. (2009). Test-retest reliability of voluntary emotion regulation. *Psychophysiology, 46*(4), 874-879. doi:10.1111/j.1469-8986.2009.00830.x

Leiberg, S., Eippert, F., Veit, R., & Anders, S. (2012). Intentional social distance regulation alters affective responses towards victims of violence: an FMRI study. *Human Brain Mapping, 33*(10), 2464-2476. doi:10.1002/hbm.21376

Leichsenring, F., Leibing, E., Kruse, J., New, A. S., & Leweke, F. (2011). Borderline personality disorder. *The Lancet, 377*(9759), 74-84. doi:10.1016/S0140-6736(10)61422-5

Lenzenweger, M. F., Lane, M. C., Loranger, A. W., & Kessler, R. C. (2007). DSM-IV personality disorders in the National Comorbidity Survey Replication. *Biological Psychiatry, 62*(6), 553-564. doi:10.1016/j.biopsych.2006.09.019

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology, 27*(4), 363-384. doi:10.1111/j.1469-8986.1990.tb02330.x

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: physiological linkage and affective exchange. *Journal of Personality and Social Psychology, 45*(3), 587-597. doi:10.1037/0022-3514.45.3.587

Lieb, K., Zanarini, M. C., Schmahl, C., Linehan, M. M., & Bohus, M. (2004). Borderline personality disorder. *The Lancet, 364*(9432), 453-461. doi:10.1016/S0140-6736(04)16770-6

Limberg, A., Barnow, S., Freyberger, H. J., & Hamm, A. O. (2011). Emotional vulnerability in borderline personality disorder is cue specific and modulated by traumatization. *Biological Psychiatry, 69*(6), 574-582. doi:10.1016/j.biopsych.2010.10.024

Linden, D. E. (2006). How psychotherapy changes the brain- the contribution of functional neuroimaging. *Molecular Psychiatry, 11*(6), 528-538. doi:10.1038/sj.mp.4001816

Linden, D. E., Habes, I., Johnston, S. J., Linden, S., Tatineni, R., Subramanian, L., . . . Goebel, R. (2012). Real-time self-regulation of emotion networks in patients with depression. *PLoS ONE, 7*(6), e38115. doi:10.1371/journal.pone.0038115

Lindquist, M. A., Loh, J. M., Atlas, L. Y., & Wager, T. D. (2009). Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. *NeuroImage, 45*(1), S187-S198. doi:10.1016/j.neuroimage.2008.10.065

Linehan, M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford Press.

Linhartová, P., Látalová, A., Kóša, B., Kašpárek, T., Schmahl, C., & Paret, C. (2019). fMRI neurofeedback in emotion regulation: A literature review. *NeuroImage, 193*, 75-92. doi:10.1016/j.neuroimage.2019.03.011

Lissek, S., Orme, K., McDowell, D. J., Johnson, L. L., Luckenbaugh, D. A., Baas, J. M., . . . Grillon, C. (2007). Emotion regulation and potentiated startle across affective picture and threat-of-shock paradigms. *Biological Psychology, 76*(1-2), 124-133. doi:10.1016/j.biopsycho.2007.07.002

Lohani, M., & Isaacowitz, D. M. (2014). Age differences in managing response to sadness elicitors using attentional deployment, positive reappraisal and suppression. *Cognition & Emotion, 28*(4), 678-697 doi:10.1080/02699931.2013.853648

Loranger, A. W. (1999). *International Personality Disorder Examination (IPDE): DSM-IV and ICD-10 modules*. Odessa, FL: Psychological Assessment Resources.

Low, C. A., Stanton, A. L., & Bower, J. E. (2008). Effects of acceptance-oriented versus evaluative emotional processing on heart rate recovery and habituation. *Emotion, 8*(3), 419-424. doi:10.1037/1528-3542.8.3.419

Lubianiker, N., Goldway, N., Fruchtman-Steinbok, T., Paret, C., Keynan, J. N., Singer, N., . . . Hendler, T. (2019). Process-based framework for precise neuromodulation. *Nature Human Behavior, 3*(5), 436-445. doi:10.1038/s41562-019-0573-y

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces (KDEF)*. CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17*(3), 354-381. doi:10.1111/j.1542-474X.1996.tb00275.x

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods, 46*(2), 596-610. doi:10.3758/s13428-013-0379-1

Martins, B., Florjanczyk, J., Jackson, N. J., Gatz, M., & Mather, M. (2018). Age differences in emotion regulation effort: Pupil response distinguishes reappraisal and distraction for older but not younger adults. *Psychology and Aging, 33*(2), 338-349. doi:10.1037/pag0000227

Mauss, I. B., Cook, C. L., Cheng, J. Y., & Gross, J. J. (2007). Individual differences in cognitive reappraisal: Experiential and physiological responses to an anger provocation. *International Journal of Psychophysiology, 66*(2), 116-124. doi:10.1016/j.ijpsycho.2007.03.017

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion, 23*(2), 209-237. doi:10.1080/02699930802204677

McRae, K. (2016). Cognitive emotion regulation: A review of theory and scientific findings. *Current Opinion in Behavioral Sciences, 10*, 119-124. doi: 10.1016/j.cobeha.2016.06.004

McRae, K., & Gross, J. J. (2020). Emotion regulation. *Emotion, 20*(1), 1-9. doi:10.1037/emo0000703

McRae, K., Ochsner, K. N., Mauss, I. B., Gabrieli, J. J. D., & Gross, J. J. (2008). Gender Differences in Emotion Regulation: An fMRI Study of Cognitive Reappraisal. *Group Process Intergroup Relations, 11*(2), 143-162. doi:10.1177/1368430207088035

McRae, K., & Shiota, M. N. (2017). Biological and physiological aspects of emotion regulation. . In C. A. Essau, S. Leblanc, & T. H. Ollendick (Eds.), *Emotion regulation and*

psychopathology in children and adolescents (pp. 43-59). New York, NY: Oxford University Press.

