Affect Regulation during Learning: The Enhancing

Effect of Cognitive Reappraisal
Abstract

Affective experiences routinely occur during learning and need to be successfully regulated. In two experiments, we used an intervention that combined elements of utility value and cognitive reappraisal to gauge its effects on engagement and performance. We predicted that participants using the reappraisal strategy would experience more engagement and higher learning outcomes than controls. Ethnically diverse adult learners ranging from 18 to 58 years of age (Experiment 1, N = 93; Experiment 2: N = 138) used affect regulation strategies or no strategy (control) in an online learning environment. Engagement and learning outcomes were measured throughout the experiment. Participants who used reappraisal generally reported more engagement and achieved higher learning outcomes than controls. A mediation analysis revealed evidence of a partial mediation effect of affective engagement between reappraisal and learning outcomes in Study 2. We discuss the implications of these findings for understanding the role of affect regulation during learning.

Keywords: affect, affect regulation, self-regulated learning.
Affect Regulation during Learning: The Enhancing Effect of Cognitive Reappraisal

Learning does not end after high school or upon completion of higher education. The process of learning takes place over a lifetime, as individuals strive to maintain the knowledge and skills that they need to be successful in an ever-changing world. Today’s technologies provide access to virtually limitless troves of information, which provide individuals with the opportunity to educate themselves on practically any topic they find worthy of understanding. With a world of information at their fingertips, it is imperative that learners of all ages are equipped with the skills to manage and regulate their own learning over the course of a lifetime. Effective self-regulated learning requires a host of skills including locating credible information sources, comprehending materials deeply, assessing understanding of new information, and even coping with the emotional (or affective) experiences that accompany learning. In this article, we explore a method for helping traditional and non-traditional learners more effectively regulate their affect so that they might become better lifelong learners.

The present focus on affect is motivated by the fact that affect is inextricably linked to both the processes and products of learning. Research has demonstrated that learners of all ages experience a wide array of positive and negative affective states including confusion, frustration, boredom, shame, engagement, anxiety, pride, hope, and many others (Csikszentmihalyi, 1990; D’Mello & Graesser, 2011; 2010; Lepper & Henderlong, 2000; Linnenbrink & Pintrich, 2004; Meyer & Turner, 2006; Nett, Goetz, & Daniels, 2010; Nett, Goetz, & Hall, 2011; Pekrun, Elliott, & Maier, 2009). These affective states can significantly facilitate or impair the cognitive, metacognitive, and motivational processes that are critical for learning.

It is important to note that with the exception of test anxiety, research on affect during learning is still in its infancy. Nevertheless, the research that does exist indicates that learners
should be equipped with the ability to regulate their affective states in order to increase task persistence and improve learning outcomes (Davis, DiStefano & Schutz, 2008; Pekrun, 2006). Affect regulation (AR) pertains to the set of processes individuals use to increase, decrease, or maintain particular affective states in order to achieve desired outcomes (Gross & Thompson, 2007). Unfortunately, learners generally lack the skills required for engaging in appropriate self-regulation (Azevedo, Guthrie, & Seibert, 2004; Greene & Land, 2000) and affect regulation (O’pt Eynde, Corte, & Verschaffel, 2007). Hence, the goal of the present research is to determine whether theoretically grounded AR interventions are effective for regulating affect in order to increase affective engagement and learning.

**Theoretical Framework**

The guiding theoretical framework consists of the control-value theory of academic emotions (Pekrun, 2006) and theories of affect regulation that emphasize the role of cognitive reappraisal strategies (among other strategies) in the up- and down-regulation of affective states (Gross, 2008; Gross & Thompson, 2007). We do this by developing and testing a strategy that combines aspects of utility value and cognitive reappraisal manipulations.

**Control-value theory.** According to appraisal theories of affect, the appraisals individuals make about their environment, and their ability to handle challenges or threats in that environment, influence the affective states that are experienced (Clore & Ortony, 2010; Lazarus, 1991; Roseman & Smith, 2001; Scherer, 2001). *Appraisal* refers to the interpretation of a particular event on the basis of a number of dimensions including novelty, goal congruence, agency, and coping potential (for a review, see Roseman & Smith, 2001). The control-value theory posits that the two types of appraisals that influence affect during learning are subjective
control and subjective value. Subjective control refers to the appraisal that one’s own behaviors will either (a) produce positive outcomes, or (b) reduce or prevent negative outcomes. Subjective value refers to the perceived importance or value of a particular learning activity.

The control-value theory suggests that successful affect regulation can be accomplished by addressing the control and value antecedents of affect. This type of regulation is referred to as appraisal-oriented regulation, because it involves restructuring one’s appraisals of control or value in the learning task. The present research focuses specifically on appraisals of value during learning of a topic that is typically perceived to be of little personal value.

