2021

Database of Picture-Based Cognitive Reappraisal Experiments: Analyses of Trial-Level Factors

Damon Abraham
University of Denver

Database of Picture-Based Cognitive Reappraisal Experiments: Analyses of Trial-Level Factors

Abstract

Cognitive reappraisal is widely recognized as an effective emotion regulation strategy for managing negative emotions. In laboratory research, reappraisal has been shown to attenuate self-reported negative affect as well as physiological and neurological markers of emotion and arousal. In these experiments, emotionally evocative images are frequently used to induce negative affect in participants. Depending on the trial condition, participants are instructed to either look and react naturally or to change their experience using reappraisal. Data are typically aggregated within trial condition, and the average difference in reported negative affect between conditions serves as the behavioral measure of reappraisal success. While reappraisal effects have been seen across multiple variations of this paradigm, there are several trial-level parameters that might contribute to the overall effectiveness of reappraisal but are currently not well-understood. We conducted a series of analyses that leverage a database of picture-based reappraisal experiments in order to examine potential triallevel factors that may promote or hinder reappraisal success. The first series of analyses examines the overall robustness of the reappraisal effect and estimates the power to detect this effect within different sample sizes. In a second series of analyses, we test what trial level factors are predictive of negative affect. Likewise, we examine whether time, in terms of a trial's ordinal position within the task, influences negative affect reported across different trial conditions. We propose and test several competing hypotheses as to whether participants habituate or sensitize to negative images over time and whether reappraisal becomes more effective with practice or less effective due to fatigue. In a third series of analyses, we examine whether the preceding trial condition influences self-reported affect on the current trial. These results will ideally contribute to a better understanding of the cognitive and affective determinants of reappraisal and may have implications for the design of future reappraisal experiments.

Document Type

Dissertation

Degree Name

Ph.D.

Department

Psychology

First Advisor

Kateri McRae

Second Advisor

Timothy Sweeny

Third Advisor

Kimberly Chiew

Keywords

Affect, Cognitive reappraisal, Emotion, Emotion regulation

Subject Categories

Cognition and Perception | Cognitive Psychology | Psychology | Social and Behavioral Sciences

Database of Picture-Based Cognitive Reappraisal Experiments: Analyses of Trial-Level Factors

---

A Dissertation

Presented to

the Faculty of the College of Arts, Humanities and Social Sciences

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Damon Abraham

June 2021

Advisor: Dr. Kateri McRae

Author: Damon Abraham
Title: Database of Picture-Based Cognitive Reappraisal Experiments: Analyses of Trial- Level Factors
Advisor: Dr. Kateri McRae
Degree Date: June 2021

Abstract

Cognitive reappraisal is widely recognized as an effective emotion regulation strategy for managing negative emotions. In laboratory research, reappraisal has been shown to attenuate self-reported negative affect as well as physiological and neurological markers of emotion and arousal. In these experiments, emotionally evocative images are frequently used to induce negative affect in participants. Depending on the trial condition, participants are instructed to either look and react naturally or to change their experience using reappraisal. Data are typically aggregated within trial condition, and the average difference in reported negative affect between conditions serves as the behavioral measure of reappraisal success. While reappraisal effects have been seen across multiple variations of this paradigm, there are several trial-level parameters that might contribute to the overall effectiveness of reappraisal but are currently not well-understood. We conducted a series of analyses that leverage a database of picture-based reappraisal experiments in order to examine potential trial-level factors that may promote or hinder reappraisal success. The first series of analyses examines the overall robustness of the reappraisal effect and estimates the power to detect this effect within different sample sizes. In a second series of analyses, we test what trial level factors are predictive of negative affect. Likewise, we examine whether time, in terms of a trial's ordinal position within the task, influences negative affect reported across different trial conditions. We propose and test several competing hypotheses as to whether participants habituate or sensitize to negative images over

time and whether reappraisal becomes more effective with practice or less effective due to fatigue. In a third series of analyses, we examine whether the preceding trial condition influences self-reported affect on the current trial. These results will ideally contribute to a better understanding of the cognitive and affective determinants of reappraisal and may have implications for the design of future reappraisal experiments.

Acknowledgments

I want to briefly acknowledge the great deal of support I have received throughout the process of preparing and writing this dissertation.

I would first like to thank my supervisor, Dr. Kateri McRae, who has offered tremendous insight, expertise, and guidance throughout this project. Moreover, the database itself reflects her own hard work over several years of research.

I wish to acknowledge my colleagues Eric Andrews and Chelsey Pan who were instrumental in helping to assemble and organize the various experiments into a central repository and to help clean the data for analyses. I would also like to thank Allison Grossberg as well, whose efforts may contribute to some of the possible future directions for this research.

In addition, I would like to thank the members of my committee, Drs. Tim Sweeny, Kimberly Chiew, and Lauren McGrath, who offered invaluable guidance on my initial proposal and feedback on the dissertation itself in preparation for publication. Likewise, I want to thank Dr. Julia Dmitrieva for her assistance with the multilevel regression analyses.

Finally, I wish to thank my wife, Bridget Abraham, for her many years of support and encouragement throughout my graduate education.

Table of Contents

List of Figures

Chapter One: Introduction

High-Level Overview of Emotion Regulation

Emotions are considered to be the short-lived adaptive patterns of perception, action, and experience that serve to orchestrate behavioral, psychophysiological, and psychological processes in response to ongoing situational demands (Cosmides & Tooby, 2000; Keltner & Gross, 1999; Levenson, 1999). Accordingly, emotions are positive or negatively valenced responses that follow the attention and appraisal of a situational antecedent (Gross, 2015; McRae & Gross, 2020). Through appraisal, emotions give meaning to situations and events and help us to orient towards salient objects and features in the environment (Pool et al., 2014; Roseman & Smith, 2001; Storbeck & Clore, 2008). Emotions are generally adaptive and beneficial for human functioning. For example, emotions can enhance consolidation and aid in the retrieval of situationally congruent information from memory (Holland & Kensinger, 2010; Lisman et al., 2011), can be informative in decision making (Mikels et al., 2011), and can serve protective functions for the survival of the organism, such as alerting one to potential threats or opportunities for reward (Brosch et al., 2008).

Yet emotions may also be incongruent with present social demands, experienced as unpleasant, overly intense, or ill-fitted to a situation, and can conflict with the individual ’s goals (Gross, 1998b). Moreover, emotions are sometimes maladaptive and can lead to suboptimal or risky decisions, distortions in memory, and can distract one away from relevant or important information in the environment (Mather & Sutherland, 2011; Porter et al., 2003; Rimmele et al., 2011; Weiss et al., 2015). Chronically

dysregulated emotions have been linked to poorer cognitive and social outcomes (John & Gross, 2004; Rubin et al., 1995; Shaw et al., 2014), increased likelihood of substance abuse (Berking et al., 2011; Fox et al., 2007), mood and anxiety disorders (D'Avanzato et al., 2013; Hofmann et al., 2012; Lukas et al., 2018), and the maintenance of depression (Abravanel & Sinha, 2015; Kiecolt-Glaser et al., 2002). The ability to regulate our experience and expression of emotion in a contextually appropriate manner is, therefore, crucial for mental health and essential for appropriate socialization and adaptive human functioning (Gross & Muñoz, 1995; John & Gross, 2004; Keltner & Gross, 1999).

Emotion Regulation Defined

Emotion regulation (ER) is an umbrella term used to describe any number of processes by which individuals attempt to manage their emotional experience and expression (Gross, 1998b; McRae & Gross, 2020). According to the process model of emotion regulation, ER can broadly be categorized into 5 families of strategies depending upon the phase of emotion generation at which the regulatory process intervenes (Gross, 2015; McRae & Gross, 2020). Situation selection involves avoiding or pursuing circumstances that are likely to give rise to a given emotional state (e.g., declining a party invitation to avoid social anxiety or watching a comedy to boost mood). Situation modification involves the behavior one engages in to alter the ongoing circumstances (e.g., having a drink to relax at the party). One may also deploy attention towards or away from emotionally evocative information. Attention deployment can be outward (e.g., looking at a smart phone to avoid eye contact with a crush) or inward focused (e.g., recalling a happy memory to avoid thinking about a stressful situation). Cognitive change involves modifying the appraisal of a situation by rethinking or reinterpreting its meaning or implications (e.g., "This party is an opportunity to make new

friends"). This family of ER strategy is the focus of the present work. Finally, emotional behaviors can be modified using response modulation (e.g., smiling to hide anxiety in a conversation).

Cognitive Reappraisal Defined

Reappraisal is a form of cognitive change involving reframing or reconstruing a situation's meaning in order to change how it is experienced emotionally (Gross & Thompson, 2007; Lazarus & Alfert, 1964). It is widely considered to be one of the more beneficial regulation strategies and is linked to multiple adaptive mental and physical health outcomes (Gross & John, 2003; John & Gross, 2004; Webb et al., 2012). How much people use reappraisal has been associated with better physical and mental health, positive social interactions, and greater academic achievement (Appleton et al., 2013; Cludius et al., 2020; English et al., 2012; Ivcevic & Brackett, 2014). Reappraisal has also been associated with greater clinical efficacy in empirically supported interventions such as cognitive behavioral therapy (CBT) and mindfulness based stress reduction (MBSR) (Aldao et al., 2014; Moscovitch et al., 2012; Wharton & Kanas, 2019). The vast majority of these findings use questionnaires to measure individual differences in how often individuals use reappraisal, sometimes called reappraisal frequency (McRae, 2013; McRae & Gross, 2020).

Reappraisal in Laboratory Experiments

In the lab, experiments on reappraisal typically involve inducing a negative emotion in the participant, either through a mood induction or in response to an emotionally evocative stimulus. The participant is instructed to change their emotional experience using reappraisal on a subset of trials. Following each trial, the participant reports their emotional experience, typically on a numerical rating scale. In this case, the regulatory goal is provided extrinsically, as the regulatory process is instructed. Here, the

output of the regulatory process is reappraisal success, which is the primary outcome measure of interest. Behaviorally, reappraisal success is often measured as the average difference between the participant's self-reported affect on regulated and non-regulated trials.

Typical Picture-Based Reappraisal Task Design

Across studies, there is considerable variation with respect to the specifics of the task, including the reappraisal instruction (e.g., "increase positivity" or "decrease negativity"), as well as the type of emotion-eliciting stimuli used (e.g., photographs, music, or video clips etc.). Our lab has predominantly employed variations of the picture-based reappraisal task, which is the focus of the present work. The basic picture-based task involves inducing a negative emotion in the participant via the presentation of negatively valenced pictorial stimuli. Likewise, in these experiments, there is always a condition in which the participant attempts to ameliorate the negativity by rethinking/reframing/reconsidering the meaning of the negative image (i.e., reappraisal). The images used in picture-based experiments are typically sourced from one of a number of affective image databases (e.g., IAPS, OASIS etc.). These databases contain normative information for each image in terms of positive or negative emotionality (valence) and motivational intensity (arousal) (Kurdi et al., 2017; Lang et al., 2008). This allows researchers to select specific stimuli that are likely to instantiate the desired emotional state in the participant.

On each trial, the participant is instructed to either allow their emotional experience to unfold naturally (i.e., "Look") or to regulate the emotion using reappraisal (i.e., "Change")[1]. The participant is then asked to report their subjective emotional

[1] We use the word "Change" here to refer to any reappraisal trial condition, however experiments may vary on the specific instruction word used for these trials (see Table 1).

experience, typically on a numerical rating scale. In addition, these experiments generally contain a neutral baseline trial condition in which the participant is presented a neutral emotional image and instructed to respond naturally ("Look").

Many picture-based reappraisal experiments are quite similar with respect to their overall task design and analyses but can nevertheless vary considerably in terms of specific task parameters. These experiments universally entail some form of regulation of negative affect. However, they may employ different reappraisal tactics and instructions for participants (McRae, Ciesielski, et al., 2012). Experiments may also vary in trial sequences, number of trials and blocks, and duration of stimulus presentations as well as the measures captured (e.g., behavioral, neural, psychophysiological, etc.,) and study environments (laboratory, fMRI, online etc.). Likewise, experiments may capture subjective ratings on various types of rating scales. While many experiments employ Likert-type rating scales, the number of ratings, specific anchor text, as well as direction of the scale (e.g., higher ratings = more positive or more negative) can differ. Regardless of these variations, the contrast between trial conditions remains the critical and essential component of the picture-based reappraisal experiment.

Outcomes of Reappraisal in Laboratory Experiments

In laboratory experiments, successful reappraisal is associated with significant decreases in self-reported negative affect (Gross, 1998a). These changes in affect persist even after accounting for experimenter demand characteristics and are corroborated by a host of other neural and physiological measures. For example, reappraisal has been associated with up-regulated activity in prefrontal neural regions associated with cognitive control and down-regulated activity in regions involved in emotional responding such as the amygdala (Buhle et al., 2014; Goldin et al., 2008; McRae et al., 2010; Ochsner et al., 2012). Likewise, reappraisal has been shown to

modulate electrocortical and peripheral physiological markers of arousal in response to stress such as skin conductance, facial electromyography (fEMG), startle eye blink, and cardiac responses (Boehme et al., 2019; Hajcak & Nieuwenhuis, 2006; Pavlov et al., 2014; Ray et al., 2010; Zaehringer et al., 2018).

Emotion Reactivity

Generally, these experiments employ repeated measures designs involving within-subject comparisons across the three trial conditions. In these analyses, a pairwise comparison between the negative and neutrally valenced “Look” trial conditions provides a measure of emotion reactivity. It is expected that negatively valenced “Look” trials will be rated as more negative than neutrally valenced trials (i.e., “Look Negative” > “Look Neutral”). This comparison also serves as a quality control check as participants who routinely rate neutral images as equivalent or more negative than negative images may be non-compliant or inattentive to task instructions. These participants are frequently excluded from the final analyses.

Regulation

According to the process model of ER, in these experiments, the outcome of the regulatory process is the consequent change in negative affect following reappraisal. However, in the picture-based reappraisal task, generally only a single affective rating is captured per participant per image. As such, a change in affect cannot be computed at the individual trial level. Therefore, regulation is operationalized as the average difference between the participant’s subjective ratings on regulated and non-regulated trials (i.e., “Look Negative” > “Change”). Behaviorally, reappraisal success is the magnitude of this difference. More specifically, reappraisal success is marked by a statistically significant lower average rating for the “Change” instructed vs. negative valence “Look” instructed trial conditions (i.e., “Look Negative”).

Gap in Present Reappraisal Research

Much of the prior research on reappraisal both in and out of the lab has centered around comparisons with other strategies in terms of how well they achieve regulatory goals (Gross, 1998a; Webb et al., 2012), the fit between goals and strategies with situational demands (Troy et al., 2013), individual differences (Gross & John, 2003; McRae, Jacobs, et al., 2012; Ray et al., 2005), and the match between individual and context (Ford & Troy, 2019; Kobylińska & Kusev, 2019) as predictors of regulatory success. Within the lab, much of the experimental research using the picture-based reappraisal task operationalizes reappraisal success globally, by averaging across all trials administered. Therefore, much less is known about more narrowly defined local contextual factors --specific to trial events -- that may be predictive of reappraisal success at the individual trial-level.

The present study aims to determine what role trial-level contextual factors might play in emotional reactivity and regulatory success on a trial level in reappraisal experiments. Of particular interest to the present study are the effects of a trial's ordinal position in the task (i.e., time), as well as the emotional valence and regulatory instruction on immediately preceding trial events.

Trial-Level Factors

Trial-Level Factors: Time on Task

Within a picture-based reappraisal experiment, time on task relates to the number of emotional stimuli presented and affect ratings made by the participant. Here, we can examine whether emotional reactivity and regulation change as a function of time, operationalized as a trial's ordinal position within the task. Therefore, reactivity and regulation, in this case, are operationalized as the subjective affect ratings during "Look Negative" and "change" trials respectively.

Framing the question in this way allows us to test several competing hypotheses, outlined in more detail below. With respect to emotional reactivity, the repeated presentations of negatively valenced images could potentially lead to a sensitization effect whereby emotional reactivity increases over time. By contrast, the repetition of negativity could lead to a habituation effect by which emotional reactivity decreases over time. In terms of emotion regulation, there may be compounding effects of repeated regulation attempts. Over the course of the experiment, a practice effect of regulation may be observed by which participants become more efficient at applying a reappraisal strategy over time. However, reappraisal is believed to be an effortful and cognitively taxing exercise. Therefore, to the extent that repeated regulation attempts consume limited cognitive resources, a fatigue effect might be observed by which participants become less efficient at reappraisal over time.

Emotion Reactivity Sensitization

Emotion reactivity sensitization effects have been seen across various measures in previous picture-based experiments that do not involve reappraisal. When presented blocks of multiple negatively valenced images, participants displayed increases in negative affect, exhibited progressive increases in corrugator EMG responses, and had potentiated eyeblink startle responses for extended periods of time (Bradley et al., 1996; Smith et al., 2005; Sutton et al., 1997).

Notably, these designs differ from the picture-based reappraisal task in that they did not incorporate a regulation condition. Moreover, the presentation of blocks of contiguous images within the same negative valence category may be a factor in whether or not sensitization effects will be exhibited. Prior studies typically incorporated intermixed designs in which images of negative, positive and neutral valence categories appear in a pseudorandom presentation sequence (Bradley et al., 2001; Lang et al.,

1993). It is assumed that intermixed designs are optimal for capturing the phasic responses to affective stimuli and should minimize mood induction effects due to the more rapid alternations of valence (Frijda, 1988). However, a recent study found that even within intermixed designs, how image valence is distributed across trials in the task can have effects on ratings of affect over time (Czekóová et al., 2015). Specifically, affect ratings for negative targets attenuated in intermixed trial sequences that included both positive and neutral stimuli.