Mehler, D. M. A., Sokunbi, M. O., Habes, I., Barawi, K., Subramanian, L., Range, M., . . . Linden, D. E. J. (2018). Targeting the affective brain-a randomized controlled trial of real-time fMRI neurofeedback in patients with depression. *Neuropsychopharmacology, 43*(13), 2578-2585. doi:10.1038/s41386-018-0126-5

Memedovic, S., Grisham, J. R., Denson, T. F., & Moulds, M. L. (2010). The effects of trait reappraisal and suppression on anger and blood pressure in response to provocation. *Journal of Research in Personality, 44*(4), 540-543. doi:10.1016/j.jrp.2010.05.002

Minzenberg, M. J., Fan, J., New, A. S., Tang, C. Y., & Siever, L. J. (2007). Fronto-limbic dysfunction in response to facial emotion in borderline personality disorder: an event-related fMRI study. *Psychiatry Research: Neuroimaging, 155*(3), 231-243. doi:10.1016/j.pscychresns.2007.03.006

Misaki, M., Phillips, R., Zotev, V., Wong, C. K., Wurfel, B. E., Krueger, F., . . . Bodurka, J. (2019). Brain activity mediators of PTSD symptom reduction during real-time fMRI amygdala neurofeedback emotional training. *Neuroimage Clinical, 24*, 102047. doi:10.1016/j.nicl.2019.102047

Moore, S. A., Zoellner, L. A., & Mollenholt, N. (2008). Are expressive suppression and cognitive reappraisal associated with stress-related symptoms? *Behaviour Research and Therapy, 46*(9), 993-1000 doi:10.1016/j.brat.2008.05.001

Morawetz, C., Bode, S., Baudewig, J., Jacobs, A. M., & Heekeren, H. R. (2016a). Neural representation of emotion regulation goals. *Human Brain Mapping, 37*(2), 600-620. doi:10.1002/hbm.23053

Morawetz, C., Bode, S., Baudewig, J., Kirilina, E., & Heekeren, H. R. (2016c). Changes in Effective Connectivity Between Dorsal and Ventral Prefrontal Regions Moderate Emotion Regulation. *Cerebral Cortex, 26*(5), 1923-1937. doi:10.1093/cercor/bhv005

Morawetz, C., Bode, S., Derntl, B., & Heekeren, H. R. (2017). The effect of strategies, goals and stimulus material on the neural mechanisms of emotion regulation: A meta-analysis of fMRI studies. *Neuroscience and Biobehavioral Reviews, 72*, 111-128. doi:10.1016/j.neubiorev.2016.11.014

Morawetz, C., Bode, S., Baudewick, & Heekeren, H. (2016b). Effective amygdala-prefrontal connectivity predicts individual differences in successful emotion regulation. *Social Cognitive and Affective Neuroscience, 12*(4), 569-585. doi: 10.1093/scan/nsw169

Murphy, J. W., & Young, M. A. (2020). Monitoring processes in extended emotion regulation. *Cognition and Emotion*, 1-9. doi:10.1080/02699931.2020.1722614

Nica, E. I., & Links, P. S. (2009). Affective instability in borderline personality disorder: Experience sampling findings. *Current Psychiatry Reports, 11*(1), 74-81. doi: 10.1007/s11920-009-0012-2

Nicholson, A. A., Rabellino, D., Densmore, M., Frewen, P. A., Paret, C., Kluetsch, R., . . . McKinnon, M. C. (2017). The neurobiology of emotion regulation in posttraumatic stress disorder: Amygdala downregulation via real‐time fMRI neurofeedback. *Human Brain Mapping, 38*(1), 541-560. doi:10.1002/hbm.23402

Niedtfeld, I., Schulze, L., Kirsch, P., Herpertz, S. C., Bohus, M., & Schmahl, C. (2010). Affect regulation and pain in borderline personality disorder: a possible link to the understanding of self-injury. *Biological Psychiatry, 68*(4), 383-391. doi:10.1016/j.biopsych.2010.04.015

Niedtfeld, I., Schulze, L., Krause-Utz, A., Demirakca, T., Bohus, M., & Schmahl, C. (2013). Voxel-based morphometry in women with borderline personality disorder with and without comorbid posttraumatic stress disorder. *PLoS ONE, 8*(6), e65824. doi:10.1371/journal.pone.0065824

Nunes, P. M., Wenzel, A., Borges, K. T., Porto, C. R., Caminha, R. M., & de Oliveira, I. R. (2009). Volumes of the hippocampus and amygdala in patients with borderline personality disorder: a meta-analysis. *Journal of Personality Disorder, 23*(4), 333-345. doi:10.1521/pedi.2009.23.4.333

Obrist, P. A. (1981). *Cardiovascular psychophysiology*. New York: Springer Science & Business Media.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences, 9*, 242-249. doi:10.1016/j.tics.2005.03.010

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage, 23*, 483-499. doi:10.1016/j.neuroimage.2004.06.030

Oei, N. Y., Veer, I. M., Wolf, O. T., Spinhoven, P., Rombouts, S. A., & Elzinga, B. M. (2012). Stress shifts brain activation towards ventral 'affective'areas during emotional distraction. *Social Cognitive and Affective Neuroscience, 7*(4), 403-412. doi:10.1093/scan/nsr024

Ohira, H., Nomura, M., Ichikawa, N., Isowa, T., Iidaka, T., Sato, A., . . . Yamada, J. (2006). Association of neural and physiological responses during voluntary emotion suppression. *NeuroImage, 29*(3), 721-733. doi:10.1016/j.neuroimage.2005.08.047

Öhman, A., & Soares, J. J. (1994). "Unconscious anxiety": phobic responses to masked stimuli. *Journal of Abnormal Psychology, 103*(2), 231-240. doi:10.1037/0021-843X.103.2.231