Learners appraise a topic as having little value when there are no immediately obvious benefits of learning about the topic. For instance, learners may feel that completing geometry homework does not have much value because being able to compute the area of a rhombus is not typically considered to be an important life skill. When learners appraise a topic as having little personal value, they may choose not to invest much effort in learning that topic, and engagement will be low. However, learners might be more likely to be actively engaged in learning the topic if they attempt to think about the utility value of the topic. Appraisals of utility value pertain to learners’ beliefs about whether the topic is relevant or useful for other aspects of their lives (Eccles et al., 1983; Simons, Vansteenkiste, Lens, & Lacante, 2004). As such, learners may perceive that a topic has utility value if they believe that learning about the topic will aid them in such endeavors as obtaining a high school degree, being accepted to a university, or acquiring a desirable job (Durik & Harackiewicz, 2007; Godes, Hulleman, & Harackiewicz, 2007; Simons et al., 2004). In the example above, learners who consider that doing well in geometry might increase their chances of being accepted to a good college might be more willing to be invested and engaged in learning the topic.
There are few experimental studies that manipulate utility value and study its effect on positive affect and engagement. The studies that do exist suggest that such interventions can increase engagement and promote positive learning outcomes (Harackiewicz, Durik, Barron, Linnenbrink, & Tauer, 2008; Hidi & Renninger, 2006; Hulleman, Durik, Schweigert, & Harackiewicz, 2008; Hulleman, Godes, & Hendricks, & Harackiezicz, 2010; Simons, Vansteenkiste, Lens, & Lacante, 2004). For example, Hulleman et al. (2010) used a utility value intervention to increase learners’ appraisals of the value of learning about math or psychology. Compared to learners in a control condition, learners who used the utility-value intervention appraised the task as being significantly more valuable and more interesting, and achieved better learning outcomes. Similar results were found by Jang (2008), and Johnson and Sinatra (2013) who both found that utility value interventions led to increased engagement and conceptual change. Utility value interventions have also been found to foster prolonged, sustained engagement in learning. Harackiewicz and colleagues (2012) found that students whose parents were instructed to emphasize the utility value of learning STEM subjects enrolled in nearly one semester more of STEM courses in their last two years of high school, compared to a control group. Considered together, these findings support the control-value theory’s claim that appraisals of utility value can significantly influence learner affect and performance.

**Theories of affect regulation.** Gross’s modal model (Gross, 2008; Gross & Barrett, 2011; Gross & Thompson, 2007) posits that affect is malleable and can be modulated in a number of ways in order to achieve desired outcomes (i.e., desirable thoughts, feelings, or behaviors). Affect regulatory processes can be conscious or unconscious, automatic or controlled, spontaneous or instructed, and can be effective or ineffective (depending on the appropriateness of the selected regulatory strategy) (Gross & Thompson, 2007).
Five classes of affect regulation strategies have been identified in the literature (Gross & Munoz, 1995; see also Gross & Thompson, 2007). These include situation selection (approaching/avoiding environments that elicit affective responses), situation modification (making systematic changes to the given environment), attentional deployment (ignoring/attending to affective stimuli), response modulation (suppressing affective responses to stimuli), and cognitive change (described below).

Of particular relevance to the present paper is cognitive change, which involves systematically changing one’s appraisals about a situation in order to alter its affective impact. The most commonly studied type of cognitive change is cognitive reappraisal (or reappraisal for short) (Giuliani & Gross, 2009; Gross, 2008; John & Gross, 2007). Since appraisals play a critical role in the generation and experience of affect, reappraisals should alter the affective experience (McRae et al., 2010). Reappraisal interventions (called instructed reappraisals from here on) typically involve instructing individuals to think of a negative affective experience in a way that makes the experience seem less negative. Instructed reappraisals often prompt participants to use their imagination by taking on the perspective of someone other than themselves. For instance, Goldin, McRae, Ramel, and Gross (2008) prompted participants to use an instructed reappraisal strategy while they watched affect-eliciting films. Participants were asked to view disgust-inducing videos of surgical procedures, vomiting, and animal slaughter and were instructed to assume the perspective of an objective medical professional whose job was to focus on the technical aspects of the film. Participants who took on the role of a medical professional experienced less negative affect than participants who used no reappraisal strategy, which suggests that instructed reappraisal can be a useful mechanism for regulating affect.
Instructed reappraisal has also been shown to improve mood and interpersonal functioning (Dennis, 2007; Gross & John, 2003; Moore, Zoellner, & Mollenholt, 2008), diminish depressive symptoms (Garnefski & Kraaj, 2006), engage executive function processes related to cognitive control (McRae et al., 2010; Ochsner & Gross, 2008), improve memory for important details (Dillon, Ritchey, Johnson, & LaBar, 2007; Richards, Butler, & Gross, 2003), and lead to decreased blood pressure and cardiovascular arousal (Richards & Gross, 1999). In addition, one recent study demonstrated that individuals, in general, consider reappraisal to be a useful and effective strategy for regulating the affective states that arise in their everyday lives (Lieberman, Inagaki, Tabibnia, & Crockett, 2011).

With such a multifaceted body of research implicating the benefits of reappraisal, it is somewhat surprising that this strategy has received considerably less attention in applied contexts such as learning. At present, contemporary research on reappraisal in learning contexts has used cross-sectional designs and self-reports to correlate the spontaneous use of reappraisal with trait-based affective states (i.e., general predispositions in affective responding) and various outcomes. However, there have been a small number of experimental studies in which instructed reappraisals have been used in learning contexts. These studies focus primarily on helping learners reappraise negative affective experiences that arise during learning. For instance, Jamieson, Mendes, Blackstock, and Schmader (2010) instructed students to reappraise feelings of anxiety while they took the GRE by telling them to think about their anxiety more positively. Specifically, the authors instructed students that whenever they experienced anxiety, they should consider that anxiety can actually be beneficial to learning. The goal was to cause learners to reappraise anxiety (a negative and harmful affective experience) as having potential to help them perform better on the Graduate Record Examination (GRE). Students who used the instructed
reappraisal strategy exhibited a significant increase in sympathetic nervous system activation and achieved better performance on the GRE than learners who received no reappraisal instructions. This finding demonstrates that instructing learners to reappraise negative affective states when they arise can help learners avoid negative affect and achieve better learning outcomes.

