Received: 31 March 2021 | Revised: 30 June 2021 | Accepted: 6 July 2021

DOI: 10.1111/psyp.13908

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Effort beats effectiveness in emotion regulation choice: Differences between suppression and distancing in subjective and physiological measures

Christoph Scheffel[1] | **Sven-Thomas Graupner**[2] | **Anne Gärtner**[1] | **Josephine Zerna**[1] | **Alexander Strobel**[1] | **Denise Dörfel**[1]

[1]Differential and Personality Psychology, Faculty of Psychology, Technische Universität Dresden, Dresden, Germany

[2]Engineering Psychology and Applied Cognitive Research, Faculty of Psychology, Technische Universität Dresden, Dresden, Germany

Correspondence
Christoph Scheffel, Differential and Personality Psychology, Faculty of Psychology, Technische Universität Dresden, Zellescher Weg 17, Dresden 01069, Germany.
Email: christoph_scheffel@tu-dresden.de

Funding information
The second study was funded by centralized funds of the Faculty of Psychology (MK201910) at the Technische Universität Dresden. Both studies were partly funded by the Deutsche Forschungsgemeinschaft (SFB 940, project B6) Open Access funding enabled and organized by Projekt DEAL.

Abstract

Emotion regulation (ER) can be implemented by different strategies which differ in their capacity to alter emotional responding. What all strategies have in common is that cognitive control must be exercised in order to implement them. The aim of the present preregistered study was to investigate whether the two ER strategies, expressive suppression and distancing, require different amounts of cognitive effort and whether effort is associated with personality traits. Effort was assessed subjectively via ratings and objectively via pupillometry and heart period. In two studies, $N = 110$ and $N = 52$ healthy adults conducted an ER paradigm. Participants used suppression and distancing during inspection of positive and negative pictures. They also had the choice to reapply either of the strategies at the end of the paradigm. Although distancing was more effective in downregulation of subjective arousal (Study 1: $p < .001, \eta_p^2 = .20$; Study 2: $p < .001$, $\eta_p^2 = .207$), about two thirds reapplied suppression, because it was perceived as less effortful. Effort was rated significantly lower for suppression compared to distancing (Study 1: $p = .042$, $\eta_p^2 = .04$; Study 2: $p = .002$, $\eta_p^2 = .13$). However, differences in effort were not reflected in pupillary data or heart period. Broad and narrow personality traits were neither associated with the preferred strategy nor with subjective or physiological effort measures. Findings suggest that people tend to use the ER strategy that is perceived as less effortful, even though it might not be the most effective strategy.

KEYWORDS
distancing, effort, emotion regulation, expressive suppression, heart rate, pupillometry

1 | THEORETICAL BACKGROUND

Emotion regulation (ER) can be conceptualized as any process by which individuals modify their emotional experiences, expressions, and physiology (Gross, 1998). Individuals encounter situations every day in which they have to regulate emotions. To achieve this, people can choose from a variety of strategies: situation selection, situation modification, attentional deployment, cognitive change, and response modulation (Gross, 1998). According to the author's process model of ER, cognitive change (often referred to as reappraisal) occurs at an early stage of the emotion generation process and is, therefore, conceptualized as antecedent-focused strategy. Response modulation occurs late (Gross, 1998, 2014) and is, therefore, a response-focused strategy. Following a recent taxonomy of Powers and LaBar (2019), these strategies can be implemented through different tactics. For example, the strategy reappraisal can be achieved using the tactic distancing, and the strategy response modulation using the tactic expressive suppression (from here on referred to as suppression). Because the present study focuses on two tactics from two different strategies, for simplicity we will refer to suppression and distancing as strategies from here on. Both strategies have been shown to reduce self-reported negative arousal and negative affect (Gross, 1998; Ray et al., 2010). Meta-analytical evidence suggests that reappraisal is superior over suppression in reducing arousal and distancing has been shown to be more effective than reinterpretation, another reappraisal tactic (Webb et al., 2012). Moreover, it has been shown that reappraisal has positive psychological, cognitive, and social outcomes, whereas suppression is associated with negative psychological outcomes (e.g., Butler et al., 2003; Gross, 2002; Gross & John, 2003; Haga et al., 2009). Still, the question remains why people tend to use suppression in specific situations, even though it is associated with psychopathology and reduced well-being and social functioning in the long term (Aldao et al., 2010; Gross & John, 2003).

Personality traits offer a possible explanation for this seemingly paradox behavior: Gross and John (2003) introduced the habitual use of reappraisal and suppression as two narrow personality traits. A high level of trait suppression makes people more likely to fall back on this strategy in specific situations. Although relations between those narrow traits and broad personality traits (i.e., Big Five) are known (Gross & John, 2003; John & Gross, 2004; Kokkonen & Pulkkinen, 2001), associations of broad personality traits (especially neuroticism and extraversion) and the effectiveness of the implementation of ER strategies are rather inconsistent (e.g., Harenski et al., 2009; Scheffel et al., 2019). Another explanation is offered by the ER flexibility approach. Findings suggest that people use different ER strategies flexibly in order to manage demands of varying situations (Aldao et al., 2015; Bonanno & Burton, 2013; Kobylinska & Kusev, 2019). Apparently, people choose the strategy that is most useful to achieve a given goal. For example, suppression—often labeled as maladaptive—seems to be beneficial in negative emotional states (Bonanno & Keltner, 1997). The adaptivity of a strategy should, therefore, be evaluated in relation to the goal orientation. Further, individuals are able to choose between ER strategies in such a way that the cost (e.g., effort)–benefit (e.g., downregulation of emotional arousal) calculation is optimal (Sheppes et al., 2014). There are indications that under highly stressful or threatening situations, effortful regulatory options become undesired (Muraven & Baumeister, 2000).

1.1 | Effort and emotion regulation

Another reason for using seemingly less effective ER strategies could be the effort required to implement them. Gross (1998) stated in the process model that the implementation of ER strategies requires cognitive control. This assumption is supported by findings regarding neural correlates of ER. Consistent across different studies, it has been shown that during intentional ER, there is increased activation in brain regions associated with cognitive control (e.g., Buhle et al., 2014; Diers et al., 2020; Dörfel et al., 2014; Ochsner et al., 2012). These cognitive control mechanisms are thought to be effortful (Hofmann et al., 2012). Schweizer et al., (2013) could show that emotional working memory (WM) training could improve affective control. Interestingly, the WM training resulted in increased activation in the subgenual anterior cingulate cortex during the ER task. This region has been shown to be involved in effortful control strategies (Wager et al., 2008). It is assumed that after the WM training, participants put sustained effort in the ER task (cortical effort model) (Engen & Kanske, 2013; Neubauer & Fink, 2009; Schweizer et al., 2013). Furthermore, it has been shown that regulation of negative emotions was perceived as more effortful than regulation of positive emotions (Gruber et al., 2012).

Not only it is known that ER requires cognitive effort, there are also studies targeting differences between ER strategies. Focusing on response modulation, Richards and Gross (1999) found that suppression impaired memory function and led to increased cardiovascular activation. This was interpreted as indication that suppression is cognitively demanding. Interestingly, memory impairments were not observed for the antecedent-focused strategy reappraisal (Richards & Gross, 2000). Suppression as a response-focused strategy on the other hand is said to require more effort because of the constant need to control automatic emotional responses (see Gross, 2002; Richards, 2004). But there are also indications that reappraisal requires cognitive effort, for example, when it is instructed relatively late (Sheppes et al., 2009) or when highly intense emotions are to be regulated (Sheppes & Gross, 2011). In both cases, the demand on control processes is high and therefore more effort is required. Cognitive control may also underlie ER flexibility (Pruessner et al., 2020).

More concrete evidence was collected in studies using pupillometry. Pupil dilation is driven by sympathetic innervation of the pupillary dilator, resulting in larger pupil diameter during different behavioral states of high alertness such as under anxiety (Andreassi, 2000; Bouffard, 2019; Larsen & Waters, 2018). For example, pupil dilation is associated with effort in cognitive control tasks (Alnaes et al., 2014; van der Wel & van Steenbergen, 2018). In addition, several studies have demonstrated the influence of emotional stimuli on pupil dilation, with emotional arousal being associated with increased dilation (e.g., Bradley et al., 2008; Henderson et al., 2014; Kinner et al., 2017). However, when looking at pupillary response during ER, inconsistent results were reported. Studies found increased pupil sizes after both instructed increasing and decreasing of negative emotions (e.g., Johnstone et al., 2007; Kinner et al., 2017; van Reekum et al., 2007; Urry et al., 2009). A possible explanation could be that during ER, both emotional arousal and cognitive effort, required to implement the strategies, influence dilation. Especially, Kinner et al., (2017) found evidence to confirm this assumption. They distinguished between components of the pupillary response that reflect emotional arousal and components that reflect cognitive effort. During an ER paradigm, pupil dilations of 30 participants were tracked. The early pupillary response (within the first 2 s) was higher during regulation and therefore reflects cognitive demands during ER. The late pupillary response was modulated by emotional arousal (Kinner et al., 2017). The authors showed that ER is effortful and that this effort can be captured by means of the early pupillary response.

Another measure that has been associated with effort and ER is heart rate variability (HRV), which refers to the variation in heart period, the interval between two consecutive heartbeats in a pre-defined time window. HRV reactivity in response to stressors and cognitive tasks remained stable across time and task type in multiple studies (Austin et al., 2007; Dragomir et al., 2014; Muhtadie et al., 2015; Sloan et al., 1995; Uchino et al., 2005), and stressors such as threat consistently led to a decrease in HRV (Balzarotti et al., 2017; Kim et al., 2018). It is not clear, however, whether this short-term change in HRV can be attributed to the threatening, engaging (arousal-related), or effort-demanding properties of the stimuli (Smith et al., 2020), as activity in brain regions associated with emotion and in those associated with cognitive control have both been related to HRV reactivity (Gianaros et al., 2004; Lane et al., 2009). Changes in heart period itself have also been identified: emotional stimuli led to increases in heart period in healthy participants but to decreases in heart period in participants with bipolar disorder (Austin et al., 2007). Decreases in heart period in healthy subjects have been found in response to memory tasks, while being associated with decreased activation in brain regions associated with cognition and emotion, such as amygdala, hippocampus, and insula (Gianaros et al., 2004). We, therefore, aim to contribute to the literature by investigating heart period during the application of ER strategies as a measure of effort and arousal.

As an interim summary, it is known that ER requires cognitive effort and can be distinguished from reactions of emotional arousal by means of pupillometry (Kinner et al., 2017). However, a direct comparison of ER strategies has not been done yet. It remains still unclear whether ER strategies require different amounts of cognitive effort. Furthermore, the subjective component of cognitive effort has been left out so far, although it is necessary to distinguish between different effort measures (Kreis et al., 2020; Steele, 2020). Last but not least, the influence of situational factors and individual differences on subjective and physiological cognitive effort during ER remain unknown.

1.2 | The present study

The present study aims to deepen the knowledge about the cognitive effort required by the implementation of two different ER strategies: Distancing as a tactic of reappraisal (antecedent-focused strategy) and suppression (response-focused strategy). Additionally, influence of situational factors and individual differences was investigated. We expected that (1). distancing leads to higher decrease in emotional arousal (subjective ratings, late pupillary response, and heart period) than suppression, (2a) people select the ER strategy that is less effortful if they have a choice, (2b) subjective and physiological cognitive effort is greater during suppression than during distancing, and these relationships are moderated by (3a) situational factors (positive vs. negative valence of the stimuli) and by (3b) individual differences (broad and narrow personality traits).

These hypotheses were targeted in two studies. In Study 1, an ER paradigm was conducted. Participants had to actively view or regulate their emotions while inspecting gray-scaled neutral, negative, and positive pictures. Individual differences in broad and narrow personality traits were assessed using questionnaires. Study 2 used the same procedure except that participants now inspected colored pictures. In a subsequent exploratory analysis, we combined the samples to examine associations between subjective and physiological effort measures and personality traits. We did this to generate a larger sample to approximate recommendations of Schönbrodt and Perugini (2013) regarding sample sizes for correlations.

2 | STUDY 1

2.1 | Method

The study procedure, all hypotheses and statistical analysis methods have been preregistered at the Open Science Framework (OSF) https://osf.io/9wjyp. Data sets, scripts, and analysis routines can be found at https://osf.io/dk4s9.

Furthermore, "we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study" in compliance with the 21-word-solution of open science (Simmons et al., 2012). A complete list of all measures assessed in the study can be found at the OSF (https://osf.io/jqy9z/). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The experimental protocol was approved by the ethics committee of the Technische Universität Dresden (EK286062019).

2.1.1 | Power analysis and participants

Power calculation was based on statistical affordances, related to both theoretical considerations and the aim to replicate the aforementioned study of Kinner et al., (2017), showing enhanced early peak dilations of the pupil during ER. The size of this effect was calculated with partial eta squared $\eta_p^2 = .13$. Calculating the sample size for an ANOVA with the given effect size and a power of 0.95 resulted in $N = 24$. However, we performed another power calculation for a separate research question not presented here (see https://osf.io/9wjyp). Power calculation for this separate research question led to a sample size of $N = 100$ to account for recent replication issues and a lack of literature regarding this research question. We oversampled by 10% to account for potential dropouts, technical failures, or performance-based exclusions. Therefore, we aimed at a sample size of $N = 110$. Thus, our statistical approach considers recent replication issues and actually lower effect sizes than that were reported in initial studies (Open Science Collaboration, 2015; Schäfer & Schwarz, 2019). It also allowed us to compute more reliable correlation coefficients between effort measures and personality traits. All computations were conducted using G*Power 3 (Faul et al., 2007).