Although picture-based reappraisal experiments do typically intermix neutral and negative valence trials, the images selected for these tasks tend to be predominantly negative (roughly 2/3rds of trials) and don't often include a positive valence category. Given the absence of positively valenced images, which might contribute to mood repair, it is possible that the picture-based reappraisal task could lead to a negative mood induction and similar sensitization effects as seen in the negative image block designs. A sensitization effect would be marked by a positive slope in participants' subjective affect ratings for "look" negative trials over the course of an experiment (see Fig. 1a).

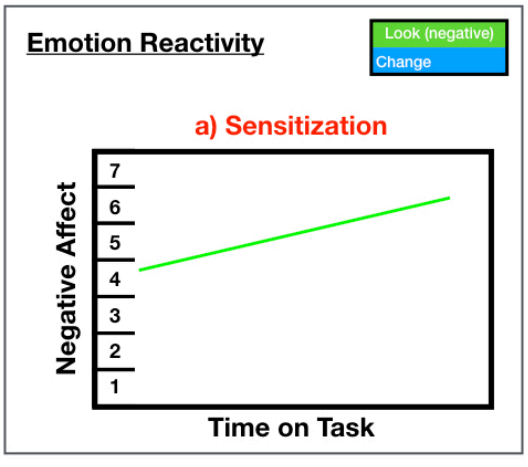

Figure 1a Emotion Reactivity Sensitization: Predicted pattern for "look" negative trials if participants become increasingly sensitized to the negative images over time. See full Figure 1 below.

Emotion Reactivity Habituation

Repeated or prolonged presentations of an emotional stressor can lead to habituation marked by a gradual attenuation in response magnitude (Harris, 1943). In terms of psychophysiological responding, habituation effects have been noted in prior picture-based experimental paradigms (Bradley et al., 1993; Codispoti et al., 2006; Wendt et al., 2012) as well as those incorporating video stimuli (Koukounas & Over, 2000). In this manner, stimulus repetition might be thought of as a form of emotion regulation in its own right such as in the case of diminishing a phobic fear response through repeated exposure (Benito & Walther, 2015).

Habituation effects tend to occur when the same affective stimulus is presented in repetition, however. For example, Bradley et al. (1993) repeatedly presented the same IAPS images to participants in blocks of intermixed valence categories (i.e., positive, neutral, and negative) (Bradley et al., 1993). Unlike Bradley et al. (1996) which presented blocks of different images of the same valence categories, the repetition of images in the 1993 study led to a nearly opposite pattern of results (Bradley et al., 1996). Skin conductance (SCR), corrugator EMG, and startle eye blink responses all exhibited a marked decline over the course of repetitions.

Similar habituation effects have been noted in subjective self-reports of negative affect in picture-based reappraisal tasks when stimuli are repeated. For example, in reappraisal experiments employing test/retest designs utilizing the same images across tests, participants tend to report a reduction in negative affect on subsequent presentations of the same stimulus particularly if the participant reappraised the stimulus during the initial presentation (Denny et al., 2015; Erk et al., 2010; MacNamara et al., 2011; Silvers, Shu, et al., 2015). Reappraisal studies generally do not repeat the same stimuli within an experiment, however. As such, it may be unlikely that a habituation of

emotional reactivity effect would be evident in most reappraisal experiments. In the present examination, an emotion reactivity habituation effect would be marked by a negative slope for subjective affect ratings for “Look Negative” trials over the course of an experiment (see Fig. 1b).

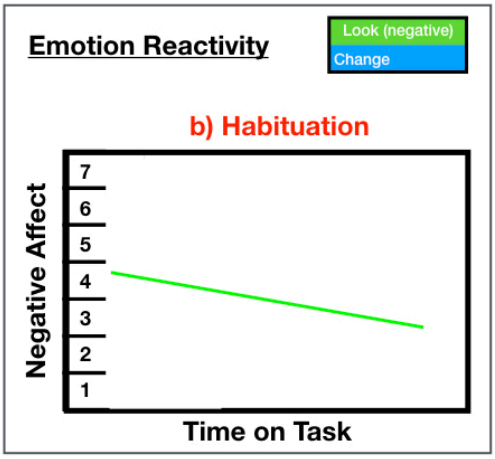

Figure 1b Emotion Reactivity Habituation: Predicted pattern for “Look Negative” trials if participants habituate to the negative images over time. See full Figure 1 below.

Reappraisal Practice Effects

Reappraisal is a cognitively demanding skill, and like any other skill should improve with practice. The number of times the participant attempts to regulate their emotions over the course of an experiment might therefore improve the efficiency and effectiveness of their reappraisal efforts. Predominantly, prior research on reappraisal practice effects has been focused on reappraisal training and intervention in longitudinal studies (Denny & Ochsner, 2014; Pogrebtsova et al., 2018), stimulus repetition practice effects (Denny et al., 2015; Silvers, Shu, et al., 2015), and individual differences in reappraisal frequency, which are assumed to reflect the effects of accumulated practice (Gross & John, 2003; McRae, Jacobs, et al., 2012). To date, very little research has

been focused on the practice effects of reappraisal (when stimuli are not repeated) within the course of a single experiment.

The experimental setting is a novel context for most research participants and the structure of the reappraisal task itself is likely somewhat novel as well, even to those who regularly practice reappraisal in daily life. The novelty of the experimental context could increase general levels of cognitive load as the participant grapples with task demands. However, as a participant becomes more comfortable with the task and acquires more experience with the reappraisal process, reappraisal efforts may become less strenuous and more productive. As such, a practice effect would be reflected in decreasing subjective self-reports of negative affect on “change” trials across repeated reappraisal attempts over the course of the experiment, as indicated by a negative slope for this trial condition (see Fig. 1c).

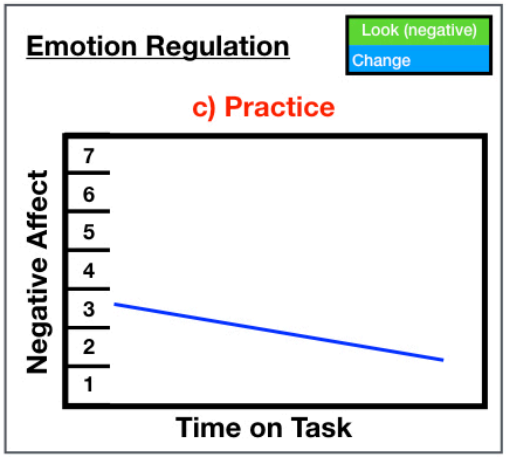

Figure 1c Emotion Regulation Practice Effect: Predicted pattern for “Change” trials if participants become more efficient at regulating over time due to a learning effect. See full Figure 1 below.

Reappraisal Fatigue Effects

In contrast to a practice effect, it is also possible that reappraisal could lead to fatigue. Reappraisal is a cognitively demanding and effortful act of self-regulation in which a person must overcome prepotent appraisals and responses to a stressful negative stimuli and events (Hofmann et al., 2012). According to ego depletion and strength models of self-control (Baumeister et al., 2007; Baumeister & Heatherton, 1996), self-regulation efforts draw from limited cognitive resources. Once self-regulatory resources are depleted, regulation becomes more effortful, less efficient, and less successful[2].

Within the reappraisal literature, the cognitive costs of reappraisal have mostly been measured indirectly. For example, studies have examined participant's choice behavior when presumably less-difficult regulation options are made available, or have evaluated the subjective difficulty of reappraisal (Milyavsky et al., 2019; Sheppes et al., 2011, 2014; Troy et al., 2018). While these studies do not speak to fatigue specifically, if regulatory resources are indeed limited then more difficult tasks should deplete resources more quickly leading to fatigue.

In other research, the cognitive costs of reappraisal were determined based upon performance on secondary cognitive reaction time tasks. One study found that a difficult cognitive task diminished regulation effectiveness in a subsequent reappraisal task (Grillon et al., 2015). However, these experiments did not relate the cognitive costs of regulation to fatigue within the reappraisal task itself. In a picture-based reappraisal task, a fatigue effect would be reflected in increasing subjective self-reports of negative

[2] It should be noted that despite numerous experiments, the reproducibility of the data supporting ego depletion models remains hotly contested (Friese et al., 2019).

affect on “Change” trials over the course of the experiment as indicated by a positive slope for this trial condition (see Fig. 1d).

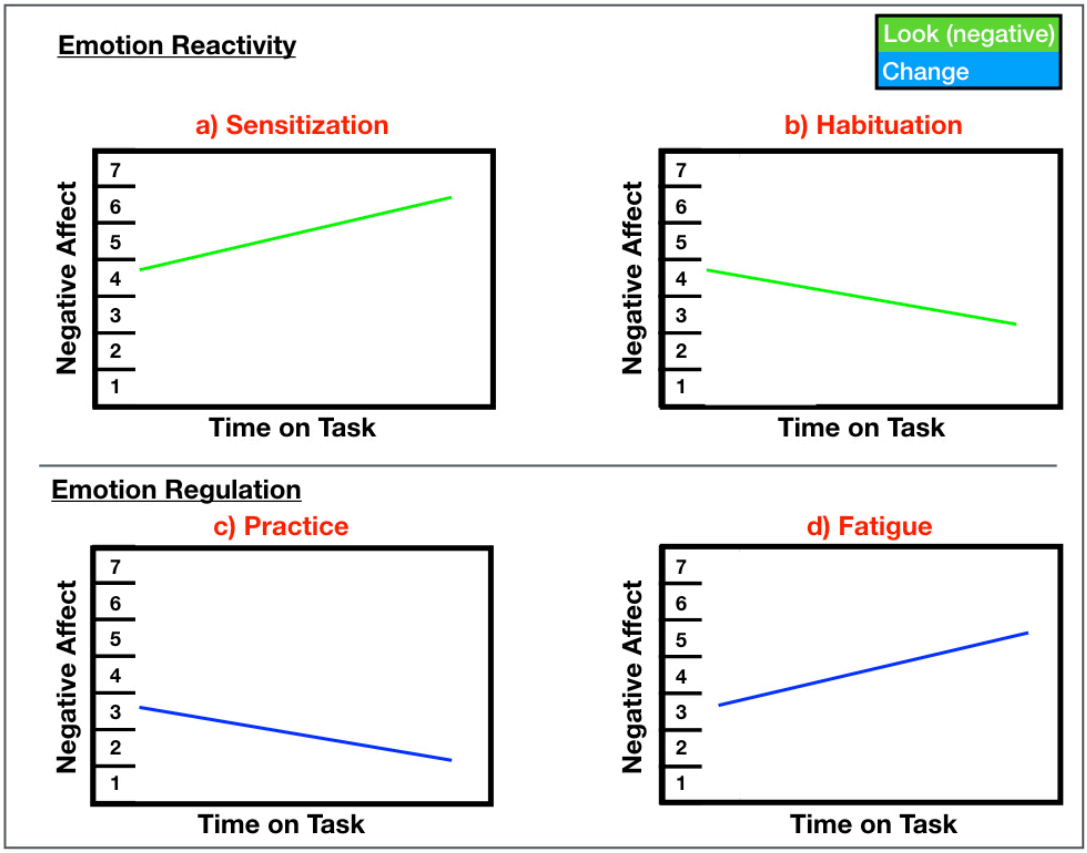

Figure 1 Predicted Patterns of Results for Time on Task: depicts patterns of predicted results for ratings of negative affect depending upon time on task broken and down by Emotion Reactivity and Emotion Regulation. 1a) Predicted pattern for “Look Negative” trials if participants become increasingly sensitized to the negative images over time. 1b) Predicted pattern for “Look Negative” trials if participants habituate to the negative images over time. 1c) Predicted pattern for “Change” trials if participants become more efficient at regulating over time due to a learning effect. 1d) Predicted pattern for “Change” trials if participants become fatigued following multiple regulation attempts.

Mixed Temporal Effects

These hypotheses are not entirely mutually exclusive, and it is possible that there may be a combination of effects. For example, sensitization combined with fatigue would be reflected in positive slopes for both “Look Negative” and “Change” trial conditions.

However, a sensitization combined with a practice effect could result in an upward slope for “Look Negative” trials, but a relatively flat slope for the “Change” trial condition as the effect of practice might be somewhat countered by the increasing difficulty to reappraise progressively more negatively seeming images as participants become sensitized. Similarly, a habituation effect combined with fatigue might result in a downward slope for “Look Negative” trials with a slightly more positive slope for “Change” trials. Finally, habituation combined with a practice effect would result in downward slopes for both trial conditions, but likely a steeper slope for the “Change” trials. Notably, I do not predict any effects of time for the “Look Neutral” trials and thereby would expect a relatively flat slope in this trial condition regardless of whether there was sensitization, habituation, practice, or fatigue.

Trial-by-Trial Sequence effects

Sequence Effects on Emotion Reactivity – Negative Valence Carryover

Picture-based reappraisal tasks typically employ pseudorandom trial sequences that intermix image valence (neutral and negative) and instruction (“Look”, “Change”) and minimize sequential repetitions of the same trial conditions. This design attempts to minimize any systematic confounding effects of the trial sequence on the primary contrast of interest. However, affective carryover effects have been reported across a variety of experimental paradigms in which preceding trial events are shown to contribute to trial-level variations in self-reported affect, electrocortical, and psychophysiological responding (Flaisch, Junghöfer, et al., 2008; Flaisch, Stockburger, et al., 2008; Larsen & Norris, 2009; Schupp et al., 2012; Waugh et al., 2011). In particular, in picture-based tasks, affective carryover tends to have an assimilation effect (as opposed to a contrast effect), by which the affective rating of a target image tends to be biased in the same direction as the valence of the preceding image (Czekóová et al.,

2015; Fujimura et al., 2013; Palumbo et al., 2017). For example, a neutral image that follows a negative image will tend to be rated as more negative than had it followed a positive or another neutral image. Hence, the preceding trial appears to establish a local context in which the appraisal of the current target image becomes embedded (Czekóová et al., 2015).

While affect carryover effects have been well-established for emotion reactivity, their effects remain relatively unexplored with respect to regulation within the reappraisal literature. Based on prior findings, valence carryover effects might be expected in the picture-based reappraisal task. Specifically, a negative valence carryover effect would be evidenced by more negative subjective ratings for current target images when preceded by a negatively valenced image as opposed to another neutral image (see Fig. 2b). However, whether there is an intervening influence of reappraisal on the negative valence carryover remains an open question.

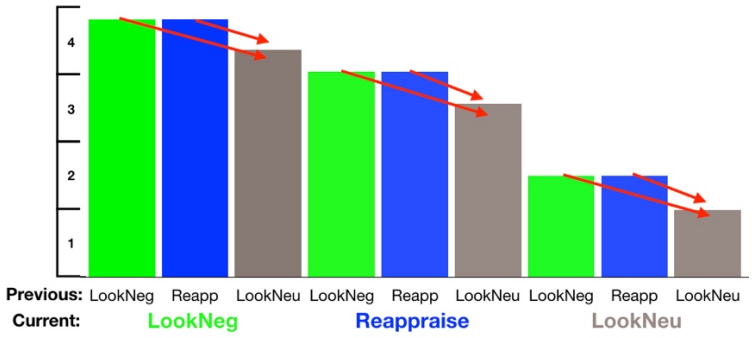

Figure 2b Negative Valence Carryover Effect: Predicted pattern of negative affect on current trial assuming a negative affect carryover effect from a preceding negative “Look Negative” or “Change” trial. See full Figure 2 below.

Sequence Effects on Regulation – Cognitive Process Carryover

The local context might also exhibit carry over effects in terms of cognitive processes. The previous instruction to “change” could instantiate a mindset whereby the

participant might reappraise the current trial despite an instruction to “look” and respond naturally. Thus, a local reappraisal context might result in a form of implicit emotion regulation (Braunstein et al., 2017; Koole & Rothermund, 2011). A recent fMRI study found that the extent to which participants engaged dorsomedial and dorsolateral prefrontal cortex (part of a core network of neural regions associated with instructed reappraisal), was predictive of lower negative affect ratings even on trials in which no “Change” instruction was given (Silvers, Wager, et al., 2015). While a cognitive process cannot be inferred based on fMRI data alone, these results suggest that participants may have been implicitly regulating their emotions even when instructed to “Look” at the negative images.

Participant responses have also been shown to be influenced by explanatory narratives (i.e., appraisal frames) that precede the image. Specifically, neutrally valenced appraisal frames about an upcoming stimulus result in reductions in self-reported negative affect and psychophysiological responding (Foti & Hajcak, 2008; Wu et al., 2012). These results suggest that the appraisal frames create an interpretative lens by which the stimulus is viewed. However, the act of reappraisal itself could potentially create a similar interpretative disposition affecting the appraisal of a subsequent stimulus. Thus, participants may be more inclined to implicitly regulate their emotions after they have just reappraised a negative image regardless of the current trial’s instruction.

A cognitive process carryover effect would likely have little influence if the current target image is neutral as these ratings should already be close to the floor. However, if participants are implicitly regulating due to a cognitive process carryover effect, then affective ratings for negative images may be selectively lessened for both current “Look

Negative" and "Change" instructed trials when preceded by a "Change" trial but not when preceded by either a "Look Negative" or "Look Neutral" trial (see Fig. 2c).

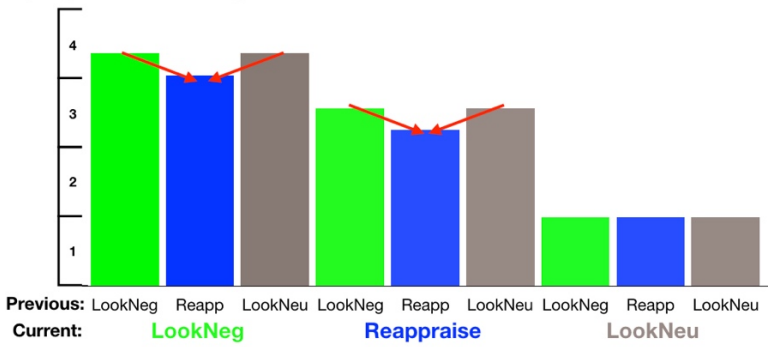

Figure 2c Cognitive Process Carryover Effect: Predicted pattern of negative affect on current trial assuming a cognitive process carryover effect from a preceding "Change" regulated trial. See full Figure 2 below.