While there is vast research on the use of reappraisal to decrease negative affect, there is considerably less research that seeks to help learners use instructed reappraisal to increase positive affect such as affective engagement during learning. As such, the present focus is to test for a causal link between instructed reappraisal, affective engagement, and subsequent learning outcomes. Affective engagement is thought to have three sub-components: a behavioral component that includes effort, persistence, question asking, and concentration; a cognitive component that includes self-regulation and strategy use; and an affective component (Fredricks, Blumenfield, & Paris, 2004). Following Baker, et al. (2010), we define affective engagement as a state of involvement in a task that is characterized by positive valence and moderate arousal. It is related to, but does not necessarily include, some aspects of Csikszentmihalyi’s (1991) conceptualization of flow, such as time distortion or less of self-consciousness. In this paper, we focus specifically on affective engagement, with the goal of expanding current literature on the relationships among instructed reappraisal, affective engagement (among other affective states), and learning outcomes.

**Current Study**

The goal of our research was to combine components of utility value and reappraisal to develop and test an instructed reappraisal strategy that may increase affective engagement in the applied context of learning. Our focus on helping learners become more actively engaged is
motivated by the fact that boredom (or the lack of engagement) is one of the most frequent affective states that students experience during learning (see D'Mello (2013) for a meta-analysis). Boredom also negatively correlates with learning outcomes (Craig et al., 2004; D’Mello & Graesser, 2011; Forbes-Riley & Litman, 2011; Schutz & Pekrun, 2007), and is a catalyst to harmful long-term outcomes such as lower self-efficacy, lack of motivation in learning, hostility and dissatisfaction towards school, abnormal behavior in school, lower work satisfaction, and diminished work output (Fogelman, 1976; McGiboney & Carter, 1988; Perkins & Hill, 1985; Robinson, 1975; Wasson, 1981). If the use of instructed reappraisal strategies can help learners become affectively engaged and improve their learning outcomes, then such strategies could be implemented in intelligent tutoring systems and web-based learning environments to help learners regulate their affect. For instance, while many learning environments provide automatic, adaptive prompting and feedback of metacognitive and cognitive strategies, there are no systems that we are currently aware of that prompt affect regulation strategies such as instructed reappraisal. These strategies could be easily implemented into learning environments whose aim is to help learners become more vigilant of their levels of affective engagement during learning, and to adaptively help learners regulate their affective states in order to improve learning.

In two experiments, we instructed a sample of participants to use instructed reappraisal strategies (IR for short) while learning about the United States Constitution and Bill of Rights. We chose to use a broad population for this study because the U.S. Constitution is an age-invariant, challenging document that is highly relevant to American citizens, yet is not well understood by most (Scaros, 2011). Therefore, we were motivated by the idea of assessing how individuals of all ages could improve their understanding of important topics like the U.S. Constitution.
The topic of the U.S. Constitution was selected because social studies is a generally disliked topic among learners of all ages, is appraised as having little value, and is associated with a high incidence of disengagement (Journell, 2009). As previously stated, the control-value theory indicates that seeing little value in a learning activity can lead to disengagement and can have negative consequences on learning outcomes.

As previously stated, pure utility value interventions have shown to be useful for promoting affective engagement during learning of topics that are perceived as being of little value. Likewise, reappraisal interventions have shown to increase positive affect, and both value and reappraisal interventions have shown to increase learning outcomes. To date, we are not aware of any existing research that attempts to combine utility value and reappraisal interventions to measure their combined effect on affective engagement and learning.

As a first step in that direction, we conducted two experiments to test an IR strategy that included components of utility value interventions (e.g., Hulleman et al., 2010; Harackiewicz et al., 2012) that have been shown to increase value, affective engagement, and learning outcomes, and IR interventions (e.g., Jamieson et al., 2010; Leroy, Grégoire, Magen, Gross, & Mikolajczak, 2012) that have been shown to increase positive affect and learning outcomes. Specifically, rather than using a traditional value intervention and instructing participants to consider how the task might be personally valuable, we prompted participants to perform a reappraisal task that involved engaging in an imaginary scenario. In previously published work on reappraisal, participants were to take on the perspective of an outside observer, or in some cases, a medical professional (for instance, while watching a gory video of a burn debridement, see Goldin et al., 2008). In the current research, participants took on the perspective of a person looking for a job at a law firm. We hypothesized that providing a context in which the task is personally valuable
(i.e., helpful in gaining a new job), would increase affective engagement and improve learning outcomes.

Research has demonstrated that reappraisals that are made early in the affect-generative process (i.e., before an affective episode has occurred) can shape the trajectory of the affective experience (Gross, 2008; Gross & Thompson, 2007). Our prediction was that prompting the use of an IR strategy at the beginning of the learning episode would be associated with maintained affective engagement (measured dimensionally by positive valence and moderate arousal, and as a discrete state by a discrete self-report measure) throughout the learning episode. Affective engagement has been associated with more task persistence and increased achievement (Csikszentmihalyi, 1997; Linnenbrink, 2007; Linnenbrink & Pintrich, 2002; Linnenbrink et al., 1999; Pekrun et al., 2010; Shernoff, Csikszentmihalyi, Schneider, & Scernoff, 2003). As such, our prediction was that the use of the IR strategy would be associated with increased affective engagement and positive learning outcomes. This hypothesis was called the facilitative reappraisal hypothesis.

On the other hand, it is possible that the use of an IR strategy could lead to negative affect due to the increased cognitive resources that are needed in order to engage in an IR strategy while learning. This hypothesis was called the harmful reappraisal hypothesis. Previous research in learning and other contexts has shown that reappraisal does not bear significant cognitive costs and does not interfere with performance (Leroy et al., 2012; Richards & Gross, 2000). Therefore, we had little reason to suspect that the harmful reappraisal hypothesis would be affirmed.