$N = 112$ participants took part in the experiment. Two participants had to be excluded because their pupils could not be tracked with the eye tracking system. The final sample consisted of $N = 110$ healthy participants, mostly psychology students (75 females; age: 24.9 ± 5.8). Please note that sample sizes of specific analyses can differ when participants had to be excluded due to technical reasons with the eye tracker or with psychophysiological measurements. All participants stated that they had normal to corrected vision and were fluent in German.

2.1.2 | Experimental procedure

All sessions took place in a shielded cabin with constant lighting. Participants received information about the experimental procedure, signed an informed consent form, and filled out a demographic questionnaire. Afterwards, the height of the table, chair, screen, eye-tracking system, and the headrest were adjusted to the person. The eye-tracking system was calibrated to make sure that the system was able to track the pupil. Participants then received written instructions on the ER paradigm and performed a training session. Next, electrodes to measure EEG (results of which will be reported elsewhere) and ECG were attached and the ER paradigm was conducted. Subsequently, participants filled out questionnaires on broad and narrow personality traits, (see below) and received financial compensation or course credit for their participation.

2.1.3 | Emotion regulation paradigm

The ER task followed a block design (see Figure 1). Participants were told to actively view neutral, negative, and positive pictures (see Section 2.1.4) or to regulate all upcoming emotions by means of distancing and expressive suppression, respectively. The experiment always started with a block "active viewing-neutral", which served as a baseline condition. During this block, 25 emotionally neutral pictures were presented and participants were asked to "actively view all pictures and permit all emotions that may arise." In the second block, participants actively viewed positive and negative pictures ("active viewing-positive", "active viewing-negative"). During the third and fourth block, participants were shown positive and negative pictures and they should regulate their emotions using distancing ("distancing-positive", "distancing-negative") and suppression ("suppression-positive", "suppression-negative"). During distancing, participants were asked to "take the position of a non-involved observer, thinking about the picture in a neutral way." Participants were told not to re-interpret the situation as made-up, attaching a different meaning to the situation, or distracting themselves. During suppression, participants were told to "suppress their emotional facial expression." To achieve this, they should imagine being observed by a third person. This person should not be able to tell just by looking at the facial expression whether the person is looking at a positive or negative picture. Participants were told not to suppress thoughts or to change their facial expression to the opposite. All participants received written instructions including examples and completed a training session which took about 15 min. After the training session, participants were asked about their applied ER strategies to resolve possible misunderstandings. The order of the two ER blocks (distancing and suppression) was completely randomized between participants. The last block contained the choice condition: Participants could choose which of the two

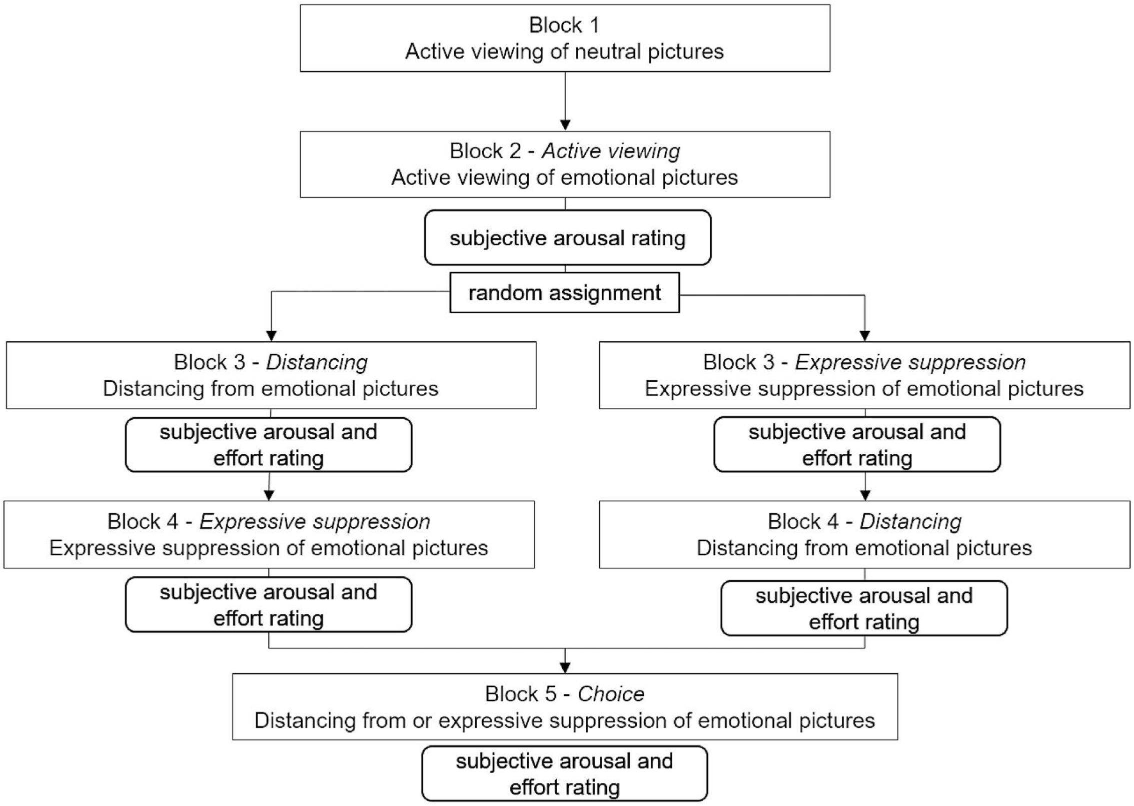

FIGURE 1 Block design of the emotion regulation paradigm. Every participant started with two "active viewing" blocks containing neutral pictures (Block 1) and positive and negative pictures (Block 2). Order of regulation blocks (Blocks 3 and 4) was random between participants. Before the last block, participants could decide whether they want to reapply suppression or distancing. Subjective arousal and effort ratings were assessed after each block using a slider on screen with a continuous scale

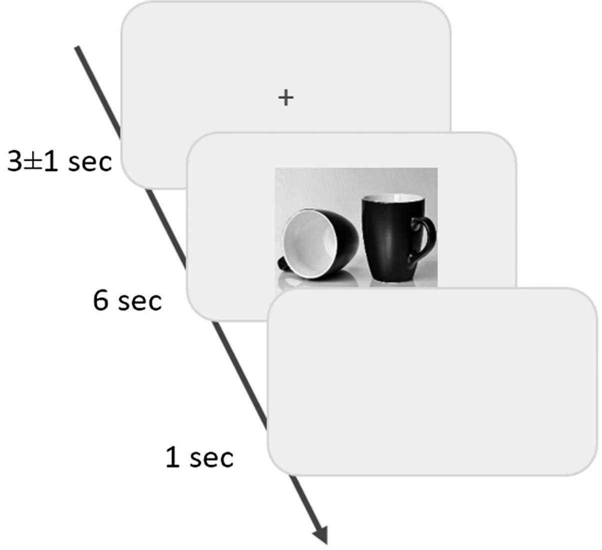

FIGURE 2 Sequence of all events within one trial. Participants were asked to actively view neutral and emotional pictures or to regulate emotional pictures via suppression or detachment. The example picture was taken from the Open Affective Standardized Image Set (OASIS) (Kurdi et al., 2017)

ER strategies they wanted to reapply, without instruction on what basis they should make their decision ("choice-positive", "choice-negative"). They stated their reasons in a questionnaire after the ER task. The ER task was developed following similar ER paradigms (Diers et al., 2020; Dörfel et al., 2014; Scheffel et al., 2019; Walter et al., 2009). Instructions for all blocks were presented on screen at the beginning of each block.

Each of the blocks 2–5 consisted of 50 trials with 25 trials showing positive pictures and 25 trials showing negative pictures. For all blocks, each trial began with a fixation cross that lasted on average 3 s (±1 s, uniform distributed). Afterwards, participants saw a neutral or emotional picture for a total of 6 s. After this, a grey screen was visible for 1 s (see Figure 2). After each block two to five participants retrospectively rated their emotional arousal on a continuous scale (ranging from "not at all aroused" to "very highly aroused"); after each block three to five participants retrospectively rated their subjective effort (ranging from "not very exhausting" to "very exhausting"). This was done separately for positive and negative pictures using a slider on the screen.

2.1.4 | Stimuli

All pictures were selected from the Emotional Picture Set (EmoPicS) (Wessa et al., 2010) and the International Affective Picture System (IAPS) (Lang et al., 2008). The 25 neutral pictures (Valence (V): $M \pm SD = 4.86 \pm 0.49$, Arousal (A): $M \pm SD = 3.01 \pm 0.61$) depicted content related to the categories objects, persons, and scenes. By using an evolutionary algorithm (Yu & Gen, 2010), 100 positive and 100 negative pictures were clustered into four sets, respectively, based on their normative valence ratings, arousal ratings, and luminescence. Positive pictures featured the categories persons, animals, scenes, objects, and food. Valence and arousal values of the four sets were comparable (set one: V: $M \pm SD = 7.22 \pm 0.34$, A: $M \pm SD = 4.87 \pm 0.71$; set two: V: $M \pm SD = 7.22 \pm 0.50$, A: $M \pm SD = 4.87 \pm 0.63$; set three: V: $M \pm SD = 7.22 \pm 0.60$, A: $M \pm SD = 4.87 \pm 0.73$; set four: V: $M \pm SD = 7.22 \pm 0.49$, A: $M \pm SD = 4.87 \pm 0.62$). Negative pictures featured the categories animals, body, disaster, disgust, injury, suffering, violence, and weapons. Again, valence and arousal values were comparable (set one: V = 2.18, A = 6.20; set two: V = 2.19, A = 6.20; set three: V = 2.19, A = 6.20; set four: V = 2.19, A = 6.19). A complete list of all pictures and their classification into sets can be found in Supporting Information 1. All pictures were displayed in grayscale. The mean luminosity of all 225 pictures was matched using MATLAB 2019a SHINE toolbox (MathWorks Inc.). Means of luminosity values did not differ between the picture sets ($F(8) = 0.99$, $p = .448$) (Willenbockel et al., 2010). To keep the mean luminosity constant during the whole experiment, the background was gray using the mean luminosity of the pictures.

2.1.5 | Pupillometry

Pupillary data were recorded with a SR Research Ltd. EyeLink 1,000 Plus system (SR Research, Ottawa, Ontario, Canada). The system tracked the pupil's position, corneal reflections, and the pupil size (in arbitrary units) of the right eye. Pupil diameter was recorded at a sampling rate of 500 Hz. Participants were placed in a distance of 0.60 m in front of the camera. This distance was kept constant during the whole experiment. The eye-tracking system was calibrated with a 9-point calibration prior to the experiment. During the ER task, a drift correction was performed before each block. The data were prepared and pre-analyzed using scripts (https://osf.io/a87sg/) following the data analysis of Kinner et al., (2017). Trials with less than 2,000 valid samples were identified and excluded using an R-based (https://www.r-project.org/) algorithm. Afterwards, participants with missing data in more than 50% of the trials in different conditions were excluded from further analysis. This affected 10 participants. The course of the pupil reaction is shown in arbitrary values. The value is based on the number of pixels in the eye camera image that were detected as pupil. Conversion into absolute measurement values is not possible. For visualization, pupillary data were baseline-corrected in the interval of 40–300 ms after stimulus onset.

For each condition ("active viewing-neutral", "active viewing-negative", "active viewing-positive", "suppression-negative", "suppression-positive", "distancing-negative", "distancing-positive", "choice-negative", "choice-positive"), two parameters were computed (Kinner et al., 2017): the early pupillary response was defined as the peak dilation in the first 2 s of picture presentation. A gradient was calculated between the minimum pupil diameter of the first second and the maximum pupil diameter between seconds 1 and 2 (Bradley et al., 2008; Henderson et al., 2014; Kinner et al., 2017). The gradient of the early pupillary response operationalizes cognitive effort. The late pupillary response was defined as the diameter increase within the interval of seconds 2–6 (Bradley et al., 2008; Henderson et al., 2014; Kinner et al., 2017). The area under the curve with respect to increase (AUC_I), which reflects the late pupillary response that is not confounded by the early pupillary response (e.g., Henderson et al., 2014; Kinner et al., 2017; van Reekum et al., 2007). The AUC_I of the late pupillary response operationalizes emotional arousal. Individual data were averaged across 25 trials per condition.

2.1.6 | Heart period

Electrocardiogram (ECG) was recorded from left and right forearms at 500 Hz sampling rate using Brain Vision Recorder (Brain Products Inc., Gilching, Germany). Passive Ag/AgCl sintered ring electrodes were used. Data were preprocessed using Brain Vision Analyzer (Brain Products Inc., Gilching, Germany). Heart period data were filtered using a zero-phase shift Butterworth filter (high pass: 5 Hz, low pass: 30 Hz, notch: 50 Hz). R peaks were detected using the HEPLAB plugin (Perakakis, 2019) for EEGLAB (Delorme & Makeig, 2004) in Matlab (MATLAB, 2019) and visual inspection. For each trial, temporal differences of R peaks were computed for the duration of each stimulus and a baseline time window (from 2,000 ms before stimulus onset). Afterwards, a baseline correction for each trial was conducted, yielding the mean change of the interbeat intervals (IBI) for each condition. For the sake of parsimony, heart period is from here on listed among the arousal measures, because our findings supported the notion that it captures both effort and arousal in ways that are difficult to distinguish (see Section 5.3).