Sequence Effects on Regulation – Cognitive Process Facilitation

Another possibility is that there is a process facilitation effect wherein reappraisal is more effective on trials that immediately follow another "Change" instructed trial. Even if participants do not implicitly regulate their emotions following reappraisal trials, the act of reappraising on a previous trial could potentially facilitate reappraisal efforts on the current trial. Research on repetition priming demonstrates that performance on cognitive interference tasks improves when interference trials are repeated (Kristjánsson & Campana, 2010). Likewise, according to the response conflict monitoring model of self-control, conflict from a preceding trial triggers enhanced top-down cognitive control on the current trial (Botvinick et al., 2001; Ullsperger et al., 2005). Similar conflict adaptation effects have been found in variations of the emotional Stroop and flanker tasks (Chechko et al., 2014; Zeng et al., 2017).

A process facilitation prediction is somewhat similar to a cognitive process carryover effect. However, a process carryover implies that participants are more likely to implicitly regulate following reappraisal and are essentially not task switching in accordance with trial instructions. Moreover, a carryover effect does not speak directly to the efficiency of the cognitive process. By contrast, a process facilitation effect is specific to reappraisal trials and suggests enhanced cognitive control. A facilitation effect would be evidenced by selectively lower negative affect ratings for current "Change" instructed trials when immediately preceded by another "Change" instruction as opposed to a "Look" instruction (see Fig. 2d).

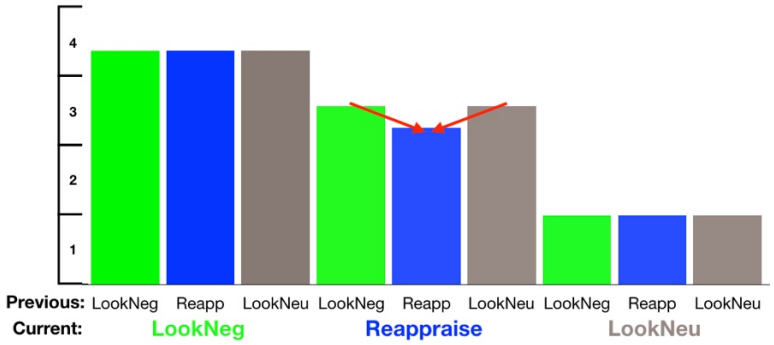

Figure 2d Cognitive Process Facilitation Effect: Predicted pattern of negative affect on current "Change" trial assuming a process facilitation from a preceding "Change" trial. See full Figure 2 below.

Sequence Effects on Regulation – Cognitive Process Fatigue Effect

In contrast to a cognitive process facilitation effect, effortful attempts at reappraisal could exhaust limited cognitive resources, thereby rendering a subsequent reappraisal attempt less efficient and effective. This would be consistent with ego depletion and strength models of self-control which suggest that engaging in effortful self-regulation leads to declining performance on subsequent regulation tasks (Hagger

et al., 2010). Evidence for trial-by-trial fatigue effects is limited; particularly in picture-based tasks. However, several studies have found that emotional and threat-related attentional biases can become more pronounced under conditions of high cognitive load (Jeong & Cho, 2020; Kobylińska & Kusev, 2019; McGuire et al., 2016; Pecchinenda & Petrucci, 2016).

To the extent that regulation temporarily taxes cognitive resources, threatening or arousing information in an immediately subsequent target may become more captivating and difficult to reappraise. Therefore, a cognitive process fatigue effect could result in selectively higher ratings of negative affect on current “Change” instructed trials that follow another “Change” trial as compared to those following “Look” instructed trials (see Fig. 2e).

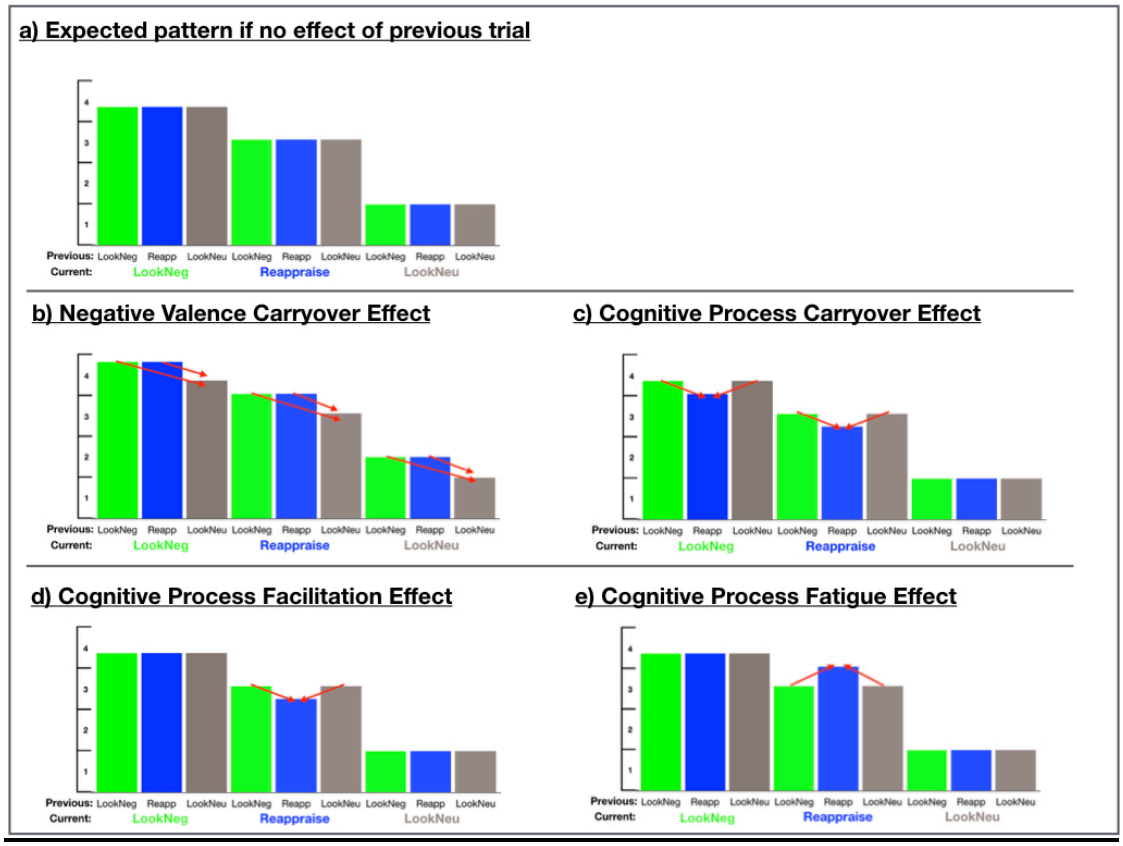

Figure 2 Predicted Patterns of Results for Trial-by-Trial Sequence Effects: depicts patterns of predicted results for ratings of negative affect depending upon the immediately preceding trial condition. The colored bars within each graph represent the 9 possible previous by current trial combinations. LookNeg = “Look” instructed trials with negative valence, LookNeu = “Look” instructed trials with neutral valence, and Reapp/Reappraise = “Change” instructed trials with negative valence. 2a) Predicted pattern of negative affect on current trial assuming no influence from preceding trial condition. 2b) Predicted pattern of negative affect on current trial assuming a negative affect carryover effect from a preceding “Look Negative” or “Change” trial. 2c) Predicted pattern of negative affect on current trial assuming a cognitive process carryover effect from a preceding “Change” regulated trial. 2d) Predicted pattern of negative affect on current “Change” trial assuming a process facilitation from a preceding “Change” trial. 2e) Predicted pattern of negative affect on current “Change” trial assuming a fatigue effect following a preceding “Change” trial.

Chapter Two: Methods

Participants

Trial-level behavioral data were aggregated from 1435 individual participants collected from 27 separate published and unpublished experiments involving a picture-based cognitive reappraisal task. Data from 22 participants were removed from the database either for having missing affect ratings on every trial (N=7) or for having too little variability in the ratings (N = 15). We applied the following criteria for exclusion based on insufficient response variability: First, at least seventy-five percent of affect ratings must have occurred within a streak of three or more trials having the same rating. Secondly, the proportion of trials endorsed with the same rating must exceed 1.96 standard deviations as compared to all other participants in the database that used the same rating scale. The final sample included 1413 participants. The specific breakdown of participant demographic information is still under preparation (See Appendix A).[3] Participant ages ranged from ~8-55 years, with the vast majority being between 18-30. Several studies contained only females, and the sample is therefore likely ~65-80% female. Two studies in the database also included developmental samples with children as young as 8-years-old, however, this is a relatively small segment of the full sample. The trial-level behavioral data from these participants were compiled into the reappraisal database.

[3] Trial-level data were stored separately from demographic and individual difference data. Given that much of these data were collected across multiple locations over a span of several years this will take more time to aggregate.

Table 1 contains the pertinent task-level and demographic information of each study. For each study, we report whether the experiment was conducted in-person (either in the lab or fMRI) or online and the rating scale that was used. All participants provided informed consent and were compensated for their participation. Compensation either took the form of cash payment (based on an hourly rate which varied somewhat depending upon the on the date and geographic location of the study) or research participation credits offered through a psychology department. The rate and manner of compensation was approved by the IRB at the institution where each study took place.

Table 1 Task and Trial-Level Event Details

Study	Included N	Scale Min	Scale Max	Num. Trials	Location	Instruction Duration	Fixation Duration	Image Duration	Rating Duration	Rest Duration	Rest cue	Look Instruction	Chage Instruction
Artist_ER	40	1	4	60	fMRI	2000	~3800 jittered	8000	4000	~3889 jittered	Fixation Cross	Look	Change
ARTIST-Online	34	1	10	60	Online	2000		8000	untimed	2000	"Relax"	Look	Change
BEER-Cartoon	49	1	5	114	Lab	2000	2000	7000	3500	2000	Fixation Cross	Look	Decrease
BEER_Task_ Final	42	1	5	126	Lab	2000	1000	7000	3500	1000	Fixation Cross	Look	Decrease
BEER-WFU	41	1	5	126	Lab	2000	1000	7000	3500	1000	Fixation Cross	Look	Decrease
CANCER-Online	32	1	10	45	Online	2000	-	8000	untimed	2000	"Relax"	Look	Change
Developmental -Behavioral	45	1	4	95	Lab	4000	-	8000	4000	4000	"Relax"	Look	Decrease
Developmental -fMRI	54	1	4	72	fMRI	4000	-	8000	4000	4000	"Relax"	Look	Decrease
Distraction-Published	18	0	9.5	168	fMRI	2000	-	8000	4000	2500	"Relax"	Attend	Decrease
EFS study_7-2-08	101	1	9	60	Lab	2000	-	7000	4000	7000	"Relax"	Look	Decrease
ELERS_1	58	1	4	140	fMRI	~2000 jittered	-	7000	3500	2000	"Relax"	Look	Change
ELERS_2	34	1	4	86	Lab	~2000 jittered	-	7000	3500	2000	"Relax"	Look	Change
Stanford EMG	39	0	7	189	Lab	3000	-	8000	3000	6000	Fixation Cross	Aware	Decrease
ERS Version B	65	1	8	132	Lab	2000	-	7000	3500	~3000 jittered	Fixation Cross	Look	Decrease
ERS Version C	22	1	8	132	Lab	2000	-	7000	3500	~3000 jittered	Fixation Cross	Look	Decrease
ERS Version D	10	1	7	180	fMRI	1000	-	7000	3500	2000	Fixation Cross	Neutral	Decrease
ERS Version E	10	1	7	180	fMRI	1000	-	7000	3500	2000	Fixation Cross	Neutral	Decrease
ERS Version F	108	1	7	180	Lab	4000	-	7000	3500	~3000 jittered	Fixation Cross	Look	Change
ERS fMRI	34	1	7	180	fMRI	4000	-	7000	3500	~3000 jittered	Fixation Cross	Look	Decrease
IGT	130	1	9	75	Lab	2000	1500	7000	untimed	1500	Fixation Cross	Look	Change to Positive
PEER	50	1	9	144	Lab	2000	-	7000	untimed	2000	"Rest"	Look	Down
POS	81	1	9	79	Lab	3000	-	7000	4000	2000	Fixation Cross	Look	Change
SOS-Online	94	1	9	192	Online	2000	-	8000	untimed	2000	"Relax"	Look	Change
ROSE Version A	30	1	9	192	Lab	2000	-	8000	3500	2500	"Relax"	Look	Change
ROSE Version B	88	1	10	60	Lab	2000	-	8000	3500	2500	"Relax"	Look	Change
RTP	85	1	7	48	Lab	2000	-	7000	untimed	untimed	-	Look	Decrease
SOS fMRI	19	1	8	120	fMRI	2000	~3800 jittered	8000	4000	~3900 jittered	"Relax"	Look	Change

The table contains the task and trial-level details for each study. N of participants is the number included in the analyses (i.e., after exclusions were removed). Event durations are shown in milliseconds (or average ms. if jittered)

General Procedure

Each of the 27 studies included in the reappraisal database involved a version of an event-related cognitive reappraisal task (Jackson et al., 2000; Ochsner et al., 2002). During these tasks, participants viewed images following an instruction to either "Look" or "Change"[4]. Under the "Look" condition, participants were asked to attend to the image, react naturally, and not try to change their emotional response, to the image presented. Under the "Change" condition, participants were trained to use reappraisal, that is, generate alternate reinterpretations or perspectives on the situation depicted, to either increase positive and/or decrease negative feelings about the image presented. See Table 1 for experiment-level specific information about the duration of trial events and specific instruction conditions. Following each image, participants rated their negative affect using a provided scale (see Figure 3). The number of response options on the rating scale varied across the experiments (see Table 1).

The images presented were either normatively negative (in both "Look" and "Change" conditions) or neutral ("Look" condition only; see 2.3.1 Stimuli). Each stimulus was presented only once per participant and in most experiments, the trial instructions appeared in a pseudorandom order with the limitation that no more than 2-to-3 of the same trial type could occur sequentially. The database does contain one study, which was an exception. This study employed a between-subjects design in which each participant only executed one type of trial instruction. Likewise, a second study did not include a Look Neutral baseline condition. Data from these studies were not included in the trial-level analyses below. The mapping of condition to the specific negative

[4] The specific instruction words ("Look", "Change" etc.) varied across studies but conveyed essentially the same meaning (See Table 1).

emotional images presented (i.e., “Look” vs. “Change”) was counterbalanced across participants within most studies with few exceptions.

Prior to the start of the task, participants were trained to use reappraisal. For in-person studies, the experimenter explained the goal of reappraisal, showed the participant sample images not used in the main task, offered examples of acceptable reappraisals, and asked participants to generate their own reappraisal, providing gentle correction when needed. For studies conducted online, the goal of reappraisal was explained to the participant, and then example reappraisals were provided with sample images not included in the main task. As practice, participants were then asked to generate reappraisals to images that were not included in the main task. Some online studies included a quiz to ensure participant understanding of instructions. For all studies, post-task questionnaires verified that participants were able to describe reappraisal, and they briefly described which specific reappraisal tactics they used most often, as well as their self-reported success in completing each trial type.

Trial Sequence

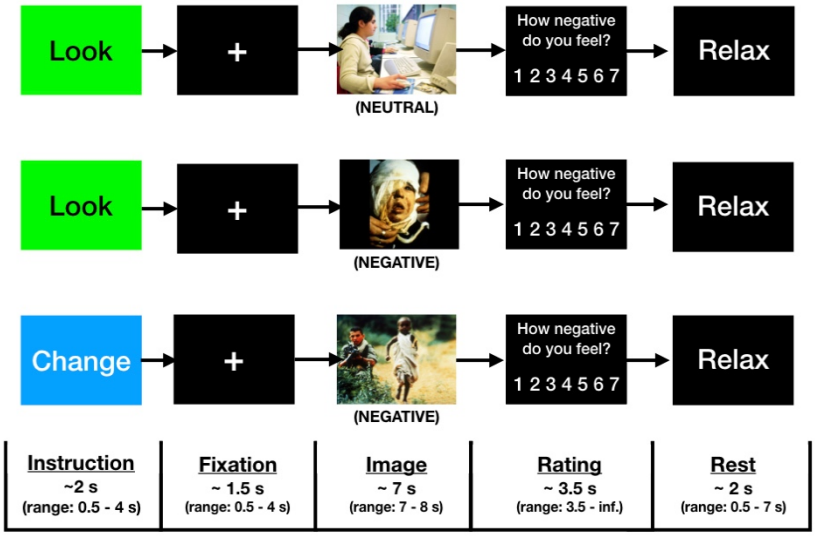

Figure 3 contains a schematic of the trial events which were nearly universal across the picture-based reappraisal experiments in the database. Each trial sequence began with an instruction to either "Look" or "Change" the following image. The instruction was followed by either a neutral or negative image. In some studies, a fixation cross appeared between the instruction and image events. Images following a "Look" instruction were roughly evenly divided between neutral and negative images. Images following a "Change" instruction were always negative. Following the image presentation, participants rated their experience of negative affect during image viewing on a numerical scale (Note: some studies had additional ratings such as arousal or positivity following the negative affect rating). After the rating/s, participants either saw another fixation cross or a cue to "relax" (or "*rest*") *before the start of the next trial. Most studies did not include a fixation after the instruction, and the specific instruction words, trial event timings, and ratings scales varied slightly depending upon experiment (see Table 1).*

Materials

The reappraisal database includes trial-level data from 148,764 unique trial events including the trial's condition (e.g., "Look", "Change", etc.), information about the stimulus that was presented, and the participant's rating of negative affect. Of these trials, 45,618, were negatively valenced "Look" trials, 41,602 were neutral "Look" trials, and 47,043 were "Change" trials. The remaining 14,501 trial events were derived from other auxiliary non-reappraisal portions of the experimental tasks.