Experiment 2 built upon the first experiment by exploring two additional questions. First, there was a question of whether learners could generate and use their own reappraisals, and whether these self-generated reappraisals could also increase affective engagement and learning.
Research indicates that learners are generally unskilled at effectively regulating their cognitive, metacognitive, and affective processes in the absence of any kind of guidance or scaffolding (Azevedo, 2009; O’pt Eynde et al., 2007), which suggests that self-generated reappraisals might be less effective than IRs. However, it is possible that self-generated reappraisals could be more relevant and personally meaningful, which might lead them to be more effective. To test these possibilities, Experiment 2 included an open-ended reappraisal condition in which participants were asked to generate and use their own reappraisal strategy during learning. An open question was whether open-ended reappraisal strategies would yield the same hypothesized benefits of IRs.

Second, we explored the question of how suppression, which is another affect regulation strategy, impacts affect and learning outcomes. Previous empirical research (not in learning contexts) has demonstrated that suppression is associated with increased physiological arousal and decreased memory for details (Campbell-Sills, Barlow, Brown, & Hoffmann; 2006; Richards & Gross, 2000). To investigate the effect of suppression during learning, Experiment 2 also included a suppression condition in which participants were asked to avoid behaviorally expressing any affective states they experienced while they learned. Our goal was to determine whether suppression would be associated with lower affective engagement and decrements in learning, when compared to the IR condition.

Lastly, in both experiments we conducted a mediation analysis to explore the causal relationships among IR, engagement, and learning. It was hypothesized that the intervention would be associated with increased affective engagement (measured by valence, arousal, and discrete affect). Engagement, in turn, was predicted to lead to improved learning outcomes.
Experiment 1

Method

Participants. Ninety-three participants ($N = 93$) volunteered for monetary compensation ($5.00$) on Amazon Mechanical Turk™ (AMT; http://www.mturk.com). AMT is a data collection tool that allows participants to receive monetary rewards for completing Human Intelligence Tasks (HITs) online, and has been demonstrated to produce reliable data in other studies, including studies in which data-intensive work was required (see Snow, O’Connor, Jurafsky, & Ng, 2008). Recent research has also confirmed that AMT samples are considerably more diverse than the typical undergraduate student population in the U.S. (Buhrmester, Kwang, & Gosling, 2011; Mason & Suri, 2012; Paolacci, Chandler, & Ipeirotis, 2010), which is well suited for the present focus on lifelong learning. The participants’ mean age was 30.5 years ($SD = 9.1$), and 53.8% were female. Participants were asked to report the highest degree they had completed in their education. Five percent had a post-secondary certificate, 18.3% had a post-secondary diploma, 19.4% had an Associate’s degree, 29.0% had a Bachelor’s degree, 12.9% had a Master’s degree, and 1% had a Ph.D., M.D, or J.D. Fourteen percent of participants did not report their highest degree of education. All participants were recruited from the United States and reported English as their primary language.

Design and Manipulations. The experiment used a between-subjects design where participants were randomly assigned to one of three conditions: instructed reappraisal ($n = 33$), error search ($n = 28$), and control ($n = 32$). Participants in all three conditions were instructed that their primary goal was to learn all they could about the U.S. Constitution. Additionally, participants in the instructed reappraisal condition were informed that learning about the Constitution can be a boring task, but that they should attempt to do their best regardless of those
feelings (should they arise). Other conditions did not receive these instructions, which was a limitation that we addressed in Experiment 2. See Appendix A for complete experimental instructions for all conditions. The instructed reappraisal and error search conditions received additional instructions regarding a secondary task that they would engage in throughout the experiment, which is described below.

**Instructed Reappraisal Condition.** The instructions given to the instructed reappraisal (IR) condition were adapted from utility value interventions used by Helleman et al. (2010) and Jang (2008), and cognitive reappraisal interventions used by Gross and colleagues (see Goldin et al., 2008; Gross & Thompson, 2007). Similar to previously published IR interventions that prompted participants to engage in a form of imagination by assuming the perspective of another person (Goldin et al., 2008), participants in the IR condition were instructed to imagine that they were applying for a job as a copy-editor in a successful law firm in their city. While they read the U.S. Constitution and Bill of Rights, participants were told to imagine that their task was to prove their skills as copy-editors by trying to “catch” all typographical and grammatical errors within the two documents. Participants were instructed to imagine that if they could prove that they were good copy-editors, they would secure the job. As we discussed in the Introduction, interventions that help participants view the task as personally valuable or relevant are associated with increased engagement and positive learning outcomes. Our rationale was that the reappraisal component of the IR strategy (imagining that they were applying for a job) would provide a context for helping learners to view the task as personally valuable (because it might lead to the attainment of a job). This, in turn, would help learners experience affective engagement, ostensibly leading to positive learning outcomes.
The rationale for assigning the task of searching for typographical and grammatical errors stemmed from the subtle changes that have occurred in written American English over the last 200 years, shifting from British spelling and grammatical rules to Americanized spelling and grammar. Over time, several words in the English language have come to be spelled differently than they were spelled during the era in which the Constitution was written. Today's modern-day American reader might come across such anomalies within the Constitution and perceive these words to be misspelled. Some examples of “misspelled” words in the Constitution are “chuse” (choose), “defence”, “controul”, and “labour”. It was important for participants to be able to locate some of these errors during the learning task so that they did not perceive that the imagined task was useless and ultimately give up.

Importantly, the primary purpose of the manipulation was not to identify whether participants could catch errors in the Constitution. Indeed, we did not measure whether participants actually found errors while they learned. The purpose was to foster a sense of value by prompting them to imagine that reading the Constitution carefully and searching for errors could help them attain an imagined job. It is important to restate here that participants in the IR condition were instructed that their primary task was to learn all they could about the Constitution. Using the IR strategy while they learned was emphasized to participants as the secondary task.