2.1.7 | Psychometric measurements

After the ER experiment, participants completed a number of questionnaires. The Big Five Inventory (BFI-2) (Soto &

TABLE 1 *M* ± *SD* of subjective arousal and effort ratings, early (gradient) and late (AUC_I) pupillary response, and heart period in response to neutral, positive, and negative pictures

	Subjective arousal	Subjective effort	Gradient (per s) phys. effort	AUC_I phys. arousal	Heart period (in ms) phys. arousal
Active viewing neutral	–	–	17.4 ± 21.0	11.00 ± 9.84	8.35 ± 17.95
Active viewing negative	266.1 ± 83.8	–	31.2 ± 18.7	6.23 ± 10.69	23.26 ± 23.81
Active viewing positive	236.7 ± 74.0	–	24.2 ± 24.8	6.90 ± 9.20	10.18 ± 19.83
Suppression negative	258.7 ± 86.7	224.1 ± 97.9	30.5 ± 17.9	3.88 ± 8.68	30.97 ± 27.87
Suppression positive	207.2 ± 84.8	174.3 ± 98.1	23.7 ± 19.5	4.88 ± 10.65	21.37 ± 25.10
Distancing negative	233.0 ± 87.0	254.0 ± 93.5	30.9 ± 19.5	5.00 ± 9.71	30.86 ± 26.89
Distancing positive	181.1 ± 85.7	179.3 ± 89.3	27.2 ± 17.8	5.50 ± 9.53	17.82 ± 25.46

Note: For descriptive purpose, gradient was multiplied with 500 to show change of pupil size per second. Heart period is reported as relative change from baseline.

John, 2017; German version: Danner et al., 2019) measured broad personality traits. Dispositional use of ER was assessed using the Emotion Regulation Questionnaire (ERQ) (Gross & John, 2003; German version: Abler & Kessler, 2009) and ER ability with the Flexible Emotion Regulation Scale (FlexER) (Dörfel et al., 2019). The short form of the Need for Cognition Scale (Cacioppo & Petty, 1982; German version: Bless et al., 1994) was used to assess need for cognition (NFC), and implicit theories of willpower in emotion control were assessed using the implicit theories questionnaire from Bernecker and Job (2017).

2.1.8 | Statistics

All statistical analyses were performed using *RStudio* (version 1.4.1103) (RStudio Team, 2020) and *R* (version 4.0.3) (R Core Team, 2020) for Windows. Please see Supporting Information 2 for a complete list of all R packages used. The level of significance was set to $\alpha = .05$. Statistical tests for all hypotheses targeted in this study were conducted as stated in our preregistration (https://osf.io/9wjyp). To examine the impact of emotional pictures on arousal, an ANOVA with the factor valence (neutral, negative, and positive) for the strategy active viewing was conducted for heart period data and late pupillary response. To examine the effect of ER strategies on emotional arousal, ANOVAs with the within-subject factors strategy (active viewing, suppression, and distancing) and valence (negative and positive) were conducted for behavioral data (subjective arousal ratings) and physiological measures (late pupillary response and heart period). To test for equivalence of the early pupillary response during active viewing of neutral, positive, and negative pictures, equivalence tests for paired samples were conducted. Equivalence was assumed if the differences of gradients between positive and neutral and negative and neutral pictures had a small or zero effect size. To examine the effect of ER strategies on effort, for behavioral data (subjective effort ratings) an ANOVA with the within-subject factors strategy (suppression and distancing) and valence (negative and positive) was conducted. For physiological data (pupillary data, heart period), an ANOVA with the within-subject factors strategy (active viewing, suppression, and distancing) and valence (negative and positive) was conducted. If the assumption of sphericity was violated, Greenhouse–Geisser-corrected *p*-values and degrees of freedom were reported. For all results, the proportion of explained variance η_p^2 served as effect size and was reported. If indicated by the data, estimated marginal means were computed as post-hoc contrasts. For analyses regarding pupillometry, we conducted performance-based exclusions (e.g., Kinner et al., 2017; Urry et al., 2009) of participants using their physiological data of the condition “active viewing-neutral”. It is expected that people show relatively few physiological responses in this condition. Therefore, participants with physiological responses higher than 1.5 interquartile ranges above the third quartile of the group mean were excluded (Kinner et al., 2017). Four participants met this criterion for the early pupillary response and four other participants for the late pupillary response.

2.2 | Results

2.2.1 | Effect of emotion regulation on arousal

Subjective arousal ratings

To explore whether the ER strategies, distancing and suppression, altered the emotional feeling, an ANOVA for arousal ratings was conducted. We found a significant main effect of valence ($F_{1,109} = 45.6, p < .001, \eta_p^2 = .30$): Negative pictures were rated as significantly more arousing than positive pictures. We further found a significant main effect of strategy ($F_{1.85,202.15} = 26.9, p < .001, \eta_p^2 = .20$). Post-hoc tests showed that both suppression and distancing significantly reduced emotional arousal compared to active viewing (both *p*s < .008). Distancing was associated with reduced emotional

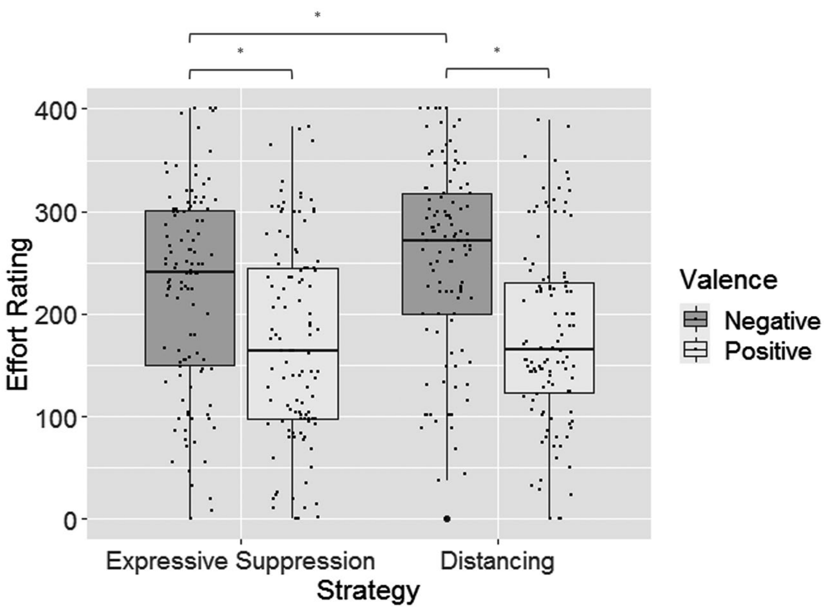

FIGURE 3 Study 1. Subjective effort ratings of the strategies suppression and distancing for negative and positive pictures, visualized as boxplots. Each dot represents the effort rating of a single subject. Bold dots represent outliers

arousal compared to suppression ($p < .001$). Moreover, we found an interaction effect ($F_{1.98,215.86} = 3.8$, $p = .024$, $\eta_p^2 = .03$). Post-hoc tests revealed a differential impact of ER strategies on emotional arousal. For negative pictures, distancing significantly reduced arousal ($p < .001$), whereas suppression did not reduce arousal ($p > .05$). For positive pictures, both distancing and suppression significantly reduced emotional arousal (all ps $< .001$; see Supporting Information 4 and Table 1).

Late pupillary response

To explore changes in AUC_I as indicator of the late pupillary response (physiological arousal), two ANOVAs were conducted. The first ANOVA only covered the strategy active viewing and the valences neutral, negative, and positive. We found a highly significant effect of valence ($F_{1.95,175.36} = 10.56$, $p < .001$, $\eta_p^2 = .11$). Post-hoc tests revealed higher AUC_I for neutral pictures compared to positive and negative pictures ($p < .001$). The second ANOVA with the factors strategy (active viewing, suppression, and distancing) and valence (negative and positive) revealed no significant main effect of valence ($F_{1,90} = 0.13$, $p > .05$, $\eta_p^2 < .01$). We further found a significant main effect of strategy ($F_{1.99,178.83} = 5.6$, $p = .004$, $\eta_p^2 = .06$), but no interaction effect ($F_{1.96,176.7} = 0.01$, $p > .05$, $\eta_p^2 < .01$). Post-hoc tests showed a significantly smaller AUC_I for suppression and distancing compared to active viewing (both ps $< .03$), but no differences between suppression and distancing ($p > .05$). This indicates a reduced arousal after regulation of emotional pictures.

Heart period

To explore changes in heart period between strategies, we conducted two ANOVAs. The first ANOVA only covered the strategy active viewing and we found a significant effect of valence (neutral, positive, and negative) ($F_{2,105} = 16.5$, $p < .001$, $\eta_p^2 = .09$). Post-hoc tests showed that viewing of negative pictures led to significantly longer IBIs, that is, an increased heart period compared to positive and neutral pictures (both ps $< .001$). There was no significant difference between neutral and positive pictures ($p > .05$). The second ANOVA covered differences between active viewing, suppression, and distancing for positive and negative pictures. Here, we found a significant main effect of valence ($F_{1,106} = 36.5$, $p < .001$, $\eta_p^2 = .026$), indicating significantly longer IBIs for positive pictures ($p < .001$). This indicates higher arousal for negative pictures. Secondly, we found a significant main effect of strategy ($F_{2,105} = 8.63$, $p < .001$, $\eta_p^2 = .026$), but no interaction effect ($F_{2,105} = 0.34$, $p > .05$, $\eta_p^2 < .01$). Post-hoc tests revealed significantly shorter IBIs, that is, a decreased heart period, during suppression and distancing compared to active viewing (both ps $< .01$). However, both regulation conditions did not differ significantly from each other ($p > .05$).

2.2.2 | Emotion regulation and effort

Subjective effort

As a first indicator for subjective effort, we inspected the choice block where participants decided their preferred strategy. We hypothesized that the participants' decisions would be driven by the effort that is required for the two strategies. A majority of $n = 60$ participants (54.5%) chose suppression, while $n = 50$ (45.5%) chose distancing. A subsequent questionnaire revealed that, indeed, 58% of participants chose the strategy that was subjectively less effortful. About 22% stated that they made the decision because the strategy was more effective for them. About 13% (n = 14) wanted a challenge and chose the strategy that felt more effortful. Interestingly, the vast majority of this group chose to reapply distancing (n = 12), although it was perceived as more effortful. An exploratory analysis showed that the choice was not influenced by order effects of the regulation blocks (see Supporting Information 5).

These patterns were also reflected in the subjective effort ratings. To evaluate effort ratings after the regulation blocks, we computed an ANOVA and found a significant main effect of valence ($F_{1,109} = 46.4$, $p < .001$, $\eta_p^2 = .30$). Participants perceived regulation of negative emotions during both strategies as more effortful (both ps < .001). Further, we found a significant main effect of strategy ($F_{1,109} = 4.25$, $p = .042$, $\eta_p^2 = .04$; see Figure 3) and a significant interaction effect of valence and strategy ($F_{1,109} = 5.6$, $p = .02$, $\eta_p^2 = .05$). Post-hoc tests revealed that for negative pictures, participants perceived suppression as less effortful than distancing ($p < .001$). For positive pictures, there was no difference between the strategies ($p > .05$).

We investigated subjective effort ratings for negative pictures exploratory for both choice groups (choice: suppression and choice: distancing). An ANOVA was computed with the within factor experimental block (suppression, distancing, and choice) and the between factor choice (repplied suppression or reapplied distancing). The ANOVA showed no significant main effect of the between factor choice ($F_{1,108} = 1.55$, $p > .05$, $\eta_p^2 = .014$), but a significant main effect of the within factor block ($F_{1.86,201.03} = 6.63$, $p = .002$, $\eta_p^2 = .058$) and a significant interaction effect ($F_{1.86,201.03} = 7.34$, $p = .001$, $\eta_p^2 = .064$). Post-hoc tests revealed that effort in the distancing block was perceived as significantly higher than in the suppression and in the choice block (both ps < .001), but only from participants that chose to reapply suppression. Participants that reapplied distancing in the choice block perceived all blocks as equally effortful (all ps > .05; see Figure 4).

Early pupillary response

Figure 5 shows the trajectory of the pupillary response over all strategies and emotions. It is clearly visible that the condition "active viewing-neutral" shows a significantly different trajectory. For possible reasons see Section 5.3.

To test whether effort of actively viewing neutral, positive, and negative pictures, was equal, equivalence tests for paired samples with gradient of the early pupillary response were conducted. Thereby, null hypotheses of statistical difference for all three comparisons (neutral—negative, neutral—positive, and negative—positive) was rejected (ps < .001), indicating equivalent gradients for all three conditions. To explore the effect of different ER strategies on effort, an ANOVA was computed with the gradient of the early pupillary response as dependent variable and strategy (active viewing, suppression, distancing) and valence (negative, positive) as independent variables. We found no significant main effect of strategy ($F_{1.81,161.300} = 1.69$, $p = .190$, $\eta_p^2 = .02$), but a highly significant effect of valence ($F_{1,90} = 15.03$, $p < .001$, $\eta_p^2 = .14$), indicating a higher gradient of the early pupillary response when viewing negative pictures compared to positive pictures.

Regarding associations with subjective effort ratings, we found a significant negative correlation between the effort rating after distancing of negative pictures and the mean gradient during distancing of both negative ($r = -.21$, $p = .044$) and positive pictures ($r = -.21$, $p = .043$).

As a brief summary of the first study: We found that distancing could significantly reduce subjective emotional arousal. However, this was not reflected in physiological measures (late pupillary response and heart period). On average, participants perceived distancing as more effortful than suppression. This was also reflected in the fact that the majority chose to reapply suppression. No significant differences between ER strategies were observed for the early pupillary response.