Stimuli

The stimuli used in the cognitive reappraisal tasks consist of 453 images that were predominately sourced from the International affective picture system (IAPS) (Lang et al., 2008). For each experiment, specific negative and neutral emotional images were selected based on their normative valence ratings. The subset of IAPS images included in the reappraisal database consist of 156 neutral (*M = 5.24, SD = .69*) and 213 negative (*M = 2.71, SD = .78*) images. The smaller subset of non-IAPS images consists of 10 neutral (*M* = 5.15, *SD* = .09) and 74 negative *(M = 3.14, SD = .82*) images. The non-IAPS images originated from various sources including the Developmental Affective Photo System (DAPS) (Cordon et al., 2013), from internet searches, and from other affective science colleagues and collaborators. Depending upon the study, these items were selected to be more appropriate for samples including young children, or to augment the existing pool of stimuli to allow for more experimental conditions. These images were normed for valence by research staff either in the PI's or collaborator's laboratories using the same or similar rating scales as the IAPS[5].

Affective Ratings

Across the 25 separate experiments included in the analyses, negative affect ratings were collected on one of 6 rating scales (See Table 1). For each version of the scale, higher numbers were used to indicate more negative affect and lower numbers indicated less negative affect or neutral affect. Within the database, the raw affect ratings for each study were transformed into normalized units centered around each scale's midpoint. This allowed for all studies within the database to be analyzed according to the same rating scale.

[5] The small number of DAPS images were normed on a 5-point scale but were converted to the same scale as the other images using a linear transformation.

To normalize the rating scale within each study, each participant's ratings were transformed to scale-centered units of the sample's standard deviation from the center of the scale. This allowed all participant's scores to be expressed in terms of the units of deviation from the center of the scale used by the sample for that study. To do this, the scale's center was subtracted from the negative affect rating on each trial. For example, if the study used a 9-point scale, the scale center of 5 was subtracted from every rating within the given study. A square root of the mean deviation from center was then calculated from the squares of the subtracted values. The subtracted values were then divided by this square root of the mean value.

Therefore, the new transformed ratings were all centered at zero with a range between 2.57 and 3.94 points between the lowest and highest values, depending upon the study. Finally, using a linear transformation, the norm-centered ratings were then refitted back to a 7-point scale. This linear transformation was implemented in order to aid data visualization and the interpretability of results while not affecting the fidelity of the ratings. The linear transformed norm-centered ratings are referred to simply as negative affect ratings from here on.

Data Analytic Strategy

Three separate series of analyses were carried out, each tailored to address a specific research question pertaining to the database. Each section below describes the specific research question and the associated analyses. Across each of these analyses, we set the criterion for significance as the conventional alpha level of $p < 0.05$. Post hoc mean comparisons were conducted using Fisher's Least Significant Difference (LSD)

tests. Statistical analyses were conducted using MATLAB software (version R2019a) with the Statistics and Machine Learning Toolbox (version 8.3)[6].

The Main Effect of Trial Condition

Following the standard method of analysis common across a variety of published papers on cognitive reappraisal, a one-way repeated-measures analysis of variance (ANOVA) was conducted in order to test the overall robustness of the main effect of reappraisal on self-reported negative affect (Hajcak & Nieuwenhuis, 2006; McRae et al., 2008; Ochsner et al., 2002). The independent variable is Trial Condition with 3 levels ("Look Negative", "Look Neutral", and "Change"). The dependent variable was Negative Affect as measured by the self-reported negative affect ratings (see 2.3.2 Affective Ratings above). This analysis was carried out on the entirety of the database collapsing across all studies. Only participants with qualifying trials for all 3 trial conditions were included (N = 1187 from 25 studies). Data from the two studies that did not include all three conditions were not included in the analysis (see Section 2.2 above).

Following the ANOVA, we explored the robustness of the trial condition effects using a bootstrap power simulation. For the simulation, we used the Power-Sim Toolbox in MATLAB (Strong & Alvarez, 2019) to estimate the power of the trial condition effect. The toolbox utilizes a bootstrap resampling method such that within each simulation participants and individual trials are selected at random with replacement from the database of 1187 qualifying participants. For each simulation, trial numbers ranged from 4 to 36 trials per each condition in steps of 4 and participant sample sizes ranged from 20 to 500 in steps of 20. For each combination of trial number and sample size, we conducted 10,000 individual one-way repeated ANOVAs or 2,250,000 total tests

[6] Outputs of the ANOVAs were also validated using SPSS.

following the same structure as above. The toolbox only counts a test as significant when there is both a significant main effect of trial condition and in which there are significant post hoc differences of Look Negative > Look Neutral and Look Negative > Change.

Predicting Negative Affect at the Trial Level

In a second series of analyses, we examined a number of potential explanatory variables that predict negative affect rating on a given trial using multiple linear regression. Unlike the ANOVA analyses, rather than modeling trial condition categorically (“Look Negative”, “Change” and “Look Neutral”), here we only modeled trial instruction (“Look” vs. “Change”) as a categorical variable and entered normative valence (“Valence”) and normative arousal (“Arousal”) for the image stimulus on the trial as separate continuous variables. By not imposing a categorical structure for image valence (“Look Negative” vs. “Look Neutral”), we allowed the models more precision in determining the weights of the coefficients. Only participants with qualifying trials in either Look or Change trial conditions were included (N = 1388 from 27 studies).

For these analyses, we were particularly interested in examining the effects of time on task (“Time”), which we operationalized as the trial number. In addition, we examined the previous trial’s instruction (“Previous Look” vs. “Previous Change”) as well as the previous trial’s negative affect rating (“Previous Negative Affect”). For these models, the previous trial variables were included as covariates for the overall improvement of model fit. However, we did not test for interactions between current and previous trial variables as this would potentially overcomplicate the model and make interpretation difficult. Due to the nature of regression, significant interactions with previous trial variables would require multiple additional models in order to test each of the post hoc comparisons sequentially. Therefore, the interpretation of any interaction

effects is better suited to a more targeted ANOVA design in which multiple post-hoc comparisons can be made within a single model (see 2.4.3 Trial sequence effects).

In order to match the size of the scaling of the negative affect outcome variable, the normative valance and arousal scores of the image stimuli were adjusted from a 9 to a 7-point scale using a linear transformation[7]. Normative valence was also reverse coded such that higher valence equated to more negativity thereby matching the direction of the negative affect outcome variable. Time was coded on a 0:1 scale reflecting the proportion of total trial events completed out of the maximum number of trials occurring in the database (across all studies). All continuous explanatory variables (Valence, Arousal, and Time) were then mean-centered. Both previous and current trial conditions were dummy coded such that "Look" was the reference category and "Change" was the predictor.

Models were fitted using MATLAB's fitlme function using the restricted maximum likelihood estimation (REML) method[8]. Beginning with a basic regression equation of Negative Affect = 1(intercept) + Valence + Arousal + Change*Time, additional explanatory variables of Previous Negative Affect, Previous Change, and interactions between variables were entered into the model stepwise. Likelihood ratio tests were conducted at each stage to determine whether the additional variables significantly improved model fit while accounting for the loss in degrees of freedom. If a variable or interaction dropped below the threshold of significance after the inclusion of additional

[7] The transformation matches the size of the scaling but not the range of the two scales as the two poles of the normative valence scale represent positive and negative affect respectively, with neutral being approximately in the middle. By contrast, the negative affect outcome variable does not contain values for positive valence.

[8] Although maximum likelihood estimation is optimal for fixed effects only models, REML was selected in order to enable direct comparisons of likelihood ratio tests between the fixed and mixed effects models.

explanatory variables, the non-significant variable was dropped from the model. The models, variables tested, and results of the likelihood ratio tests for all linear fixed and mixed effects are reported in Appendix B.

Once a final fixed effects model was determined, we then conducted a series of linear mixed effects models maintaining the same fixed effects structure. Random slopes were calculated for participant nested within study and for image (i.e., the stimulus presented on a given trial). Following the same stepwise procedure above, random effects were entered into the model iteratively and likelihood ratio tests were performed at each stage. If a fixed effects explanatory variable dropped below the threshold of significance with the inclusion of the random effects, the non-significant variable was dropped from the model (See Appendix B for details).

Trial Sequence Effects

In a third series of analyses, we tested for a main effect of the preceding trial event (i.e., the condition of the previous trial), and whether the preceding trial condition interacts with the current trial condition. For this analysis we conducted a 3 (Current Trial: “Look Neutral” vs. “Look Negative” vs. “Change”) by 3 (Previous Trial: “Look Neutral” vs. “Look Negative” vs. “Change”) within-subjects repeated measures ANOVA. The dependent measure was again Negative Affect as measured by self-reported negative affect ratings (see 2.3.3 Affective Ratings above). The analysis was conducted on the entirety of the database collapsing across study. Only participants with a minimum of 4 qualifying trials for each of the 9 trial conditions were included in the analysis (N =638 in 17 studies).

Notably, we opted not to combine this analysis with the repeated measures ANOVA described in Section 2.4.1 above, even though both analyses test the main effect of current trial condition. By limiting this analysis to only participants with a

sufficient number of qualifying trials in order to assess the interaction effect and conduct post hoc tests, the sample size was greatly reduced in the present analysis. As the aim of the previously described ANOVA is to test the robustness of the main effect of reappraisal it is more appropriate to run the first test on the largest available sample.

Following the ANOVA, we explored the robustness of the previous trial condition effects using a series of bootstrap analyses. In the first series, following the same procedure as described in section 2.4.1 above we again used the Power-Sim Toolbox to estimate the power of the main effect of previous trial condition as well as the power of the interaction.

In a second series of bootstrap analyses, we used the trial condition means for each participant and conducted separate one-way ANOVAs for the same 25 separate sample sizes ranging from 20 to 500 in steps of 20. At each sample size, 10,000 separate tests were conducted, and participants were selected randomly from the pool of 638 possible for each test. Unlike the power simulation above, this method does not use within-sample replacement such that each sample was constructed of a unique set of participants. Likewise, the power simulation resamples trials with replacement such that trial-level means may reflect trial subsets or duplicate trials within the condition. Here, the participant's trial-level means were constructed from every available trial within each condition prior to conducting the simulation.

We then examined the patterns of post hoc comparisons for the previous trial condition and the interaction effects. Post hoc comparisons were made for the 3 previous trial conditions as well as the interaction (i.e., the 9 previous by current trial conditions) for every test. We then calculated the proportion of significant mean differences for each comparison out of the 10,000 tests conducted within each sample size. To filter out noise and provide a more conservative estimation of the effect, only

significant comparisons in which the mean differences followed in the same direction as the ANOVA were counted in the proportion.

Chapter Three: Results

Trial Condition

A repeated measures ANOVA was conducted on the entire database collapsing across study to compare the effect of trial condition on negative affect rating. The ANOVA indicated a significant main effect of trial condition on negative affect ($F(2, 2372) = 5562$, $p < 0.001$, $\eta p^2 = 0.82$). LSD post hoc tests indicated that Look Negative trials ($M = 4.74$, $SD = .99$) were rated as significantly more negative than Look Neutral ($M = 1.66$, $SD = 0.71$, $p < 0.001$, $d = 2.92$) and Change trials ($M = 3.87$, $SD = 1.16$, $p < 0.001$, $d = 0.99$). Affect ratings on Change trials were also significantly more negative than Look Neutral trials ($p < 0.001$, $d = 1.91$) (see Figure 4a).

Power Simulation

The results of a bootstrap power simulation suggest that the effect of trial condition is highly robust. The simulation which randomly samples from the trial-level data of the 1187 qualifying participants indicated that even within a smaller sample size of only 20 participants, a power of .8 could be achieved with as few as 4 trials per condition. Power approaches .99 by increasing the trial numbers to 32 per condition, or by increasing the sample size to 40 participants at 4 trials per condition (see Figure 4b).

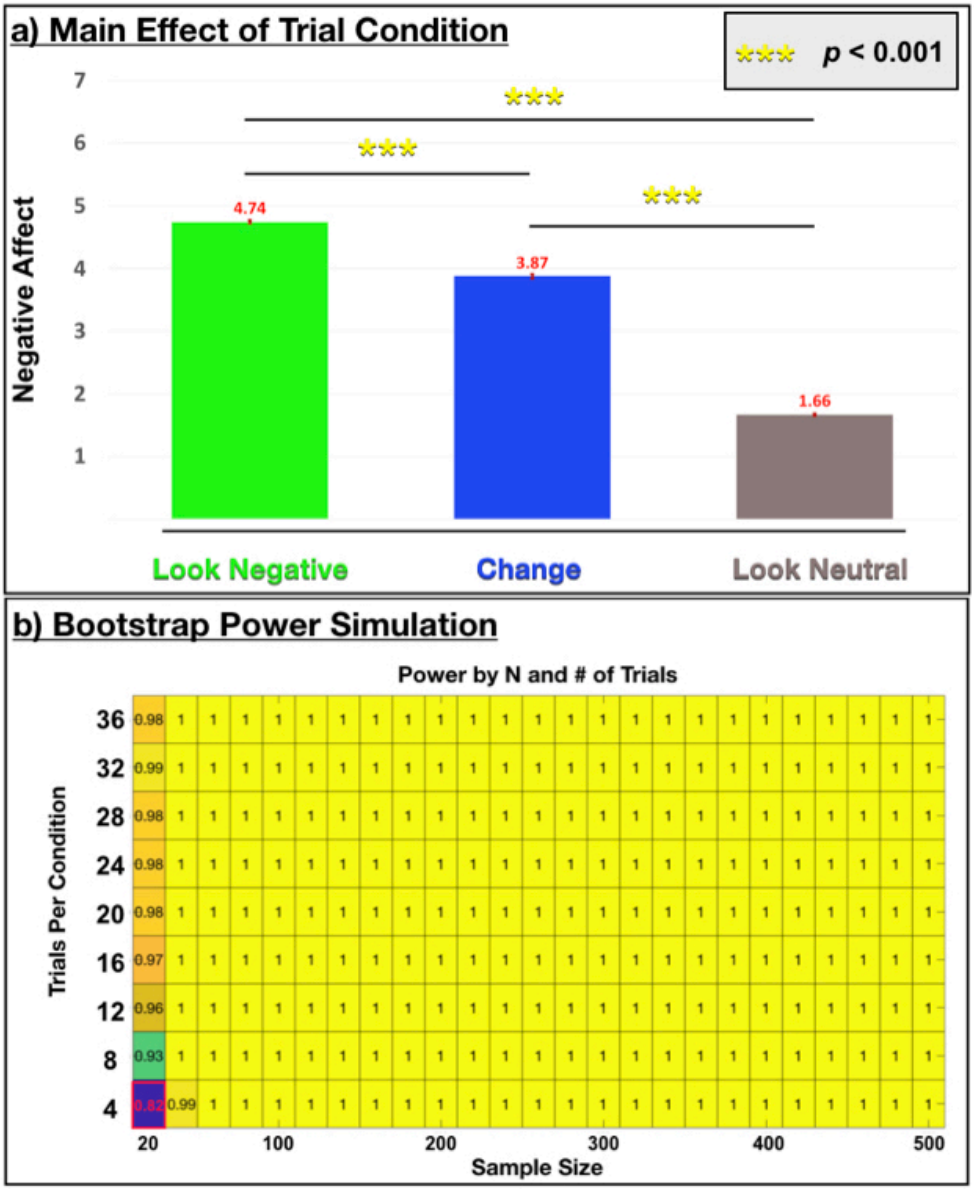

Figure 4 The Main Effect of Trial Condition: (a) Contains the Post Hoc trial condition comparisons for negative affect. Trial condition means are indicated in red text above the bar graphs (b) Contains the results of a bootstrap power simulation. Warmer colors indicate higher power.

The Effect of Time on Task

Fixed Effects Multiple Linear Regression Models

A series of fixed effects multiple linear regression models were conducted to predict reported negative affect based on several potential explanatory variables. Following several model iterations (See Appendix B), a significant regression equation was found for the final selected fixed effects model ($F(10, 105862) = 8749$, $p < .001$), with an adjusted R^2 of .452. Participants' predicted negative affect is equal to the Intercept ($\beta = 2.89$, $p < .001$) + Valence ($\beta = 1.21$, $p < .001$) + Arousal ($\beta = .34$, $p < .001$) – Change ($\beta = -.47$, $p < .001$) – Time ($\beta = -.18$, $p < .001$) – Previous Change ($\beta = -.17$, $p < .001$) + Previous Negative Affect ($\beta = .13$, $p < .001$) – Valence by Change ($\beta = -.17$, $p < .001$) – Arousal by Change ($\beta = -.31$, $p < .001$) – Valence by Time ($\beta = -.20$, $p < .001$) + Arousal by Time ($\beta = .32$, $p < .001$).

Linear Mixed Effects Models

After the fixed effects model was selected, random effects were iteratively added for individual participants nested within study and for image. Likelihood ratio tests were conducted at each stage to determine whether the addition of each random effect variable significantly improved model fit. In some instances, the random effects accounted for the variance explained by a fixed effects explanatory variable. Fixed effects variables that fell below significance with the addition of random effects to the model were subsequently dropped. The mixed effects models, variables tested, and results of the likelihood ratio tests are reported in Appendix B.

A significant regression equation was found for the final linear mixed effects model ($F(8, 105864) = 302.08$, $p < .001$), with an adjusted R^2 of .694. For the fixed effects parameters, predicted Negative Affect is equal to the Intercept ($\beta = 3.33$, $p < .001$) + Valence ($\beta = 1.23$, $p < .001$) + Arousal ($\beta = 0.16$, $p < .001$) – Change ($\beta = -0.65$, $p < .001$) – Time ($\beta = -0.15$, $p < .001$) – Previous Change ($\beta = -0.03$, $p < .01$) + Previous

Negative Affect (β = 0.04, p < .001) – Valence by Change (β = -0.20, p < .001) – Arousal by Change (β = -.09, p < .01). For the random effects of Participant nested within Study, the final model included the Intercept + Valence + Arousal + Change + Time + Previous Change + Previous Negative Affect. For the random effects of Image, the final model included the Intercept + Change. The random effects covariance parameters of the final model are reported in Appendix C. The slopes of Look and Change over Time are shown in Figure 5.