Oldham, J. M. (2006). Borderline personality disorder and suicidality. *American Journal of Psychiatry, 163*(1), 20-26. doi:10.1176/appi.ajp.163.1.20

Olsson, A., & Phelps, E. A. (2004). Learned fear of “unseen” faces after Pavlovian, observational, and instructed fear. *Psychological Science, 15*(12), 822-828. doi:10.1111/j.0956-7976.2004.00762.x

Opitz, P. C., Lee, I. A., Gross, J. J., & Urry, H. L. (2014). Fluid cognitive ability is a resource for successful emotion regulation in older and younger adults. *Frontiers in Psychology, 5*, 609. doi:10.3389/fpsyg.2014.00609

Ortner, C. N. M. (2015). Divergent effects of reappraisal and labeling internal affective feelings on subjective emotional experience. *Motivation and Emotion, 39*(4), 563-570. doi:10.1007/s11031-015-9473-2

Paret, C., Brenninkmeyer, J., Meyer, B., Yuen, K. S., Gartmann, N., Mechias, M.-L., & Kalisch, R. (2011). A test for the implementation–maintenance model of reappraisal. *Frontiers in Psychology, 2*, 216. doi:10.3389/fpsyg.2011.00216

Paret, C., Goldway, N., Zich, C., Keynan, J. N., Hendler, T., Linden, D., & Kadosh, K. C. (2019). Current progress in real-time functional magnetic resonance-based neurofeedback: Methodological challenges and achievements. *NeuroImage*, 116-107. doi:10.1016/j.neuroimage.2019.116107

Paret, C., & Hendler, T. (2020). Live from the "regulating brain": Harnessing the brain to change emotion. *Emotion, 20*(1), 126-131. doi:10.1037/emo0000674

Paret, C., Kluetsch, R., Ruf, M., Demirakca, T., Hoesterey, S., Ende, G., & Schmahl, C. (2014). Down-regulation of amygdala activation with real-time fMRI neurofeedback in a healthy female sample. *Frontiers in Behavioral Neuroscience, 8*, 299. doi:10.3389/fnbeh.2014.00299

Paret, C., Kluetsch, R., Zaehringer, J., Ruf, M., Demirakca, T., Bohus, M., . . . Schmahl, C. (2016). Alterations of amygdala-prefrontal connectivity with real-time fMRI neurofeedback in BPD patients. *Social Cognitive and Affective Neuroscience, 11*(6), 952-960. doi:10.1093/scan/nsw016

Paret, C., Zaehringer, J., Ruf, M., Gerchen, M. F., Mall, S., Hendler, T., . . . Ende, G. (2018). Monitoring and control of amygdala neurofeedback involves distributed information processing in the human brain. *Human Brain Mapping, 39*(7), 3018-3031. doi:10.1002/hbm.24057

Parkinson, B., & Totterdell, P. (1999). Classifying affect-regulation strategies. *Cognition & Emotion, 13*(3), 277-303. doi:10.1080/026999399379285

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nature Review Neuroscience, 11*(11), 773-783. doi:10.1038/nrn2920

Peters, B. J., & Jamieson, J. P. (2016). The consequences of suppressing affective displays in romantic relationships: A challenge and threat perspective. *Emotion, 16*(7), 1050-1066. doi:10.1037/emo0000202

Peters, B. J., Overall, N. C., & Jamieson, J. P. (2014). Physiological and cognitive consequences of suppressing and expressing emotion in dyadic interactions. *International Journal of Psychophysiology, 94*(1), 100-107. doi:10.1016/j.ijpsycho.2014.07.015

Phan, K. L., Taylor, S. F., Welsh, R. C., Ho, S. H., Britton, J. C., & Liberzon, I. (2004). Neural correlates of individual ratings of emotional salience: a trial-related fMRI study. *NeuroImage, 21*(2), 768-780. doi:10.1016/j.neuroimage.2003.09.072

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron, 48*(2), 175-187. doi:10.1016/j.neuron.2005.09.025

Piper, M. E., & Curtin, J. J. (2006). Tobacco withdrawal and negative affect: An analysis of initial emotional response intensity and voluntary emotion regulation. *Journal of Abnormal Psychology, 115*(1), 96-102. doi:10.1037/0021-843X.115.1.96

Plieger, T., Melchers, M., Vetterlein, A., Görtz, J., Kuhn, S., Ruppel, M., & Reuter, M. (2017). The serotonin transporter polymorphism (5-HTTLPR) and coping strategies influence successful emotion regulation in an acute stress situation: Physiological evidence. *International Journal of Psychophysiology, 114*, 31-37. doi:10.1016/j.ijpsycho.2017.02.006

Putnam, K. M., & Silk, K. R. (2005). Emotion dysregulation and the development of borderline personality disorder. *Developmental Psychopathology, 17*(4), 899-925. doi:10.10170S0954579405050431

Raugh, I. M., Chapman, H. C., Bartolomeo, L. A., Gonzalez, C., & Strauss, G. P. (2019). A comprehensive review of psychophysiological applications for ecological momentary

assessment in psychiatric populations. *Psychological Assessment, 31*(3), 304-317. doi:10.1037/pas0000651

Reinecke, A., Filippini, N., Berna, C., Western, D. G., Hanson, B., Cooper, M. J., . . . Harmer, C. J. (2015). Effective emotion regulation strategies improve fMRI and ECG markers of psychopathology in panic disorder: implications for psychological treatment action. *Translational Psychiatry, 5*, e673. doi:10.1038/tp.2015.160

Reitz, S., Kluetsch, R., Niedtfeld, I., Knorz, T., Lis, S., Paret, C., . . . Schmahl, C. (2015). Incision and stress regulation in borderline personality disorder: neurobiological mechanisms of self-injurious behaviour. *British Journal of Psychiatry, 207*(2), 165-172. doi:10.1192/bjp.bp.114.153379