Given that the current manipulation was adapted from several utility value manipulations that were designed to increase participants’ appraisals of value in the task, (Hulleman et al., 2007; Leroy et al., 2012), these appraisals were not of primary interest to the current research. Therefore, participants’ appraisals of value were not measured in the current experiments. Our primary focus was whether the IR strategy helped learners achieve and maintain affective
engagement by providing a context in which the learning activity could be of personal value. As such, we measured learners’ self-reported affective engagement along the dimensions of valence and arousal on the affect grid (described below), and as a discrete affective state (among other affective states) through forced-choice affect judgments throughout the experiment, rather than perceptions of value.

*Error Search Condition.* There was a concern that the imagined IR strategy used by the IR condition might cause participants to read the Constitution and Bill of Rights more carefully and perhaps process the information more deeply. Decades of research have suggested that information that is attended to, elaborated upon, and processed deeply is more likely to be successfully encoded and recalled than information that receives little attention, elaboration, and processing (Craik & Lockhart, 1972; Craik & Tulving, 1975; Otten, Henson, & Rugg, 2001; Paller, Kutas, & Mayes, 1987; Schulman, 1971; Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004; Watkins, 1983). The imagined error search task may also have the unintended effect of causing learners to adopt a performance goal of looking for errors, which might encourage learners to attend more carefully to the text and process the information more deeply. These behaviors might facilitate encoding and recall. Alternatively, it is also possible that the IR strategy might cause participants to process the text at a more shallow level. That is, participants might scan the texts in search of errors, rather than reading the text deeply to extract the more important information. If either of these hypotheses is correct, then any differences in affective engagement and learning outcomes between participants who used IR and those who did not might be due to the incidental depth of processing (or lack thereof) that accompanies error searching, rather than the use of reappraisal.
To alleviate this concern, it was necessary to include a condition that received the same error search manipulation as the IR condition, but that was not instructed to use an IR strategy. Hence, participants in the Error Search (ES) condition were instructed to attempt to catch all typographical and grammatical errors in the document. However, unlike the IR condition, the ES condition did not receive IR instructions. The important distinction between these two conditions is that, although participants in both conditions were given the additional imagined task of “error searching”, only participants in the IR condition were given the IR strategy of imagining they were applying for a job. By making this distinction, we presumed that any differences found between conditions could be primarily attributed to the IR strategy, rather than the task of error searching. Importantly, we did not measure how many errors participants “caught”, and they were not asked to keep track of the errors. Participants were instructed only to be on the lookout for as many errors as they could find.

**Control Condition.** The control condition received no instructed reappraisal strategy. The control condition’s only task, then, was to learn as much as possible about the Constitution.

**Learning and Testing Interface.** All materials were presented in AMT™ using Adobe Flash Player®. The U.S. Constitution and Bill of Rights were presented in pages across 18 trials, with each page containing an average of 508 words ($SD = 132$). Participants were allowed to read each page at their own pace. The interface tracked the amount of time spent (in milliseconds) on each content page as a measure of reading time. The first 10 pages of content covered the U.S. Constitution, and the final 8 pages covered the Bill of Rights.

**Affect Measures.** We examined affect in two ways. First, we measured participants’ dimensional affect (valence and arousal) using the Affect Grid (Russell, Weiss, & Mendelsohn, 1989). The Affect Grid is a single item affect measurement instrument consisting of a $9 \times 9$
(valence × arousal) grid, and is a validated measure of affect with adequate reliability (Cronbach’s $\alpha = .90$), convergent validity (correlations of .90 or higher with similar scales of affect), and discriminant validity (correlations of .20 or less with dissimilar scales of affect). The arousal dimension ranges from low arousal (1) to high arousal (9), while the valence dimension ranges from unpleasant feelings (1) to pleasant feelings (9). The Affect Grid was presented on the computer screen and participants responded by using the mouse to select the box anywhere on the grid that best represented their current affective state. In the present two experiments, the Affect Grid achieved excellent reliability across all trials (Experiment 1, $\alpha = .96$; Experiment 2, $\alpha = .94$).

This measure is useful for gauging affect because it allows participants to report their affective states by placing an X in a box in one of four quadrants (pleasant-activating, pleasant-deactivating, unpleasant-activating, unpleasant-deactivating). The pleasant-activating quadrant is composed of positive valence and high arousal, and includes affective states such as joy and engagement. Based on our conceptualization, engagement should be associated with mild positive valence and moderate arousal on the Affect Grid. The pleasant-deactivating quadrant is composed of positive valence and low arousal and represents affective states such as relaxation and calmness. The unpleasant-activating quadrant is composed of negative valence and high arousal, and includes affective states such as frustration and confusion. Finally, the unpleasant-deactivating quadrant is composed of negative valence and low arousal, and includes affective states such as boredom. The benefit of using the Affect Grid is that participants can report their affective state without the presence of affect labels, which can sometimes be confusing or interpreted differently across individuals.
Participants were also asked to report their discrete affective states by selecting one of six affective terms: happy, bored, frustrated, confused, engaged, and neutral (no affect). These are the dominant affective states that learners experience in similar tasks (see D’Mello, 2013 for a meta-analysis on affect during learning with technology). Though our primary focus was on affective engagement, it was also of interest to determine whether the IR intervention was associated with changes in other positive (e.g., happiness) and negative (e.g., frustration, confusion, boredom) affective states, when compared to controls. Due to the fact that multiple affective states fall within each of the four quadrants on the Affect Grid (for example, confusion and frustration) it was important that we asked participants to report the specific affective state they were experiencing so that we could obtain a more clear picture of their affective experiences.