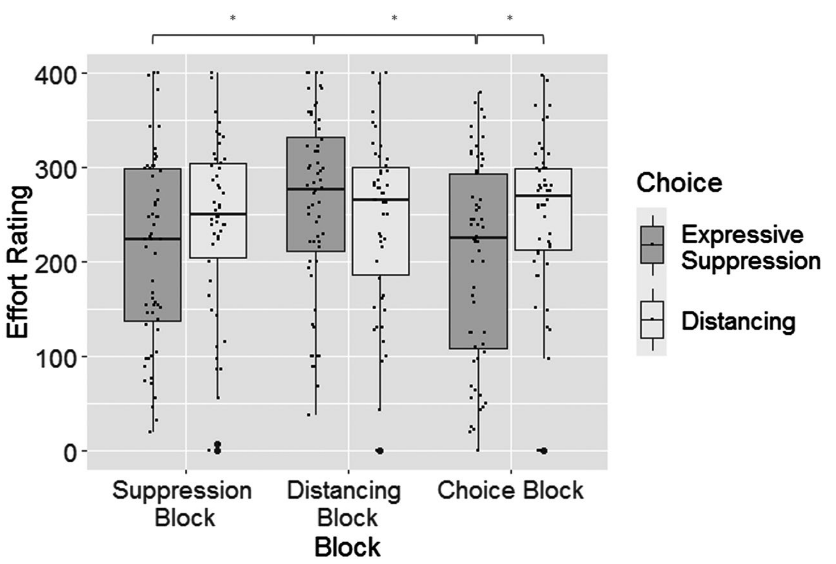

FIGURE 4 Study 1. Subjective effort ratings of the blocks suppression, distancing, and choice for negative pictures. The groups represent the choice of the participants in the last block: either to reapply suppression or to reapply distancing. Each dot represents the effort rating of a single subject. Bold dots represent outliers

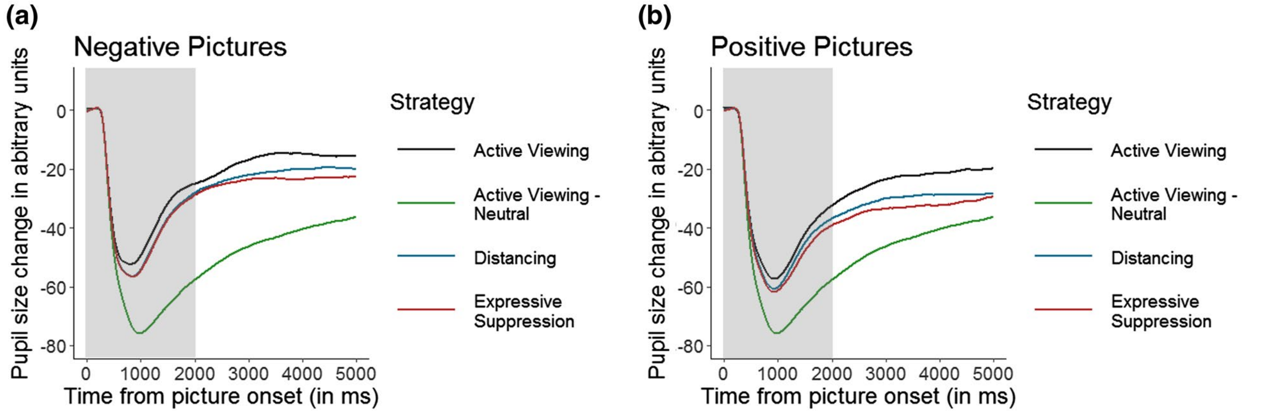

FIGURE 5 Mean pupillary responses to neutral, positive, and negative pictures in Study 1. (a) Trajectory of all the conditions “active viewing-negative”, “suppression-negative”, “distancing-negative”, and “active viewing-neutral" in arbitrary values. The gray area shows the time window of the early pupillary reaction. (b) Trajectory of all the conditions “active viewing-positive”, “suppression-positive”, “distancing-positive”, and “active viewing-neutral" in arbitrary values. The gray area shows the time window of the early pupillary reaction

3 | STUDY 2

The main goals in Study 2 were to attempt to replicate the results of Study 1 and to investigate peripheral physiological effort measures more deeply. In the first study, we used gray-scaled pictures because we wanted to control for other potential influences on the pupillary responses than arousal and effort. However, most ER research is done with colored pictures (e.g., Diers et al., 2020; Dörfel et al., 2020; Scheffel et al., 2019) and valence and arousal of colored pictures are perceived as higher than for gray-scaled pictures (Bekhtereva & Muller, 2017). Therefore, we conducted a second study investigating whether results on subjective arousal ratings and heart period also hold for colored pictures. Furthermore, we wanted to explore whether differences between ER strategies are reflected in peripheral physiological arousal and effort measures when observing colored pictures.

The study procedure, all hypotheses, and statistical analysis methods have been preregistered at https://osf.io/ytujb/. Data sets, scripts, and analysis routines can be found at https://osf.io/dk4s9. Again, “we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study” in compliance with the 21-word-solution of open science (Simmons et al., 2012). A complete list of all measures assessed in the study can be found on OSF (https://osf.io/jqy9z/).

3.1 | Method

3.1.1 | Power analysis and participants

As already mentioned, Kinner et al., (2017) found effect sizes of η_p^2 = .13 for differences in ER strategies on the early pupillary response (physiological effort). Calculation of the sample size for an ANOVA with the given effect size and a power of 0.95 resulted in $N = 24$. To prevent against replication issues and following current recommendations (Gignac & Szodorai, 2016; Open Science Collaboration, 2015; Schäfer & Schwarz, 2019), sample size calculations were conducted with half of the reported effect size (η_p^2 = .065) and resulted in a sample size of $N = 48$. To prevent against dropouts due to technical failures, a sample size of $N > 50$ was aimed at. All computations were conducted using G*Power 3 (Faul et al., 2007).

In the second study, $N = 54$ participants took part in the experiment. Pupils of two participants could not be captured with the eye-tracking system. These participants had to be excluded from further analyses. The final sample consisted of $N = 52$ healthy participants (36 females; age: 24.3 ± 6.0 years). Again, please note that sample sizes of specific analyses can differ when participants had to be excluded due to technical reasons with the eye tracker or with psychophysiological measurements. All participants stated that they had normal to corrected vision and were fluent in German.

3.1.2 | Experimental procedure

All sessions took place in a shielded cabin with constant lighting. Participants received information about the experimental procedure, signed an informed consent form, and filled out a general questionnaire on demographic data. Afterwards, the height of the table, chair, screen, eye-tracking system, and the headrest were adjusted to the person. Electrodes to measure ECG were attached and a 6-min resting measurement was conducted. Participants then received written instructions on the ER paradigm and performed a training session. The eye-tracking system was calibrated and the ER paradigm was run. Afterwards, participants filled out a variety of questionnaires and received financial compensation or course credit for their participation. The local ethics committee of the Technische Universität Dresden approved the experimental protocol (reference number: EK 286062019).

3.1.3 | Emotion regulation paradigm

The ER task followed the same block design as used in Study 1 (see Section 2.1.3). Participants were told to actively view neutral, negative, and positive pictures (see Section 3.1.4) or to regulate all upcoming emotions by means of two different ER strategies: distancing and suppression. The ER paradigm was the exact same block design as used in Study 1 (see Supporting Information 6). Additionally, in Study 2, participants rated their subjective emotional arousal for active viewing of neutral pictures and their subjective effort also for active viewing of positive and negative pictures.

3.1.4 | Stimuli

The pictures in Study 2 were again selected from the Emotional Picture Set (EmoPicS) (Wessa et al., 2010) and the International Affective Picture System (IAPS) (Lang et al., 2008). The 25 neutral pictures (V: $M \pm SD = 4.86 \pm 0.49$, A: $M \pm SD = 3.01 \pm 0.61$) remained unchanged compared to Study 1. Again, 100 positive and 100 negative pictures were clustered into four sets based on normative valence and arousal ratings and RGB color space. Therefore, the classification into the sets differed (see Supporting Information 7 for a complete list of pictures and their classification into sets). Valence and arousal were comparable for both positive pictures (set one: V: $M \pm SD = 7.30 \pm 0.46$, A: $M \pm SD = 4.73 \pm 0.69$; set two: V: $M \pm SD = 7.16 \pm 0.57$, A: $M \pm SD = 4.87 \pm 0.67$; set three: V: $M \pm SD = 7.15 \pm 0.51$, A: $M \pm SD = 5.06 \pm 0.80$; set four: V: $M \pm SD = 7.25 \pm 0.46$, A: $M \pm SD = 4.86 \pm 0.49$) and negative pictures (set one: V: $SD = 2.13 \pm 0.61$, A: $M \pm SD = 6.13 \pm 0.66$; set two: V: $M \pm SD = 2.21 \pm 0.52$, A: $M \pm SD = 6.23 \pm 0.73$; set three: V: $M \pm SD = 2.23 \pm 0.66$, A: $M \pm SD = 6.12 \pm 0.82$; set four: V: $M \pm SD = 2.18 \pm 0.57$, A: $M \pm SD = 6.19 \pm 0.66$). All pictures were displayed in color. The mean luminosity was matched separately for negative and positive pictures using an adaption of the MATLAB 2019a SHINE toolbox (MathWorks Inc.)—SHINE_color (https://osf.io/auzjy/). Because some pictures looked unreal after the colors were adjusted, they were replaced. Mean of luminosity values did not differ between the picture sets ($F_8 = 0.04$, $p > .99$). Please note that the RGB values between picture sets of different valences differed significantly (all $p < .001$), but not between picture sets of same valences. To keep the mean luminosity constant during the whole experiment, the background was gray with the mean luminosity of the pictures.

3.1.5 | Pupillometry

To record pupillary data, an SR Research Ltd. EyeLink 1000 Plus system was used (SR Research, Ottawa, Ontario, Canada). The recording and analysis of pupillary data followed the same procedure as described in Study 1 (see Section 2.1.5). In Study 2, eight participants had to be excluded because of more than 50% missing trials in one condition.

3.1.6 | Heart period

Electrocardiogram (ECG) was recorded from left and right forearms at 500 Hz sampling rate using Brain Vision Recorder (Brain Products Inc., Gilching, Germany). Recording and analysis of heart period data followed the same procedure as described in Study 1 (see Section 2.1.6)

3.1.7 | Psychometric measurements

After the ER experiment, participants completed the Big Five Inventory (BFI-2) (Soto & John, 2017; German version: Danner et al., 2019), the Emotion Regulation Questionnaire (ERQ) (Gross & John, 2003; German version: Abler & Kessler, 2009), the Flexible Emotion Regulation Scale (FlexER) (Dörfel et al., 2019), the short form of the Need for Cognition Scale (Cacioppo & Petty, 1982; German version: Bless et al., 1994), and the questionnaire for implicit theories of willpower in emotion control (Bernecker & Job, 2017).

3.1.8 | Statistics

All statistical analyses were performed using RStudio (version 1.4.1103) (RStudio Team, 2020) and R (version 4.0.3) (R Core Team, 2020) for Windows. The level of significance was set to $\alpha = .05$. Statistical tests for all hypotheses targeted in this study were conducted as stated in our preregistration (https://osf.io/9wjyp). To examine the impact of emotional pictures on arousal, an additional ANOVA with the factor valence (neutral, negative, and positive) for the strategy active viewing was conducted for behavioral data (subjective arousal) and physiological data (heart period and late pupillary response). To examine the effect of ER strategies on emotional arousal, ANOVAs with the within-subject factors strategy (active viewing, suppression, and distancing) and valence (negative and positive) were conducted for behavioral (subjective arousal ratings) and physiological measures (late pupillary response and heart period). To test for equivalence of the early pupillary response during active viewing of neutral, positive, and negative pictures, equivalence tests for paired samples were conducted. Equivalence was assumed

TABLE 2 *M* ± *SD* of self-reported arousal and effort ratings, early (gradient) and late (AUC_I) pupillary response, and heart period in response to neutral, positive, and negative pictures

	Subjective arousal	Subjective effort	Gradient (per s) phys. effort	AUC_I phys. arousal	Heart period (in ms) phys. arousal
Active viewing neutral	63.8 ± 68.1	107.0 ± 112.7	24.4 ± 20.5	10.83 ± 17.2	−26.13 ± 110.3
Active viewing negative	289.3 ± 79.1	260.4 ± 99.4	26.0 ± 40.1	10.00 ± 9.84	−1.05 ± 72.4
Active viewing positive	197.6 ± 83.7	64.7 ± 61.1	29.3 ± 21.9	9.84 ± 15.73	−9.48 ± 72.9
Suppression negative	273.8 ± 85.9	222.5 ± 111.7	25.9 ± 26.9	9.04 ± 11.56	5.27 ± 69.1
Suppression positive	184.0 ± 88.7	147.1 ± 96.0	24.8 ± 29.7	8.70 ± 12.47	−2.22 ± 51.2
Distancing negative	242.0 ± 79.2	253.8 ± 99.9	28.1 ± 24.3	6.04 ± 11.19	16.14 ± 42.5
Distancing positive	149.2 ± 90.2	159.6 ± 105.9	24.8 ± 22.1	9.23 ± 11.93	−7.65 ± 75.1

Note: For descriptive purpose, gradient was multiplied with 500 to show change of pupil size per second. Heart period is reported as relative change from baseline.

if the differences of gradients between positive and neutral and negative and neutral pictures had a small or zero effect size. To examine the effect of ER strategies on effort, for behavioral (subjective effort ratings) and for physiological data (pupillary response, heart period), ANOVAs with the within-subject factors strategy (active viewing, suppression, and distancing) and valence (negative and positive) were conducted. If the assumption of sphericity was violated, Greenhouse–Geisser-corrected *p*-values and degrees of freedom were reported. For significant results, the proportion of explained variance η_p^2 served as effect size. If indicated by the data, estimated marginal means were computed as post-hoc contrasts. Significance levels were Bonferroni corrected. For analyses regarding pupillometry, we conducted performance-based exclusions (e.g., Kinner et al., 2017; Urry et al., 2009) of participants using their physiological data of the condition active viewing-neutral. People are expected to show relatively few physiological responses in this condition. Therefore, participants with physiological responses that were more higher than 1.5 interquartile ranges above the third quartile of the group mean were excluded (Kinner et al., 2017). No participants met this criterion for the early pupillary response, but two participants for the late pupillary response.

