Cognitive reappraisal changes cognitive evaluations more than affective experiences

Henna I. Vartiainen[1] & Erik C. Nook[1]

[1] Department of Psychology, Princeton University, NJ 08544, United States

Corresponding author: Henna Vartiainen

Email: hennavartiainen@princeton.edu

ORCID iDs

Henna I. Vartiainen (0000-0002-5416-898X)

Erik C. Nook (0000-0001-7967-0792)

Abstract word count: 250

Intro + Discussion word count: 1,951

Statements and Declarations

Funding: This work was supported by a seed grant from Princeton Precision Health to H.I.V and E.C.N. and a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (31386) to E.C.N.

Competing interests: The authors declare no competing interests.

Availability of Data and Material: Datafiles related to this manuscript have been uploaded to the Open Science Framework (osf.io/uw5c2)

Code Availability: R Markdown files containing the analytic code and results have been uploaded to the Open Science Framework (osf.io/uw5c2).

Authors' Contributions: H.I.V. and E.C.N. designed research; H.I.V. performed research; H.I.V. analyzed data; H.I.V. and E.C.N. wrote the paper.

Ethics Approval: Approval was granted by the Ethics Committee of Princeton University (IRB #15223)

Consent to Participate: All participants provided informed consent.

Consent for Publication: Not applicable.

Abstract

Cognitive reappraisal is an emotion regulation strategy that involves changing one's interpretation of a situation to change its emotional impact. Cognitive reappraisal studies typically assess changes in valence ratings following reappraisal as the primary dependent variable, but scholars have recently documented that valence ratings can take two forms: affective valence (i.e., the positivity/negativity of one's hedonic internal feelings) and semantic valence (i.e., the positivity/negativity of one's cognitive evaluation of the stimulus). Because typical rating instructions do not distinguish between these forms of valence, it remains unknown how strongly cognitive reappraisal shifts either one. We address this gap through two preregistered studies. In Study 1, N=155 participants completed a classic cognitive reappraisal task in which they either responded naturally to or cognitively reappraised negative images. Critically, we manipulated whether participants rated either affective valence, semantic valence, or "default" valence (conventional instructions). Reappraisal significantly improved valence ratings in all conditions, but reappraisal most strongly influenced semantic valence and least strongly affective valence. Interestingly, default valence behaved most similarly to semantic valence, suggesting that existing reappraisal studies largely document changes in semantic evaluations, more than affective experiences. Study 2 (N=105) replicated these effects within-subjects. These results extend research on affective and semantic valence by demonstrating that they vary in their responsiveness to regulation. Additionally, they suggest that existing research likely overestimates the impact of reappraisal on felt emotional experiences, prompting new perspectives on past findings and more focused methods for assessing the hedonic impact of emotion regulation.

Keywords: valence, semantic valence, affective valence, reappraisal, emotion regulation

Several affective and clinical theories argue that our thoughts and feelings are tightly intertwined, such that changing how we think changes how we feel. For example, constructionist theories posit that emotions arise from a combination of physiological arousal and its cognitive interpretation (Schacter & Singer, 1962; Barrett, 2006), and appraisal theories contend that cognitive appraisals of events, agents, or objects trigger emotional experiences (Clore, Ortony, & Collins, 1988). Gross's (1998; 2015) process model of emotion regulation extends these ideas by illustrating how reinterpreting (i.e., *re*appraising) one's situation can alter the downstream experience and expression of emotion. Clinically, Beck's (1967) cognitive theory argues that maladaptive *thought processes* (especially about oneself, the world, and the future) contribute to depression.

Substantial empirical data support each of these theoretical accounts. Cognitions are indeed more negative in depressed than non-depressed individuals (Gotlib & Joormann, 2010), identical physiological arousal can yield different emotional and behavioral reactions depending on one's cognitive interpretations (Schachter & Singer, 1962; Lindquist & Barrett, 2008), and cognitive reappraisal is highly effective at reducing negative emotions across self-reported, neural, and physiological levels of analysis (Lazarus & Alfert, 1964; Gross, 2002, 2015; Webb et al, 2012; Ochsner et al, 2002; 2004; Buhle et al, 2014; Morawetz et al, 2016). Indeed, use of reappraisal has been linked to greater psychological wellbeing, positive affect, and mental health (Gross & John 2003; Dawel et al, 2024; Kam et al, 2024), and it is a key therapeutic strategy used in Cognitive Behavioral Therapy (Beck, 2020).

Despite this robust evidence, recent work has raised important questions about how emotional change is measured in reappraisal studies. More specifically, our typical strategies for measuring affective states may insufficiently attend to subtle distinctions between different types of *valence* (i.e., the negativity or positivity of one's affective state, a core dimension of emotion; Barrett, 2006; Russell et al, 1980; Watson & Tellegen, 1985). Recent work has decomposed valence into two empirically dissociable constructs: *Affective valence*, or the hedonic quality of

an emotional experience, and *semantic valence*, or the cognitive evaluation of a situation as positive or negative (Itkes et al, 2017; Itkes et al, 2019; Hamzani et al, 2020). Notably, when participants are explicitly asked to rate either affective or semantic valence, their responses diverge. For example, ratings of affective valence better predict physiological responses (e.g. heart rate, facial EMG, and electrodermal activity) to affective stimuli than ratings of semantic valence (Hamzani et al, 2020), and ratings of affective valence (as well as associated physiological responses) habituate over repeated exposure to emotional images, whereas semantic ratings do not (Itkes et al, 2017).

To date, cognitive reappraisal studies have not attended to these different constructs. In fact, there is high variability in how studies measure reappraisal effectiveness (**Table 1**). Many studies ask participants to report “how they feel” or “how negative their emotional experience is” (Ochsner et al, 2002; 2004; Adam et al, 2014; Morawetz et al, 2016). Although this may seem to capture affective valence, this instruction unfortunately may not sufficiently coach participants to rate only their emotional experiences rather than their cognitive evaluations of stimuli. In fact, when valence rating instructions are “ambiguous” as to semantic or affective valence, ratings fall between affective and semantic ratings rather than aligning with affective valence (Hamzani et al, 2020). In contrast, some cognitive reappraisal studies seem to assess semantic (rather than affective) valence, as they ask the participants to rate how positively/negatively they evaluate the stimulus (Phan et al 2005; Van Reekum et al, 2007; Winecoff et al, 2010).

Given that these subtle differences in wording track empirically dissociable constructs (e.g. Itkes et al, 2017; 2019; Hamzani et al, 2020), this methodological variability is not trivial in the conclusions we draw from these studies. Specifically, we do not know 1) whether reappraisal changes both affective and semantic valence, or 2) the extent to which default valence rating instructions in reappraisal tasks measure affective experiences or cognitive evaluations. This draws into question how strongly we can assume that cognitive reappraisal

designs assess changes in participants' actual emotional experiences, the construct that most studies are presumed to test.

Here we address this gap through two preregistered studies (Study 1: osf.io/varht, Study 2: osf.io/6etx8) in which participants completed classic cognitive reappraisal tasks (Ochsner et al, 2002; 2004; Nook et al, 2020; 2021), but we manipulated (between-subjects in Study 1 and within-subjects in Study 2) whether participants reported on their affective valence, semantic valence, or default/ambiguous valence (using instructions commonly used in cognitive reappraisal research). We preregistered the hypotheses that cognitive reappraisal would reduce negative and increase positive valence ratings across all conditions, but reappraisal will more strongly influence semantic valence ratings than affective valence ratings. Addressing how reappraisal affects these different types of valence ratings can guide methods used in reappraisal studies, provide further evidence that affective and semantic valence are dissociable constructs, and eventually improve our understanding of which interventions might shift internal feeling states.

Table 1. Selected valence rating instructions used in cognitive reappraisal studies.

Study	Valence rating description
Troy et al, 2013	Participants rated the greatest amount of sadness they experienced.
Kim & Hamann, 2007	Participants rated their degree of success at carrying out each regulation task.
Urry et al, 2006	Participants rated how successful they were at reappraising.
Harenski & Hamann, 2006	Participants rated their level of emotional arousal.
Phan et al, 2005	Participants rated the intensity of their negative affect.
Ochsner et al, 2002, 2004	Participants rated the strength of their current negative affect.
Milyavski et al, 2019	Participants rated their affect.
Bomyea et al, 2021	Participants rated their current distress level.
Morawetz et al, 2016	Participants rated the extent of their negative emotions.
Hayes et al, 2010	Participants rated emotional response to the picture, and a confidence

	rating regarding how well they were able to reappraise.
Krendl et al, 2011	Participants rated the relative strength of their negative emotions to the image.
McRae et al, 2012	Participants rated the strength of their negative affect.
Modinos et al, 2010	Participants rated the strength of their current negative affect.
van Reekum et al, 2007	Participants rated how pleasant or unpleasant they found each picture.
Winecoff et al, 2010	Participants rated how negative or positive the image was.
Nook et al, 2017	Participants rated how they feel.

Study 1

Method

Participants. Two hundred and five participants were recruited on Prolific. Fifty participants were excluded due to failing preregistered attention checks, leaving *N* = 155 usable participants (see **Table 2** for demographics). However, for transparency, results using the full sample are shown in **Supplemental Materials.** The pattern of results was replicated using the full sample, mitigating concerns that these exclusions biased results. Preregistered attention checks consisted of randomized scale instructions (*e.g.*, “Please select the rating all the way to the left”) embedded within the valence rating trials. Participants were eligible if they were 18 years or older, fluent in English, located in the United States, and had at least a 95% approval rating on previous tasks. Participants were randomly assigned to the affective (N = 50), default (N = 54), or semantic (N = 51) conditions (detailed below).