Reyes, B. A., Carvalho, A. F., Vakharia, K., & Van Bockstaele, E. J. (2011). Amygdalar peptidergic circuits regulating noradrenergic locus coeruleus neurons: linking limbic and arousal centers. *Experimental Neurology, 230*(1), 96-105. doi:10.1016/j.expneurol.2011.04.001

Richards, J. M., & Gross, J. J. (1999). Composure at any cost? The cognitive consequences of emotion suppression. *Personality and Social Psychology Bulletin, 25*(8), 1033-1044. doi:10.1177/01461672992511010

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology, 79*(3), 410-424. doi:10.1037/0022-3514.79.3.410

Roberts, N. A., Levenson, R. W., & Gross, J. J. (2008). Cardiovascular costs of emotion suppression cross ethnic lines. *International Journal of Psychophysiology, 70*(1), 82-87. doi:10.1016/j.ijpsycho.2008.06.003

Robinson, J. L., & Demaree, H. A. (2009). Experiencing and regulating sadness: Physiological and cognitive effects. *Brain and Cognition, 70*(1), 13-20. doi:10.1016/j.bandc.2008.06.007

Rohrmann, S., Hopp, H., Schienle, A., & Hodapp, V. (2009). Emotion regulation, disgust sensitivity, and psychophysiological responses to a disgust-inducing film. *Anxiety, Stress and Coping: An International Journal, 22*(2), 215-236. doi:10.1080/10615800802016591

Rosen, J. B., & Davis, M. (1988). Enhancement of acoustic startle by electrical stimulation of the amygdala. *Behavioral Neuroscience, 102*(2), 195-202. doi:10.1037/0735-7044.102.2.195

Rosen, J. B., Hitchcock, J. M., Sananes, C. B., Miserendino, M. J., & Davis, M. (1991). A direct projection from the central nucleus of the amygdala to the acoustic startle pathway: anterograde and retrograde tracing studies. *Behavioral Neuroscience, 105*(6), 817-825. doi:10.1037/0735-7044.105.6.817

Rosenthal, M. Z., Gratz, K. L., Kosson, D. S., Cheavens, J. S., Lejuez, C., & Lynch, T. R. (2008). Borderline personality disorder and emotional responding: A review of the research literature. *Clinical Psychology Review, 28*(1), 75-91. doi:10.1016/j.cpr.2007.04.001si

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin, 92*(2), 500-504. doi:10.1037/0033-2909.92.2.500

Roth, G., Benita, M., Amrani, C., Shachar, B. H., Asoulin, H., Moed, A., . . . Kanat-Maymon, Y. (2014). Integration of negative emotional experience versus suppression: addressing the question of adaptive functioning. *Emotion, 14*(5), 908-919. doi:10.1037/a0037051

Ruiz, S., Lee, S., Soekadar, S. R., Caria, A., Veit, R., Kircher, T., . . . Sitaram, R. (2013). Acquired self-control of insula cortex modulates emotion recognition and brain network connectivity in schizophrenia. *Human Brain Mapping, 34*(1), 200-212. doi:10.1002/hbm.21427

Ruocco, A. C., Amirthavasagam, S., Choi-Kain, L. W., & McMain, S. F. (2013). Neural correlates of negative emotionality in borderline personality disorder: an activation-likelihood-estimation meta-analysis. *Biological Psychiatry, 73*(2), 153-160. doi:10.1016/j.biopsych.2012.07.014

Ruocco, A. C., Amirthavasagam, S., & Zakzanis, K. K. (2012). Amygdala and hippocampal volume reductions as candidate endophenotypes for borderline personality disorder: a meta-analysis of magnetic resonance imaging studies. *Psychiatry Reseach, 201*(3), 245-252. doi:10.1016/j.pscychresns.2012.02.012

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178. doi: 10.1037/h0077714

Sabatinelli, D., Bradley, M. M., Fitzsimmons, J. R., & Lang, P. J. (2005). Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *NeuroImage, 24*(4), 1265-1270. doi:10.1016/j.neuroimage.2004.12.015

Salsman, N. L., & Linehan, M. M. (2012). An investigation of the relationships among negative affect, difficulties in emotion regulation, and features of borderline personality disorder. *Journal of Psychopathology and Behavioral Assessment, 34*(2), 260-267. doi:10.1007/s10862-012-9275-8

Sanislow, C. A., Grilo, C. M., Morey, L. C., Bender, D. S., Skodol, A. E., Gunderson, J. G., . . . McGlashan, T. H. (2002). Confirmatory factor analysis of DSM-IV criteria for borderline personality disorder: findings from the collaborative longitudinal personality disorders study. *American Journal of Psychiatry, 159*(2), 284-290. doi:10.1176/appi.ajp.159.2.284

Santangelo, P., Bohus, M., & Ebner-Priemer, U. W. (2014). Ecological momentary assessment in borderline personality disorder: a review of recent findings and methodological challenges. *Journal of Personality Disorders, 28*(4), 555-576. doi: 10.1521/pedi_2012_26_067

Santangelo, P., Koenig, J., Funke, V., Parzer, P., Resch, F., Ebner-Priemer, U. W., & Kaess, M. (2017). Ecological momentary assessment of affective and interpersonal instability in adolescent non-suicidal self-injury. *Journal of Abnormal Child Psychology, 45*(7), 1429-1438. doi:10.1007/s10802-016-0249-2.