3.2 | Results

3.2.1 | Effect of emotion regulation on arousal

Subjective arousal ratings

In the second study, we computed an ANOVA for the active viewing condition to explore the impact of valence on subjective arousal ratings. We found a highly significant effect of valence ($F_{1.97,100.63} = 128.84$, $p < .001$, $\eta_p^2 = .716$), indicating that both positive and negative pictures were perceived as more arousing than neutral pictures (both *p*s < .001). Additionally, negative pictures were perceived as more arousing than positive pictures ($p < .001$; see Supporting Information 8).

To explore changes in arousal ratings over strategies, a second ANOVA was computed. We found a significant main effect of valence ($F_{1,51} = 66.5$, $p < .001$, $\eta_p^2 = .566$). Negative pictures were rated as significantly more arousing than positive pictures. Additionally, we found a significant main effect of strategy ($F_{1.93,98.57} = 13.33$, $p < .001$, $\eta_p^2 = .207$). No interaction effect between strategy and valence was found ($p = .975$, $\eta_p^2 < .001$). Post-hoc test revealed a differential impact of ER strategies on emotional arousal. For both negative and positive pictures, distancing significantly reduced arousal (both *p*s < .001), whereas suppression did not reduce arousal (both *p*s > .05; see Supporting Information 10 and Table 3).

Late pupillary response

An ANOVA was computed to explore a possible modulation of the late pupillary response (AUC_I) by valence (neutral, positive, and negative) during active viewing. We found no significant effect of valence ($F_{1.96,80.2} = 0.10$, $p = .899$, $\eta_p^2 < .01$). To determine the modulation of the late pupillary response by strategy and valence, an ANOVA with the AUC_I of each condition was computed. We found no significant main effects of strategy or valence (*p*s ≥ .23 and $\eta_p^2 \leq .035$). Additionally, no interaction effect was observed ($p > .05$ and $\eta_p^2 = .033$). This indicates that ER had no effect on the late pupillary response (see Table 2).

Heart period

Two ANOVAS were conducted to explore effects of valence and strategies on heart period. When considering the strategy active viewing only, we once again found a significant effect of valence ($F_{2,50} = 5.5$, $p = .005$, $\eta_p^2 = .067$). During active viewing of negative pictures, heart period significantly increased as indicated by longer IBIs compared to positive and neutral pictures ($p_{pos\text{-}neg} = .011$ and $p_{neu\text{-}neg} = .016$). The second ANOVA covered changes of heart period over the strategies active viewing, suppression, and distancing for positive and negative pictures. Here, we found no effect of strategy ($F_{2,50} = 1.87$, $p = .167$, $\eta_p^2 = .035$), but again an effect of valence ($F_{1,51} = 11.49$, $p = .001$, $\eta_p^2 = .184$), indicating longer IBIs and therefore increased heart period for positive pictures ($p < .05$; see Table 2).

3.2.2 | Emotion regulation and effort

Subjective effort

Similar to Study 1, we examined the participants' choice for the last experimental block. The choice pattern was even more pronounced in Study 2. Of all $N = 52$ participants, 67.3% ($n = 35$) chose suppression and 32.7% ($n = 17$) chose distancing. This time, 63.4% chose their ER strategy because it was perceived as less effortful, while 23.1% chose it because it was perceived as more effective. Participants that rated suppression as less effortful chose suppression more often in the choice block (see Figure 7). Again, the choice was not influenced by order effects of the regulation blocks (see Supporting Information 10).

The decisions and respective reasons were again reflected in subjective effort ratings. Again, an ANOVA revealed a significant main effect of valence ($F_{1,51} = 75.14$, $p < .001$, $\eta_p^2 = .596$) indicating higher effort ratings for negative pictures for all strategies (active viewing, suppression, and distancing; all ps $< .001$). The main effect of strategy was significant ($F_{1.89,96.20} = 7.69$, $p = .001$, $\eta_p^2 = .131$; see Figure 6). Finally, we found a significant interaction effect of valence and strategy ($F_{1.77,90.07} = 25.08$, $p < .001$, $\eta_p^2 = .330$; see Figure 6). Post-hoc tests revealed significantly higher effort ratings for the strategies suppression and distancing compared to active viewing (both ps $< .001$), but no difference between suppression and distancing ($p > .05$). This only applied to positive pictures. For negative pictures, significantly higher effort ratings for active viewing ($p = .029$) compared to suppression were found, indicating a moderating role of valence on subjective effort, especially for the strategy active viewing.

Similar to Study 1, differences in effort ratings were only found for negative pictures. We, therefore, again investigated exploratory effort ratings for negative pictures more detailled for both choice groups (Choice: suppression and choice: distancing). An ANOVA was computed with the within factor experimental block (suppression, distancing, and choice) and the between factor choice (repplied suppression or reapplied distancing). The ANOVA showed neither a significant main effect of the between factor choice ($F_{1,50} = 0.28$, $p > .05$, $\eta_p^2 = .006$) nor a significant main effect of the within factor block ($F_{1.82,91.06} = 1.22$, $p > .05$, $\eta_p^2 = .024$). However, a significant interaction effect was found ($F_{1.82,91.06} = 4.70$, $p = .014$, $\eta_p^2 = .086$). Post-hoc tests revealed that effort in the distancing block was perceived as significantly higher than in the suppression and in the choice block (both ps $< .02$), but only for participants that chose to reapply suppression. Participants that reapplied distancing in the choice block perceived all blocks as equally effortful (all ps $> .05$; see Figure 7).

Early pupillary response

Figure 8 shows the trajectory of the pupillary response of all strategies and valences. It is clearly visible that the condition "active viewing-neutral" shows a significantly different trajectory. For possible reasons please see Section 5.3.

Equivalence tests were conducted to test for equal gradients of the early pupillary reaction during active viewing of neutral, negative, and positive pictures. Comparable to

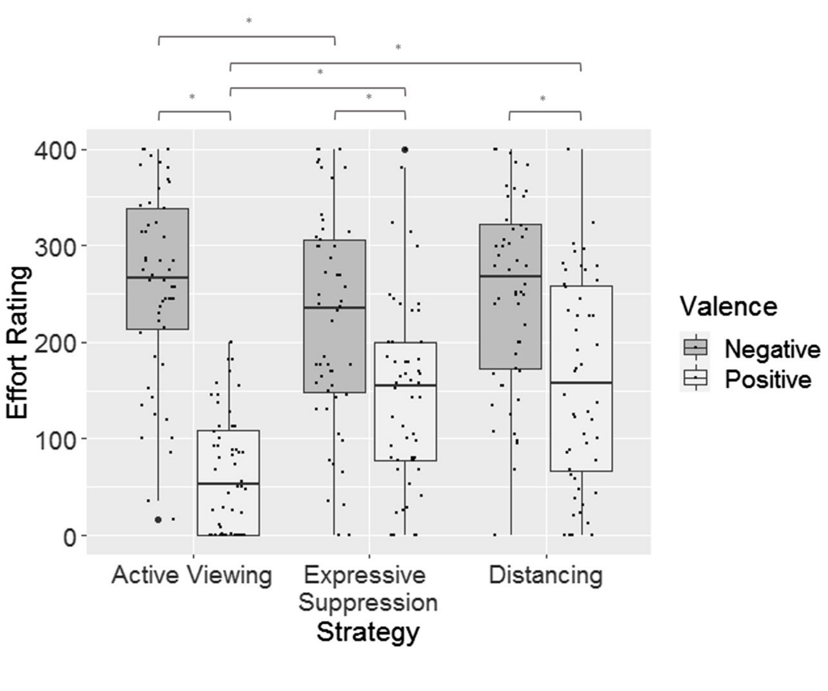

FIGURE 6 Study 2. Subjective effort ratings of the strategies suppression and distancing for negative and positive pictures, visualized as boxplots. Each dot represents the effort rating of a single subject. Bold dots represent outliers

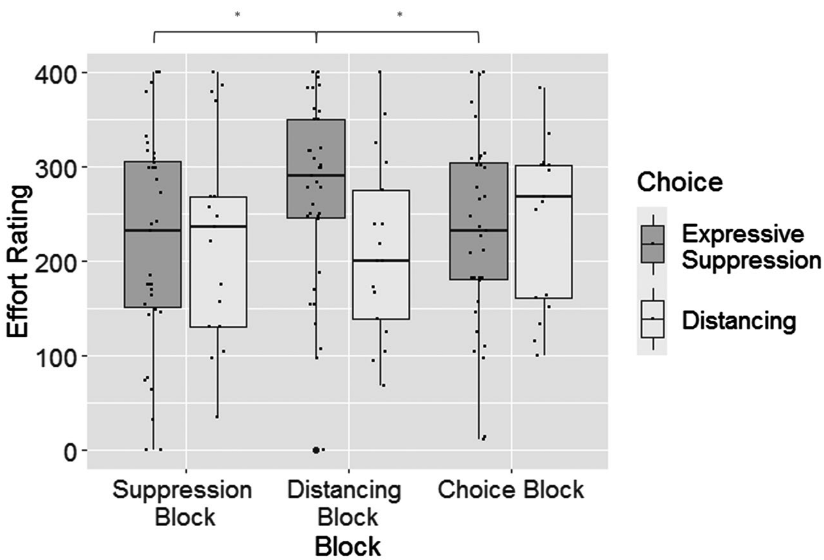

FIGURE 7 Study 2. Subjective effort ratings of the blocks suppression, distancing, and choice for negative pictures. The groups represent the choice of the participants in the last block: either to reapply suppression or to reapply distancing. Each dot represents the effort rating of a single subject. Bold dots represent outliers

Study 1, null hypotheses of statistical difference for all three comparisons were rejected (*p*s < .001), indicating equivalent gradients for all three conditions. As in Study 1, an ANOVA was computed to explore the change of the early pupillary response across strategies and valences. We neither found a significant main effect of valence ($F_{1,41} = 0.00$, $p = .960$, $\eta_p^2 < .01$), nor a significant main effect of strategy ($F_{1.72,70.50} = 0.53$, $p = .566$, $\eta_p^2 = .01$), nor a significant interaction effect ($F_{1.56,63.86} = 0.44$, $p = .596$, $\eta_p^2 = .01$).

4 | EXPLORATORY ANALYSES

The following exploratory analyses investigate associations between individual differences in arousal and effort measures (subjective ratings, choice behavior, pupillometry, and heart rate) and broad (Big Five) and narrow (habitual use of reappraisal and suppression, need for cognition, emotion regulation flexibility, implicit theories about emotion control) personality traits. We stated in the preregistrations for each study to analyze the associations separately for each sample. However, we decided to deviate from this approach. Sample sizes of $N = 110$ (Study 1) and $N = 52$ (Study 2) would be too small to calculate reliable correlations (Schönbrodt & Perugini, 2013). In the following, we therefore used combined samples to be closer to the recommended sample size for correlations of $N = 250$ and thus obtain more reliable results. The combined sample size now comprised $N = 162$ participants (112 females; age: 24.8 ± 5.9 years), $n = 95$ of whom chose suppression (64 females; age: 24.6 ± 6.2 years) and $n = 67$ of whom chose distancing (48 females, age: 24.9 ± 5.6 years) in the choice block of the experiments.

4.1 | Statistics

As a prerequisite, normal distribution and equivalence of variance between Study 1 and Study 2 of all variables of

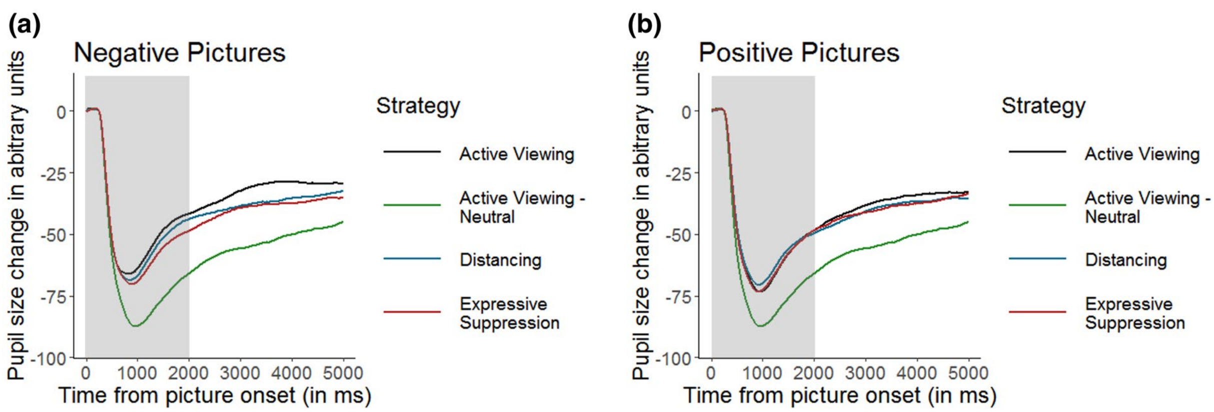

FIGURE 8 Mean pupillary responses to neutral, positive, and negative pictures in Study 2. (a) Trajectory of all the conditions “active viewing-negative”, “suppression-negative”, “distancing-negative”, and “active viewing-neutral" in arbitrary values. The gray area shows the time window of the early pupillary reaction. (b) Trajectory of all the conditions “active viewing-positive”, “suppression-positive”, “distancing-positive”, and “active viewing-neutral" in arbitrary values. The gray area shows the time window of the early pupillary reaction

interest (Supporting Informations 11, 12, and 13) were tested. Further, we computed equivalence tests (for two independent samples) to see if values of variables (subjective arousal and effort ratings and questionnaire data) were comparable in both studies. Because the effect of picture coloring on physiological measures (pupillometry and heart period) remains unknown, an equivalence test for the baseline period was calculated for these measures. Equivalence for all measures was assumed if the differences in means between studies had a small or zero effect size (for results see Supporting Information 14).