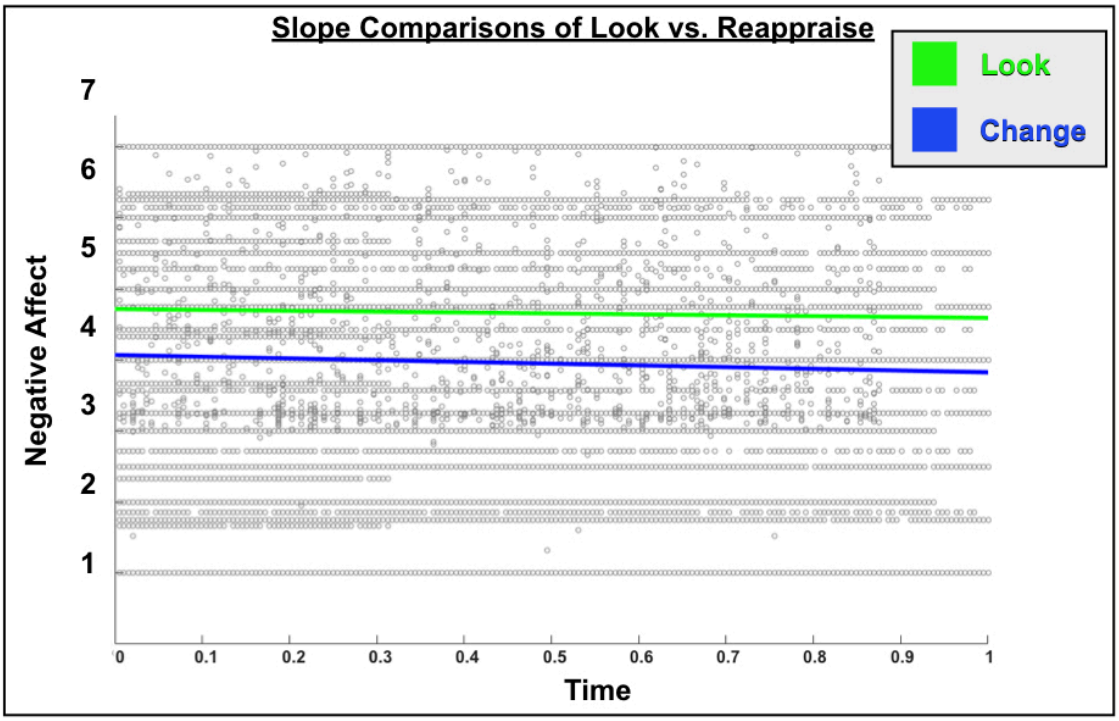

Figure 5 The Effect of Time on Task: Depicts the slopes of the Look and Change trial conditions over Time. Intercepts are adjusted for normative Valence and Arousal.

The Effect of Previous Trial Condition

Repeated Measures ANOVA

As seen in the trial condition effects in section 3.1, a repeated measures ANOVA indicated a significant main effect of current trial condition on negative affect (F(2, 1274) = 3413, $p < 0.001$, $\eta p^2 = 0.84$). Refer to section 3.1 for post hoc comparisons of current trial condition with a more complete data set. There was also a significant main effect of previous trial condition on negative affect (F(2, 1274) = 37, $p < 0.001$, $\eta p^2 = 0.05$). LSD post hoc comparisons indicated that negative affect ratings were significantly more negative when preceded by a Look Negative instruction on the previous trial (M = 3.47, SD = 1.67) as compared to a previous Change instruction (M = 3.39, SD = 1.7, $p <$ 0.001, $d = 0.07$) or a previous Look Neutral instruction (M = 3.36, SD = 1.67, $p < 0.001$, $d = 0.1$). Negative affect ratings were also significantly more negative when preceded by a Change instruction compared to a previous Look Neutral instruction ($p < 0.01$, $d = 0.06$).

The current by previous trial condition interaction was also significant (F(4, 2548) = 4.39, $p < .01$, $\eta p^2 = 0.007$). There were dissociable effects of the previous trial at different levels of the current trial condition. For the current Look Negative trial condition, LSD post hoc comparisons indicated no significant differences across the three previous trial conditions. However, there were significant differences across previous trial conditions for both the current Change trials and current Look Neutral trials.

For the current Change trial condition, LSD post hoc comparisons indicated that trials which were preceded by a Look Negative trial (M = 4.0, SD = 1.19) were more negative than when preceded by a Look Neutral trial (M = 3.82, SD = 1.16, $p < 0.001$, d = 0.33) or by a Change trial (M = 3.91, SD = 1.23, $p < 0.001$, d = 0.15). Current Change trials that were preceded by a Change trial were also more negative than when

preceded by a Look Neutral trial ($p < .01$, $d = 0.13$). For the current Look Neutral Condition, post hoc comparisons indicated that trials were more negative when preceded by a Look Negative trial ($M = 1.64$, $SD = .72$) compared with either a Change trial ($M = 1.55$, $SD = .66$, $p < 0.001$, $d = 0.23$) or a Look Neutral trial ($M = 1.53$, $SD = .67$, $p < 0.001$, $d = 0.24$). Post hoc comparisons for the interaction effect are shown in Figure 6a.

Power Simulations

The results of separate bootstrap power simulations suggest that the power to detect either a significant main effect of previous trial condition or a significant current by previous interaction effect requires large sample sizes and trial numbers in order to achieve a power of .80. For the first simulation, a significant main effect of previous trial condition had to be found in order for a test be considered a “success”. Assuming a relatively large number of trials per condition and a larger sample, the power estimate remains small. In order to achieve a power of .80, the simulation projects that 28 trials would be needed for each condition for a sample size of 280 participants or 24 trials per condition for a sample size of 300.

For the second simulation, a significant current by previous interaction had to be found in order for a test be considered a “success”. In this case, the simulation estimates that power was a bit stronger than for the main effect of previous condition. The power estimates for the current by previous interaction approaches .80 (.79) with 36 trials per condition in a sample size of 100 participants and achieves .80 with only 24 trials per condition in a sample size of 120 participants. The results of both power simulations can be seen in Figure 7.

Bootstrap Resampling Simulations

The bootstrap resampling simulations largely supported the findings from the ANOVA conducted on full sample of 638 participants (See 3.3.1).[9] The proportion of significant main effects for previous trial condition and for the current by previous interaction both increased linearly in step with increasing sample sizes. For the previous condition main effect, the proportion of significant tests out of 10,000 was .78 in sample sizes of 80 participants. The proportion of significant previous by trial interaction effects also increased linearly with sample size but not as steeply. This proportion only reached .80 in larger a sample size of 440 participants. Figure 6 (b through d) displays the proportion of significant tests found for every sample size that was tested at each level of the current trial condition.

The patterns of post hoc comparisons lend further support of the results of the ANOVA. For both the main effect of previous trial condition and the current by previous interaction, the patterns of simulated post hoc mean comparisons closely mirrored the overall effects seen in the ANOVA. Significant mean comparisons in the ANOVA were similarly reflected by linear increases in the proportion of significant tests with increasing sample sizes. Likewise, nonsignificant mean comparisons exhibited only a small or flat and non-increasing proportion of significant tests with increasing sample sizes (see Figure 6).

[9] The simulation also indicated highly significant main effects for current trial in all tests (all p's < .0001). See section 3.1 for current trial condition main effects analysis conducted on the full sample.

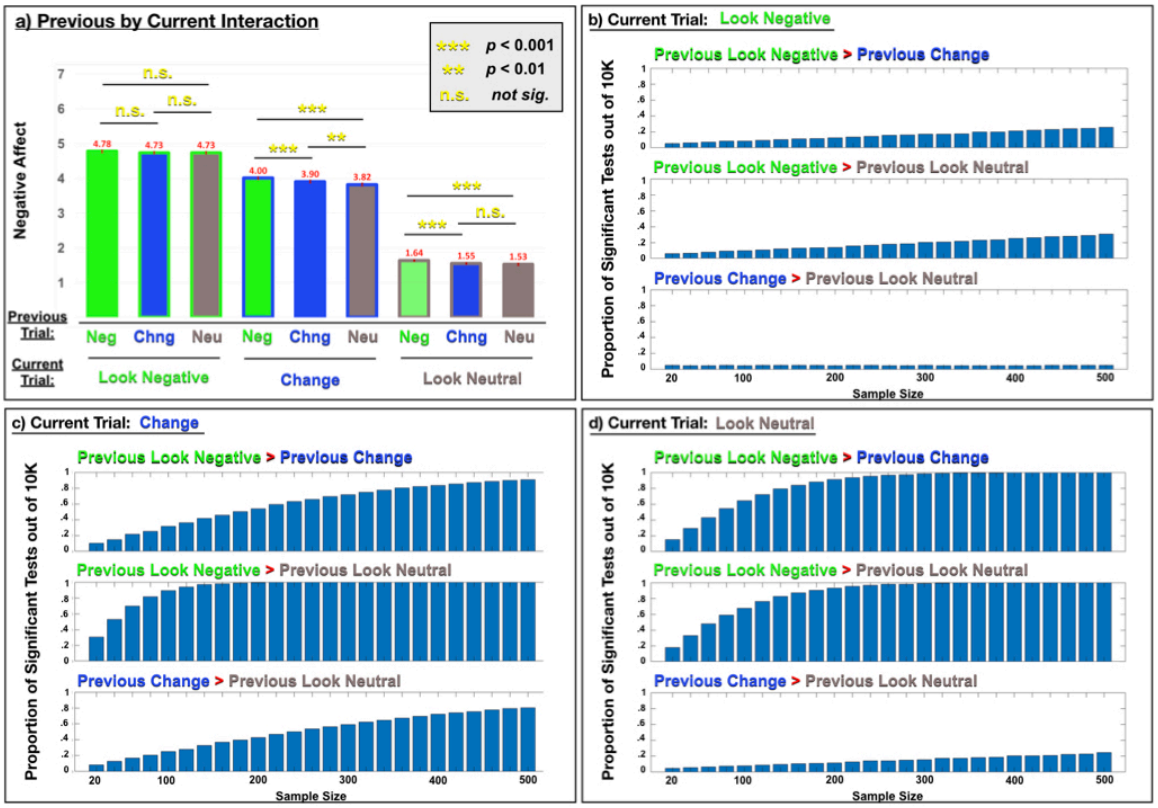

Figure 6 The Effect of Previous Trial Condition: (a) Post Hoc comparisons for previous trial condition at each level of current trial condition. Trial condition means are indicated in red text above the bar graphs. (b) Proportion of significant tests out of 10,000 found in the bootstrap simulation by sample size for each previous trial condition at the current Look Negative trial condition. (c) current Change trial condition. (d) current Look Neutral trial condition.

Chapter Four: Discussion

Summary of Findings

The results point strongly towards a robust effect of cognitive reappraisal. Specifically, participants reported lower levels of negative affect for trials in which they were instructed to change how they feel about a negative image using reappraisal (Change) compared to when instructed to respond naturally (Look Negative). The average difference in reported negative affect in these two conditions approached nearly 1-point (0.87) on a 7-point scale. Compared with responding naturally, reappraisal appears to reduce negative affect by about 18% on average[10]. The reappraisal effect holds even after accounting for the normative Valence and Arousal of the image stimuli. Likewise, both Valence and Arousal had less influence over negative affect ratings during reappraisal as compared to the Look instructed trials.

Of particular interest was whether emotional reactivity and regulation change as a function of Time. The observed effect of Time was most consistent with an emotion habituation explanation (see Figure 1). Over the course of trials, participant's negative affect ratings tended to decrease. However, this effect does not appear to be particular to any trial condition as there were no differential decreases in ratings over Time between the Look and Change instructed trial conditions.

We also found evidence of a negative valence carryover effect (see Figure 2). Higher negative affect ratings on the immediately preceding trial were predictive of more

[10] The mean negative affect rating on Change trials was 3.87, which is 18.35% lower than the mean negative affect rating of 4.74 on the Look Negative trials.

negative ratings on the current trial. The carryover effect was most impactful for current Change and Look Neutral trials. However, the carryover effect also appears to be at least partially mitigated by the participant's reappraisal efforts. Negative affect ratings on the current trial were slightly lower overall when participants had just reappraised a negative image on the immediately preceding trial as compared to when instructed to Look.

The power and bootstrap resampling analyses lent further support for these interpretations. While these analyses affirmed the overall robustness of the effects of reappraisal, they also suggest that detecting the more subtle effects of the previous trial condition likely require very large samples and numerous trials per condition.

Implications

Our results replicate previous findings of the effects of a reappraisal manipulation on negative affect and established these effects as incredibly robust. The effect of reappraisal was highly significant and had a large effects size in spite of numerous study-level idiosyncratic characteristics such as the number and duration of trials, specific instructions and study manipulations, testing environments (e.g., laboratory, fMRI, online), auxiliary measures taken (e.g., psychophysiological etc.), other secondary task parameters, and period of time in which the study was conducted.

These results not only support previous findings in the reappraisal literature but have implications for future research. Given the large sample and number of studies that were aggregated into these analyses, these results should help to establish a precedent for future studies. By leveraging the database, researchers can make reasonable projections for the number of participants needed to achieve an expected power and set expectations for the magnitude of effect sizes when designing experiments. For example, future studies might set out to test variations on different reappraisal tactics or

task instructions and can compare their results with the effect sizes seen here. Furthermore, researchers might consult these results when comparing the effectiveness of reappraisal with other emotion regulation strategies.

The sheer size of the database in terms of the number of studies and participants provided an opportunity to test competing hypothesis regarding the influence of Time on emotional reactivity and emotion regulation. With respect to emotional reactivity, our results were more indicative of an emotion habituation than a sensitization explanation. Notably, the habituation effect of Time did not appear to be impacted by trial instruction as Change instructed trials exhibited similar decreases in negative affect ratings over the course of trials. Therefore, habituation did not have an additive or interactive effect on reappraisal. Moreover, these findings also do not support an emotion regulation fatigue effect. Despite the decreasing negative affect on Change trials over Time, the results do not necessarily support a practice effect of regulation either. Had a practice effect been evident, there should be greater relative decreases in negative affect over Time for the Change trials compared with either a relatively flat or increasing negative affect over Time for Look instructed trials.

Habituation effects have been more commonly reported in experiments utilizing repetitions of the same stimulus (Denny et al., 2015; Erk et al., 2010; Lang et al., 1993; MacNamara et al., 2011; Silvers, Shu, et al., 2015). However, we found evidence for habituation that generalized across the negative valence category of images that consisted of differing themes and contents. This may have implications in clinical settings such as seen with repeated exposure therapy for phobias and post-traumatic stress disorder (PTSD) (Benito & Walther, 2015; Rauch et al., 2012).

During exposure therapy, fearful emotions or traumatic memories are instantiated while patients actively engage with the evoking stimulus. The effectiveness of this

approach is attributed to habituation, as the patients experience a diminishing negative response while confronted with the stimulus and not given the option to reduce negative affect through avoidance etc. The current findings bring into question the need for specificity of the evoking stimulus or whether the beneficial habituation effects might be achieved within a wider range of stimuli that differ in content but are matched in their dimensions of arousal and valence.

Interestingly, while these data do not conform with a reappraisal practice effect, contrary to what the ego depletion and strength models of self-control would predict, reappraisal does not appear to be a muscle that is easily fatigued (Baumeister et al., 2007; Baumeister & Heatherton, 1996). Yet, a number of previous studies have found that reappraisal was associated with cognitive costs on secondary memory, cognitive control, and attention tasks conducted either subsequently or simultaneous with reappraisal (Ortner et al., 2013; Ortner & de Koning, 2013; Ortner et al., 2016). However, these studies did not test for diminishing effects of reappraisal over Time. Nor do they suggest that the secondary cognitive task had a negative impact on reappraisal success. Likewise, concurrent or immediately proximal secondary tasks may add additional cognitive load as participants are required to switch between them. It may also be that the nature of the temporal dynamics in the present study allows for greater recovery between trials than when there is an additional cognitive demand.

The absence of a fatigue effect in our results has some important implications for emotion regulation outside of the lab. For someone who is having a bad day, reappraising may be just as effective even after several events which elicit reappraisal attempts. This also suggests that people might benefit from opting for reappraisal, despite its perceived difficulty as compared with other regulation strategies (Troy et al., 2018). At least on the timescale we studied, reappraisal may not sap cognitive resources

significantly, which might impact people's decisions about using reappraisal (Milyavsky et al., 2019).

In addition to the effects of Time over the course of the study, there was a more localized influence of the immediately preceding trial's condition and affective rating. The observed negative affect carryover effect was consistent with previous studies that have found that residual affect (both negative and positive) from a previous stimulus can bias the affective responding to the current stimulus (Czekóová et al., 2015; Fujimura et al., 2013; Palumbo et al., 2017).

The current findings have implications for appraisal theories which postulate that emotions follow from one's subjective evaluation of an external situation (Roseman & Smith, 2001). The process model of elicitation contends that appraisals derive from both automatic associative and more deliberative cognitive processes, which can sometimes come into conflict with one another (Smith & Kirby, 2000, 2001). One possible explanation for the observed carryover is that the ongoing experience of residual negative affect may be informing and biasing participants' deliberative cognitive appraisals of the current stimulus. Hence, more negative interpretations might be ascribed to ambiguous or otherwise innocuous image contents thus altering its perceived meaning.

According to the affect-as-information hypothesis, emotions provide a source of information about situational contexts (Clore et al., 2001; Clore & Huntsinger, 2007; Storbeck & Clore, 2008). This seems ecologically tractable in that people tend to rely on incidental affect informed by prior experiences to predict immediately proximal contexts (Wilson-Mendenhall, 2017; Wormwood et al., 2019). Therefore, participants might refer to their present state of affect that resulted from the previous trial in order to establish the context in which the present target image is appraised.