We set our target sample size through a power analysis using the pwr package in R (Champely, 2020). To achieve 80% power (at α = .05), prior work suggested that a sample size of 19 would be required to detect differences in self-reports of affective and semantic valence in response to repeated exposure (Itkes et al, 2017; $\eta^2p = 0.314$), and a sample size of 15 would be required to detect a main effect of reappraisal (Nook et al, 2021; $\eta^2p = 0.37$). However, to

provide additional power for the key result of interest (that reappraisal efficacy differs across valence instruction conditions), we preregistered that we would recruit participants until we obtained 50 usable participants per condition. Target sample size, inclusion and exclusion criteria, study methods, analytic plan, and hypotheses were preregistered prior to data collection at osf.io/varht.

Materials and procedure. Study 1 employed a mixed within-between design with a 2-level Reappraisal factor [Look vs Change] manipulated within-subjects and a 3-level Valence Instructions factor [Affective vs Semantic vs Default] manipulated between-subjects. At the start of the experiment, participants were randomly assigned to one of three (affective, semantic, or default) Valence Instructions conditions. The affective and semantic instructions were adapted from Itkes et al (2017), and the default instructions were adapted from Nook et al (2021). Excerpts of these instructions are provided in **Table 3**, and full materials are available in the **Supplemental Materials.**

Table 2. Participant demographics and dataset characteristics for Studies 1 and 2.

		Study 1			Study 2
Condition		Affective	Default	Semantic	
N		50	54	51	105
Age M (SD)		38.5 (10.8)	41.4 (14.1)	39.4 (12.9)	37.53 (11.83)
Sex N (%)	Male	28 (56%)	25 (46%)	23 (45%)	40 (38%)
	Female	20 (40%)	28 (52%)	27 (53%)	64 (61%)
	Other	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	Not reported	2 (4%)	1 (2%)	1 (2%)	1 (1%)
Gender N (%)	Male	28 (56%)	24 (44%)	23 (45%)	42 (40%)
	Female	21 (42%)	27 (50%)	26 (51%)	62 (59%)
	Nonbinary	0 (0%)	0 (0%)	1 (2%)	1 (1%)
	Agender	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	Unsure	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	Other	0 (0%)	0 (0%)	0 (0%)	0 (0%)

	Not reported	1 (2%)	3 (6%)	1 (2%)	0 (0%)
Race N (%)	White	38 (76%)	43 (80%)	39 (76%)	75 (71%)
	Black	11 (22%)	7 (13%)	12 (24%)	27 (26%)
	Asian	2 (4%)	2 (4%)	2 (4%)	2 (2%)
	Native American	1 (2%)	1 (2%)	0 (0%)	2 (2%)
	Middle Eastern	1 (2%)	0 (0%)	1 (2%)	0 (0%)
	Pacific Islander	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	Other	0 (0%)	1 (2%)	0 (0%)	1 (1%)
	Not reported	0 (0%)	0 (0%)	0 (0%)	1 (1%)
Ethnicity N (%)	Hispanic	2 (4%)	1 (2%)	1 (2%)	8 (8%)
	Not Hispanic	45 (90%)	53 (98%)	50 (98%)	95 (90%)
	Not reported	2 (4%)	0 (0%)	0 (0%)	2 (2%)
Missing trials		11.20%	9.49%	7.25%	8.20%

In the affective condition, participants were presented with the definition of affective valence (i.e. the hedonic value of an image, or how the image makes subjects feel). Participants were instructed to distinguish between the positivity/negativity of their feelings and the positivity/negativity of the image's content and to report specifically on their feelings while viewing the image. For example, if participants knew that the picture had negative content but they did not feel anything, they were instructed to report that they did not experience strong feelings. In the semantic condition, participants were presented with the definition of semantic valence (i.e. a cognitive evaluation of how positive or negative the content of the image is). Again, participants were instructed to distinguish between what they felt and what the image's content and to report only how positively/negatively they judged the content of the image to be, regardless of how they felt. In the Default condition, participants did not receive specific instructions on either of these two constructs. Instead, they were simply asked to rate how positive or negative they felt in response to each image, which is a common way of measuring valence in reappraisal studies (e.g., Krendl et al, 2011; Hayes et al, 2010; Nook et al, 2021).

After receiving their assigned Valence Instructions, all participants completed a classic cognitive reappraisal task where they rated their affective responses to pictorial stimuli from the Open Affective Standardized Image Set (OASIS; Kurdi et al, 2017; Ochsner et al, 2002; 2004; Nook et al, 2021; **Figure 1A**). Two image sets (20 images each) were used, and we counterbalanced assignment of list to conditions. These sets were reproduced from Nook et al (2020) and were matched on valence ($p = .899$), arousal ($p = .750$), and content (for more information, see Table S1 of Nook et al, 2020). As such, each participant rated 40 images in total (20 in the Reappraisal condition and 20 in the Look condition), and no participant viewed the same image more than once.

On Look trials, participants saw the word LOOK above an image, instructing them to simply view the images and provide valence ratings according to their assigned valence condition. On reappraisal trials, participants saw the word CHANGE above the image, instructing them to reinterpret the meaning of each image in a way that made them feel better. Each image was displayed for 10 seconds, after which participants had 4 seconds to provide their negative rating, and 4 seconds to provide their positive rating. The screen proceeded if they had not responded in this time, leading to some missing data (average of 9.29% of trials in Study 1, **Table 2**). Between trials, participants viewed a fixation cross for 5 seconds. The order of the images, Look/Change conditions, and rating scales (positive/negative) was randomized. Following the experiment, all participants completed a set of questionnaires and demographic questions.

Table 3. Excerpts of the valence instructions for the three conditions.

Affective valence	Semantic valence	Default valence
“For this experiment, we want to distinguish between what you feel inside and the contents of the picture. Today, we want you to rate the feeling you experienced inside, NOT the contents of the picture… For example, you can look at a picture of a car accident and not feel strong emotions inside, even though you know that car accidents are negative. If that happens, you should rate that you do not feel strong feelings… After each image, you'll see two rating scales, one asking how positive you feel, and the other asking how negative you feel. Order will vary across trials, so pay attention to which one you're answering… Remember: you are supposed to rate how you actually feel, NOT how positive or negative the image is.”	“For this experiment, we want to distinguish between what you feel inside and the contents of the picture. Today, we want you to rate the contents of the picture, NOT the feelings you experienced inside… For example, you can look at a picture of a car accident and not feel strong emotions inside, even though you know that car accidents are negative. If that happens, you should rate how negative the image's scene is, regardless of how you feel.. After each image, you'll see two rating scales, one asking how positive the image is, and the other asking how negative the image is. Order will vary across trials, so pay attention to which one you're answering… Remember: you are supposed to rate how positive or negative the event in the image is, NOT how positive or negative you feel.”	“After each image, you'll see two rating scales, one asking how positive you feel, and the other asking how negative you feel. Order will vary across trials, so pay attention to which one you're answering.”

Note. Full instructions can be found in the Supplemental Materials.