Sauer, C., Sheppes, G., Lackner, H. K., Arens, E. A., Tarrasch, R., & Barnow, S. (2016). Emotion regulation choice in female patients with borderline personality disorder: Findings from self-reports and experimental measures. *Psychiatry Reseach, 242*, 375-384. doi:10.1016/j.psychres.2016.04.113

Schächinger, H., Weinbacher, M., Kiss, A., Ritz, R., & Langewitz, W. (2001). Cardiovascular indices of peripheral and central sympathetic activation. *Psychosomatic Medicine, 63*(5), 788-796. doi:10.1097/00006842-200109000-00012

Schaefer, S. M., Jackson, D. C., Davidson, R. J., Aguirre, G. K., Kimberg, D. Y., & Thompson-Schill, S. L. (2002). Modulation of amygdalar activity by the conscious regulation of negative emotion. *Journal of Cognitive Neuroscience, 14*(6), 913-921. doi:10.1162/089892902760191135

Scharnowski, F., & Weiskopf, N. (2015). Cognitive enhancement through real-time fMRI neurofeedback. *Current Opinion in Behavioral Sciences, 4*, 122-127. doi:10.1016/j.cobeha.2015.05.001

Schmahl, C., Elzinga, B. M., Ebner, U. W., Simms, T., Sanislow, C., Vermetten, E., . . . Bremner, J. D. (2004). Psychophysiological reactivity to traumatic and abandonment scripts in borderline personality and posttraumatic stress disorders: a preliminary report. *Psychiatry Reseach, 126*(1), 33-42. doi:10.1016/j.psychres.2004.01.005

Schmahl, C., Herpertz, S. C., Bertsch, K., Ende, G., Flor, H., Kirsch, P., . . . Schneider, M. (2014). Mechanisms of disturbed emotion processing and social interaction in borderline personality disorder: state of knowledge and research agenda of the German

Clinical Research Unit. *Borderline Personality Disorder and Emotion Dysregulation, 1*(1), 12. doi:10.1186/2051-6673-1-12

Schmahl, C., Vermetten, E., Elzinga, B. M., & Douglas Bremner, J. (2003). Magnetic resonance imaging of hippocampal and amygdala volume in women with childhood abuse and borderline personality disorder. *Psychiatry Reseach, 122*(3), 193-198. doi:10.1016/s0925-4927(03)00023-4

Schmitt, R., Winter, D., Niedtfeld, I., Herpertz, S. C., & Schmahl, C. (2016). Effects of psychotherapy on neuronal correlates of reappraisal in female patients with borderline personality disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(6), 548-557. doi:10.1016/j.bpsc.2016.07.003

Schnell, K., & Herpertz, S. C. (2007). Effects of dialectic-behavioral-therapy on the neural correlates of affective hyperarousal in borderline personality disorder. *Journal of Psychiatric Research, 41*(10), 837-847. doi:10.1016/j.jpsychires.2006.08.011

Schonfelder, S., Kanske, P., Heissler, J., & Wessa, M. (2014). Time course of emotion-related responding during distraction and reappraisal. *Social Cognitive and Affective Neuroscience, 9*(9), 1310-1319. doi:10.1093/scan/nst116

Schulze, L., Domes, G., Krüger, A., Berger, C., Fleischer, M., Prehn, K., . . . Herpertz, S. C. (2011). Neuronal correlates of cognitive reappraisal in borderline patients with affective instability. *Biological Psychiatry, 69*(6), 564-573. doi:10.1016/j.biopsych.2010.10.025

Schulze, L., Schmahl, C., & Niedtfeld, I. (2016). Neural correlates of disturbed emotion processing in borderline personality disorder: a multimodal meta-analysis. *Biological Psychiatry, 79*(2), 97-106. doi:10.1016/j.biopsych.2015.03.027

Schulze, L., Schulze, A., Renneberg, B., Schmahl, C., & Niedtfeld, I. (2019). Neural Correlates of Affective Disturbances: A Comparative Meta-analysis of Negative Affect Processing in Borderline Personality Disorder, Major Depressive Disorder, and Posttraumatic Stress Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4*(3), 220-232. doi:10.1016/j.bpsc.2018.11.004

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Birbaumer, N., & Lang, P. J. (1997). Probe P3 and blinks: two measures of affective startle modulation. *Psychophysiology, 34*(1), 1-6. doi:10.1111/j.1469-8986.1997.tb02409.x

Schwartz, M., Collura, T., Kamiya, J., & Schwartz, N. (2016). The history and definitions of biofeedback and applied psychophysiology. In *Biofeedback a practitioners guide* (Vol. 4, pp. 3-23). New York: The Guilford Press.

Shafir, R., Guarino, T., Lee, I. A., & Sheppes, G. (2017). Emotion regulation choice in an evaluative context: the moderating role of self-esteem. *Cognition & Emotion, 31*(8), 1725-1732. doi:10.1080/02699931.2016.1252723

Sheppes, G., Catran, E., & Meiran, N. (2009). Reappraisal (but not distraction) is going to make you sweat: Physiological evidence for self-control effort. *International Journal of Psychophysiology, 71*(2), 91-96. doi:10.1016/j.ijpsycho.2008.06.006

Sheppes, G., Scheibe, S., Suri, G., Radu, P., Blechert, J., & Gross, J. J. (2014). Emotion regulation choice: a conceptual framework and supporting evidence. *Journal of Experimental Psychology: General, 143*(1), 163-181. doi:10.1037/a0030831

Sheppes, G., Suri, G., & Gross, J. J. (2015). Emotion Regulation and Psychopathology. *Annual Review of Clinical Psychology, 11*, 379-405. doi: 10.1146/annurev-clinpsy-032814-112739

Shermohammed, M., Mehta, P. H., Zhang, J., Brandes, C. M., Chang, L. J., & Somerville, L. H. (2017). Does psychosocial stress impact cognitive reappraisal? Behavioral and neural evidence. *Journal of Cognitive Neuroscience, 29*(11), 1803-1816. doi:10.1162/jocn_a_01157

Shiota, M. N., & Levenson, R. W. (2009). Effects of aging on experimentally instructed detached reappraisal, positive reappraisal, and emotional behavior suppression. *Psychology of Aging, 24*(4), 890-900. doi:10.1037/a0017896