Another possible explanation may be that the cognitive interpretation of the previous image's valence, rather than the subjective experience of negative affect, is responsible for the carryover effect. It could be that, rather than an incidental effect of emotion, the previous cognition (negative appraisal) of a previous image itself establishes the context in which the current image is appraised. For example, studies that have employed facial stimuli have found that neutral facial expressions are interpreted as expressing an emotion that is congruent with adjacent contextual information (Aviezer et al., 2008; Barrett et al., 2011; Mullennix et al., 2019). Likewise, the interpretation of surprise, a high-arousal, ambivalent expression, is influenced by immediate context (Neta et al., 2011).

In film theory, this is known as the Kuleshov effect; a montage technique in which disconnected shots are stitched together to create a coherent connotation (Mullennix et al., 2019; Pudovkin & Montagu, 1958). This technique is frequently employed in order to cause the audience to draw inferences about the meaning of a scene or a character's internal state (Li, 2014). In the original Kuleshov experiment, participant's judged an actor's neutral expression as displaying either hunger, sorrow, or happiness when juxtaposed with a scene of a bowl of soup, a woman in a coffin, or a child playing respectively (Pudovkin & Montagu, 1958).

Extending upon previous research, a major contribution of the present work is the novel finding that reappraisal partially mitigated the extent to which negative affect was carried over. This has important implications for the field emotion regulation, both in and out of the lab. To the extent that the appraisal of the current situation is biased by the residual negativity experienced in a previous situation, this can lead to suboptimal behaviors. For example, becoming short-tempered with your family because you had a difficult day at work etc. Our results indicate that by regulating negative emotions in the

present situation, the lingering effects of reappraisal might help to reduce the undue influence of said negative emotional situations on future situations. Likewise, by minimizing the bias of previous negative affect on current appraisals, one may have greater success in appropriately calibrating their behavior to the current situation.

The specific mechanism by which reappraisal mitigates the carryover effect is unclear. It could be that reappraisal itself diminishes the amount of residual negative affect carried over to a subsequent trial, irrespective of the degree of negative affect on the current trial. For this explanation to be accurate would mean that the reappraisal process either continued after the participants made their affective ratings or promoted a larger post-rating recovery compared to when reappraisal did not occur. While this explanation cannot be entirely ruled out, the data herein were not collected with these temporal dynamics in mind.

Another perhaps more plausible explanation is that the lesser degree of residual affect carried over following reappraisal is simply due to the fact that reappraisal trials, on average, are rated as less negative. To the extent that reappraisal successfully reduces the participants' experience of negative affect, there is less negativity to carry over to the subsequent trial. A third possibility is that reappraisal somehow influences the subsequent appraisal. If negative valence from the preceding trial biases the interpretation and appraisal on the current trial, perhaps this bias is partially influenced by being in a reappraisal mindset.

In terms of emotion regulation research, this finding raises some other important questions as to the mechanisms of reappraisal. First, it does not appear that reappraisal becomes more effective or more efficient on repeated subsequent regulation attempts (beyond the habituation-related decrease also observed in the look condition).

Therefore, there was no evidence for a cognitive facilitation effect of reappraisal in the present study.

We had anticipated that by engaging in an overlapping cognitive process, reappraisal might metaphorically “prime the pump” for the subsequent reappraisal attempt. Similar process overlap effects have been found in other cognitive domains such as with transfer-appropriate-processing, in which memory is enhanced to the extent that the cognitive processes during retrieval match those during encoding (Roediger et al., 1989). One study found that reappraisal success was facilitated when there was a greater degree of overlap in neural processes involved with both emotion generation and regulation (Otto et al., 2014). In another study, lower negative affect ratings across trial conditions were predicted by the extent to which brain network activity resembled the patterns of activity seen during reappraisal (Silvers, Wager, et al., 2015). While we did not find evidence for facilitation in the present examination, this question may be better addressed in future neuroimaging work that aims to test whether reappraisal success is predicted by the extent to which there is an overlap of network activity with an immediately preceding reappraisal trial.

Similarly, the results do not support a cognitive process carryover effect. Despite the mitigating effects of reappraisal on negative affect carryover, the observed effect does not appear to be accounted for by participants reappraising in spite of current instructions. Had there been a process carryover, then current negative affect ratings should have been lower following a previous Change trial as compared with either previous Look conditions. One positive outcome of this is that participants on the whole appeared to be compliant with trial instructions. However, these results also suggest that researchers may wish to be mindful of the structure of trials sequences in their studies and of the potential corrupting influence from the previous trial’s affect.

Limitations and Future Directions

While the reappraisal effect appears quite robust, there are some limitations that should be noted. First, with respect to the power simulation, the analyses only included participants that were compliant with the task's instructions and offered quality responses. Therefore, the projected number of participants needed to achieve a given power is likely to be underestimated within a noisier sample. Researchers need to account for a proportion of non-compliant participants within study samples when leveraging the database to derive power estimates for future experiments.

Secondly, these data were predominantly collected within the laboratory of a single PI. Hence, despite the variations across each of the studies included in the analyses, there are likely to be many similarities and overlapping features in terms of overall study design such as the style of instruction, etc. It is possible that had the database included more data collected from other laboratories, there would be a larger degree of between-study variability, which could affect the magnitude of the reappraisal effect observed or the estimation of power. However, as these studies already vary considerably, it is likely that the influence of Principal Investigator may be relatively marginal as compared with the other task-level parameters. Likewise, as our intention is to make the database an open-sourced repository, we hope that researchers from other labs will contribute to it in the future.

While the volume and variation of data within the database was instrumental for examining emotional reactivity and emotion regulation effects over Time, the individual experiments in the database were not designed with this test in mind. It is possible that an emotion regulation practice effect was indeed present but obscured by the observed habituation effect. Future studies should explore these questions more directly. For example, trial sequences could be designed such that the normative negative valence of

the images increases linearly with trial number. By progressively increasing stimulus negativity, a habituation effect might be partially offset. If an emotion regulation practice effect occurred, then the regulated Change instructed trials should have a more negative slope than the unregulated Look instructed trials, even if both slopes were slightly positive. At present, while practice effects cannot be fully ruled out, these data are most consistent with emotional habituation.

The multilevel model also included random effects for the intercept and trial instruction for each image stimulus. Although the means of the normative Valance and Arousal for the images were already included as fixed effects in the model, the model fit was significantly improved with the inclusion of the image-level random effects. This suggests that some of the variation in negative affect rating was explained by idiosyncratic image properties not captured by the normative Valence and Arousal information. Moreover, there was variation in negative affect across images for Change instructed trials. This suggests that the images differed in their overall reappraisability, even while controlling for Valence and Arousal.

While out of scope of our current aims, a potentially promising avenue for future research will be to explore what properties of the image stimuli might be predictive of reappraisal success. Just as previous work has established normative databases for affective images, one outcome of the database is to establish reappraisal norms for these images. Following from recent work on qualifying reappraisal affordances, a future extension of the present work will be to determine what properties of images, including low-level visual characteristics, and higher-level emotional and cognitive characters, are predictive of greater reappraisal success or variance in success (Suri et al., 2018; Uusberg et al., 2019). This line of questioning may begin to tell us more about the

appraisal process itself and under what circumstances may reappraisal be more or less effective.

Another promising direction for future research might be to gather more information about the appraisals themselves. In a series of recent studies, participants transcribed their thoughts while either looking at or reappraising images (Nook et al., 2017, 2020). In these studies, certain linguistic patterns were found to be predictive of reappraisal success. A similar approach might be employed to gain insight as to whether and how residual negative affect influences the current appraisal as well as how reappraisal might intervene on this process. An examination of the linguistic patterns captured on Look Neutral trials that follow Look Negative and Change trials might be informative as to the previous trial's influence on the current appraisal.

Unlike the robust effects seen for reappraisal on the current trial, the results of the power and bootstrap simulations suggest that the effects of previous trial condition and the interaction with the current trial condition are quite subtle, requiring much larger sample sizes to detect. This may be a limiting factor for future research; particularly for resource-intensive approaches such as neuroimaging. Likewise, the examination of these effects requires trial-level resolution and thereby does not lend itself to most meta-analytic approaches in which the data have already been aggregated and summarized.

Our ability to probe the previous trial effect was also somewhat restricted by the limited number of participants with qualifying trials per each condition. The studies in the database were not designed to examine these trial-by-trial effects. Consequently, more than one half of participants did not have a sufficient number of qualifying trials per trial condition to be included in the analyses. It is possible that future studies designed specifically to examine the effects of the previous trial might be able to achieve sufficient

power with smaller samples by optimizing trial sequences and ensuring a larger number of qualifying trials for the analyses.

Conclusions

The past two decades have seen continued growth and interest in reappraisal research and applications for mental health interventions (Gross, 1998b; John & Gross, 2004; McRae, Ciesielski, et al., 2012; Ochsner et al., 2002). The present investigation compliments this growing body of research by examining reappraisal within a very large sample. While the main effects of reappraisal itself may not be very novel or surprising, this approach affords a window into much more subtle processes and points towards mechanisms not easily assayed within a single experiment or with smaller sample sizes.

This research also raises important questions as to the mechanisms of both the negative valence carryover effect, and reappraisal processes. The field of emotion regulation and appraisal theory more broadly would benefit from a better understanding of the mechanisms by which reappraisal influences this carryover process. Future investigations should leverage paradigms that are specifically tailored to address these questions.

Aside from these novel findings and the questions identified for future research, a major outcome of the present work was to establish a database of negative affect ratings during the reappraisal task. It is our hope not only that this database will be leveraged by other researchers, but that it will become open to outside contributions and continue to grow. The database is a potentially invaluable tool for running simulations, generating and evaluating other novel questions not addressed here, and potentially examining the role of individual differences.

References

Abravanel, B. T., & Sinha, R. (2015). Emotion dysregulation mediates the relationship between lifetime cumulative adversity and depressive symptomatology. *Journal of Psychiatric Research*, *61*, 89–96. https://doi.org/10.1016/j.jpsychires.2014.11.012

Aldao, A., Jazaieri, H., Goldin, P. R., & Gross, J. J. (2014). Adaptive and maladaptive emotion regulation strategies: Interactive effects during CBT for social anxiety disorder. *Journal of Anxiety Disorders*, *28*(4), 382–389. https://doi.org/10.1016/j.janxdis.2014.03.005

Appleton, A. A., Buka, S. L., Loucks, E. B., Gilman, S. E., & Kubzansky, L. D. (2013). Divergent associations of adaptive and maladaptive emotion regulation strategies with inflammation. *Health Psychology*, *32*(7), 748–756. https://doi.org/10.1037/a0030068

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*(7), 724–732.

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, *20*(5), 286–290.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*(1), 1–15. https://doi.org/10.1207/s15327965pli0701_1

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351–355. https://doi.org/10.1111/j.1467-8721.2007.00534.x

Benito, K. G., & Walther, M. (2015). Therapeutic process during exposure: Habituation model. *Journal of Obsessive-Compulsive and Related Disorders*, *6*, 147–157. https://doi.org/10.1016/j.jocrd.2015.01.006

Berking, M., Margraf, M., Ebert, D., Wupperman, P., Hofmann, S. G., & Junghanns, K. (2011). Deficits in emotion-regulation skills predict alcohol use during and after cognitive–behavioral therapy for alcohol dependence. *Journal of Consulting and Clinical Psychology*, *79*(3), 307–318. https://doi.org/10.1037/a0023421

Boehme, S., Biehl, S. C., & Mühlberger, A. (2019). Effects of Differential Strategies of Emotion Regulation. *Brain Sciences*, *9*(9), 225. https://doi.org/10.3390/brainsci9090225

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

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. https://doi.org/10.1037/1528-3542.1.3.276

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1996). Picture media and emotion: Effects of a sustained affective context. *Psychophysiology*, *33*(6), 662–670. https://doi.org/10.1111/j.1469-8986.1996.tb02362.x

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. https://doi.org/10.1037/0735-7044.107.6.970

Braunstein, L. M., Gross, J. J., & Ochsner, K. N. (2017). Explicit and implicit emotion regulation: A multi-level framework. *Social Cognitive and Affective Neuroscience*, *12*(10), 1545–1557. https://doi.org/10.1093/scan/nsx096

Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: rapid spatial orienting toward positive emotional stimuli. *Psychological Science*, 19(4), 362-370. https://journals.sagepub.com/doi/10.1111/j.1467-9280.2008.02094.x

Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., Weber, J., & Ochsner, K. N. (2014). Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. *Cerebral Cortex*, *24*(11), 2981–2990. https://doi.org/10.1093/cercor/bht154

Chechko, N., Kellermann, T., Schneider, F., & Habel, U. (2014). Conflict adaptation in emotional task underlies the amplification of target. *Emotion*, *14*(2), 321–330. https://doi.org/10.1037/a0035208

Clore, G. L., Gasper, K., & Garvin, E. (2001). Affect as information. *Handbook of Affect and Social Cognition*, 121–144.

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, *11*(9), 393–399.

Cludius, B., Mennin, D., & Ehring, T. (2020). Emotion regulation as a transdiagnostic process. *Emotion*, *20*(1), 37–42. https://doi.org/10.1037/emo0000646

Codispoti, M., Ferrari, V., & Bradley, M. M. (2006). Repetitive picture processing: Autonomic and cortical correlates. *Brain research*, *1068*(1), 213–220. https://doi.org/10.1016/j.brainres.2005.11.009

Cordon, I. M., Melinder, A. M. D., Goodman, G. S., & Edelstein, R. S. (2013). Children's and adults' memory for emotional pictures: Examining age-related patterns using the Developmental Affective Photo System. *Journal of Experimental Child Psychology*, *114*(2), 339–356. https://doi.org/10.1016/j.jecp.2012.08.004

Cosmides, L., & Tooby, J. (2000). *Evolutionary psychology and the emotions* (M. Lewis & J. M. Haviland-Jones, Eds.; p. 720). Guilford Publications.

http://books.google.com/books?id=SQ8F7zdhORwC&dq=isbn:1572305290&hl=&cd=1&source=gbs_api

Czekóová, K., Shaw, D. J., Janoušová, E., & Urbánek, T. (2015). It's all in the past: Temporal-context effects modulate subjective evaluations of emotional visual stimuli, regardless of presentation sequence. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00367

D'Avanzato, C., Joormann, J., Siemer, M., & Gotlib, I. H. (2013). Emotion regulation in depression and anxiety: Examining diagnostic specificity and stability of strategy use. *Cognitive Therapy and Research*, *37*(5), 968–980. https://doi.org/10.1007/s10608-013-9537-0

Denny, B. T., Inhoff, M. C., Zerubavel, N., Davachi, L., & Ochsner, K. N. (2015). Getting over it: Long-lasting effects of emotion regulation on amygdala response. *Psychological Science*, *26*(9), 1377–1388. https://doi.org/10.1177/0956797615578863

Denny, B. T., & Ochsner, K. N. (2014). Behavioral effects of longitudinal training in cognitive reappraisal. *Emotion*, *14*(2), 425–433. https://doi.org/10.1037/a0035276

English, T., John, O. P., Srivastava, S., & Gross, J. J. (2012). Emotion regulation and peer-rated social functioning: A 4-year longitudinal study. *Journal of Research in Personality*, *46*(6), 780–784. https://doi.org/10.1016/j.jrp.2012.09.006

Erk, S., Mikschl, A., Stier, S., Ciaramidaro, A., Gapp, V., Weber, B., & Walter, H. (2010). Acute and sustained effects of cognitive emotion regulation in major depression. *Journal of Neuroscience*, *30*(47), 15726–15734. https://doi.org/10.1523/JNEUROSCI.1856-10.2010

Flaisch, T., Junghöfer, M., Bradley, M. M., Schupp, H. T., & Lang, P. J. (2008). Rapid picture processing: Affective primes and targets. *Psychophysiology*, *45*(1), 1–10. https://doi.org/10.1111/j.1469-8986.2007.00600.x

Flaisch, T., Stockburger, J., & Schupp, H. T. (2008). Affective prime and target picture processing: An ERP analysis of early and late interference effects. *Brain Topography*, *20*(4), 183–191. https://doi.org/10.1007/s10548-008-0045-6

Ford, B. Q., & Troy, A. S. (2019). Reappraisal reconsidered: A closer look at the costs of an acclaimed emotion-regulation strategy. *Current Directions in Psychological Science*, *28*(2), 195–203. https://doi.org/10.1177/0963721419827526

Foti, D., & Hajcak, G. (2008). Deconstructing reappraisal: Descriptions preceding arousing pictures modulate the subsequent neural response. *Journal of Cognitive Neuroscience*, *20*(6), 977–988. https://doi.org/10.1162/jocn.2008.20066

Fox, H. C., Axelrod, S. R., Paliwal, P., Sleeper, J., & Sinha, R. (2007). Difficulties in emotion regulation and impulse control during cocaine abstinence. *Drug and Alcohol Dependence*, *89*(2), 298–301. https://doi.org/10.1016/j.drugalcdep.2006.12.026

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, *23*(2), 107–131. https://doi.org/10.1177/1088868318762183

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349–358.

Fujimura, T., Katahira, K., & Okanoya, K. (2013). Contextual modulation of physiological and psychological responses triggered by emotional stimuli. *Frontiers in Psychology*, *4, 212*. https://doi.org/10.3389/fpsyg.2013.00212

Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*(6), 577–586. https://doi.org/10.1016/j.biopsych.2007.05.031

Grillon, C., Quispe-Escudero, D., Mathur, A., & Ernst, M. (2015). Mental fatigue impairs emotion regulation. *Emotion*, *15*(3), 383–389. https://doi.org/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. https://doi.org/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*. https://journals-sagepub-com.du.idm.oclc.org/doi/10.1037/1089-2680.2.3.271

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, *26*(1), 1–26. https://doi.org/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. https://doi.org/10.1037/0022-3514.85.2.348

Gross, J. J., & Muñoz, R. F. (1995). Emotion regulation and mental health. *Clinical Psychology: Science and Practice*, *2*(2), 151–164. https://doi.org/10.1111/j.1468-2850.1995.tb00036.x

Gross, J. J., & Thompson, R. A. (2007). Emotion Regulation: Conceptual Foundations. *Handbook of Emotion Regulation,* 3–24. The Guilford Press.