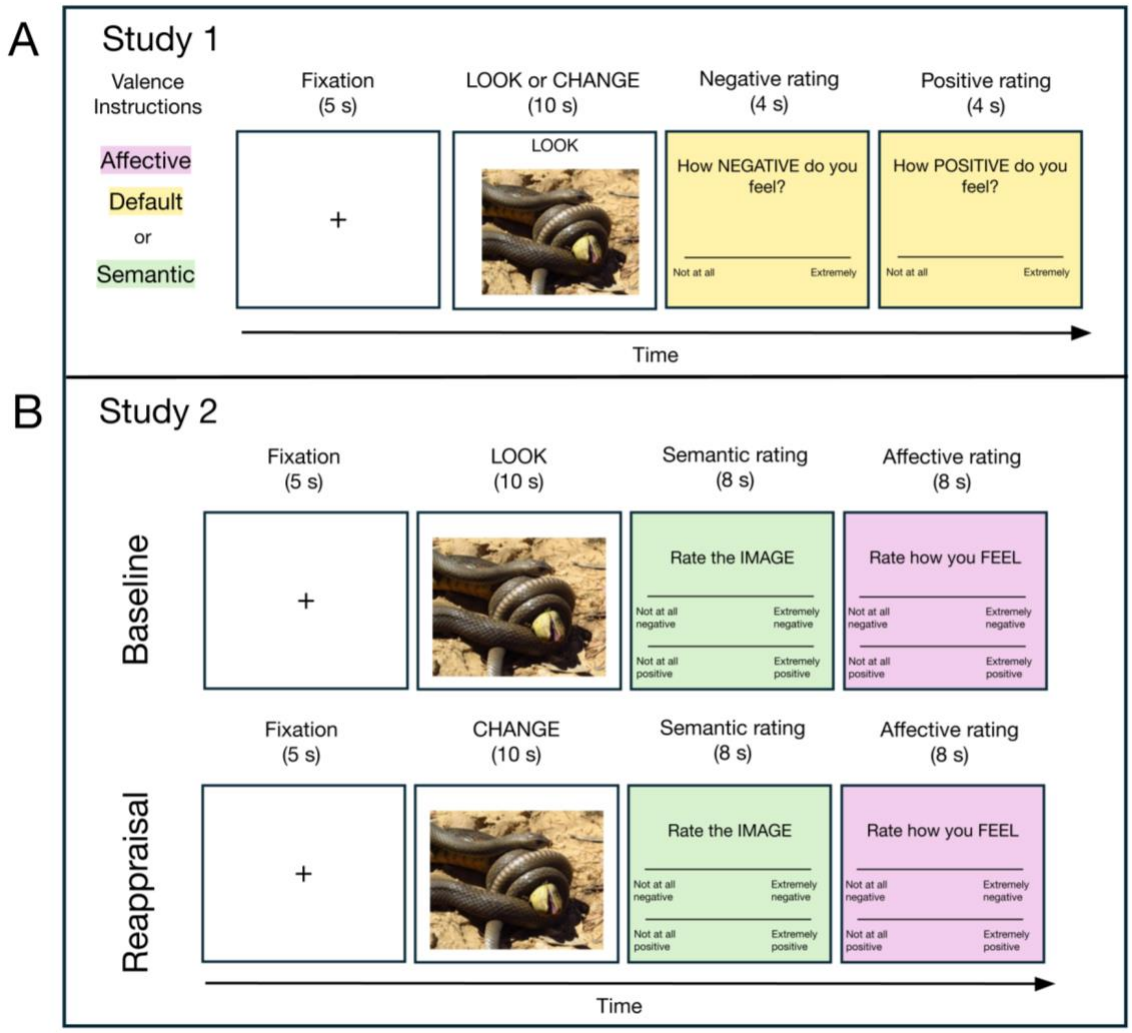

Figure 1. Experimental design for both studies. A) In Study 1, participants were instructed to provide either affective, semantic, or default valence ratings throughout the experiment (manipulated between-subjects). Participants responded naturally when the word LOOK was presented above the image, and they reappraised the meaning of the image when the word CHANGE was displayed. Following image presentation, participants provided negative and positive valence ratings according to their assigned valence instruction condition. The depicted instructions represent the Default condition. The order of Look/Change instructions and positive/negative ratings were randomized for each participant. B) In Study 2, participants first completed a baseline phase where they provided both semantic and affective valence ratings to a set of images. Then, in the experimental phase, participants saw the same images a second time and reappraised all images before again providing affective and semantic ratings. Order of semantic/affective ratings were randomized for each trial.

Analyses. We preregistered the following analyses. We conducted linear mixed effects analyses of trial-level affect ratings, with Reappraisal (Look vs Change; within-subjects) and Valence Instructions (Affective vs Semantic vs Default; between-subjects) as predictors. These

analyses were conducted separately for positive ratings and negative ratings. Both models included a random intercept for participant ID to nest trial-level ratings within participants. We then subjected this model to an ANOVA to determine main effects and interactions of each of our factors. Although we preregistered analyzing negative and positive valence ratings separately, we observed that these were strongly correlated within-subject (r = -.70, p < .001) and that results overall did not differ across them. We thus present these results in the **Supplemental Materials** and in the main text present analyses of "unidimensional valence", which was calculated by subtracting trial-level negative ratings from positive ratings. Thus more positive unidimensional ratings indicate more positive affect and more negative unidimensional ratings indicate more negative affect.

We hypothesized a significant main effect of Reappraisal, such that ratings would be less negative/more positive in the Change condition than the Look condition. In addition, we expected to see a significant Reappraisal x Valence Instructions interaction, indicating that the framing of the instructions impacted the outcomes of reappraisal. We interpreted the findings of the Reappraisal x Valence Instructions interaction with a specific focus on whether reappraisal had a greater impact on changing semantic ratings over affective ratings following reappraisal. We conducted contrast analyses to compare reappraisal effects across conditions. We hypothesized that reappraisal would lead to the greatest changes in Semantic valence, smallest changes in Affective valence, and that the Default condition would be between these two. We based this latter hypothesis on the notion that cognitive reappraisal primarily involves changing the way participants construe the meaning of the image (Gross, 1998), but reappraisal may be less effective at modulating core physiological responding that forms the ingredients of hedonic experiences (Troy et al, 2018; Hamzani et al, 2020).

Results

Results are presented in **Figure 2**. As hypothesized, we found a significant main effect of Reappraisal on unidimensional valence ratings, $F(1, 5484.5) = 991.07$, $p < .001$, indicating that participants reported feeling worse on Look ($M = -30.40$, $SD = 55.90$) than Change ($M = 13.60$, $SD = 56.70$) trials, consistent with prior reappraisal research. Follow-up contrasts within each condition showed that ratings became significantly more positive from Look to Change in all three conditions: affective ($b = 36.60$, $SE = 2.49$, $z = 14.67$, 95% CI [31.70, 41.50], $p < .001$), default ($b = 46.90$, $SE = 2.37$, $z = 19.81$, 95% CI [42.30, 51.50], $p < .001$), and semantic ($b = 48.70$, $SE = 2.41$, $z = 20.21$, 95% CI [44.00, 53.40], $p < .001$).

There was also a significant main effect of Valence Instructions on unidimensional valence ratings, $F(2, 150) = 3.92$, $p = .022$, indicating that rating intensity overall varied across the affective ($M = -10.80$, $SD = 61.30$), semantic ($M = -13.00$, $SD = 60.00$), and default ($M = -1.55$, $SD = 59.40$) conditions. Follow-up pairwise contrasts showed that this was due to more negative semantic compared to default ratings ($b = 11.17$, $SE = 4.33$, 95% CI [1.02, 21.33], $p = .027$). Other pairwise comparisons were not significant (ps > .070).

The interaction between Reappraisal and Valence Instructions was also significant, $F(2, 5484.2) = 7.05$, $p < .001$, indicating that the effect of reappraisal differed across the three valence instruction conditions (**Figure 2B**). Follow-up contrasts showed that the magnitude of difference between Look and Change trials was smaller in the affective valence condition compared to both the default condition ($b = -10.31$, $SE = 3.44$, $z = -3.00$, 95% CI [-18.52, -2.10], $p = .008$) and the semantic condition ($b = -12.12$, $SE = 3.47$, $z = -3.49$, 95% CI [-20.40, -3.84], $p = .001$). The impact of reappraisal on valence ratings did not differ between semantic and default conditions ($b = -1.81$, $SE = 3.38$, $z = -0.54$, 95% CI [-9.88, 6.26], $p = .932$).

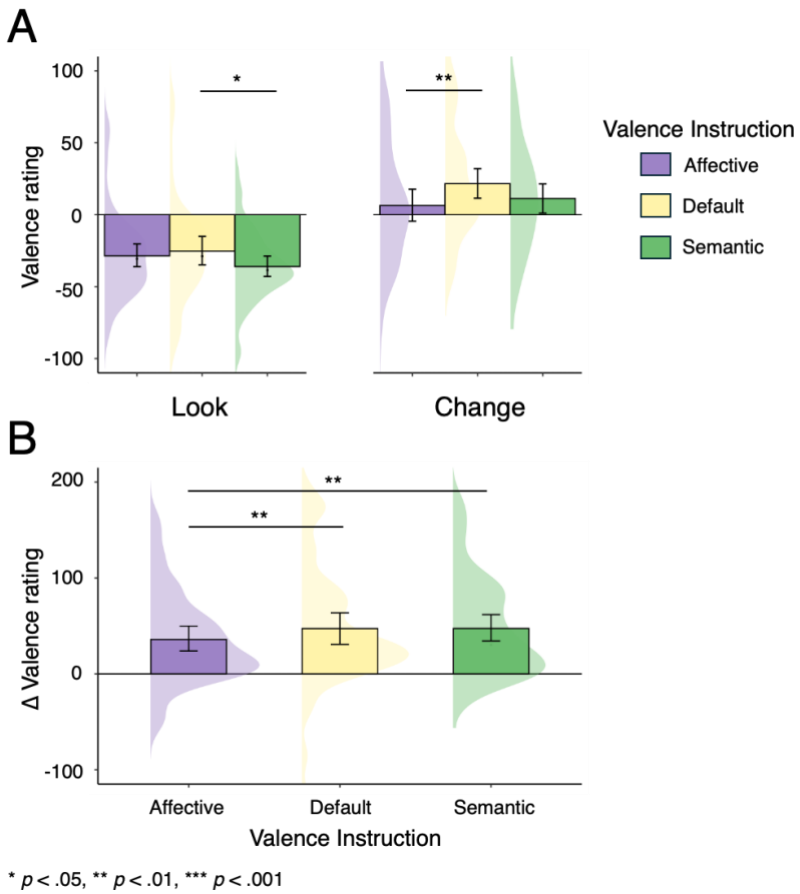

Figure 2. Study 1 results: The effect of reappraisal on affective, semantic, and default valence ratings. A) Averages for each condition showing both a main effect of reappraisal (i.e., ratings were more negative for Look than Change trials) and the interaction between reappraisal and valence instructions. Pairwise contrasts showed that semantic ratings were more negative than default ratings for Look trials and that default ratings were more positive than affective ratings on Change trials. B) Change scores showing differences between Look and Change trials within each Valence Instruction condition. The magnitude of valence change following reappraisal varied across Valence Instructions conditions. Affective valence ratings were significantly less affected by reappraisal than default and semantic ratings. These change scores did not differ in magnitude between default and semantic ratings.

Study 2

Method

Participants. One hundred and forty-five participants were recruited on Prolific. Forty participants were excluded due to preregistered attention checks, leaving N = 105 usable participants (see **Table 2** for demographics). However, for transparency, the results using the

full sample are shown in **Supplemental Materials.** The overall pattern of results was the same in the unfiltered sample, however, the interaction only marginally significant ($p = .054$), likely reflecting the poor data quality of participants who failed attention checks (see **Supplemental Materials** for more information). As in Study 1, attention checks consisted of randomized scale instructions (*e.g.*, "Please select the rating all the way to the left") embedded within the valence rating trials. Participants were eligible if they were 18 years or older, fluent in English, located in the United States, and had at least a 95% approval rating on previous tasks.

Using 80% power and $\alpha = .05$, our *a priori* power analyses indicated that a sample size of 35 would be required to detect within-subjects differences in self-reports of affective and semantic valence (Olteanu et al, 2023; $\eta^2p = 0.19$), and a sample size of 15 would be required to detect within-subjects differences due to reappraisal (Nook et al, 2021; $\eta^2p = 0.37$). However, to account for individual variability and to ensure sufficient power to detect individual differences, we preregistered that we would recruit participants until we obtained 100 usable participants. This larger sample size granted us more confidence in results and provided more power for detecting potentially smaller effects. As in Study 1, target sample size, methods, analyses, and hypotheses for Study 2 were preregistered prior to data collection at osf.io/6etx8.