Shiota, M. N., & Levenson, R. W. (2012). Turn down the volume or change the channel? Emotional effects of detached versus positive reappraisal. *Journal of Personality and Social Psychology, 103*(3), 416-429. doi:10.1037/a0029208

Shiota, M. N., Neufeld, S. L., Yeung, W. H., Moser, S. E., & Perea, E. F. (2011). Feeling good: autonomic nervous system responding in five positive emotions. *Emotion, 11*(6), 1368-1378. doi:10.1037/a0024278

Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., . . . Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin, 144*(4), 343-393. doi:10.1037/bul0000128

Simons, R. F., Detenber, B. H., Roedema, T. M., & Reiss, J. E. (1999). Emotion processing in three systems: The medium and the message. *Psychophysiology, 36*(5), 619-627. doi:10.1111/1469-8986.3650619

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., . . . Sulzer, J. (2017). Closed-loop brain training: the science of neurofeedback. *Nature Review Neuroscience, 18*(2), 86-100. doi:10.1038/nrn.2016.164

Sloan, E., Hall, K., Moulding, R., Bryce, S., Mildred, H., & Staiger, P. K. (2017). Emotion regulation as a transdiagnostic treatment construct across anxiety, depression, substance, eating and borderline personality disorders: A systematic review. *Clinical Psychology Review, 57*, 141-163. doi:10.1016/j.cpr.2017.09.002

Soeteman, D. I., Hakkaart-van Roijen, L., Verheul, R., & Busschbach, J. J. (2008). The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry, 69*(2), 259-265. doi:10.4088/jcp.v69n0212

Sorger, B., Scharnowski, F., Linden, D. E. J., Hampson, M., & Young, K. D. (2019). Control freaks: Towards optimal selection of control conditions for fMRI neurofeedback studies. *NeuroImage, 186*, 256-265. doi:10.1016/j.neuroimage.2018.11.004

Soto, J. A., Lee, E. A., & Roberts, N. A. (2016). Convergence in feeling, divergence in physiology: How culture influences the consequences of disgust suppression and amplification among European Americans and Asian Americans. *Psychophysiology, 53*(1), 41-51. doi:10.1111/psyp.12579

Spetter, M. S., Malekshahi, R., Birbaumer, N., Lührs, M., van der Veer, A. H., Scheffler, K., . . . Hallschmid, M. (2017). Volitional regulation of brain responses to food stimuli in overweight and obese subjects: A real-time fMRI feedback study. *Appetite, 112*, 188-195. doi:10.1016/j.appet.2017.01.032

Stemmler, G. (2004). Physiological processes during emotion. In P. Philippot, R. S. Feldman, P. Philippot, & R. S. Feldman (Eds.), *The regulation of emotion.* (pp. 33-70). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*(3736), 652-654. doi:10.1126/science.153.3736.652

Stiglmayr, C., Schmahl, C., Bremner, J. D., Bohus, M., & Ebner-Priemer, U. (2009). Development and psychometric characteristics of the DSS-4 as a short instrument to assess dissociative experience during neuropsychological experiments. *Psychopathology, 42*(6), 370-374. doi:10.1159/000236908

Stiller, A.-K., Kattner, M. F., Gunzenhauser, C., & Schmitz, B. (2019). The effect of positive reappraisal on the availability of self-control resources and self-regulated learning. *Educational Psychology, 39*(1), 86-111. doi:10.1080/01443410.2018.1524851

Stoffers, J. M., Völlm, B. A., Rücker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *The Cochrane Library*.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. *Annals of Behavioral Medicine, 16*(3), 199-202. doi:10.1093/abm/16.3.199

Strauss, G. P., Ossenfort, K. L., & Whearty, K. M. (2016). Reappraisal and distraction emotion regulation strategies are associated with distinct patterns of visual attention and differing levels of cognitive demand. *PLoS ONE, 11*(11), e0162290. doi:10.1371/journal.pone.0162290

Strehl, U. (2014). What learning theories can teach us in designing neurofeedback treatments. *Frontiers in Human Neuroscience, 8*, 894. doi:10.3389/fnhum.2014.00894

Subramanian, L., Hindle, J. V., Johnston, S., Roberts, M. V., Husain, M., Goebel, R., & Linden, D. (2011). Real-time functional magnetic resonance imaging neurofeedback for treatment of Parkinson's disease. *Journal of Neuroscience, 31*(45), 16309-16317. doi:10.1523/jneurosci.3498-11.2011

Sulzer, J., Haller, S., Scharnowski, F., Weiskopf, N., Birbaumer, N., Blefari, M. L., . . . Sitaram, R. (2013). Real-time fMRI neurofeedback: Progress and challenges. *NeuroImage, 76*, 386-399. doi:10.1016/j.neuroimage.2013.03.033

Sutton, S. K., Davidson, R. J., Donzella, B., Irwin, W., & Dottl, D. A. (1997). Manipulating affective state using extended picture presentations. *Psychophysiology, 34*(2), 217-226. doi:10.1111/j.1469-8986.1997.tb02135.x

Svaldi, J., Caffier, D., & Tuschen-Caffier, B. (2010). Emotion suppression but not reappraisal increases desire to binge in women with binge eating disorder. *Psychotherapy and Psychosomatics, 79*(3), 188-190. doi:10.1159/000296138

Svaldi, J., Tuschen-Caffier, B., Lackner, H. K., Zimmermann, S., & Naumann, E. (2012). The effects of emotion regulation on the desire to overeat in restrained eaters. *Appetite, 59*(2), 256-263. doi:10.1016/j.appet.2012.04.016

Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-Aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV–heart rate variability analysis software. *Computer Methods and Programs in Biomedicine, 113*(1), 210-220. doi:10.1016/j.cmpb.2013.07.024

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2007). The skeletomotor system: Surface electromyography. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 267-299): Cambridge University Press.