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, *136*(4), 495–525. https://doi.org/10.1037/a0019486

Hajcak, G., & Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 291–297. https://doi.org/10.3758/CABN.6.4.291

Harris, J. D. (1943). Habituatory response decrement in the intact organism. *Psychological Bulletin*, *40*(6), 385–422. https://doi.org/10.1037/h0053918

Hofmann, S. G., Sawyer, A. T., Fang, A., & Asnaani, A. (2012). Emotion dysregulation model of mood and anxiety disorders. *Depression and anxiety*, *29*(5), 409–416. https://doi.org/10.1002/da.21888

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180. https://doi.org/10.1016/j.tics.2012.01.006

Holland, A. C., & Kensinger, E. A. (2010). Emotion and autobiographical memory. *Physics of Life Reviews*, *7*(1), 88–131. https://doi.org/10.1016/j.plrev.2010.01.006

Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality*, *52*, 29–36. https://doi.org/10.1016/j.jrp.2014.06.005

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. https://doi.org/10.1111/1469-8986.3740515

Jeong, H. J., & Cho, Y. S. (2020). The effects of induced and trait anxiety on the sequential modulation of emotional conflict. *Psychological research*. https://doi.org/10.1007/s00426-020-01289-1

John, O. P., & Gross, J. J. (2004). Healthy and unhealthy emotion regulation: Personality processes, individual differences, and life span development. *Journal of personality*, *72*(6), 1301–1334. https://doi.org/10.1111/j.1467-6494.2004.00298.x

Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition & emotion*, *13*(5), 467–480. https://doi.org/10.1080/026999399379140

Kiecolt-Glaser, J. K., McGuire, L., Robles, T. F., & Glaser, R. (2002). Emotions, morbidity, and mortality: new perspectives from psychoneuroimmunology. *Annual review of psychology*, *53*(1), 83–107. https://doi.org/10.1146/annurev.psych.53.100901.135217

Kobylińska, D., & Kusev, P. (2019). Flexible emotion regulation: how situational demands and individual differences influence the effectiveness of regulatory strategies. *Frontiers in Psychology*, *10, 72*. https://doi.org/10.3389/fpsyg.2019.00072

Koole, S. L., & Rothermund, K. (2011). “I feel better but I don’t know why”: The psychology of implicit emotion regulation. *Cognition and emotion*, *25*(3), 389–399. https://doi.org/10.1080/02699931.2010.550505

Koukounas, E., & Over, R. (2000). Changes in the magnitude of the eyeblink startle response during habituation of sexual arousal. *Behaviour research and therapy*, *38*(6), 573–584. https://doi.org/10.1016/S0005-7967(99)00075-3

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: A review of repetition priming in visual search tasks. *Attention, perception, & psychophysics*, *72*(1), 5–18. https://doi.org/10.3758/APP.72.1.5

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior research methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

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

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Larsen, J. T., & Norris, J. I. (2009). A facial electromyographic investigation of affective contrast. *Psychophysiology*, *46*(4), 831–842. https://doi.org/10.1111/j.1469-8986.2009.00820.x

Lazarus, R. S., & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. *The journal of abnormal and social psychology*, *69*(2), 195–205. https://doi.org/10.1037/h0044635

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition & emotion*, *13*(5), 481–504. https://doi.org/10.1080/026999399379159

Li, X. (2014). *Research on the montage technique in the film and television*. In International Conference on Education, Language, Art and Intercultural Communication (ICELAIC-14) (pp. 596-600). Atlantis Press.

Lisman, J., Grace, A. A., & Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in Neurosciences*, *34*(10), 536–547. https://doi.org/10.1016/j.tins.2011.07.006

Lukas, C. A., Ebert, D. D., Fuentes, H. T., Caspar, F., & Berking, M. (2018). Deficits in general emotion regulation skills–Evidence of a transdiagnostic factor. *Journal of clinical psychology*, *74*(6), 1017–1033. https://doi.org/10.1002/jclp.22565

MacNamara, A., Ochsner, K. N., & Hajcak, G. (2011). Previously reappraised: The lasting effect of description type on picture-elicited electrocortical activity. *Social cognitive and affective neuroscience*, *6*(3), 348–358. https://doi.org/10.1093/scan/nsq053

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

McGuire, A. B., Gillath, O., & Vitevitch, M. S. (2016). Effects of mental resource availability on looming task performance. *Attention, perception, & psychophysics*, *78*(1), 107–113. https://doi.org/10.3758/s13414-015-1006-2

McRae, K. (2013). Emotion regulation frequency and success: separating constructs from methods and time scale. *Social and personality psychology compass*, *7*(5), 289–302. https://doi.org/10.1111/spc3.12027

McRae, K., Ciesielski, B., & Gross, J. J. (2012). Unpacking cognitive reappraisal: goals, tactics, and outcomes. *Emotion*, *12*(2), 250–255. https://doi.org/10.1037/a0026351

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

McRae, K., Hughes, B., Chopra, S., Gabrieli, J. D. E., Gross, J. J., & Ochsner, K. N. (2010). The neural bases of distraction and reappraisal. *Journal of Cognitive Neuroscience*, *22*(2), 248–262. https://doi.org/10.1162/jocn.2009.21243

McRae, K., Jacobs, S. E., Ray, R. D., John, O. P., & Gross, J. J. (2012). Individual differences in reappraisal ability: Links to reappraisal frequency, well-being, and cognitive control. *Journal of Research in Personality*, *46*(1), 2–7. https://doi.org/10.1016/j.jrp.2011.10.003

McRae, K., Ochsner, K. N., Mauss, I. B., Gabrieli, J. J., & Gross, J. J. (2008). Gender differences in emotion regulation: An fMRI study of cognitive reappraisal. *Group processes & intergroup relations*, *11*(2), 143–162.

Mikels, J. A., Maglio, S. J., Reed, A. E., & Kaplowitz, L. J. (2011). Should I go with my gut? Investigating the benefits of emotion-focused decision making. *Emotion*, *11*(4), 743–753. https://doi.org/10.1037/a0023986

Milyavsky, M., Webber, D., Fernandez, J. R., Kruglanski, A. W., Goldenberg, A., Suri, G., & Gross, J. J. (2019). To reappraise or not to reappraise? Emotion regulation choice and cognitive energetics. *Emotion*, *19*(6), 964–981. https://doi.org/10.1037/emo0000498

Moscovitch, D. A., Gavric, D. L., Senn, J. M., Santesso, D. L., Miskovic, V., Schmidt, L. A., McCabe, R. E., & Antony, M. M. (2012). Changes in judgment biases and use of emotion regulation strategies during cognitive-behavioral therapy for social anxiety disorder: distinguishing treatment responders from nonresponders. *Cognitive therapy and research*, *36*(4), 261–271. https://doi.org/10.1007/s10608-011-9371-1

Mullennix, J., Barber, J., & Cory, T. (2019). An examination of the Kuleshov effect using still photographs. *PLoS One*, *14*(10), e0224623.

Neta, M., Davis, F. C., & Whalen, P. J. (2011). Valence resolution of ambiguous facial expressions using an emotional oddball task. *Emotion*, *11*(6), 1425.

Nook, E. C., Schleider, J. L., & Somerville, L. H. (2017). A linguistic signature of psychological distancing in emotion regulation. *Journal of Experimental Psychology: General*, *146*(3), 337.

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, *20*(5), 773.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, *14*(8), 1215–1229. https://doi.org/10.1162/089892902760807212

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: A synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, *1251*(1), E1–E24. https://doi.org/10.1111/j.1749-6632.2012.06751.x

Ortner, C. N., & de Koning, M. (2013). Effects of regulating positive emotions through reappraisal and suppression on verbal and non-verbal recognition memory. *PLoS One*, *8*(4), e62750.

Ortner, C. N. M., Ste Marie, M., & Corno, D. (2016). Cognitive costs of reappraisal depend on both emotional stimulus intensity and individual differences in habitual reappraisal. *PLoS One*, *11*(12), e0167253.

Ortner, C. N., Zelazo, P. D., & Anderson, A. K. (2013). Effects of emotion regulation on concurrent attentional performance. *Motivation and Emotion*, *37*(2), 346–354.

Otto, B., Misra, S., Prasad, A., & McRae, K. (2014). Functional overlap of top-down emotion regulation and generation: An fMRI study identifying common neural

substrates between cognitive reappraisal and cognitively generated emotions. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(3), 923–938. https://doi.org/10.3758/s13415-013-0240-0

Palumbo, R., D'Ascenzo, S., Quercia, A., & Tommasi, L. (2017). Adaptation to complex pictures: exposure to emotional valence induces assimilative aftereffects. *Frontiers in Psychology*, *8, 54*. https://doi.org/10.3389/fpsyg.2017.00054

Pavlov, S. V., Reva, N. V., Loktev, K. V., Tumyalis, A. V., Korenyok, V. V., & Aftanas, L. I. (2014). The temporal dynamics of cognitive reappraisal: Cardiovascular consequences of downregulation of negative emotion and upregulation of positive emotion. *Psychophysiology*, *51*(2), 178–186. https://doi.org/10.1111/psyp.12159

Pecchinenda, A., & Petrucci, M. (2016). Emotion unchained: Facial expression modulates gaze cueing under cognitive load. *PLoS One*, *11*(12). https://doi.org/10.1371/journal.pone.0168111

Pogrebtsova, E., Craig, J., Chris, A., O'Shea, D., & González-Morales, M. G. (2018). Exploring daily affective changes in university students with a mindful positive reappraisal intervention: A daily diary randomized controlled trial. *Stress and Health*, *34*(1), 46–58. https://doi.org/10.1002/smi.2759

Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2014). Where is the chocolate? Rapid spatial orienting toward stimuli associated with primary rewards. *Cognition*, *130*(3), 348–359. https://doi.org/10.1016/j.cognition.2013.12.002

Porter, S., Spencer, L., & Birt, A. R. (2003). Blinded by emotion? Effect of the emotionality of a scene on susceptibility to false memories. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *35*(3), 165–175. https://doi.org/10.1037/h0087198

Pudovkin, V. I., & Montagu, I. G. S. (1958). *Film technique and film acting*. Vision London.

Rauch, S., Eftekhari, A., & Ruzek, J. I. (2012). Review of exposure therapy: A gold standard for PTSD treatment. *Journal of Rehabilitation Research and Development*, *49*(5), 679–688.

Ray, R. D., McRae, K., Ochsner, K. N., & Gross, J. J. (2010). Cognitive reappraisal of negative affect: Converging evidence from EMG and self-report. *Emotion*, *10*(4), 587–592. https://doi.org/10.1037/a0019015

Ray, R., Ochsner, K. N., Cooper, J., Robertson, E., Gabrieli, J., & Gross, J. J. (2005). Individual differences in trait rumination and the neural systems supporting cognitive reappraisal. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(2), 156.

Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion*, *11*(3), 553–562. https://doi.org/10.1037/a0024246

Roediger, H. L., Srinivas, K., & Weldon, M. S. (1989). Dissociations between implicit measures of retention. *Implicit Memory: Theoretical Issues*, 67–84.

Roseman, I. J., & Smith, C. A. (2001). Appraisal theory. *Appraisal Processes in Emotion: Theory, Methods, Research*, 3–19.

Rubin, K. H., Coplan, R. J., Fox, N. A., & Calkins, S. D. (1995). Emotionality, emotion regulation, and preschoolers' social adaptation. *Development and Psychopathology*, *7*(1), 49–62. https://doi.org/10.1017/S0954579400006337

Schupp, H. T., Schmälzle, R., Flaisch, T., Weike, A. I., & Hamm, A. O. (2012). Affective picture processing as a function of preceding picture valence: An ERP analysis. *Biological Psychology*, *91*(1), 81–87. https://doi.org/10.1016/j.biopsycho.2012.04.006

Shaw, P., Stringaris, A., Nigg, J., & Leibenluft, E. (2014). Emotion dysregulation in attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *171*(3), 276–293. https://doi.org/10.1176/appi.ajp.2013.13070966

Sheppes, G., & Meiran, N. (2008). Divergent cognitive costs for online forms of reappraisal and distraction. *Emotion*, *8*(6), 870–874. https://doi.org/10.1037/a0013711

Sheppes, G., Scheibe, S., Suri, G., & Gross, J. J. (2011). Emotion-regulation choice. *Psychological Science*, *22*(11), 1391–1396. https://doi.org/10.1177/0956797611418350

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. https://doi.org/10.1037/a0030831

Silvers, J. A., Shu, J., Hubbard, A. D., Weber, J., & Ochsner, K. N. (2015). Concurrent and lasting effects of emotion regulation on amygdala response in adolescence and young adulthood. *Developmental Science*, *18*(5), 771–784. https://doi.org/10.1111/desc.12260

Silvers, J. A., Wager, T. D., Weber, J., & Ochsner, K. N. (2015). The neural bases of uninstructed negative emotion modulation. *Social Cognitive and Affective Neuroscience*, *10*(1), 10–18. https://doi.org/10.1093/scan/nsu016

Smith, C. A., & Kirby, L. D. (2000). *Consequences require antecedents: Toward a process model of emotion elicitation. In J. P. Forgas (Ed.), Studies in emotion and social interaction, second series. Feeling and thinking: The role of affect in social cognition (p. 83–106). Cambridge University Press.*

Smith, C. A., & Kirby, L. D. (2001). Toward delivering on the promise of appraisal theory. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (p. 132). Oxford University Press.

Smith, J. C., Bradley, M. M., & Lang, P. J. (2005). State anxiety and affective physiology: Effects of sustained exposure to affective pictures. *Biological Psychology*, *69*(3), 247–260. https://doi.org/10.1016/j.biopsycho.2004.09.001

Storbeck, J., & Clore, G. L. (2008). Affective arousal as information: how affective arousal influences judgments, learning, and memory. *Social and Personality Psychology Compass*, *2*(5), 1824–1843. https://doi.org/10.1111/j.1751-9004.2008.00138.x

Strong, R. W., & Alvarez, G. (2019). Using simulation and resampling to improve the statistical power and reproducibility of psychological research. https://doi.org/10.31234/osf.io/2bt6q

Suri, G., Sheppes, G., Young, G., Abraham, D., McRae, K., & Gross, J. J. (2018). Emotion regulation choice: The role of environmental affordances. *Cognition & Emotion*, *32*(5), 963–971. https://doi.org/10.1080/02699931.2017.1371003

Sutton, S. K., Davidson, R. J., Donzella, B., Irw1n, W., & Dottl, D. A. (1997). Manipulating affective state using extended picture presentations. *Psychophysiology*, *34*(2), 217–226. https://doi.org/10.1111/j.1469-8986.1997.tb02135.x

Troy, A. S., Shallcross, A. J., Brunner, A., Friedman, R., & Jones, M. C. (2018). Cognitive reappraisal and acceptance: Effects on emotion, physiology, and perceived cognitive costs. *Emotion*, *18*(1), 58–74. https://doi.org/10.1037/emo0000371

Troy, A. S., Shallcross, A. J., & Mauss, I. B. (2013). A person-by-situation approach to emotion regulation: Cognitive reappraisal can either help or hurt, depending on the context. *Psychological Science*, *24*(12), 2505–2514. https://doi.org/10.1177/0956797613496434

Ullsperger, M., Bylsma, L., & Botvinick, M. (2005). The conflict adaptation effect: It's not just priming. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(4), 467–472. https://doi.org/10.3758/CABN.5.4.467

Uusberg, A., Taxer, J. L., Yih, J., Uusberg, H., & Gross, J. J. (2019). Reappraising reappraisal. *Emotion Review*, *11*(4), 267–282.

Waugh, C. E., Thompson, R. J., & Gotlib, I. H. (2011). Flexible emotional responsiveness in trait resilience. *Emotion*, *11*(5), 1059–1067. https://doi.org/10.1037/a0021786

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. https://doi.org/10.1037/a0027600

Weiss, N. H., Sullivan, T. P., & Tull, M. T. (2015). Explicating the role of emotion dysregulation in risky behaviors: A review and synthesis of the literature with directions for future research and clinical practice. *Current Opinion in Psychology*, *3*, 22–29. https://doi.org/10.1016/j.copsyc.2015.01.013

Wendt, J., Schmidt, L. E., Lotze, M., & Hamm, A. O. (2012). Mechanisms of change: Effects of repetitive exposure to feared stimuli on the brain's fear network. *Psychophysiology*, *49*(10), 1319–1329. https://doi.org/10.1111/j.1469-8986.2012.01451.x

Wharton, E., & Kanas, N. (2019). Mindfulness-based stress reduction for the treatment of anxiety disorders. *International Journal of Group Psychotherapy*, *69*(3), 362–372. https://doi.org/10.1080/00207284.2019.1599289

Wilson-Mendenhall, C. D. (2017). Constructing emotion through simulation. *Current Opinion in Psychology*, *17*, 189–194.

Wormwood, J. B., Siegel, E. H., Kopec, J., Quigley, K. S., & Barrett, L. F. (2019). You are what I feel: A test of the affective realism hypothesis. *Emotion*, *19*(5), 788.