4.1.1 | Associations between measures of arousal, effort, and personality traits

Because equivalence tests did not confirm equivalence of all measures for both experiments (see Supporting Informations 14, 15, and 16), partial correlations were computed to control for differences between Study 1 and Study 2. Further, correlations were computed separately for associations between subjective (ratings) and physiological (late pupillary response and heart period) arousal measures and for associations between the subjective effort measure (ratings), the physiological effort measure (early pupillary response), and personality traits. Significance levels were Bonferroni corrected for multiple comparisons (for arousal measures: $\alpha = .05/6 = .008$ because of six conditions; for effort measures: $\alpha = .05/4 = .012$ because of four conditions; for associations with personality traits: $\alpha = .05/10 = .005$ because of ten personality measures). Please note that the sample size for correlations was somewhat smaller ($N = 145$), because only cases with complete data were included.

4.1.2 | Characterization of choice groups

We wanted to examine how individuals differ according to their decision for distancing or suppression in the last block. Therefore, we explored whether the two groups differed in broad (Big Five) and narrow (habitual use of reappraisal and suppression, and need for cognition) personality traits. For normally distributed variables, independent samples *t* tests (choice: suppression and choice: distancing) were computed. For variables that were not normally distributed, Mann–Whitney *U* tests were conducted. Holm correction (Holm, 1979) was used to correct for multiple comparisons.

Lastly, we explored whether the decision in the last ER block could be predicted by effort measures for positive and negative pictures or personality traits. Because there were many predictor variables and overfitting of the regression model was likely, we computed a full model and compared it with a stepwise logistic regression model. Therefore, we randomly divided the sample into a training set (80%) and a test set (20%).

4.2 | Results

4.2.1 | Associations between measures of arousal, effort, and personality traits

We found no associations between subjective arousal ratings and physiological arousal measures (late pupillary response: all $ps > .008$, heart period: all $ps > .008$). Heart period and late pupillary response were also unrelated (all $ps > .008$). We also found no significant associations between subjective effort ratings and physiological effort measures (early pupillary response: all $ps > .012$). Moreover, we found no significant associations between personality traits and subjective effort ratings (all $ps > .005$). In contrast, the association between heart period (physiological arousal) during distancing of negative pictures and lay beliefs was significant ($r = .25$, $p = .002$) and we found a positive correlation between Neuroticism and physiological effort (early pupillary response) for suppression of positive emotions ($r = .24$, $p = .004$).

4.2.2 | Characterization of choice groups

There were no significant differences in the mean values of different personality traits between groups (see Table 3). That is, participants who chose suppression did not differ in relevant personality traits from participants who chose distancing.

When predicting the participants' choice by personality traits, subjective effort ratings, and physiological effort (early pupillary response), we found the following: The decision was significantly predicted by the effort rating for suppression of negative pictures ($p = .024$) and the effort rating for distancing of negative pictures ($p = .034$). Although the physiological effort for suppression of negative pictures ($p = .057$), the physiological effort for distancing of negative pictures ($p = .152$), and ERQ reappraisal score ($p = .157$) were not significant, those variables are represented in the stepwise logistic regression (see Table 4).

A positive estimate for effort rating "suppression-negative" indicated that with higher self-reported effort after using suppression the participant was more likely to choose distancing for the last block. Vice versa, a negative estimate for effort rating "distancing-negative" indicated that the participant was more likely to choose suppression if self-reported effort after the distancing block was higher. However, the model fit was relatively low with an accuracy of 37.5%. Please note that the model fit and the obtained significant predictors were limited to the random selection of the training sample.

TABLE 3 Comparison of personality traits between groups choice: suppression and choice: distancing

Variable	Choice: suppression		Choice: distancing				
	M	*SD*	*M*	*SD*	*t*(160)	*W*	p_{cor}
ERQ-suppression	3.46	1.23	3.54	1.04	–	2,966.5	1.0
ERQ-reappraisal	4.75	0.94	4.74	0.82	−0.05	–	1.0
BFI-openness	3.80	0.60	3.80	0.63	–	3,135	1.0
BFI-conscientiousness	3.56	0.67	3.50	0.61	–	3,429.5	1.0
BFI-extraversion	3.44	0.68	3.33	0.60	0.99	–	1.0
BFI-agreeableness	3.85	0.54	3.85	0.56	–	3,038	1.0
BFI-neuroticism	2.69	0.73	2.71	0.65	–	3,049	1.0
Need for cognition	14.4	13.9	16.5	12.8	–	2,889	1.0
Flexible ER	3.08	0.33	3.13	0.24	–	3,025	1.0
Implicit theories	3.31	0.53	3.23	0.48	–	3,471	1.0

Note: Test to compare differences between groups was either independent samples *t* test or Mann–Whitney *U* test, depending on normal distribution of variables. *P*-values were corrected with Holm correction.

5 | DISCUSSION

The present research investigated the effort required by two ER strategies: suppression and distancing. Effort and arousal were studied comprehensively in two studies and from multiple perspectives—subjectively, and objectively via psychophysiological measures (pupillometry and heart period). In Study 1, emotional arousal during viewing negative and positive pictures was successfully reduced by the use of suppression and distancing as indicated by subjective arousal ratings and, in part, by physiological arousal. Subjectively, distancing reduced arousal more effectively than suppression, which is in line with existing literature (Webb et al., 2012). Subjective effort differed between strategies and picture valence: Regulation of negative pictures was perceived as more effortful than regulation of positive pictures. On average, suppression was perceived as more effortful compared to distancing, as indicated by subjective effort ratings. This subjective perception was not confirmed by physiological measures (early pupillary reaction). Moreover, about two thirds of the participants chose to reapply suppression, mainly because it was perceived as less effortful. In Study 2, these results were fully replicated. Again, both strategies reduced subjective arousal, with distancing being more effective. On average, suppression was perceived less effortful. This was also reflected in the fact that two thirds of the participants chose to reapply suppression; the most frequently cited reason again was effort. The results of the subjective ratings were again not reflected in physiological measures (pupillometry and heart period). However, there were large individual differences in the perceived effort. These differences were not explained by broad and narrow personality traits in an exploratory analysis.

5.1 | Effort plays a crucial role in the selection of an ER strategy

In line with our hypotheses, most of the participants chose the strategy that was less effortful, but not the strategy that could more effectively regulate their emotional arousal. Against our expectations, for almost two thirds in both samples, the less effortful alternative was suppression. Based on the process model of ER (Gross, 1998, 2014), suppression should

TABLE 4 Results of step wise logistic regression for predicting the choice in the last regulation block

	b (SE)	*SD*	*Z*	*p*
Effort rating "suppression-negative"	0.005	0.002	2.26	.024*
Effort rating "distancing-negative"	−0.005	0.002	−2.12	.034*
Early pupillary response "suppression-negative"	−13.98	7.330	−1.91	.057
Early pupillary response "distancing-negative"	10.84	7.576	1.43	.152
ERQ reappraisal	−0.39	0.219	−1.42	.157

Note: McFadden $R^2 = 0.1306$.

*Significant ($p < .05$).

be more effortful, because it takes place relatively late in the ER process. Individuals that use suppression have to alter their automatic emotional responding constantly. Supporting evidence was found by Richards and Gross (2000), describing poorer memory performance after suppression compared to cognitive reappraisal. This poorer performance was interpreted as a result of higher cognitive demands and effort. In study 2 of Richards and Gross (2000), low- and high-emotion slide sets were used. On closer inspection, the authors found overall poorer memory performance on high emotional picture sets, indicating higher required effort for both strategies. This corresponds to later considerations by Sheppes and Gross (2011). Highly intense emotional information is not fully modified by a neutral interpretation. As specified in the extended process model of ER (Gross, 2015a, 2015b), during ER a cycling process is run through. A constant valuation is made. This means it is being assessed whether a neutral interpretation of the emotional information succeeded. If not, another action is implemented. Typically, if a strategy was not successful, a new strategy should be selected (Gross, 2015b). However, in our investigation, this was not possible for the participants as they had to use a specific strategy in each block. Their only possibility was to continue to use the instructed strategy and try harder. This could have led to increased effort associated with distancing that was afterwards reported in subjective ratings. Moreover, the implementation of distancing was even more difficult because we used high-intense negative picture sets in order to reliably evoke peripheral physiological responses.

Difficulties in the implementation of reappraisal have at least been reported for patients with bipolar disorder (Gruber et al., 2012). Bipolar participants reported more effort in regulation attempts, but lower success than healthy controls. In line with the argumentation of Gruber et al., (2012), it is possible that participants in our study were not able to fully regulate their emotions by means of distancing, because intensity of our pictures was relatively high (Sheppes & Gross, 2011). Thus, they overestimated the effort for distancing, which resulted in higher effort ratings for this strategy compared to suppression. It is worth mentioning that effort itself is perceived as aversive (Inzlicht et al., 2018; Kurzban, 2016) and is known to evoke self-reported negative emotions like anxiety, frustration, stress, or fatigue (Inzlicht & Al-Khindi, 2012; Inzlicht et al., 2018; Saunders et al., 2015). Therefore, it is possible that first, subjective arousal ratings are confounded with aversive consequences of the spent effort. However, this is contradicted by the fact that the arousal ratings of distancing were lower, even though this strategy was perceived as more effortful.

In their current review and meta-analysis on ER choice by Matthews et al., (2021), the authors examined 18 determinants of emotion regulation choice that were categorized as being affective, cognitive, motivational, individual and sociocultural in nature. In line with their proposed framework, effort belongs to cognitive determinants, and intensity, arousal and valence of our stimuli to affective determinants. In addition, motivational determinants such as goals and anticipation, effects of age as individual determinants, and social and cultural context seem to be associated with whether and how people chose to regulate their emotions. Some of those factors have not yet received sufficient empirical attention. Clearly, more research is needed to disentangle the aspects of the individual doing the regulation, the emotion that is being regulated, the immediate situation and broader context of the regulation. We add to this research and provide further evidence, especially for the cognitive determinants, by showing that people chose their ER strategy not primarily based on the effectiveness, but the perceived cognitive effort of a given strategy.

5.2 | Individual differences in personality traits do not explain individual differences in arousal and effort

We did not find any associations between individual differences in arousal, effort, and personality traits. It has been reported before that personality traits do not predict individual differences in neural and behavioral correlates of ER (Scheffel et al., 2019). Interestingly, not even habitual use of reappraisal and suppression as narrow personality traits (Gross & John, 2003) were associated with the decision to reapply one of the strategies. A possible reason is the experimental setting. In everyday life, people can flexibly use a variety of ER strategies to achieve a goal in a given situation (Aldao et al., 2015; Kobylinska & Kusev, 2019). Even further, recent evidence suggests that it is useful to combine two ER strategies, especially for suppression (Thuillard & Dan-Glauser, 2020). However, flexible choice of strategies or a combination of ER strategies was not possible in our experimental setting. A defined ER strategy was specified ahead of each block and instructions on how to regulate have to be followed. A "strong" situation is created, minimizing individual differences (Mischel, 1968; Scheffel et al., 2019). However, this does not hold for considerations regarding the last choice block. A situation was created where participants could freely decide between two strategies. And indeed, there were individual differences between participants resulting in about one third of the participants choosing distancing. But again, the two choice groups did not differ in any of the broad or narrow personality traits and a prediction of the decision was only reasonable with the effort ratings of the previous blocks. However, a brief calculation did not reveal any plausible and meaningful associations between subjective effort and arousal ratings of the choice block and personality traits. Based on considerations from flexible ER (Aldao et al., 2015; Bonanno & Burton, 2013; Kobylinska & Kusev, 2019), ER strategies have to be useful for a given goal.

In our context, the external goal was only to use one of the two strategies. Although the internal goal of most people is to avoid negative emotions (English et al., 2017; Tamir, 2016), at the end of a rather long paradigm, the internal goal of most of the participants may have been to expend as little effort as possible. Participants may, therefore, have chosen a strategy that, on the one hand, regulates emotions reasonably well, but, on the other hand, requires as little effort as possible. For most participants, strategy was suppression.

Our research suggests that distancing is more effortful than suppression, although it is more effective in reducing subjective arousal. This provides important directions for future research. As effort is often experienced as aversive (Inzlicht et al., 2018), this could influence the choice of strategies in a given situation. A solution is offered by the cortical effort model (Neubauer & Fink, 2009): Well-trained individuals are more capable to recruit the required neural circuits to perform a task. This should be considered for more effortful strategies like distancing, so that aversive effort does not interfere with regulatory goals. Specifically, training ER strategies that are perceived as more effortful but reduce subjective arousal more effectively might benefit ER by reducing perceived effort over time.

5.3 | Subjective arousal and effort are not reflected in physiological arousal and effort

5.3.1 | Late pupillary response

In the present studies, results regarding the late pupillary response were ambiguous. In Study 1, the late pupillary response was lower during ER compared to active viewing. In Study 2, we found no differences between strategies and thus were not able to replicate this effect. These results are in line with findings of Kinner et al., (2017). They only found larger pupil sizes after increasing emotions, but no changes after downregulation of emotions. In the literature, even increased pupil sizes during cognitive reappraisal have been reported (Johnstone et al., 2007; van Reekum et al., 2007; Urry et al., 2009). We assume that the late pupillary response is not suitable to capture physiological regulation success. We speculate that the late pupillary response also reflects a certain amount of cognitive, regulatory effort that arises during ER and overrides the effect of reduced arousal during ER. This assumption is supported by the consistent finding that unregulated emotional arousal is associated with larger pupillary response (e.g., Bradley et al., 2008; Kinner et al., 2017). However, we could not confirm this in our data, which is probably due to the fixed order of the first two blocks resulting in a clearly altered pupillary reaction to neutral pictures. Depending on the research question, future studies should consider changing the block order or abandoning the block design altogether.