Materials and procedure. This study employed a 2 x 2 design with both the Reappraisal factor [Look vs Change] and the Valence Instructions factor [Semantic vs Affective] manipulated within-subjects. *All* participants were presented with the definitions of affective and semantic valence used in Study 1 (**Table 3**). However, in Study 2 participants were instructed to provide *both* affective and semantic ratings for all stimuli.

Only one set of OASIS images (Kurdi et al, 2017) from Study 1 (20 images) was used for this study, as all participants saw these 20 images twice, once in a baseline passive viewing phase and once in a reappraisal phase (adapted from Nook et al., 2021; **Figure 1B**). In the Passive Viewing phase, participants were instructed to simply view the images and provide affective and semantic valence ratings (one for negative, one for positive). After participants

provided their baseline ratings, they completed the second Reappraisal phase. Participants were instructed to reinterpret the meaning of each image in a way that made them feel better before providing their ratings. Each image was displayed for 10 seconds, after which participants had 8 seconds to provide their affective or semantic ratings (separate scales for positive and negative valence), and 8 seconds to provide the remaining valence ratings (semantic/affective; separate scales for positive and negative valence). The interstimulus interval was 5 seconds. Trial order and order of affective/semantic ratings were randomized for each subject within each condition. Following the experiment, all participants completed a set of questionnaires and demographic questions.

Analyses. To test our preregistered research questions, we conducted linear mixed effects analyses of unidimensional trial-level valence ratings, with Reappraisal (Look vs Change: within-subjects) and Valence Instructions (Affective vs Semantic: within-subjects) as predictors. Given that results did not differ across negative and positive ratings in Study 1, we shifted our preregistered analyses to focus on unidimensional ratings, calculated by subtracting the negative ratings from positive ratings. We again expected to see a significant main effect of Reappraisal, indicating less negative/higher positive ratings in the Change condition than the Look condition across Valence Instructions. In addition, we expected to see a significant interaction between Reappraisal and Valence Instructions, showing that the framing of the instructions impacts the outcomes of reappraisal. We focused our interpretation of this interaction on our hypothesis that reappraisal would have a greater impact on changing semantic ratings over affective ratings. We conducted contrast analyses to compare reappraisal effects across conditions. For thoroughness, we present exploratory analyses of negative and positive ratings separately in the **Supplemental Materials**.

Results

There was a significant main effect of reappraisal on valence ratings, $F(1, 7614.00) = 907.50$, $p < .001$, indicating that participants reported overall improved valence on reappraisal trials ($M = 1.93$, $SD = 59.50$) compared to baseline trials ($M = -33.90$, $SD = 54.60$). Follow-up contrasts within each condition showed that the effect of reappraisal was significant for both affective ratings ($b = 32.3$, $SE = 1.68$, $z = 19.24$, 95% CI [29.10, 35.60], $p < .001$) and semantic ratings ($b = 39.5$, $SE = 1.69$, $z = 23.40$, 95% CI [36.2, 42.9], $p < .001$; **Figure 3A**).

We also observed a significant main effect of Valence Instructions, $F(1, 7608.10) = 144.41$, $p < .001$, indicating that valence ratings were overall more positive for affective ratings ($M = -8.59$, $SD = 59.90$) than semantic ratings ($M = -22.90$, $SD = 59.10$) (**Figure 3B**).

Finally, the interaction between Reappraisal and Valence Instructions was significant, $F(1, 7607.60) = 9.14$, $p = .003$, indicating that the effect of reappraisal differed across valence rating conditions. Follow-up contrasts showed that the magnitude of change was lower in the affective condition compared to the semantic condition ($b = -7.20$, $SE = 2.38$, $z = -3.02$, 95% CI [-11.90, -2.53], p = .003) (**Figure 3B**).

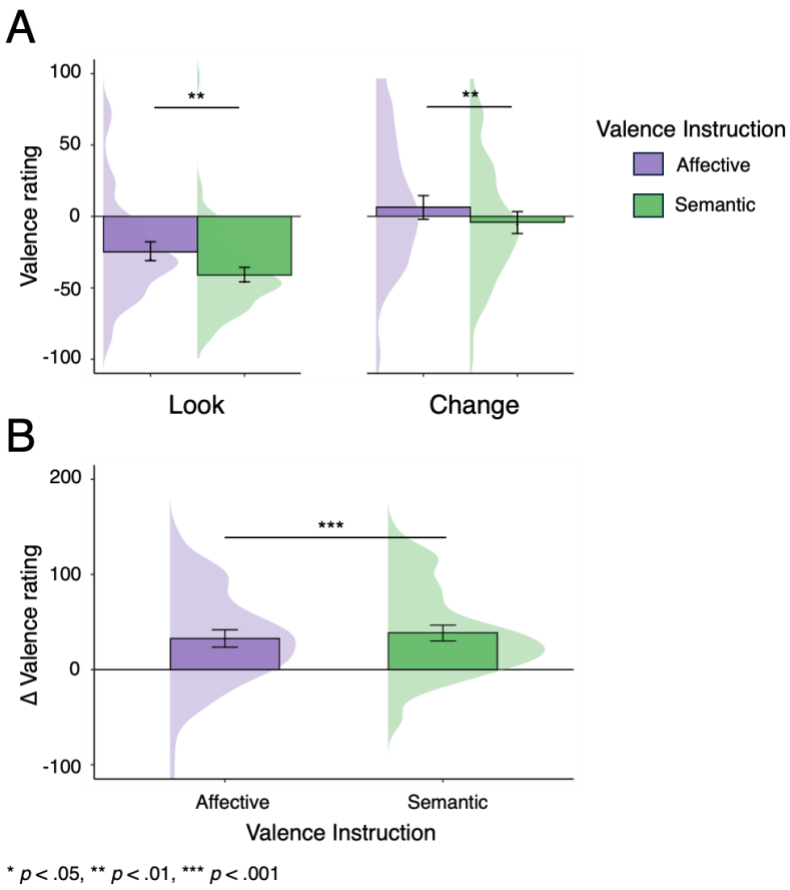

Figure 3. Study 2 results: The effect of reappraisal on affective and semantic ratings of valence, when manipulated within-subjects. A) Average valence ratings in each condition. Reappraisal significantly improved valence in both conditions. Additionally, valence instructions influenced valence ratings: Both before and after reappraisal, semantic ratings were more negative compared to affective ratings. B) Valence instructions influenced the magnitude of valence change following reappraisal. Affective ratings were significantly less influenced by reappraisal than semantic ratings.

Discussion

When we ask people to change their thoughts to change their feelings (i.e., use cognitive reappraisal), does it matter what we measure when determining their regulatory success? Here we integrated cognitive reappraisal research with recent insights into the complexities of valence (Itkes et al, 2017; 2019; Hamzani et al, 2020; Wang et al, 2021). Across two preregistered studies, we found that reappraisal more strongly influenced semantic valence ratings (i.e., cognitive evaluations of images) than affective valence ratings (i.e., hedonic

experiences). Additionally, “default” valence ratings (i.e., instructions used in prior studies) behaved more like semantic than affective valence ratings. Effects replicated across within-person and between-person designs. These results (i) provide additional evidence that cognitive evaluations and affective experiences are distinct facets of valence, (ii) demonstrate that this distinction has deep implications for how we measure and interpret valence change in reappraisal studies and (iii) spur new directions for emotion regulation research.

One possible explanation for increased malleability of semantic evaluations is that reappraisal primarily operates through reinterpretation processes which directly target semantic representations. Thus, reappraisal may be more effective at changing what an event means than how it feels, highlighting a potential dissociation between affective and cognitive aspects of emotion regulation. This aligns with recent work distinguishing between generation and implementation of reappraisals (Waugh et al, 2022). Reappraisal generation involves creating new interpretations of a situation, whereas reappraisal implementation involves applying them to modify one's emotional response. Our findings may reflect these stages, such that participants more easily and effectively generate reappraisals that shift how they *think* about stimuli, but the implementation of reappraisals to change their feelings is on average harder, slower, and/or less effective. Future work could more directly test whether reappraisal generation affects only semantic representations, while reappraisal implementation also extends to affective experiences.

A second explanation for our results is that affective ratings tap participants’ *core affect* (i.e., one’s constantly internal affective state integrating bodily, contextual, and cognitive inputs; Russell, 2003; Barrett, 2017). If affective ratings reflect the sum of many different contributors beyond one’s appraisal of just one visual stimulus, we would expect that changing just that one input would yield a smaller change in affect than in their evaluation of that stimulus. Indeed, the process model of emotion regulation emphasizes reappraisal as changing the stimulus meaning, but also acknowledges that affective experience is influenced by current experience

(attention, bodily state; Gross, 2015). Consistent with this, research indicates that affective experiences are not only shaped by cognitive processes, but also by trait factors, stress, and situational demands (Aldao & Dixon-Gordon, 2014). Together, these accounts provide initial explanations why reappraisal may more readily influence semantic evaluations than hedonic affect, but these prompt further research on the mechanisms by which reinterpretations translate (or fail to translate) to core affective states.