**Comprehension and Learning Measures.** An 18-item embedded test (Cronbach’s α = .67) was used to assess participants’ trial-by-trial comprehension of the content, and an 18-item posttest (Cronbach’s α = .60) was presented at the end of the learning session to assess learning.

The embedded measure consisted of nine text-based questions (i.e., *What percentage of Congress must be in agreement for a new amendment to be made to the Constitution?*) and nine inference questions (i.e., *In order for a Bill to become law, it must go through several checks and balances between the President and Congress. Why do you think this important?*). These questions were presented immediately after each content page. Participants responded to each question by clicking one of four multiple choice options. These four options consisted of the target (the correct response to the question) and three foils: a near-miss (an option that sounded correct but was not), a thematic miss (an option that followed the theme of the content but was not actually related to the question), and a miss (an option that was not at all related).
The posttest measure was presented immediately after the learning session. Similar to the embedded measure, there were nine text-based and nine inference questions. Participants responded to these questions by selecting from one of four multiple choice options.

**Manipulation Check.** A manipulation check was developed to ensure that participants used the prescribed IR strategy in the IR condition. Following the posttest, all participants were instructed to answer the following question, “*In general, how did you regulate the emotions that arose while you learned?*” This question was answered by selecting from one of three options: (1) *I tried to think about my emotions in a different way so that they did not affect me as much* (reappraisal), (2) *I tried to not pay any attention to or express the emotions I experienced* (suppression), or (3) *I did nothing to regulate my emotions.* If participants in the IR condition followed the instructions, then they would be expected to select Option 1. The ES and control conditions were expected to select Option 2 or 3.

**Procedure.** Participants began the experiment by reading standard instructions about how to use the Affect Grid (see Russell et al., 1980). Participants were then given the primary instructions (all conditions) and secondary instructions (IR and ES only) before proceeding to the learning task.

In each of the 18 trials, participants read a segment of the U.S. Constitution or Bill of Rights. Next, participants were asked to report their current dimensional affect by completing one Affect Grid. Participants were then presented with a question about what they had just read. Answering the question was the final step in most trials, and after answering the question, participants navigated to the next content page to begin a new trial. The progress through most trials was Read Text → Complete Affect Grid → Answer Question → Move to Next Trial.
However, after every third trial, participants were asked to report their current discrete affective state by selecting one affect term from the following list: engaged, bored, frustrated, confused, happy, or neutral. Thus, the progress for every third trial was: Read Text → Complete Affect Grid → Answer Question → Discrete Affect Rating → Move to Next Trial. Upon completion of all 18 trials, participants were given as much time as needed to complete the posttest, followed by a brief demographics survey. See Figure 1 for a flow chart of the experimental procedure.

**Results**

The analyses examined the effect of condition on participants’ self-reported affective states and performance on embedded and posttest questions. Affective engagement was measured by self-reports of valence, arousal, and our discrete affect measure. Participants’ accuracy on the embedded and posttest questions were used as a measure of their performance.

For all variables, data points that exceeded 3 standard deviations above or below the mean were considered outliers and were not included in the analyses. Statistical testing included between-subjects MANOVAs and ANOVAs, and an alpha level of 0.05 was adopted for significance testing. We were specifically interested in comparisons between the IR condition and the other conditions (ES and control). As such, we focused on two comparisons: the IR condition to the ES condition, and the IR condition to the control condition. We made the prediction that the IR condition would report significantly higher affective engagement, and would achieve higher learning outcomes, compared to the other two conditions. We report effect sizes between all significant comparisons. Descriptive statistics for all variables across condition can be found in Table 1.
Manipulation Check. A chi-square analysis revealed that the IR, ES, and control conditions differed significantly in their use of the three affect regulation strategies (reappraisal, suppression, nothing), $\chi^2 (4, N = 80) = 18.8, p < .001, \phi = .486$. In the IR condition, 59.3% used reappraisal, 22.2% used suppression, and 18.5% used no strategy. In the ES condition, 30.4% used reappraisal, 30.4% used suppression, and 39.1% used no strategy. Finally, in the control condition, 10.0% used reappraisal, 23.3% used suppression, and 66.7% used no strategy.

Dimensional Affect. We first measured engagement by calculating participants’ dimensional affect. This was accomplished by aggregating arousal and valence scores on the Affect Grid across all 18 trials. As previously stated, engagement is hypothesized to be a state of involvement in an activity with positive valence and moderate arousal (Baker et al., 2010).

A one-way multivariate analysis of variance (MANOVA) on the joint effect of condition on arousal and valence was significant, Wilks’ $\lambda = .873, F (4, 178) = 3.14$, partial $\eta^2 = .066$. Therefore, we conducted two univariate ANOVAs to explore the effect of condition on arousal and valence separately.

There was a significant main effect of condition on arousal, $F (2, 90) = 3.36$, $MSE = 2.40$, partial $\eta^2 = .069$. The IR condition reported significantly higher arousal than the control condition ($d = .66$), but not the ES condition ($d = .26$). The ANOVA for valence was also significant, $F (2, 90) = 4.14$, $MSE = 2.13$, partial $\eta^2 = .084$. Results indicated that the IR condition reported significantly higher valence than the control ($d = .69$) and ES ($d = .61$) conditions.