Wu, L., Winkler, M. H., Andreatta, M., Hajcak, G., & Pauli, P. (2012). Appraisal frames of pleasant and unpleasant pictures alter emotional responses as reflected in self-report and facial electromyographic activity. *International Journal of Psychophysiology*, *85*(2), 224–229. https://doi.org/10.1016/j.ijpsycho.2012.04.010

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. https://doi.org/10.1037/bne0000271

Zeng, Q., Qi, S., Li, M., Yao, S., Ding, C., & Yang, D. (2017). Enhanced conflict-driven cognitive control by emotional arousal, not by valence. *Cognition and Emotion*, *31*(6), 1083–1096. https://doi.org/10.1080/02699931.2016.1189882

Appendix A

Study	Excluded N	Included N	# Males	# Females	# Other / Unreported	Age Range	Mean Age Males	Mean Age Females
Artist_ER	3	40	14	25	4	19 : 43	32.57 (6.1)	29.96 (7.3)
ARTIST-Online	3	34	-	-	-	-	-	-
BEER-Cartoon	0	49	-	-	-	-	-	-
BEER_Task_Final	0	42	6	32	4	18 : 22	19.83 (1)	19.31 (1.3)
BEER-WFU	0	41	13	26	2	18 : 22	19.77 (1.5)	18.85 (1.1)
CANCER-Online	2	32	2	5	27	38 : 66	48 (14.1)	59.6 (7.2)
Developmental-Behavioral	0	45	12	10	23	10 : 23	17.56 (3.8)	16.5 (4.6)
Developmental-fMRI	0	54	24	23	7	8 : 22	15.56 (4.2)	16.68 (3.8)
Distraction-Published	0	18	0	18	0	-	24.4 (3.5)	-
EFstudy_7-2-08	6	101	41	59	7	18 : 36	-	-
ELERS_1	1	58	0	55	4	19 : 35	n/a	27.85 (4)
ELERS_2	2	34	-	-	-	-	-	-
Standford EMG	0	39	0	39	0	18 : 29	n/a	18.9 (1.6)
ERS Version B	0	65	17	47	1	17 : 27	19.06 (1.2)	19.49 (1.7)
ERS Version C	0	22	7	12	3	18 : 22	19.86 (1.6)	20 (1)
IGT	4	130	41	90	3	18 : 35	20.07 (2.6)	19.66 (1.5)
PEER	0	50	25	24	1	18 : 47	21.84 (5.8)	22.54 (4)
POS	0	81	15	56	10	18 : 35	19.93 (1.6)	23.41 (5)
SOS-Online	1	94	-	-	-	-	-	-
ERS Version D	0	10	4	5	1	19 : 24	20.5 (2.4)	19.8 (0.8)
ERS Version E	0	10	4	3	3	19 : 21	19.5 (0.6)	20 (1)
ROSE Version A	0	30	1	16	13	18 : 29	19 (0)	19.69 (2.8)
ROSE Version B	0	88	11	48	29	18 : 27	20.18 (2.6)	19.4 (1.2)
ERS Version F	0	108	27	75	6	17 : 28	20.07 (2.2)	19.12 (1.3)
RTP	0	85	22	63	0	18 : 23	-	-
ERS fMRI	0	34	13	21	0	18 : 37	24.08 (4.5)	26.24 (5.7)
SOS fMRI	0	19	0	18	1	19 : 31	n/a	24 (3.2)

Demographics breakdown for studies in the database. Red dashes indicate data unavailable or is still in preparation. Age data in parentheses indicates standard deviation

Appendix B

Fixed Effects Model 1		2	3	4	5	6	7	8
Name	Estimate	SE	tStat	DF	pValue	Lower	Upper	
(Intercept)	3.29	0.01	602.48	122867	<.001	3.27	3.30	
Valence	1.17	0.01	158.42	122867	<.001	1.15	1.18	
Arousal	0.29	0.01	46.46	122867	<.001	0.28	0.31	
Change Instruction	-0.84	0.01	-80.76	122867	<.001	-0.86	-0.82	
Time	-0.27	0.02	-12.06	122867	<.001	-0.31	-0.23	
Change Instruction:Time	0.26	0.04	6.67	122867	<.001	0.18	0.34	
Fixed Effects Model 2								
Name	Estimate	SE	tStat	DF	pValue	Lower	Upper	
(Intercept)	2.90	0.01	277.70	105865	<.001	2.87	2.92	
Valence	1.22	0.01	152.94	105865	<.001	1.21	1.24	
Arousal	0.27	0.01	40.13	105865	<.001	0.26	0.28	
Change Instruction	-0.85	0.01	-76.21	105865	<.001	-0.87	-0.83	
Time	-0.14	0.02	-5.92	105865	<.001	-0.19	-0.10	
Previous Change	-0.17	0.01	-17.35	105865	<.001	-0.19	-0.15	
Previous Negative Affect	0.13	0.00	55.52	105865	<.001	0.12	0.13	
Change Instruction:Time	0.13	0.04	2.91	105865	<.01	0.04	0.21	
Theoretical Likelihood Ratio Test								
Model	DF	AIC	BIC	LogLik	Adj Rsqrd	LRStat	deltaDF	pValue
Fixed Effects Model 1	7.00	456690.00	456760.00	-228340.00	0.4231			
Fixed Effects Model 2	9.00	390020.00	390110.00	-195000.00	0.4465	66671.0000	2.00	0
Fixed Effects Model 3	justed R-squared:							
Name	Estimate	SE	tStat	DF	pValue	Lower	Upper	
(Intercept)	2.89	0.01	276.75	105863	<.001	2.87	2.91	
Valence	1.23	0.01	153.06	105863	<.001	1.21	1.24	
Arousal	0.28	0.01	40.54	105863	<.001	0.26	0.29	
Change Instruction	-0.85	0.01	-76.05	105863	<.001	-0.87	-0.82	
Time	-0.15	0.02	-6.32	105863	<.001	-0.20	-0.11	
Previous Change	-0.17	0.01	-17.07	105863	<.001	-0.19	-0.15	
Previous Negative Affect	0.13	0.00	55.58	105863	<.001	0.12	0.13	
Valence:Time	-0.22	0.03	-7.47	105863	<.001	-0.28	-0.16	
Arousal:Time	0.33	0.03	12.20	105863	<.001	0.28	0.38	
Change Instruction:Time	0.06	0.05	1.22	105863	0.2230	-0.04	0.15	
Theoretical Likelihood Ratio Test								
Model	DF	AIC	BIC	LogLik	Adj Rsqrd	LRStat	deltaDF	pValue
Fixed Effects Model 2	9.00	390022.88	390109.01	-195002.44	0.4465			
Fixed Effects Model 3	11.00	389886.86	389992.13	-194932.43	0.4473	140.0133	2.00	0

Fixed Effects Model 4								
Name	Estimate	SE	tStat	DF	pValue	Lower	Upper	
(Intercept)	2.89	0.01	277.21	105864	<.001	2.87	2.91	
Valence	1.23	0.01	153.16	105864	<.001	1.21	1.24	
Arousal	0.28	0.01	40.52	105864	<.001	0.26	0.29	
Change Instruction	-0.85	0.01	-76.56	105864	<.001	-0.87	-0.83	
Time	-0.14	0.02	-6.68	105864	<.001	-0.18	-0.10	
Previous Change	-0.17	0.01	-17.24	105864	<.001	-0.19	-0.15	
Previous Negative Affect	0.13	0.00	55.60	105864	<.001	0.12	0.13	
Valence:Time	-0.21	0.03	-7.41	105864	<.001	-0.27	-0.16	
Arousal:Time	0.33	0.03	12.30	105864	<.001	0.28	0.38	

Theoretical Likelihood Ratio Test								
Model	DF	AIC	BIC	LogLik	Adj Rsqrd	LRStat	deltaDF	pValue
Fixed Effects Model 4	10.00	389882.10	389977.80	-194931.05	0.4473			
Fixed Effects Model 3	11.00	389886.86	389992.13	-194932.43	0.4473	-2.7590	1.00	1

Fixed Effects Model 5	justed R-squared:							
Name	Estimate	SE	tStat	DF	pValue	Lower	Upper	
(Intercept)	2.89	0.01	277.54	105862	<.001	2.87	2.91	
Valence	1.21	0.01	130.54	105862	<.001	1.19	1.23	
Arousal	0.34	0.01	42.47	105862	<.001	0.32	0.36	
Change Instruction	-0.47	0.02	-27.22	105862	<.001	-0.50	-0.43	
Time	-0.18	0.02	-8.85	105862	<.001	-0.22	-0.14	
Previous Change	-0.17	0.01	-17.03	105862	<.001	-0.19	-0.15	
Previous Negative Affect	0.13	0.00	56.72	105862	<.001	0.13	0.14	
Valence:Change Instruction	-0.17	0.02	-8.88	105862	<.001	-0.21	-0.13	
Arousal:Change Instruction	-0.31	0.02	-20.24	105862	<.001	-0.34	-0.28	
Valence:Time	-0.20	0.03	-6.92	105862	<.001	-0.26	-0.14	
Arousal:Time	0.32	0.03	12.00	105862	<.001	0.27	0.37	

Theoretical Likelihood Ratio Test								
Model	DF	AIC	BIC	LogLik	Adj Rsqrd	LRStat	deltaDF	pValue
Fixed Effects Model 4	10.00	389882.10	389977.80	-194931.05	0.4473			
Fixed Effects Model 5	12.00	388901.52	389016.36	-194438.76	0.4524	984.5869	2.00	0

RANDOM EFFECTS MODELS	*Random Effects for: Intercept + Change Instruction + Time + Valence + Arousal + Previous Negative Affect + Previous Change for Participant nested within Study*							
Linear Mixed Effects Model 1								
Name	**Estimate**	**SE**	**tStat**	**DF**	**pValue**	**Lower**	**Upper**	
(Intercept)	3.19	0.02	148.00	105862	<.001	3.14	3.23	
Valence	1.21	0.01	80.87	105862	<.001	1.18	1.24	
Arousal	0.31	0.01	30.98	105862	<.001	0.29	0.33	
Change Instruction	-0.37	0.03	-13.12	105862	<.001	-0.42	-0.31	
Time	-0.07	0.03	-2.21	105862	0.027	-0.14	-0.01	
Previous Change	-0.03	0.01	-3.35	105862	<.001	-0.05	-0.01	
Previous Negative Affect	0.03	0.00	10.30	105862	<.001	0.02	0.03	
Valence:Change Instruction	-0.26	0.02	-14.98	105862	<.001	-0.29	-0.22	
Arousal:Change Instruction	-0.26	0.01	-19.42	105862	<.001	-0.28	-0.23	
Valence:Time	-0.33	0.03	-11.40	105862	<.001	-0.39	-0.27	
Arousal:Time	0.29	0.02	11.62	105862	<.001	0.24	0.34	
Theoretical Likelihood Ratio Test								
Model	**DF**	**AIC**	**BIC**	**LogLik**	**Adj Rsqrd**	**LRStat**	**deltaDF**	**pValue**
Fixed Effects Model 5	12.00	388901.52	389016.36	-194438.76	0.4524			
Linear Mixed Effects Model 1	40.00	356110.19	356492.99	-178015.10	0.6400	32847.3267	28.00	0

	Random Effects for: Intercept + Change Instruction + Time + Valence + Arousal + Previous Negative Affect + Previous Change for Participant nested within Study AND Change Instruction for Image							
Linear Mixed Effects Model 2								
Name	**Estimate**	**SE**	**tStat**	**DF**	**pValue**	**Lower**	**Upper**	
(Intercept)	3.33	0.04	84.87	105862	<.001	3.25	3.41	
Valence	1.23	0.04	27.37	105862	<.001	1.14	1.32	
Arousal	0.16	0.04	4.26	105862	<.001	0.08	0.23	
Change Instruction	-0.64	0.04	-15.85	105862	<.001	-0.72	-0.56	
Time	-0.15	0.03	-4.50	105862	<.001	-0.22	-0.09	
Previous Change	-0.03	0.01	-2.98	105862	<.01	-0.05	-0.01	
Previous Negative Affect	0.04	0.00	15.36	105862	<.001	0.04	0.05	
Valence:Change Instruction	-0.21	0.04	-5.12	105862	<.001	-0.28	-0.13	
Arousal:Change Instruction	-0.09	0.03	-3.11	105862	<.01	-0.14	-0.03	
Valence:Time	**-0.05**	**0.04**	**-1.31**	**105862**	**0.192**	**-0.12**	**0.02**	
Arousal:Time	0.07	0.03	2.10	105862	0.036	0.00	0.13	
Theoretical Likelihood Ratio Test								
Model	**DF**	**AIC**	**BIC**	**LogLik**	**Adj Rsqrd**	**LRStat**	**deltaDF**	**pValue**
Linear Mixed Effects Model 1	40.00	356110.19	356492.99	-178015.10	0.6444			
Linear Mixed Effects Model 2	43.00	343035.96	343447.47	-171474.98	0.6939	13080.2288	3.00	0

Linear Mixed Effects Model 3	Random Effects for: Intercept + Change Instruction + Time + Valence + Arousal + Previous Negative Affect + Previous Change for Participant nested within Study AND Change Instruction for Image							
Name	Estimate	SE	tStat	DF	pValue	Lower	Upper	
(Intercept)	3.33	0.04	84.80	105863	<.001	3.25	3.41	
Valence	1.23	0.04	27.30	105863	<.001	1.14	1.31	
Arousal	0.16	0.04	4.27	105863	<.001	0.09	0.23	
Change Instruction	-0.64	0.04	-15.87	105863	<.001	-0.72	-0.56	
Time	-0.16	0.03	-4.73	105863	<.001	-0.23	-0.09	
Previous Change	-0.03	0.01	-2.99	105863	<.01	-0.05	-0.01	
Previous Negative Affect	0.04	0.00	15.36	105863	<.001	0.04	0.05	
Valence:Change Instruction	-0.20	0.04	-5.07	105863	<.001	-0.28	-0.13	
Arousal:Change Instruction	-0.09	0.03	-3.15	105863	<.01	-0.14	-0.03	
Arousal:Time	0.03	0.02	1.77	105863	0.077	0.00	0.07	
Theoretical Likelihood Ratio Test								
Model	DF	AIC	BIC	LogLik	Adj Rsqrd	LRStat	deltaDF	pValue
Linear Mixed Effects Model 3	42.00	343030.91	343432.84	-171473.45	0.6938	0.0000	0.00	0
Linear Mixed Effects Model 2	43.00	343035.96	343447.47	-171474.98	0.6939	-3.0539	1.00	1

Linear Mixed Effects Model 4	Random Effects for: Intercept + Change Instruction + Time + Valence + Arousal + Previous Negative Affect + Previous Change for Participant nested within Study AND Change Instruction for Image							
Name	Estimate	SE	tStat	DF	pValue	Lower	Upper	
(Intercept)	3.33	0.04	84.77	105864	<.001	3.26	3.41	
Valence	1.23	0.04	27.27	105864	<.001	1.14	1.31	
Arousal	0.16	0.04	4.23	105864	<.001	0.08	0.23	
Change Instruction	-0.65	0.04	-15.88	105864	<.001	-0.72	-0.57	
Time	-0.15	0.03	-4.54	105864	<.001	-0.22	-0.09	
Previous Change	-0.03	0.01	-2.99	105864	<.01	-0.05	-0.01	
Previous Negative Affect	0.04	0.00	15.36	105864	<.001	0.04	0.05	
Valence:Change Instruction	-0.20	0.04	-5.04	105864	<.001	-0.28	-0.12	
Arousal:Change Instruction	-0.09	0.03	-3.15	105864	<.01	-0.14	-0.03	
Theoretical Likelihood Ratio Test								
Model	DF	AIC	BIC	LogLik	Adj Rsqrd	LRStat	deltaDF	pValue
Linear Mixed Effects Model 4	41.00	343025.97	343418.33	-171471.98	0.6939			
Linear Mixed Effects Model 3	42.00	343030.91	343432.84	-171473.45	0.6938	-2.9416	1.00	1

Fixed effects coefficients for each model tested. Models are compared using Likelihood Ratio Tests. Simpler models are selected provided they do not significantly differ from more complicated models in terms of the Likelihood Ratio Test. Non-significant variables are dropped from the model at each stage.

Appendix C

Random effects covariance parameters (95% CIs): Group: Participant nested within Study (1383 Levels)					
Name1	**Name2**	**Type**	**Estimate**	**Lower**	**Upper**
(Intercept)	(Intercept)	std	0.6727851	0.641907	0.705149
Valence	(Intercept)	corr	0.3808934	0.313753	0.444244
Arousal	(Intercept)	corr	0.3577156	0.331386	0.383488
Change	(Intercept)	corr	-0.208513	-0.26995	-0.14538
Time	(Intercept)	corr	0.1990874	0.100937	0.293399
Previous Change	(Intercept)	corr	-0.759637	-0.78738	-0.72883
Previous Negative Affect	(Intercept)	corr	-0.246408	-0.32211	-0.16757
Valence	Valence	std	0.3899413	0.368653	0.412459
Arousal	Valence	corr	0.0703228	-0.00103	0.140959
Change	Valence	corr	-0.244105	-0.29966	-0.1869
Time	Valence	corr	0.1330937	0.038793	0.225045
Previous Change	Valence	corr	0.0028793	-0.39467	0.399519
Previous Negative Affect	Valence	corr	-0.503367	-0.59594	-0.39758
Arousal	Arousal	std	0.2175961	0.201676	0.234773
Change	Arousal	corr	-0.131333	-0.2189	-0.04167
Time	Arousal	corr	-0.036501	-0.16147	0.089616
Previous Change	Arousal	corr	-0.367512	NaN	NaN
Previous Negative Affect	Arousal	corr	-0.192906	-0.28001	-0.10265
Change	Change	std	0.7985537	0.762052	0.836804
Time	Change	corr	0.1152722	NaN	NaN
Previous Change	Change	corr	-0.362304	NaN	NaN
Previous Negative Affect	Change	corr	0.2814214	0.199642	0.359308
Time	Time	std	0.7163038	0.644089	0.796615
Previous Change	Time	corr	-0.435585	-0.7176	-0.03086
Previous Negative Affect	Time	corr	0.1384908	0.022546	0.250759
Previous Change	Previous Change	std	0.0380394	0.023466	0.061663
Previous Negative Affect	Previous Change	corr	-0.166724	-0.2299	-0.10215
Previous Negative Affect	Previous Negative Affect	std	0.0561484	0.050439	0.062504
Random effects covariance parameters (95% CIs): Group: Image (425 Levels)					
Name1	**Name2**	**Type**	**Estimate**	**Lower**	**Upper**
(Intercept)	(Intercept)	std	0.665263	0.617606	0.716597
Change	(Intercept)	corr	-0.367518	-0.52061	-0.19152
Change	Change	std	0.2419576	0.205528	0.284844

Random effects covariance parameters for the final selected linear mixed effects model