We also found in Study 1 that "default" instructions behave more like semantic than affective ratings. Even though the instructions we used asked participants to report how they *feel*, when we did not specify between affective and semantic valence, participants' responses seem to reflect semantic more than affective ratings. This suggests that prior studies that did not provide such a distinction for their participants likely overestimated how strongly reappraisal modifies actual affect. This has many important implications. For example, clinical and at-risk samples have shown small or inconsistent differences from healthy controls in their reappraisal success in tasks like these (Visted et al, 2018; Picó-Pérez et al, 2017), and recent work has suggested that reappraisal capacity is unrelated to depressive and anxiety symptoms (Andrews et al, 2023). This contrasts with theoretical accounts that emphasize reappraisal as a key vulnerability factor in psychopathology and assume reappraisal efficacy differs between clinical and healthy populations (Beck, 1967; Beck & Clark, 1997; Aldao, Nolen-Hoeksema, & Schweizer, 2010). If default valence ratings primarily reflect cognitive evaluations, these discrepancies may stem from measurement issues. Recognizing this distinction is critical for interpreting past findings and improving future research on emotion regulation in clinical samples.

Our results also add to the growing body of work showing that semantic and affective valence are dissociable (*e.g.,* Itkes et al, 2017, 2019; Hamzani et al, 2020; Wang et al, 2021). Thus, reappraisal appears to be a "wedge" scholars can use to dissociate these constructs. In addition, in Study 2 we observed that semantic valence ratings were overall more negative than

affective valence ratings, revealing that at baseline, images were evaluated as more negative than they were experienced. These findings suggest that affective/semantic distinctions are important even when estimating normative emotional responses to stimuli: Images may "appear" to have greater impact on felt emotional experiences than they actually do, depending on rating instructions.

Although our findings highlight the importance of distinguishing semantic from affective valence in emotion regulation, our study is not without limitations. First, we used staged, standardized images as stimuli. Although this approach gives us strong experimental control, it does not reflect the complexity and personal meaning of real-life situations that are often autobiographical and socially situated. In addition, we used Prolific to recruit our participants. Although our sample was well-balanced on gender, it lacked diversity with regard to characteristics such as age, nationality, cultural background, race, and ethnicity. Therefore, our findings may not generalize to broader populations. Finally, our results relied on self-report measures, which are subject to biases in introspection and demand characteristics. Even though self-report is the most commonly assessed measure of emotion, incorporating physiological and behavioral indices could provide a more comprehensive image of how reappraisal influences affective experiences.

These findings open promising directions for future research. Although not directly tested here, our results raise intriguing questions about how semantic and affective representations of valence interact at the individual level. In particular, the "alignment" of semantic and affective representations may have implications for emotion regulation and mental health. For example, individuals with depression and anxiety tend to maintain both persistent negative affect *and* negatively biased interpretations (Gotlib & Joormann, 2010; Everaert et al, 2012; Hirsch et al, 2016), suggesting that coupling between feeling and evaluation could be relevant to psychopathology. Moreover, individual differences in interoceptive accuracy and focus (Garfinkel et al, 2015; Murphy et al, 2019; Ainley & Tsakiris, 2013) may influence how closely

these processes align. Clarifying the role of semantic-affective alignment thus represents an important new direction for understanding variability in emotion regulation efficacy and vulnerability to affective disorders.

To conclude, recognizing that semantic and affective shifts in emotion regulation can dissociate underscores the need for more precise assessments of reappraisal outcomes and for interventions that target both affective and cognitive components of emotion. Future research should examine how individual differences, such as psychopathology or semantic-affective coupling, influence reappraisal ability and whether modifying the alignment between semantic and affective representations of valence serves as a mechanism of effective emotion regulation.

References

Adam, R., Schönfelder, S., Forneck, J., & Wessa, M. (2014). Regulating the blink: cognitive reappraisal modulates attention. Frontiers in psychology, 5, 143.

Ainley, V., & Tsakiris, M. (2013). Body conscious? Interoceptive awareness, measured by heartbeat perception, is negatively correlated with self-objectification. PloS one, 8(2), e55568.

Aldao, A., & Dixon-Gordon, K. L. (2014). Broadening the scope of research on emotion regulation strategies and psychopathology. Cognitive behaviour therapy, 43(1), 22-33.

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. Clinical psychology review, 30(2), 217-237.

Andrews, J. L., Dalgleish, T., Stretton, J., & Schweizer, S. (2023). Reappraisal capacity is unrelated to depressive and anxiety symptoms. Scientific Reports, 13(1), 7189.

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. Personality and social psychology review, 10(1), 20-46.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. Social cognitive and affective neuroscience, 12(1), 1-23.

Beck, A. T. (1967). Depression: Clinical, Experimental, and Theoretical Aspects. University of Pennsylvania Press.

Beck, J. S. (2020). Cognitive behavior therapy: Basics and beyond. Guilford Publications.

Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. Behaviour research and therapy, 35(1), 49-58.

Bomyea, J., Ball, T. M., Simmons, A. N., Campbell-Sills, L., Paulus, M. P., & Stein, M. B. (2020). Change in neural response during emotion regulation is associated with symptom reduction in cognitive behavioral therapy for anxiety disorders. Journal of affective disorders, 271, 207.

Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., ... & Ochsner, K. N. (2014). Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. Cerebral cortex, 24(11), 2981-2990.

Champely S (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0, https://github.com/heliosdrm/pwr.

Clore, G., Ortony, A. & Collins, A. (1988). The Cognitive Structure of Emotion. 10.2307/2074241.

Dawel, A., Mewton, P., Gulliver, A., Farrer, L. M., Calear, A. L., Newman, E., & Cherbuin, N. (2024). For whom and what does cognitive reappraisal help? A prospective study. Cognitive Therapy and Research, 48(4), 687-703.

Everaert, J., Koster, E. H., & Derakshan, N. (2012). The combined cognitive bias hypothesis in depression. Clinical psychology review, 32(5), 413-424.

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: distinguishing interoceptive accuracy from interoceptive awareness. Biological psychology, 104, 65-74.

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. Annual review of clinical psychology, 6(1), 285-312.

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.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. Review of general psychology, 2(3), 271-299.

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

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. Psychological inquiry, 26(1), 1-26.

Hamzani, O., Mazar, T., Itkes, O., Petranker, R., & Kron, A. (2020). Semantic and affective representations of valence: Prediction of autonomic and facial responses from feelings-focused and knowledge-focused self-reports. Emotion, 20(3), 486.

Harenski, C. L., & Hamann, S. (2006). Neural correlates of regulating negative emotions related to moral violations. Neuroimage, 30(1), 313-324.

Hayes, J. P., Morey, R. A., Petty, C. M., Seth, S., Smoski, M. J., McCarthy, G., & LaBar, K. S. (2010). Staying cool when things get hot: Emotion regulation modulates neural mechanisms of memory encoding. Frontiers in human neuroscience, 4, 230.

Hirsch, C. R., Meeten, F., Krahé, C., & Reeder, C. (2016). Resolving ambiguity in emotional disorders: The nature and role of interpretation biases. Annual review of clinical psychology, 12(1), 281-305.

Itkes, O., & Kron, A. (2019). Affective and semantic representations of valence: A conceptual framework. Emotion Review, 11(4), 283-293.

Itkes, O., Kimchi, R., Haj-Ali, H., Shapiro, A., & Kron, A. (2017). Dissociating affective and semantic valence. Journal of Experimental Psychology: General, 146(7), 924.

Kam, J. W., Wan-Sai-Cheong, L., Zuk, A. A. O., Mehta, A., Dixon, M. L., & Gross, J. J. (2024). A brief reappraisal intervention leads to durable affective benefits. Emotion, 24(7), 1676.

Kim, S. H., & Hamann, S. (2007). Neural correlates of positive and negative emotion regulation. Journal of cognitive neuroscience, 19(5), 776-798.

Krendl, A. C., Kensinger, E. A., & Ambady, N. (2012). How does the brain regulate negative bias to stigma?. Social Cognitive and Affective Neuroscience, 7(6), 715-726.

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). Behavior research methods, 49(2), 457-470.

Lazarus, R. S., & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. The Journal of Abnormal and Social Psychology, 69(2), 195.

Lindquist, K. A., & Barrett, L. F. (2008). Constructing emotion: The experience of fear as a conceptual act. Psychological science, 19(9), 898-903.

McRae, K., Gross, J. J., Weber, J., Robertson, E. R., Sokol-Hessner, P., Ray, R. D., ... & Ochsner, K. N. (2012). The development of emotion regulation: an fMRI study of cognitive reappraisal in children, adolescents and young adults. Social cognitive and affective neuroscience, 7(1), 11-22.

Milyavsky, M., Webber, D., Fernandez, J. R., Kruglanski, A. W., Goldenberg, A., Suri, G., & Gross, J. J. (2019). To reappraise or not to reappraise? Emotion regulation choice and cognitive energetics. Emotion, 19(6), 964.

Modinos, G., Ormel, J., & Aleman, A. (2010). Individual differences in dispositional mindfulness and brain activity involved in reappraisal of emotion. Social cognitive and affective neuroscience, 5(4), 369-377.

Morawetz, C., Bode, S., Baudewig, J., Kirilina, E., & Heekeren, H. R. (2016). Changes in effective connectivity between dorsal and ventral prefrontal regions moderate emotion regulation. Cerebral Cortex, 26(5), 1923-1937.

Murphy, J., Catmur, C., & Bird, G. (2019). Classifying individual differences in interoception: Implications for the measurement of interoceptive awareness. Psychonomic bulletin & review, 26(5), 1467-1471.

Nook, E. C., Satpute, A. B., & Ochsner, K. N. (2021). Emotion naming impedes both cognitive reappraisal and mindful acceptance strategies of emotion regulation. Affective Science, 2(2), 187-198.

Nook, E. C., Vidal Bustamante, C. M., Cho, H. Y., & Somerville, L. H. (2020). Use of linguistic distancing and cognitive reappraisal strategies during emotion regulation in children, adolescents, and young adults. Emotion, 20(4), 525.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. (2002). Rethinking feelings: an FMRI study of the cognitive regulation of emotion. Journal of cognitive neuroscience, 14(8), 1215-1229.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down-and up-regulation of negative emotion. Neuroimage, 23(2), 483-499.