Taken together, there are three important findings. First, participants who used an IR strategy (IR condition) or merely engaged in an error search task (ES condition) experience equal levels of arousal. Second, participants who used IR or who engaged in the error search task experienced more arousal than learners who did nothing. Lastly, participants who used IR experienced more
positive valence than learners who searched for errors or did nothing. Therefore, our findings suggest that participants in the IR condition exhibited significantly more affective engagement than the control condition, because they experienced more positive valence and more arousal. Participants in the IR condition reported more positive valence than participants in the ES condition, but did not report higher arousal. However, participants in the IR condition may have experienced more affective engagement than the ES condition because they reported positive valence and moderate arousal (which is in the positive-activating quadrant of the affect grid and is indicative of engagement), while the ES condition reported negative valence and moderate arousal (which is in the negative-activating quadrant of the affect grid and is indicative of negative affective states like frustration).

**Discrete Affective States.** Proportional occurrence of each discrete affective state was calculated for each participant. A MANOVA to examine the joint effect of condition on the self-reported affective states of *engagement, boredom, frustration, confusion, happiness,* and *neutral* yielded a significant overall model, Wilks’ $\lambda = .710$, $F(12, 152) = 2.19$, partial $\eta^2 = .158$.

Subsequent ANOVAs revealed a significant main effect of condition for engagement, $F(2, 81) = 3.89$, $MSE = .078$, partial $\eta^2 = .088$. Comparisons indicated that the IR condition reported significantly more engagement than the control condition ($d = .63$) but not the ES condition.

There was also a significant main effect of condition on confusion, $F(2, 81) = 7.20$, $MSE = .063$, partial $\eta^2 = .151$, in that the IR condition reported significantly less confusion than the ES ($d = .86$), and control ($d = .69$) conditions. A similar pattern was found for frustration, $F(2, 81) = 3.15$, $MSE = .107$, partial $\eta^2 = .072$, with the IR condition reporting significantly less frustration.
than the ES \( (d = .54) \) and control \( (d = .58) \) conditions. There were no significant differences in boredom, happiness, and neutral across conditions.

It is possible that the measure of discrete affect was not sensitive enough to detect significant differences between conditions for happiness and boredom. Despite the lack of statistical significance, the trend for happiness and boredom was in the predicted direction. Specifically, participants in the IR condition reported more happiness than the ES and control conditions, as well as less boredom. The ES and control conditions reported equivalent levels of happiness and boredom (see Table 1). Given the trend between the IR condition and the ES and control conditions, it is possible that a more sensitive measure of affect might have detected significant differences for happiness and boredom.

Taken together with the results from the affect grid, these results indicate that participants in the IR condition reported significantly more affective engagement than the control condition, as well as less confusion and frustration. The results also indicated that the use of IR has benefits over an error search task (that did not use a utility value reappraisal intervention) in terms of increasing positive valence and decreasing negative affective states.

**Reading Time.** It was important to confirm that there were no differences in reading time since spending additional time attending to the material could have influenced learning outcomes, thus making it difficult to interpret the effect of condition on learning. We computed the mean reading time per page (in seconds) for each participant. A one-way ANOVA indicated that there were no statistically significant differences in reading time across conditions, \( F (2, 87) = .220, \ MSE = 1527.7, p > .05 \), partial \( \eta^2 = .005 \).

**Learning.** Scores on embedded and posttest measures were calculated for each participant by determining the proportion of correct responses out of the total number of
Running Head: Affect regulation during learning

questions. Participants’ accuracy on text-based questions were highly correlated with accuracy on inference questions ($r = .87$). Thus, we analyzed all 18 items together rather than analyzing text-based and inference items separately.

We conducted a 2 × 3 repeated measures ANOVA using test as a two-level within subjects factor (embedded test and posttest) and condition as a three level between subjects factor (instructed reappraisal, error search, and control). There was a significant main effect of test, $F (1, 70) = 17.87$, $MSE = .006$, partial $n^2 = .203$. Overall, participants performed significantly better on the embedded test ($M = 0.75$, $SD = 0.12$) than the posttest ($M = 0.69$, $SD = 0.10$). We also found a significant main effect for condition, $F (2, 70) = 3.77$, $MSE = .016$, partial $n^2 = .097$, with the IR condition ($M = 0.76$, $SD = 0.11$) significantly outperforming the ES condition ($M = .69$, $SD = 0.10$), but not the control condition ($M = .74$, $SD = 0.09$). The finding that participants tended to achieve better scores on the embedded test than the posttest indicates that although participants were able to retain the information long enough to answer a question immediately after each trial, they were less able to retain what they had learned at the end of the experiment.

There was also a significant test × condition interaction, $F (2, 70) = 3.60$, $MSE = 0.22$, partial $n^2 = 0.93$ (see Figure 2). On the embedded test, participants in the IR condition performed significantly higher than both the ES ($d = .74$) and control ($d = .48$) conditions. A slightly different pattern emerged for the posttest. Here, participants in the IR condition performed significantly higher than the ES condition ($d = .63$), but not the control condition.

The Mediating Role of Affect in Experiment 1
Findings from Experiment 1 revealed that *condition* had a significant impact on affective engagement (between the IR and all other conditions), positive valence and negative affective states such as confusion and frustration (between the IR and ES condition) and learning outcomes (between the IR and all other conditions on the embedded test, and between the IR and ES conditions on the posttest). This result aligns well with the control-value theory and theories of affect regulation, which propose the facilitative effect of IR on affect and learning. Together these theories suggest a plausible causal relationship between: (a) IR strategies that contain a utility value and reappraisal component, (b) affective engagement and other affective experiences, and (c) learning outcomes, such that the impact of IR on learning is mediated by learners’ affective states. To explore this hypothesis, we conducted a mediation analysis to test the mediation of affective engagement (measured by valence and arousal on the affect grid and self-reported engagement on the discrete affect measure) between condition (IR, ES, and control) and learning outcomes (embedded scores and posttest).