5.3.2 | Early pupillary response

Against our expectations, analyses of physiological effort measurements (pupillometry and heart period) yielded no significant results. Not only were differences between ER strategies in subjective effort not reflected in physiological effort measures, but we also failed to replicate findings of Kinner et al., (2017)The sentence "Not only were differences…findings of Kinner et al., (2017)" is not clear. Please check. They showed that the early pupillary response reflected cognitive effort of ER. Emotions unfold after seconds (Gross, 2015a), so the time period of the early pupillary response should be unaffected by the arousal. However, in our investigation, there were no significant differences between conditions. Further factors besides cognitive effort and emotional arousal influence pupillary responses: The pupil diameter is also sensitive to habituation and fatigue (Tryon, 1975). Therefore, our block design may have had an influence on the early pupillary response. One indication is the trajectory of the pupillary response to neutral pictures, which were always presented at the beginning of the experiment. Regulation conditions took place after the two active viewing blocks that served as baseline conditions. For example, Querino et al., (2015) showed that differences in pupil diameter between automatic and effortful parts of the Five Digit Task (FDT) diminish when considering only the second blocks of these parts. In other words, the pupillary response could not be sensitive for differences in effort in later parts of the experiment (for further evidence see Takeuchi et al., 2011). Additionally, subjective cognitive effort decreases over time (Yeo & Neal, 2008). Moreover, habituation effects on effort mobilization have been described for cardiovascular responses in the context of affective stimuli (Silvestrini & Gendolla, 2011). However, with our design and our database, we cannot make a well-founded statement about whether habituation effects may have played a role here. This should be investigated with an appropriate design.

A further influence on the pupillary response is memory load (Peysakhovich et al., 2015). In our experiments, participants saw the instructions only at the beginning of each block. For all following pictures, the strategies and their possible implementation had to be kept in mind. Therefore, the influence of memory load on the early pupillary response remains unclear specifically in our study.

5.3.3 | Heart period

During active viewing, both studies found longer IBIs during negative than positive or neutral pictures. However, when including all strategies, both studies found longer IBIs during positive than negative pictures. Furthermore, an effect of strategy on HP could not be replicated in Study 2. Since we aimed to contribute to a distinction of HP as a measure of effort and arousal, we will focus on the findings of subjective effort an arousal that were homogenous between both studies. There are two possible explanations for our findings: First, both emotional arousal (Bradley et al., 2008) and cognitive effort (Hajcak et al., 2003; Hoshikawa & Yamamoto, 1997; Inzlicht et al., 2018; Silvestrini, 2017) evoke activity in the sympathetic nervous system (SNS), complicating a distinction of both measures. This is supported by our finding that distancing was generally perceived as more effortful and more arousal-reducing than suppression, which suggests that a possible effect of strategy on HP is overlaid by how effortful *and* arousal reducing a strategy is, not how effortful *or* arousal reducing. Furthermore, both studies found higher effort, higher arousal, and shorter IBIs for negative pictures when including all strategies. This supports the argument of a joint effect of effort and arousal on HP, since here, both effort *and* arousal are high, increasing the heart rate. And secondly, when only including active viewing, negative pictures might have led to longer IBIs due to an initial 'freezing' response, a period of orientation towards aversive stimuli, which has been found by Bradley et al., (2001), Lang et al., (2000), and Adenauer et al., (2010) before. According to these studies, a decrease in IBIs sets in after several seconds of longer IBIs, which would also exacerbate the difference between positive and negative pictures in our studies due to the randomized order of the stimuli. Adenauer et al., (2010) even suggest that this 'freezing' response is a marker of healthy stimulus processing, because resilient individuals showed the same pattern of HP change as healthy controls, whereas PTSD patients showed an immediate increase in IBIs.

6 | LIMITATIONS

Several limitations should be considered when interpreting findings of the present study. First, the results regarding subjective arousal and effort ratings must be viewed with caution. Subjective ratings were recorded retrospectively at the end of each block. It is known that affect labeling can attenuate emotional experiences (Lieberman et al., 2007; Torre & Lieberman, 2018). Thus, we chose our approach to ensure that self-evaluative cognitive processes do not interfere with the experience of arousal during ER. However, memory processes might have an influence on subjective ratings in an unknown manner. The same could apply to retrospective effort ratings. Furthermore, we did not compare the subjective arousal ratings with the normative ratings of the IAPS and EmoPicS, so we cannot exclude the possibility that the lack of associations with subjective arousal result from a different perception of the stimuli in our sample. A decrease in arousal ratings just a few years after the normative ratings were collected could be shown (Libkuman et al., 2007), as well as cultural differences in arousal perception of the IAPS pictures have been shown before (Ribeiro et al., 2005).

Second, as already discussed, the block design with the fixed two active viewing blocks at the beginning could have influenced subjective and physiological measures, for example, the trajectory of the pupillary response. However, the ER choice block ought to take place immediately after the regulation blocks in order to have the impression of effectiveness and effort of both strategies in mind. Further research should examine whether order effects influence subjective and physiological measures. In the study of Kinner et al., (2017), an event-related design was used. Additionally, the pupillary response in later blocks could be influenced by habituation and fatigue (Tryon, 1975). However, the block design created a situation where participants could make a decision before the last block, thereby influencing the procedure and creating a sense of novelty. Moreover, we compared trajectories during distancing and expressive suppression not only with active viewing, but also among each other. The influence of fatigue was mitigated by the random order of the blocks. Furthermore, the sample must be taken into account when interpreting the results. Our sample primarily consisted of psychology students, so variance restrictions in ER might be possible. Moreover, in the study of Kinner et al., (2017), the sample consisted only of women because they respond more strongly to emotional stimuli (Kinner et al., 2017; Lithari et al., 2010). We included women and men, which may have led to smaller effects. However, a brief exploratory analysis showed no effect of gender on AUC_I or gradients in our data (please see OSF analyses scripts for more details).

Finally, as Sheppes et al., (2011) found, the emotional intensity of stimuli has an impact on which strategy is preferably used. In the present study, the arousal values of the picture sets were comparably high. Therefore, future studies should investigate whether the choice of ER strategies changes when using picture sets with lower arousal values.

7 | CONCLUSION

Against our hypotheses, distancing as the most effective form of reappraisal was not perceived as less effortful. Most participants felt less exhausted after applying suppression which resulted in the decision of about two thirds of the participants to reapply this strategy. However, differences in subjective effort were not reflected in the early pupillary response. Future research should elaborate the influence of design on the early pupillary response and identify possible moderators. Interestingly, individual differences in subjective effort ratings and physiological effort measures could not be explained by broad or narrow personality traits.

The present investigation adds evidence on how ER strategies differ in their required effort. More important, however, is the observation that the decision for or against an ER strategy was not primarily based on the effectiveness of an ER strategy, but by the perceived cognitive effort. A more detailed exploration of the effort of ER strategies, therefore, seems promising. Our work highlights the importance of adding subjective components to objective measures of effort in ER research. Future research should include further ER strategies and focus on the influence of high and low stimulus arousal on required effort.

CONFLICTS OF INTEREST

The authors have nothing to disclose.

AUTHOR CONTRIBUTIONS

Christoph Scheffel: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Software; Visualization; Writing-original draft; Writing-review & editing. **Sven-Thomas Graupner:** Formal analysis; Software; Visualization; Writing-review & editing. **Anne Gärtner:** Conceptualization; Writing-review & editing. **Josephine Zerna:** Formal analysis; Investigation; Writing-review & editing. **Alexander Strobel:** Conceptualization; Methodology; Writing-review & editing. **Denise Dörfel:** Conceptualization; Methodology; Writing-review & editing.

ORCID

Christoph Scheffel https://orcid.org/0000-0001-5963-9229
Anne Gärtner https://orcid.org/0000-0003-4296-963X
Josephine Zerna https://orcid.org/0000-0003-2892-884X
Alexander Strobel https://orcid.org/0000-0002-9426-5397
Denise Dörfel https://orcid.org/0000-0002-5632-419X

REFERENCES

Abler, B., & Kessler, H. (2009). Emotion Regulation Questionnaire—A German version of the ERQ by Gross and John. *Diagnostica*, *55*(3), 144–152. https://doi.org/10.1026/0012-1924.55.3.144

Adenauer, H., Catani, C., Keil, J., Aichinger, H., & Neuner, F. (2010). Is freezing an adaptive reaction to threat? Evidence from heart rate reactivity to emotional pictures in victims of war and torture. *Psychophysiology*, *47*(2), 315–322. https://doi.org/10.1111/j.1469-8986.2009.00940.x

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

Aldao, A., Sheppes, G., & Gross, J. J. (2015). Emotion regulation flexibility. *Cognitive Therapy and Research*, *39*(3), 263–278. https://doi.org/10.1007/s10608-014-9662-4

Alnaes, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4), 1–20. https://doi.org/10.1167/14.4.1

Andreassi, J. L. (2000). Pupillary response and behavior. In J. L. Andreassi (Ed.), *Psychophysiology: Human behavior & physiological response* (pp. 218–233). Lawrence Erlbaum Associates.

Austin, M. A., Riniolo, T. C., & Porges, S. W. (2007). Borderline personality disorder and emotion regulation: Insights from the Polyvagal Theory. *Brain and Cognition*, *65*(1), 69–76. https://doi.org/10.1016/j.bandc.2006.05.007

Balzarotti, S., Biassoni, F., Colombo, B., & Ciceri, M. R. (2017). Cardiac vagal control as a marker of emotion regulation in healthy adults: A review. *Biological Psychology*, *130*, 54–66. https://doi.org/10.1016/j.biopsycho.2017.10.008

Bekhtereva, V., & Müller, M. M. (2017). Bringing color to emotion: The influence of color on attentional bias to briefly presented emotional images. *Cognitive Affective & Behavioral Neuroscience*, *17*(5), 1028–1047. https://doi.org/10.3758/s13415-017-0530-z

Bernecker, K., & Job, V. (2017). Implicit theories about willpower in resisting temptations and emotion control. *Zeitschrift für Psychologie*, *225*(2), 157–166. https://doi.org/10.1027/2151-2604/a000292

Bless, H., Wanke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for cognition—a scale measuring engagement and happiness in cognitive tasks. *Zeitschrift für Sozialpsychologie*, *25*(2), 147–154.

Bonanno, G. A., & Burton, C. L. (2013). Regulatory flexibility: An individual differences perspective on coping and emotion regulation. *Perspectives on Psychological Science*, *8*(9), 591–612. https://doi.org/10.1177/1745691613504116

Bonanno, G. A., & Keltner, D. (1997). Facial expressions of emotion and the course of conjugal bereavement. *Journal of Abnormal Psychology*, *106*(1), 126–137. https://doi.org/10.1037//0021-843x.106.1.126

Bouffard, M. A. (2019). The pupil. *Continuum: Lifelong Learning in Neurology*, *25*(5), 1194–1214. https://doi.org/10.1212/CON.0000000000000771

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., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607. https://doi.org/10.1111/j.1469-8986.2008.00654.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

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

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037//0022-3514.42.1.116

Danner, D., Rammstedt, B., Bluemke, M., Lechner, C., Berres, S., Knopf, T., & John, O. P. (2019). The German Big Five Inventory 2: Measuring five personality domains and 15 facets. *Diagnostica*, *65*(3), 121–132. https://doi.org/10.1026/a000002

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Diers, K., Dörfel, D., Gärtner, A., Schönfeld, S., Walter, H., Strobel, A., & Brocke, B. (2020). Should we keep some distance from distancing? Regulatory and post-regulatory effects of emotion downregulation. *PsyArXiv*. https://doi.org/10.31234/osf.io/chw4v

Dörfel, D., Gärtner, A., & Scheffel, C. (2020). Resting state cortico-limbic functional connectivity and dispositional use of emotion regulation strategies: A replication and extension study. *Frontiers in Behavioral Neuroscience*, *14*, 128. https://doi.org/10.3389/fnbeh.2020.00128

Dörfel, D., Gärtner, A., & Strobel, A. (2019). A new self-report instrument for measuring emotion regulation flexibility. In *Society for Affective Science (SAS) Annual Conference*.

Dörfel, D., Lamke, J. P., Hummel, F., Wagner, U., Erk, S., & Walter, H. (2014). Common and differential neural networks of emotion regulation by detachment, reinterpretation, distraction, and expressive suppression: A comparative fMRI investigation. *NeuroImage*, *101*, 298–309. https://doi.org/10.1016/j.neuroimage.2014.06.051

Dragomir, A. I., Gentile, C., Nolan, R. P., & D'Antono, B. (2014). Three-year stability of cardiovascular and autonomic nervous system responses to psychological stress. *Psychophysiology*, *51*(9), 921–931. https://doi.org/10.1111/psyp.12231

Engen, H., & Kanske, P. (2013). How working memory training improves emotion regulation: Neural efficiency, effort, and transfer effects. *Journal of Neuroscience*, *33*(30), 12152–12153. https://doi.org/10.1523/JNEUROSCI.2115-13.2013

English, T., Lee, I. A., John, O. P., & Gross, J. J. (2017). Emotion regulation strategy selection in daily life: The role of social context and goals. *Motivation and Emotion*, *41*(2), 230–242. https://doi.org/10.1007/s11031-016-9597-z

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Gianaros, P. J., Van Der Veen, F. M., & Jennings, J. R. (2004). Regional cerebral blood flow correlates with heart period and high-frequency heart period variability during working-memory tasks: Implications for the cortical and subcortical regulation of cardiac autonomic activity. *Psychophysiology*, *41*(4), 521–530. https://doi.org/10.1111/1469-8986.2004.00179.x

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

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

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

Gross, J. J. (2014). Emotion regulation: Conceptual and empirical foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (2nd ed., pp. 3–22). The Guilford Press.