Olteanu, L., Salama, K., Kimchi, R., & Kron, A. (2023). A divergent effect of stimulus perceptual details on affective and semantic representations of valence. Emotion, 23(5), 1236.

Phan, K. L., Fitzgerald, D. A., Nathan, P. J., Moore, G. J., Uhde, T. W., & Tancer, M. E. (2005). Neural substrates for voluntary suppression of negative affect: a functional magnetic resonance imaging study. Biological psychiatry, 57(3), 210-219.

Picó-Pérez, M., Radua, J., Steward, T., Menchón, J. M., & Soriano-Mas, C. (2017). Emotion regulation in mood and anxiety disorders: A meta-analysis of fMRI cognitive reappraisal studies. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 79, 96-104.

Russell, J. A. (1980). A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. Psychological review, 110(1), 145.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. Psychological review, 69(5), 379.

Troy, A. S., Shallcross, A. J., & Mauss, I. B. (2013). A person-by-situation approach to emotion regulation: Cognitive reappraisal can either help or hurt, depending on the context. Psychological science, 24(12), 2505-2514.

Urry, H. L., Van Reekum, C. M., Johnstone, T., Kalin, N. H., Thurow, M. E., Schaefer, H. S., ... & Davidson, R. J. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. Journal of Neuroscience, 26(16), 4415-4425.

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.

Visted, E., Vøllestad, J., Nielsen, M. B., & Schanche, E. (2018). Emotion regulation in current and remitted depression: a systematic review and meta-analysis. Frontiers in psychology, 9, 756.

Wang, L., Li, X., Pi, Z., Xiang, S., Yao, X., & Qi, S. (2021). Spatiotemporal dynamics of affective and semantic valence among women. Frontiers in Human Neuroscience, 15, 602192.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. Psychological bulletin, 98(2), 219.

Waugh, C. E., Vlasenko, V. V., & McRae, K. (2022). What parts of reappraisal make us feel better? Dissociating the generation of reappraisals from their implementation. Affective Science, 3(3), 653-661.

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.

Winecoff, A., LaBar, K. S., Madden, D. J., Cabeza, R., & Huettel, S. A. (2011). Cognitive and neural contributors to emotion regulation in aging. Social cognitive and affective neuroscience, 6(2), 165-176.

Wolgast, M., Lundh, L. G., & Viborg, G. (2011). Cognitive reappraisal and acceptance: An experimental comparison of two emotion regulation strategies. Behaviour research and therapy, 49(12), 858-866.

Supplemental Materials

To accompany

Cognitive reappraisal changes cognitive evaluations more than affective experiences

Vartiainen & Nook

Contents

I. Study 1 Task Instructions

Default condition	Affective condition	Semantic condition

Practice instructions

For this experiment, you will see a series of images and be asked to rate **[how they make you feel** / **how they make you feel** / **what you think of them.]** Each image will be presented along with either the word LOOK or the word CHANGE above the image. These are instructions that ask you to interact with the image in a certain way, and we'll teach you what they mean. Let's walk through an example!

Look/Change practice

[put up picture with LOOK above]
Here is a picture with the word LOOK above it. LOOK means you should simply observe the image naturally. Don't try to change your reactions at all.

[put up picture with CHANGE above]
Here is another picture with the word CHANGE above it. CHANGE means you should try to make yourself feel better about the image by reinterpreting the scene so that it means something different to you. For example, you can create a story or context around the image that changes how you feel about it. Try doing that now.

Look/Change attention check

In this experiment, you will be shown images. What will you do during the LOOK and CHANGE trials?

LOOK: simply observe the image, CHANGE: reinterpret the image to make myself feel better
LOOK: simply observe the image, CHANGE: simply observe the image
LOOK: reinterpret the image to make myself feel better, CHANGE: simply observe the image
LOOK: reinterpret the image to make myself feel better, CHANGE: reinterpret the image to make myself feel better

Rating scales practice

Great! After you see each picture you will then have the chance to rate your response. Now we'll teach you how to use the rating scales.

After each image, you'll see two rating scales, one asking **[how positive you feel, and the other asking how negative you feel** / **how positive you feel, and the other asking how negative you feel** / **how positive the image is, and the other asking how negative the image is.]** Order will vary across trials, so pay attention to which one you're answering.
Important: We want to know your ratings after you've followed either the LOOK or the CHANGE instruction. This means that your ratings after CHANGE trials should reflect **[how you feel / how you feel / what you think]** AFTER you've changed how you think about it. Your

ratings after LOOK trials should reflect how you **[feel / feel / think]** when you see the image, without trying to change your **[feelings / feelings / thoughts]**. You can click on the scales to give your response. You will have 5 seconds to respond.

[put up a negative rating scale]
Here's an example of a scale. Click on the scale to give a rating.
[put up other positive rating scale]
Here is the other rating scale for you to try out!

AFFECTIVE/SEMANTIC INSTRUCTIONS

In this experiment you will see a set of pictures. Each picture contains a scene or event, and we will ask you to report how positive and how negative **[you feel / the event in the picture is].**

Now, making reports like these are not always easy, as sometimes it is not clear what to report about when asked this question. In these instructions we will explain to you what we mean when we ask you to report about **[your feelings / the picture].**

For this experiment, we want to distinguish between what you feel inside and the contents of the picture. Today, we want you to **[rate the feeling you experienced inside, NOT the contents of the picture/ rate the contents of the picture, NOT the feelings you experienced inside.]**

It is possible that you will look at a picture you know has negative content but you won't feel anything inside. For example, you can look at a picture of a car accident and not feel strong emotions inside, even though you know that car accidents are negative. If that happens, you should rate **[that you do not feel strong feelings / how negative the image's scene is, regardless of how you feel].**

We are NOT asking you **[if the picture was pleasant or unpleasant, your opinion about the picture, what you expected yourself to feel or what you thought we expected you to feel. / if the picture made you feel pleasant or unpleasant, what you expected yourself to think of the picture, or what you thought we expected you to think of the picture.] We only want you to rate [the feeling that you experienced inside while viewing the picture / the positivity/negativity of the picture's contents, not the feeling you experienced inside while viewing the picture.]**

AFFECTIVE/SEMANTIC ATTENTION CHECK

In the following section, you will be shown images. What will you report?

- How positive or negative I actually feel
- How positive or negative I think I should feel
- How positive or negative the image is

FINAL PRACTICE (All conditions)

Ok let's put that all together. You'll now complete two practice trials back-to-back.

For each trial, you will see a picture on the screen. While you see the image, LOOK means you should respond naturally to the picture, CHANGE means you should reinterpret the picture to make yourself feel better. After the image disappears, you will have 5 seconds rate how positive and negative **[you feel / you feel / the image is]** on the scales. Your ratings after CHANGE trials should reflect what you **[feel / feel / think]** after you have reinterpreted the picture. Your ratings after LOOK trials should reflect how you **[feel / feel / think]** when you see the image, without trying to change your **[feelings / feelings / thoughts]**. The order will vary across trials, so pay attention to which one you're answering!
[one look, one change trial]

Main task instructions

Final instructions (All conditions)
Great! Now we will move on to the actual experiment.

As a reminder, if the word LOOK is presented under the image, simply observe the image naturally. If the word CHANGE is presented under the image, try to make yourself feel better about the image by reinterpreting the scene so that it means something different to you, something that makes you feel better.

After each image, you'll have two rating scales, one asking **[how positive you feel, and the other asking how negative you feel. / how positive you feel, and the other asking how negative you feel. / how positive the image is, and the other one asking how negative the image is.]** Your ratings after CHANGE trials should reflect what you **[feel / feel / think]** after you have reinterpreted the picture. Your ratings after LOOK trials should reflect how you **[feel / feel / think]** when you see the image, without trying to change your **[feelings / feelings / thoughts]**. You will have 5 seconds to respond. The order will vary across trials, so pay attention to which one you're answering! Please be as honest as you can.

[Remember: you are supposed to rate how you actually feel, NOT how positive or negative the image is. / Remember: you are supposed to rate how positive or negative the event in the image is, NOT how positive or negative you feel.]

Please press the button below to begin the task.

II. Study 2 Task Instructions

Practice instructions

1) Valence instructions (affective/semantic)

In this experiment you will see two sets of pictures. In this first set, we want you to make a set of ratings about the picture. Each picture contains a scene or event, and we will ask you to report 1) how positive or negative you feel when looking at the picture and 2) how positive or negative the event in the picture is.

Your ratings should reflect how you feel and think when you see the image, without trying to change your feelings or thoughts.

Now, making ratings like these are not always easy, as it's not always clear what to report about when asked these questions. In these instructions we will explain to you what we mean when we ask you to report about your feelings and the picture.

We want you to rate two different things: 1) how you feel inside and 2) what you think about the meaning of the picture. We want you to distinguish between these things and rate each of them on different scales.

1) When we ask you to rate **how negative/positive you feel**, we only want you to rate the actual feeling that you experienced inside while viewing the picture, not what you expected to feel or what you think about the contents of the image.
2) When we ask you how **negative/positive the image is**, we only want you to rate the picture's contents, regardless of your feelings about those contents.

Let's walk through an example. You might look at a picture of a car accident and not feel strong emotions inside, but you might think that car accidents are negative scenes. On this trial, then, you should rate little negativity when we ask how you feel, but you would give a higher negativity rating when we ask what you think about the picture.

In addition to rating negativity, you will also be rating positivity. So you will make 4 ratings for each picture.

2) Rating scales practice

Now we'll let you practice using the rating scales. You will have 8 seconds to respond to each.
[put up rating scales]

3) Attention checks

In the following section, you will be shown images. When you are asked to report how positive/negative you feel, what will you report?