Temes, C. M., Frankenburg, F. R., Fitzmaurice, G. M., & Zanarini, M. C. (2019). Deaths by Suicide and Other Causes Among Patients With Borderline Personality Disorder and Personality-Disordered Comparison Subjects Over 24 Years of Prospective Follow-Up. *The Journal of Clinical Psychiatry, 80*(1). doi:10.4088/jcp.18m12436

Thibault, R. T., Lifshitz, M., & Raz, A. (2016). The self-regulating brain and neurofeedback: Experimental science and clinical promise. *Cortex, 74*, 247-261. doi:10.1016/j.cortex.2015.10.024

Thibault, R. T., Lifshitz, M., & Raz, A. (2018). The climate of neurofeedback: scientific rigour and the perils of ideology. *Brain, 141*(2), e11. doi:10.1093/brain/awx330

Thibault, R. T., MacPherson, A., Lifshitz, M., Roth, R. R., & Raz, A. (2018). Neurofeedback with fMRI: A critical systematic review. *NeuroImage, 172*, 786-807. doi:10.1016/j.neuroimage.2017.12.071

Thompson, K. N., Allen, N. B., Chong, S., & Chanen, A. M. (2018). Affective startle modulation in young people with first-presentation borderline personality disorder. *Psychiatry Research, 263*, 166-172. doi:10.1016/j.psychres.2018.02.049.

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta‐analysis: a comparison of methods. *Statistics in Medicine, 18*(20), 2693-2708. doi:10.1002/(SICI)1097-0258(19991030)18:20<2693::AID-SIM235>3.0.CO;2-V

Torgersen, S., Kringlen, E., & Cramer, V. (2001). The prevalence of personality disorders in a community sample. *Archives of General Psychiatry, 58*(6), 590-596. doi:10.1001/archpsyc.58.6.590

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. *Psychological Assesssment, 21*(4), 457-462. doi:10.1037/a0017653

Trull, T. J., Jahng, S., Tomko, R. L., Wood, P. K., & Sher, K. J. (2010). Revised NESARC personality disorder diagnoses: gender, prevalence, and comorbidity with substance dependence disorders. *Journal of Personality Disorders, 24*(4), 412-426. doi:10.1521/pedi.2010.24.4.412

Trull, T. J., Solhan, M. B., Tragesser, S. L., Jahng, S., Wood, P. K., Piasecki, T. M., & Watson, D. (2008). Affective instability: measuring a core feature of borderline personality disorder with ecological momentary assessment. *Journal of Abnormal Psychology, 117*(3), 647-661. doi:10.1037/a0012532

Ulrich-Lai, Y. M., & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience, 10*(6), 397-409. doi:10.1038/nrn2647

Urry, H. L. (2009). Using Reappraisal To Regulate Unpleasant Emotional Episodes: Goals and Timing Matter. *Emotion, 9*(6), 782-797. doi:10.1037/a0017109

Urry, H. L. (2010). Seeing, Thinking, and Feeling: Emotion-Regulating Effects of Gaze-Directed Cognitive Reappraisal. *Emotion, 10*(1), 125-135. doi:10.1037/a0017434

Urry, H. L., van Reekum, C. M., Johnstone, T., & Davidson, R. J. (2009). Individual differences in some (but not all) medial prefrontal regions reflect cognitive demand while regulating unpleasant emotion. *NeuroImage, 47*(3), 852-863. doi:10.1016/j.neuroimage.2009.05.069

Urry, H. L., van Reekum, C. M., Johnstone, T., Kalin, N. H., Thurow, M. E., Schaefer, H. S., . . . Davidson, R. J. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. *Journal of Neuroscience, 26*(16), 4415-4425. doi:10.1523/jneurosci.3215-05.2006

Uy, C. C., Jeffery, I. A., Wilson, M., Aluru, V., Madan, A., Lu, Y., & Raghavan, P. (2013). Autonomic Mechanisms of Emotional Reactivity and Regulation. *Scientific Research, 43*(8), 669-675. doi:10.4236/psych.2013.48095

van Ginkel, J. R., Sijtsma, K., van der Ark, L. A., & Vermunt, J. K. (2010). Incidence of missing item scores in personality measurement, and simple item-score imputation. *Methodology, 6*, 17-30. doi:10.1027/1614-2241/a000003

van Reekum, C. M., Johnstone, T., Urry, H. L., Thurow, M. E., Schaefer, H. S., Alexander, A. L., & Davidson, R. J. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. *NeuroImage, 36*(3), 1041-1055. doi:10.1016/j.neuroimage.2007.03.052

van Zutphen, L., Siep, N., Jacob, G. A., Domes, G., Sprenger, A., Willenborg, B., . . . Arntz, A. (2018). Always on guard: emotion regulation in women with borderline personality disorder compared to nonpatient controls and patients with cluster-C personality disorder. *Journal of Psychiatry and Neuroscience, 43*(1), 37-47. doi:10.1503/jpn.170008

VanOyen Witvliet, C., & Vrana, S. R. (1995). Psychophysiological responses as indices of affective dimensions. *Psychophysiology, 32*(5), 436-443. doi:10.1111/j.1469-8986.1995.tb02094.x

Vitale, J. E., & Newman, J. P. (2012). Affective startle modulation in incarcerated women with borderline personality disorder features. *Personality Disorders, 3*(2), 155-166. doi:10.1037/a0025554

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: a new measure of emotion? *Journal of Abnormal Psychology, 97*(4), 487-491. doi:10.1037/0021-843X.97.4.487

Wangelin, B. C., Bradley, M. M., Kastner, A., & Lang, P. J. (2012). Affective engagement for facial expressions and emotional scenes: the influence of social anxiety. *Biological Psychology, 91*(1), 103-110. doi:10.1016/j.biopsycho.2012.05.002