Gross, J. J. (2015a). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, *26*(1), 1–26. https://doi.org/10.1080/1047840x.2014.940781

Gross, J. J. (2015b). The extended process model of emotion regulation: Elaborations, applications, and future directions REPLY. *Psychological Inquiry*, *26*(1), 130–137. https://doi.org/10.1080/1047840x.2015.989751

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

Gruber, J., Harvey, A. G., & Gross, J. J. (2012). When trying is not enough: Emotion regulation and the effort-success gap in bipolar disorder. *Emotion*, *12*(5), 997–1003. https://doi.org/10.1037/a0026822

Haga, S. M., Kraft, P., & Corby, E. K. (2009). Emotion regulation: Antecedents and well-being outcomes of cognitive reappraisal and expressive suppression in cross-cultural samples. *Journal of Happiness Studies*, *10*(3), 271–291. https://doi.org/10.1007/s10902-007-9080-3

Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology*, *40*(6), 895–903. https://doi.org/10.1111/1469-8986.00107

Harenski, C. L., Kim, S. H., & Hamann, S. (2009). Neuroticism and psychopathy predict brain activation during moral and nonmoral emotion regulation. *Cognitive Affective & Behavioral Neuroscience*, *9*(1), 1–15. https://doi.org/10.3758/CABN.9.1.1

Henderson, R. R., Bradley, M. M., & Lang, P. J. (2014). Modulation of the initial light reflex during affective picture viewing. *Psychophysiology*, *51*(9), 815–818. https://doi.org/10.1111/psyp.12236

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

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. http://www.jstor.org/stable/4615733

Hoshikawa, Y., & Yamamoto, Y. (1997). Effects of Stroop color-word conflict test on the autonomic nervous system responses. *American Journal of Physiology*, *272*(3 Pt 2), H1113–H1121. https://doi.org/10.1152/ajpheart.1997.272.3.H1113

Inzlicht, M., & Al-Khindi, T. (2012). ERN and the placebo: A misattribution approach to studying the arousal properties of the error-related negativity. *Journal of Experimental Psychology-General*, *141*(4), 799–807. https://doi.org/10.1037/a0027586

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Science*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

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), 1301–1333. https://doi.org/10.1111/j.1467-6494.2004.00298.x

Johnstone, T., van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. (2007). Failure to regulate: Counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *Journal of Neuroscience*, *27*(33), 8877–8884. https://doi.org/10.1523/JNEUROSCI.2063-07.2007

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investigation*, *15*(3), 235–245. https://doi.org/10.30773/pi.2017.08.17

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

Kobylinska, 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

Kokkonen, M., & Pulkkinen, L. (2001). Examination of the paths between personality, current mood, its evaluation, and emotion regulation. *European Journal of Personality*, *15*(2), 83–104. https://doi.org/10.1002/per.397

Kreis, I., Moritz, S., & Pfuhl, G. (2020). Objective versus subjective effort in schizophrenia. *Frontiers in Psychology*, *11*, 1469. https://doi.org/10.3389/fpsyg.2020.01469

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

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology*, *7*, 67–70. https://doi.org/10.1016/j.copsyc.2015.08.003

Lane, R. D., McRae, K., Reiman, E. M., Chen, K., Ahern, G. L., & Thayer, J. F. (2009). Neural correlates of heart rate variability during emotion. *NeuroImage*, *44*(1), 213–222. https://doi.org/10.1016/j.neuroimage.2008.07.056

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

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*(3), 137–159. https://doi.org/10.1016/s0165-0327(00)00343-8

Larsen, R. S., & Waters, J. (2018). Neuromodulatory correlates of pupil dilation. *Frontiers in Neural Circuits*, *12*, 21. https://doi.org/10.3389/fncir.2018.00021

Libkuman, T. M., Otani, H., Kern, R., Viger, S. G., & Novak, N. (2007). Multidimensional normative ratings for the International Affective Picture System. *Behavior Research Methods*, *39*(2), 326–334. https://doi.org/10.3758/bf03193164

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words—Affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science*, *18*(5), 421–428. https://doi.org/10.1111/j.1467-9280.2007.01916.x

Lithari, C., Frantzidis, C. A., Papadelis, C., Vivas, A. B., Klados, M. A., Kourtidou-Papadeli, C., Pappas, C., Ioannides, A. A., & Bamidis, P. D. (2010). Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. *Brain Topography*, *23*(1), 27–40. https://doi.org/10.1007/s10548-009-0130-5

MATLAB. (2019). *9.6.0.1072779 (R2019a)*. The Mathworks Inc.

Matthews, M., Webb, T. L., Shafir, R., Snow, M., & Sheppes, G. (2021). Identifying the determinants of emotion regulation choice: A systematic review with meta-analysis. *Cognition and Emotion*, *1–29*, https://doi.org/10.1080/02699931.2021.1945538

Mischel, W. (1968). *Personality and assessment*. Wiley.

Muhtadie, L., Koslov, K., Akinola, M., & Mendes, W. B. (2015). Vagal flexibility: A physiological predictor of social sensitivity. *Journal of Personality and Social Psychology*, *109*(1), 106–120. https://doi.org/10.1037/pspp0000016

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, *126*(2), 247–259. https://doi.org/10.1037/0033-2909.126.2.247

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews*, *33*(7), 1004–1023. https://doi.org/10.1016/j.neubiorev.2009.04.001

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. *Annuals of the New York Academy of Sciences*, *1251*, E1–E24. https://doi.org/10.1111/j.1749-6632.2012.06751.x

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Perakakis, P. (2019). HEPLAB: A Matlab graphical interface for the preprocessing of the heartbeat-evoked potential (Version v1.0.0). *Zenodo*. https://doi.org/10.5281/zenodo.2649943

Peysakhovich, V., Causse, M., Scannella, S., & Dehais, F. (2015). Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort. *International Journal of Psychophysiology*, *97*(1), 30–37. https://doi.org/10.1016/j.ijpsycho.2015.04.019

Powers, J. P., & LaBar, K. S. (2019). Regulating emotion through distancing: A taxonomy, neurocognitive model, and supporting meta-analysis. *Neuroscience and Biobehavioral Reviews*, *96*, 155–173. https://doi.org/10.1016/j.neubiorev.2018.04.023

Pruessner, L., Barnow, S., Holt, D. V., Joormann, J., & Schulze, K. (2020). A cognitive control framework for understanding emotion regulation flexibility. *Emotion*, *20*(1), 21–29. https://doi.org/10.1037/emo0000658

Querino, E., Dos Santos, L., Ginani, G., Nicolau, E., Miranda, D., Romano-Silva, M., & Malloy-Diniz, L. (2015). Cognitive effort and pupil dilation in controlled and automatic processes. *Translational Neuroscience*, *6*(1), 168–173. https://doi.org/10.1515/tnsci-2015-0017

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

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

Ribeiro, R. L., Pompeia, S., & Bueno, O. F. (2005). Comparison of Brazilian and American norms for the International Affective Picture System (IAPS). *Brazilian Journal of Psychiatry*, *27*(3), 208–215. https://doi.org/10.1590/s1516-44462005000300009

Richards, J. M. (2004). The cognitive consequences of concealing feelings. *Current Directions in Psychological Science*, *13*(4), 131–134. https://doi.org/10.1111/j.0963-7214.2004.00291.x

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

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

RStudio Team. (2020). *RStudio: Integrated Development for R*. PBC. http://www.rstudio.com/

Saunders, B., Milyavskaya, M., & Inzlicht, M. (2015). What does cognitive control feel like? Effective and ineffective cognitive control is associated with divergent phenomenology. *Psychophysiology*, *52*(9), 1205–1217. https://doi.org/10.1111/psyp.12454

Schäfer, T., & Schwarz, M. A. (2019). The Meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, *10*, 813. https://doi.org/10.3389/fpsyg.2019.00813

Scheffel, C., Diers, K., Schönfeld, S., Brocke, B., Strobel, A., & Dörfel, D. (2019). Cognitive emotion regulation and personality: An analysis of individual differences in the neural and behavioral correlates of successful reappraisal. *Personality Neuroscience*, *2*, e11. https://doi.org/10.1017/pen.2019.11

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schweizer, S., Grahn, J., Hampshire, A., Mobbs, D., & Dalgleish, T. (2013). Training the emotional brain: Improving affective control through emotional working memory training. *Journal of Neuroscience*, *33*(12), 5301–5311. https://doi.org/10.1523/JNEUROSCI.2593-12.2013

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

Sheppes, G., & Gross, J. J. (2011). Is timing everything? Temporal considerations in emotion regulation. *Personality and Social Psychology Review*, *15*(4), 319–331. https://doi.org/10.1177/1088868310395778

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

Silvestrini, N. (2017). Psychological and neural mechanisms associated with effort-related cardiovascular reactivity and cognitive control: An integrative approach. *International Journal of Psychophysiology*, *119*, 11–18. https://doi.org/10.1016/j.ijpsycho.2016.12.009

Silvestrini, N., & Gendolla, G. H. (2011). Do not prime too much: Prime frequency effects of masked affective stimuli on effort-related cardiovascular response. *Biological Psychology*, *87*(2), 195–199. https://doi.org/10.1016/j.biopsycho.2011.01.006

Simmons, J. P., Nelson, L. D., & Simonsohn, U. A. (2012). A 21 word solution. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2160588

Sloan, R. P., Shapiro, P. A., Bagiella, E., Gorman, J. M., & Bigger, J. T. Jr (1995). Temporal stability of heart period variability during a resting baseline and in response to psychological challenge. *Psychophysiology*, *32*(2), 191–196. https://doi.org/10.1111/j.1469-8986.1995.tb03311.x

Smith, T. W., Deits-Lebehn, C., Williams, P. G., Baucom, B. R. W., & Uchino, B. N. (2020). Toward a social psychophysiology of vagally mediated heart rate variability: Concepts and methods in self-regulation, emotion, and interpersonal processes. *Social and Personality Psychology Compass*, *14*(3), e12516. https://doi.org/10.1111/spc3.12516

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Steele, J. (2020). What is (perception of) effort? Objective and subjective effort during task performance. *PsyArXiv*. https://doi.org/10.31234/osf.io/kbyhm

Takeuchi, T., Puntous, T., Tuladhar, A., Yoshimoto, S., & Shirama, A. (2011). Estimation of mental effort in learning visual search by measuring pupil response. *PLoS One*, *6*(7), e21973. https://doi.org/10.1371/journal.pone.0021973

Tamir, M. (2016). Why do people regulate their emotions? A taxonomy of motives in emotion regulation. *Personality and Social Psychology Review*, *20*(3), 199–222. https://doi.org/10.1177/1088868315586325

Thuillard, S., & Dan-Glauser, E. S. (2020). The simultaneous use of emotional suppression and situation selection to regulate emotions incrementally favors physiological responses. *BMC Psychology*, *8*(1), 133. https://doi.org/10.1186/s40359-020-00495-1

Torre, J. B., & Lieberman, M. D. (2018). Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review*, *10*(2), 116–124. https://doi.org/10.1177/1754073917742706

Tryon, W. W. (1975). Pupillometry: A survey of sources of variation. *Psychophysiology*, *12*(1), 90–93. https://doi.org/10.1111/j.1469-8986.1975.tb03068.x

Uchino, B. N., Holt-Lunstad, J., Bloor, L. E., & Campo, R. A. (2005). Aging and cardiovascular reactivity to stress: Longitudinal evidence for changes in stress reactivity. *Psychology and Aging*, *20*(1), 134–143. https://doi.org/10.1037/0882-7974.20.1.134

Urry, H. L., van Reekum, C. M., Johnstone, T., & Davidson, R. J. (2009). Individual differences in some (but not all) medial prefrontal regions reflect cognitive demand while regulating unpleasant emotion. *NeuroImage*, *47*(3), 852–863. https://doi.org/10.1016/j.neuroimage.2009.05.069

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review.*, *26*(6), 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

van Reekum, C. M., Johnstone, T., Urry, H. L., Thurow, M. E., Schaefer, H. S., Alexander, A. L., & Davidson, R. J. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. *NeuroImage*, *36*(3), 1041–1055. https://doi.org/10.1016/j.neuroimage.2007.03.052

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, *59*(6), 1037–1050. https://doi.org/10.1016/j.neuron.2008.09.006

Walter, H., von Kalckreuth, A., Schardt, D., Stephan, A., Goschke, T., & Erk, S. (2009). The temporal dynamics of voluntary emotion regulation. *PLoS One*, *4*(8), e6726. https://doi.org/10.1371/journal.pone.0006726

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

Wessa, M., Kanske, P., Neumeister, P., Bode, K., Heissler, J., & Schönfelder, S. (2010). EmoPics: Subjektive und psychophysiologische Evaluation neuen Bildmaterials für die klinisch-biopsychologische Forschung. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *39*(Suppl. 1/11), 77.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Yeo, G., & Neal, A. (2008). Subjective cognitive effort: A model of states, traits, and time. *Journal of Applied Psychology*, *93*(3), 617–631. https://doi.org/10.1037/0021-9010.93.3.617

Yu, X., & Gen, M. (2010). Introduction to Evolutionary Algorithms. *Introduction to Evolutionary Algorithms*, *1–418*, https://doi.org/10.1007/978-1-84996-129-5

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Supplementary Material

How to cite this article: Scheffel, C., Graupner, S.-T., Gärtner, A., Zerna, J., Strobel, A., & Dörfel, D. (2021). Effort beats effectiveness in emotion regulation choice: Differences between suppression and distancing in subjective and physiological measures. *Psychophysiology*, 00e1–24. https://doi.org/10.1111/psyp.13908