- How positive or negative I actually feel
- How positive or negative I think I should feel or think
- How positive or negative the image is, regardless of how I feel

In the following section, you will be shown images. When you are asked to report how positive/negative the image is, what will you report?

- How positive or negative I actually feel
- How positive or negative I think I should feel or think
- How positive or negative the image is

4) Baseline practice

Ok let's put that all together. You'll now complete two practice trials back-to-back.

For each trial, you will see a picture on the screen. After the image disappears, you will answer two questions: One about how positive/negative you feel, and one about how positive/negative the image is. Your ratings should reflect how you naturally feel and think when you see the image, without trying to change your feelings or thoughts. You will have 8 seconds to answer each question.

The order will vary across trials, so pay attention to which one you're answering!

[two practice trials]

5) Baseline final instructions

Great! Now we will move on to the actual experiment.

After each image, you will answer two questions: how positive/negative you feel, and how positive/negative the image is. Your ratings should reflect how you feel and think when you see the image, without trying to change your feelings or thoughts. You will have 8 seconds to respond. The order will vary across trials, so pay attention to which one you're answering! Please be as honest as you can.

Remember:

1) When we ask you to rate how you feel, you are supposed to rate how you actually feel, NOT how positive or negative the image is.
2) When we ask you to rate the image, you are supposed to rate how positive or negative the event in the image is, NOT how positive or negative you feel.

Please press the button below to begin the task.

[baseline task]

6) Reappraisal instructions

You have now completed the first set of ratings.

In the next part of the experiment, you will be shown the same images. However, this time we will ask you to do something different! When you see each image this time, you should try to

make yourself feel better about it by reinterpreting the scene so that it means something different to you. You should create a story or context around the image that changes how you feel about it.

You'll then rate the image on the same scales as before. However, your ratings should reflect how you feel and how you think AFTER you've changed how you think about it. Tell us how negative/positive your feelings inside are and how negative/positive the meaning of the image is AFTER you've reinterpreted it.

7) Reappraisal practice

You'll now complete two practice trials back-to-back.

For each trial, you will see a picture on the screen. While you see the image, you should reinterpret the picture to make yourself feel better. After the image disappears, you will have 8 seconds rate how positive/negative you feel, and how positive/negative the image is on the scales. Your ratings should reflect what you feel and think AFTER you have reinterpreted the picture. The order will vary across trials, so pay attention to which one you're answering!

[two practice trials]

8) Reappraisal final instructions

Great! Now we will move on to the actual experiment.

After each image, you will answer two questions: how positive/negative you feel, and how positive/negative the image is. Your ratings should reflect what you feel and think AFTER you have reinterpreted the picture. You will have 8 seconds to respond. The order will vary across trials, so pay attention to which one you're answering! Please be as honest as you can.

Remember:

1) When we ask you to rate how you feel, you are supposed to rate how you actually feel, NOT how positive or negative the image is.
2) When we ask you to rate the image, you are supposed to rate how positive or negative the event in the image is, NOT how positive or negative you feel.

Please press the button below to begin the task.

III. Data Quality and Exclusions

Table S1. Characteristics of the main dataset, full dataset, and excluded dataset.

	Main dataset	Full dataset	Excluded dataset
Study 1			
% missing			
Affective	11.2	12.30	17.30
Default	9.49	10.50	13.00
Semantic	7.25	9.49	16.20
Look rating Mean (SD)	-30.40 (55.90)	-28.10 (57.80)	-20.50 (62.90)
Change rating Mean (SD)	13.60 (56.70)	13.20 (58.40)	11.80 (63.90)
Average rating Mean (SD)			
Affective	-10.80 (61.30)	-6.00 (62.80)	10.10 (66.40)
Default	-1.55 (59.40)	-2.79 (59.70)	-6.11 (60.40)
Semantic	-13.00 (60.00)	-12.80 (62.20)	-12.10 (69.30)
Study 2			
% missing	8.20	8.20	15.40
Look rating Mean (SD)	-33.90 (54.60)	-33.40 (54.40)	-31.90 (53.80)
Change rating Mean (SD)	1.93 (59.50)	1.71 (58.40)	1.11 (55.00)
Average rating Mean (SD)			
Affective	-8.59 (59.90)	-9.83 (59.10)	-13.40 (56.60)
Semantic	-22.90 (54.60)	-21.20 (58.60)	-16.40 (57.00)

IV. Study 1 Full Sample Results

In our main analyses, we excluded participants who failed our preregistered attention checks. As shown in **Table S1**, excluded participants had a notably higher amount of missing data (13-17.30%) and higher variability (*SDs* 60.40-69.30) compared to our main dataset (7.25-11.2%; *SDs* 55.90-61.30). Nonetheless including all *N* = 205 participants did not change the overall pattern of results observed in the main sample (see below for detailed results). In the full dataset, the main effects and the interaction remained significant (p = .029), and the direction of effects was consistent with our main analyses. However, individual pairwise contrasts were only marginally significant (*p*s .05-.06), likely reflecting increased noise introduced by lower-quality data. These results indicate that our findings are reliable, and that excluding lower quality data primarily reduced noise rather than driving results reported in the manuscript.

We found a significant main effect of Reappraisal on valence ratings, $F(1, 7135.90) = 1051.80$, $p < .001$, indicating that participants reported feeling worse on Look ($M = -28.10$, $SD = 57.80$) than Change ($M = 13.20$, $SD = 58.40$) trials. Follow-up contrasts within each condition showed that ratings became significantly more positive from Look to Change in all three conditions: affective ($b = 36.30$, $SE = 2.33$, $z = 15.59$, 95% CI [31.70, 40.80], $p < .001$), default ($b = 43.50$, $SE = 2.09$, $z = 20.86$, 95% CI [39.40, 47.60], $p < .001$), and semantic ($b = 43.80$, $SE = 2.18$, $z = 20.09$, 95% CI [39.50, 48.10], $p < .001$).

There was also a significant main effect of Valence Instructions on unidimensional valence ratings, $F(2, 197.40) = 3.34$, $p = .038$, indicating that rating intensity overall varied across the affective ($M = -6.99$, $SD = 62.80$), semantic ($M = -12.80$, $SD = 62.20$), and default ($M = -2.79$, $SD = 59.70$) conditions. Follow-up pairwise contrasts showed that this was due to more positive default compared to semantic ratings ($b = 10.02$, $SE = 3.89$, $z = 2.58$, 95% CI [0.90, 19.13], $p = .027$). Other pairwise comparisons were not significant (*p*s > .318).

The interaction between Reappraisal and Valence Instruction was also significant, $F(2, 7135.30) = 3.55$, $p = .029$. Valence change due to reappraisal was marginally smaller in the affective valence condition compared to both the default ($b = -7.25$, $SE = 3.13$, $z = -2.32$, 95% CI [-14.70, 0.21], $p = .060$) and semantic conditions ($b = -7.54$, $SE = 3.19$, $z = -2.36$, 95% CI [-15.20, 0.08], $p = .053$). The impact of reappraisal on valence ratings did not differ between the semantic and default conditions ($b = -0.29$, $SE = 3.02$, $z = -0.10$, 95% CI [-7.50, 6.92], $p = .999$).

V. Study 1 Positive and Negative Valence Results

Here we report results from just the sample that passed inclusion criteria (N = 155), but we report separate analyses of negative affect and positive affect ratings (as pre-registered).

Negative valence. We found a significant main effect of reappraisal on negative valence, $F(1, 5660.0) = 671.85$, $p < .001$, indicating that negative valence ratings differed between Look (M = 60.00, SD = 33.30) and Change (M = 39.80, SD = 32.20) trials. Follow-up contrasts within each condition showed that negative ratings decreased significantly from Look to Change in all three conditions: affective (b = 16.00, SE = 1.39, z = 11.50, *95% CI* [13.20, 18.70], $p < .001$), default (b = 20.70, SE = 1.33, z = 15.55, *95% CI* [18.10, 23.30], $p < .001$), and semantic (b = 24.20, SE = 1.35, z = 17.96, *95% CI* [21.60, 26.90], $p < .001$).

There was also no significant main effect of Valence Instructions, $F(2, 151.70) = 2.31$, $p = .0103$, indicating that negative ratings did not significantly differ between conditions.

Critically, there was a significant interaction between Reappraisal and Valence Instructions, $F(2, 5660.00) = 9.15$, $p < .001$, indicating that the effect of reappraisal differed across the three valence instruction conditions. Follow-up contrasts showed that the magnitude of change for negative ratings was lower in the affective condition compared to default (b = 4.73, SE = 1.92, z = 2.46, *95% CI* [0.137, 9.31], p = .041) and semantic conditions (b = 8.25, SE = 1.93, z = 4.27, *95% CI* [3.64, 12.87], $p < .001$). However, changes in negative valence due to reappraisal did not differ across semantic and default conditions (b = 3.53, SE = 1.89, z = 1.86, *95% CI* [-0.99, 8.05], p = .176).

Positive valence. There was a significant main effect of Reappraisal for positive valence ratings as well, $F(1, 5675.40) = 960.19$, $p < .001$, as Look trials (M = 29.90, SD = 29.70) were less positive than Change trials (M = 52.80, SD = 32.40). Positive ratings increased significantly from Look to Change in all three conditions: affective (b = 19.30, SE = 1.30, z = 14.79, *95% CI*

[16.70, 21.80], $p < .001$), default ($b = 25.00$, $SE = 1.25$, $z = 20.04$, *95% CI* [22.60, 27.50], $p < .001$), and semantic ($b = 24.10$, $SE = 1.27$, $z = 18.96$, *95% CI* [21.60, 26.60], $p < .001$).

Interestingly, there was no significant main effect of condition for positive affect ratings in Study 1, $F(2, 151.50) = 2.466$, $p = .088$.