Webb, T. L., Miles, E., & Sheeran, P. (2012). Dealing with feeling: A meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychological Bulletin, 138*(4), 775-808. doi:10.1037/a0027600

Weinberg, A., Klonsky, E. D., & Hajcak, G. (2009). Autonomic impairment in borderline personality disorder: a laboratory investigation. *Brain and Cognition, 71*(3), 279-286. doi:10.1016/j.bandc.2009.07.014

Weiskopf, N. (2012). Real-time fMRI and its application to neurofeedback. *NeuroImage, 62*(2), 682-692. doi:10.1016/j.neuroimage.2011.10.009

Weiskopf, N., Veit, R., Erb, M., Mathiak, K., Grodd, W., Goebel, R., & Birbaumer, N. (2003). Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. *NeuroImage, 19*(3), 577-586. doi:10.1016/S1053-8119(03)00145-9

Wessa, M., Kanske, P., Neumeister, P., Bode, K., Heissler, J., & Schönfelder, S. (2010). EmoPics: Subjektive und psychophysiologische Evaluation neuen Bildmaterials für die klinisch-bio-psychologische Forschung. *Zeitschrift für Klinische Psychologie und Psychotherapie, 39*(Suppl. 1/11), 77.

Williams, D. P., Cash, C., Rankin, C., Bernardi, A., Koenig, J., & Thayer, J. F. (2015). Resting heart rate variability predicts self-reported difficulties in emotion regulation: a focus on different facets of emotion regulation. *Frontierts in Psychology, 6*, 261. doi:10.3389/fpsyg.2015.00261

Williams, L. E., Bargh, J. A., Nocera, C. C., & Gray, J. R. (2009). The unconscious regulation of emotion: nonconscious reappraisal goals modulate emotional reactivity. *Emotion, 9*(6), 847-854. doi:10.1037/a0017745

Wolgast, M., Lundh, L.-G., & Viborg, G. (2011). Cognitive reappraisal and acceptance: An experimental comparison of two emotion regulation strategies. *Behaviour Research and Therapy, 49*(12), 858-866. doi:10.1016/j.brat.2011.09.011

Wu, M., Liang, Y., Wang, Q., Zhao, Y., & Zhou, R. (2016). Emotion dysregulation of women with premenstrual syndrome. *Sciientific Reports, 6*, 38501. doi:10.1038/srep38501

Young, K. D., Misaki, M., Harmer, C. J., Victor, T., Zotev, V., Phillips, R., . . . Bodurka, J. (2017). Real-time functional magnetic resonance imaging amygdala neurofeedback changes positive information processing in major depressive disorder. *Biological Psychiatry, 82*(8), 578-586. doi:10.1016/j.biopsych.2017.03.013

Young, K. D., Siegle, G. J., Zotev, V., Phillips, R., Misaki, M., Yuan, H., . . . Bodurka, J. (2017). Randomized clinical trial of real-time fMRI amygdala neurofeedback for major depressive disorder: effects on symptoms and autobiographical memory recall. *American Journal of Psychiatry, 174*(8), 748-755. doi:10.1176/appi.ajp.2017.16060637

Young, K. D., Zotev, V., Phillips, R., Misaki, M., Yuan, H., Drevets, W. C., & Bodurka, J. (2014). Real-time FMRI neurofeedback training of amygdala activity in patients with major depressive disorder. *PLoS ONE, 9*(2), e88785. doi:10.1371/journal.pone.0088785

Yuan, J., Liu, Y., Ding, N., & Yang, J. (2014). The regulation of induced depression during a frustrating situation: Benefits of expressive suppression in Chinese individuals. *PLoS ONE, 9*(5), e97420. doi:10.1371/journal.pone.0097420

Zaehringer, J., Schmahl, C., Ende, G., & Paret, C. (2018). Emotion-modulated startle reflex during reappraisal: Probe timing and behavioral correlates. *Behavioral Neuroscience, 132*(6), 573-579. doi:10.1037/bne0000271

Zanarini, M. C., Frankenburg, F. R., Khera, G. S., & Bleichmar, J. (2001). Treatment histories of borderline inpatients. *Comprehensive Psychiatry, 42*(2), 144-150. doi:10.1053/comp.2001.19749

Zanarini, M. C., Vujanovic, A. A., Parachini, E. A., Boulanger, J. L., Frankenburg, F. R., & Hennen, J. (2003). Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *Journal of Personality Disorders, 17*(3), 233-242. doi:10.1521/pedi.17.3.233.22147

Zeier, P., Sandner, M., & Wessa, M. (2019). Script-based Reappraisal Test introducing a new paradigm to investigate the effect of reappraisal inventiveness on reappraisal effectiveness. *Cognition & Emotion, 9*, 1-7. doi:10.1080/02699931.2019.1663153

Zilverstand, A., Parvaz, M. A., & Goldstein, R. Z. (2017). Neuroimaging cognitive reappraisal in clinical populations to define neural targets for enhancing emotion regulation. A systematic review. *NeuroImage, 151*, 105-116. doi:10.1016/j.neuroimage.2016.06.009

Zotev, V., Krueger, F., Phillips, R., Alvarez, R. P., Simmons, W. K., Bellgowan, P., . . . Bodurka, J. (2011). Self-regulation of amygdala activation using real-time fMRI neurofeedback. *PLoS ONE, 6*(9), e24522. doi:10.1371/journal.pone.0024522

Zotev, V., Phillips, R., Misaki, M., Wong, C. K., Wurfel, B. E., Krueger, F., . . . Bodurka, J. (2018). Real-time fMRI neurofeedback training of the amygdala activity with simultaneous EEG in veterans with combat-related PTSD. *Neuroimage Clinical, 19*, 106-121. doi:10.1016/j.nicl.2018.04.010

8 LIST OF TABLES