The interaction between reappraisal and condition was also significant, $F(2, 5675.40) = 5.834$, $p = .003$, indicating that the effect of reappraisal on positive ratings differed across the three conditions. Follow-up contrasts showed that the magnitude of change was lower in the affective condition compared to default ($b = -5.77$, $SE = 1.80$, $z = -3.20$, *95% CI* [-10.08, -1.46], $p = .004$) and semantic conditions ($b = -4.88$, $SE = 1.82$, $z = -2.68$, *95% CI* [-9.22, -0.53], $p = .022$). There were no significant differences between semantic and default conditions ($b = 0.90$, $SE = 1.78$, $z = 0.50$, *95% CI* [-3.36, 5.16], $p = .943$).

VI. Study 2 Full Sample Results

As for Study 1, we excluded participants who failed the preregistered attention checks from our main analyses. In Study 2, the excluded participants again had a higher proportion of missing data (15.40% compared to 8.20%). Below we present results from all *N* = 145 participants in Study 2 who completed the study, even if they did not pass pre-registered attention checks. Including these participants led the interaction between Reappraisal and Valence Instructions to fall to marginal significance (p = .054), although valence change was still greater for semantic than affective valence. Further analysis suggests this lack of significance is likely due to non-attentive participants not following task instructions and thereby eliminating the subtle within-subjects manipulation in Study 2. In fact, we depict the Study 2 valence change results for just the included *N* = 105 participants, the full *N* = 145 participants, and just the *N* = 40 participants we excluded in **Figure S1**. As can be seen, excluded participants did not show any difference in regulatory efficacy across Valence Instructions Conditions, in line with the idea that these participants were non-attentive to Valence Rating instructions, weakening the overall effect. Given that the direction of effects was replicated across datasets (and fully replicated in Study 1), we believe our pre-registered exclusion criteria were appropriate and remain confident in the conclusions presented in the main text.

There was a significant main effect of reappraisal on valence ratings, $F(1, 10306) = 1186.63$, $p < .001$, indicating that participants reported overall improved valence on reappraisal trials ($M = 1.71$, $SD = 58.40$) compared to baseline trials ($M = -33.40$, $SD = 54.40$). Follow-up contrasts within each condition showed that this effect of reappraisal was significant for both affective ratings ($b = 33.40$, $SE = 1.45$, $z = 23.09$, 95% CI [30.60, 36.30], $p < .001$) and semantic ratings ($b = 37.50$, $SE = 1.45$, $z = 25.75$, 95% CI [34.50, 40.20], $p < .001$).

We also observed a significant main effect of Valence Instructions, $F(1, 10277) = 123.01$, $p < .001$, indicating that valence ratings were overall more positive for affective ratings ($M = -9.83$, $SD = 59.10$) than semantic ratings ($M = -22.10$, $SD = 58.60$).

Finally, the interaction between Reappraisal and Valence Instructions was marginally significant, $F(1, 10276) = 3.73$, $p = .054$, and showed a similar trend where the magnitude of change for valence ratings was lower in the affective condition compared to semantic ($b = -3.95$, $SE = 2.05$, $z = -1.93$, *95% CI* [-7.96, 0.06], $p = .054$).

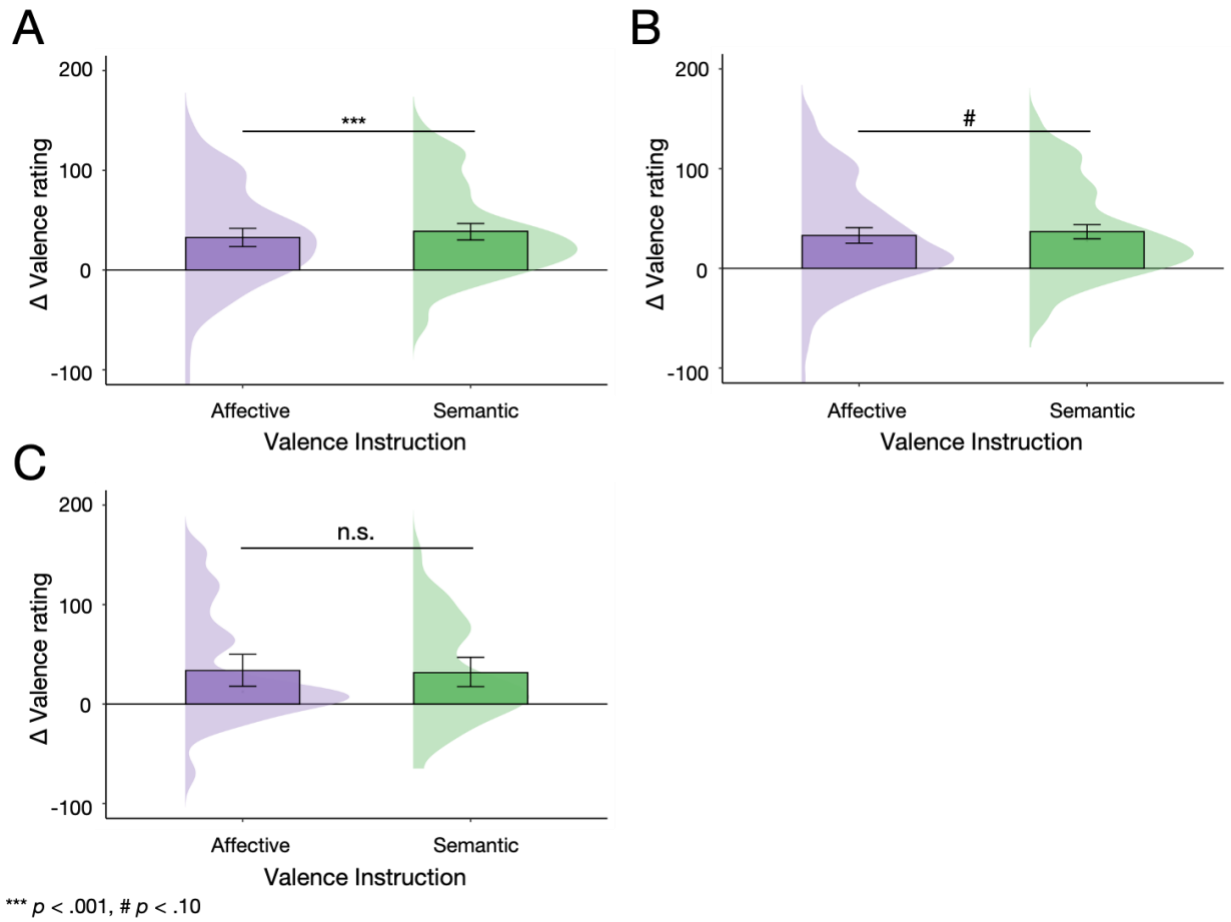

Figure S1. Study 2 results using the main, full, and excluded datasets. The effect of reappraisal on affective and semantic ratings of valence, when manipulated within-subjects. A) In the main results (N=105), Valence instructions influenced the magnitude of valence change following reappraisal. Affective ratings were significantly less influenced by reappraisal than semantic ratings. B) In the full dataset (N=145), Valence instructions had a marginal effect on the magnitude of change following reappraisal ($p = .054$), with affective ratings significantly less influenced by reappraisal than semantic ratings. C) In the excluded dataset (N=40), Valence instructions did not influence the magnitude of valence change following reappraisal.

VI. Study 2 Positive and Negative Valence Results

Here we report results from just the sample that passed inclusion criteria (N = 105), but we report separate analyses of negative affect and positive affect ratings.

Negative valence. There was a significant main effect of reappraisal on negative valence ratings, $F(1, 8004.90) = 673.82$, $p < .001$, indicating that ratings were more negative for baseline (M = 60.90, SD = 33.20) than reappraisal (M = 43.70, SD = 33.30) trials. Follow-up contrasts within each condition showed that negative ratings decreased significantly from baseline to reappraisal in both affective (b = -15.10, SE = 0.95, z = -15.80, *95% CI* [-16.90, -13.20], $p < .001$) and semantic (b = -19.90, SE = 0.95, z = -20.92, *95% CI* [-21.80, -18.00], $p < .001$) conditions.

There was also a significant main effect of Valence Instructions on negative ratings, $F(1, 8004.00) = 149.27$, $p < .001$, showing that participants provided less negative ratings on the affective (M = 48.10, SD = 34.20) than the semantic (M = 56.40, SD = 34.10) scales.

The interaction between reappraisal and condition was also significant for negative ratings, $F(1, 8003.40) = 13.03$, $p < .001$, indicating that the effect of reappraisal on negative affect differed across conditions. As for unidimensional ratings, follow-up contrasts showed that the magnitude of change in negative affect was lower in the affective condition compared to the semantic condition (b = -4.86, SE = 1.35, z = -3.61, *95% CI* [-7.50, -1.51], p = .009).

Positive valence. There was a significant main effect of Reappraisal on positive affect ratings, $F(1, 7746.60) = 823.41$, $p < .001$, such that ratings were less positive on baseline (M = 28.90, SD = 28.90) than reappraisal (M = 47.20, SD = 32.40) trials. Follow-up contrasts within each condition showed that ratings increased significantly from baseline to reappraisal in both affective (b = -32.30, SE = 1.68, z = -19.24, *95% CI* [-35.60, -29.10], $p < .001$) and semantic (b = -39.50, SE = 1.69, z = -23.40, *95% CI* [-42.90, -36.20], $p < .001$) conditions.

There was also a significant main effect of Valence Rating Instructions, $F(1, 7745.40) = 80.66$, $p < .001$, showing that positive ratings were higher in the affective ($M = 41.00$, $SD = 32.50$) than semantic ($M = 35.30$, $SD = 31.40$) conditions. The interaction between reappraisal and condition was not significant $F(1, 7744.10) = 2.13$, $p = .144$.