We followed the mediation procedure recommended by Preacher and Haynes (2004). For mediation to occur, the IV (condition) should significantly predict the DVs (embedded and posttest scores in separate analyses) (Condition 1), and the mediator (affective engagement) (Condition 2). When the DV is regressed on both the IV and the mediator, the mediator should be a significant predictor (Condition 3a) but the IV should not (Condition 3b) for full mediation. If Condition 3b is not satisfied (meaning that the IV is a significant predictor), but the relationship between IV and DV is reduced by a non-trivial amount due to the mediator, then there is evidence of partial mediation. For each of the following analyses, we adopted a one-tailed alpha for significance testing of $p < .10$, because we are predicting the direction of the relationship (all positive).
Embedded scores. We began by testing the relationship between condition, valence, and embedded scores (Condition → Valence → Embedded). The IR condition was the reference group, and was coded as 0. The remaining groups (error search and control) were coded as 1. There was a significant effect of condition on the embedded test, $B= 0.06, p < .05$ (Condition 1). Condition also significantly predicted valence, $B = -.04, p < .05$ (Condition 2). However, Condition 3a was not met, which suggests that there was no evidence of mediation. Identical patterns were found when we assessed arousal (Condition 3a was unmet) and engagement on the discrete self-report measure (Condition 3b was unmet) as mediators. Thus, there appears to be no evidence of mediation on participants’ embedded scores.

Posttest Scores. Condition 1 was violated in each of the three models (Condition → Valence → Posttest; Condition → Arousal → Posttest; Condition → Engagement on the discrete self-report measure → Posttest). Therefore, the mediation analysis was not pursued further. Taken together, these findings suggest that there was no mediation between condition and learning in Experiment 1.

Discussion

Experiment 1 provided initial insights into the effect of IR on participants’ affective engagement and learning outcomes. Compared to the control condition, learners in the IR condition experienced higher affective engagement on the affect grid and the discrete self-report measure, less confusion and frustration, and significantly higher learning outcomes on the embedded questions, but not the posttest.

The IR and ES conditions did not differ in affective engagement measured by the discrete affect self-reports, but the IR condition reported significantly more positive valence, less
confusion and less frustration. In addition, findings from the affect grid revealed that learners in the IR condition reported affect that fell in the positive-activating quadrant which is indicative of affective engagement. In contrast, the ES condition reported affect that fell in the negative-activating quadrant, which is indicative of confusion or frustration. The IR condition also significantly outperformed the ES condition on both embedded questions and the posttest. This suggests the improved performance of the IR condition over the control condition was attributable to the use of the IR strategy, and not the task of error searching. Taken together, results from Experiment 1 offer tentative support for the facilitative reappraisal hypothesis.

There were five limitations in Experiment 1 that we addressed in a second experiment. First, in Experiment 1 we did not administer a pretest at the beginning of the learning session to gauge participants’ prior knowledge. Without a measure of participants’ prior knowledge, we were unable to assess the extent to which individual participants gained knowledge from pretest to posttest. In addition, it is plausible that the inability to detect mediation effects in Experiment 1 might have been partially due to the fact that we did not account for participants’ prior knowledge. Tests of mediation may have been more fruitful if participants’ learning gains were measured, and this was addressed in Experiment 2 with an identical pretest and posttest.

Second, there was a concern that the questions on the knowledge tests might not be sufficiently sensitive to detect small differences. There might have also been a ceiling effect since several questions were answered correctly by 75% or more of all participants, which indicates that these questions were not challenging enough. To address this problem, we replaced the easier embedded and posttest questions with more challenging questions. This was accomplished by conducting three small pilot studies to test new questions that were related to the same content as each discarded question. Each of the new questions that were selected for
Experiment 2 were answered correctly by less than 50% of pilot study participants (who had not yet read the U.S. Constitution).

Third, participants’ discrete affective states (happy, bored, frustrated, confused, engaged, and neutral) were only sampled after every third trial. It would have been beneficial to obtain these reports after every trial to get a more fine-grained assessment of how these states unfolded throughout the learning session. Thus, in Experiment 2 we presented the discrete affect self-report measure in every trial (immediately following the Affect Grid), rather than in every third trial.

Fourth, Experiment 1 was potentially confounded by mentioning boredom to the IR condition and not the ES and control conditions. Priming participants in the IR condition to the notion that the task might be boring could have altered their affective responses and behavior during learning. This concern was addressed in the second experiment by ensuring that instructions pertaining to boredom were identical across conditions (described in the next section below).

Lastly, Experiment 1 was limited by its use of only one type of performance assessment (multiple choice embedded questions and posttest). In Experiment 2, we included a 10-item essay posttest that measured participants’ ability to articulate what they learned about the Constitution. Because adding this measure would significantly increase the amount of time needed to complete the experiment, we decided to remove the Bill of Rights from Experiment 2 and ask participants to learn only about the Constitution. This reduced the number of trials in Experiment 2 from 18 to 10.

As mentioned in the Introduction, two additional affect regulation strategies were implemented in Experiment 2. These included an open-ended reappraisal (OE) condition and a
suppression condition, which are described in detail in the next section. Furthermore, the ES condition was not included because Experiment 1 appears to have ruled out concerns that the effects between the instructed reappraisal condition and other conditions was due to the secondary error searching task (see Introduction).

Experiment 2

Method

Participants. One hundred and thirty eight participants ($N = 138$) volunteered for monetary compensation ($4.00) on AMT. The participants’ mean age was 30.4 years ($SD = 10.64$) and 59.5% were female. Eight percent had a post-secondary certificate, 15.2% had a post-secondary diploma, 23.2% had an Associate’s degree, 34.8% had a Bachelor’s degree, 2.2% had a Master’s degree, and less than 1% had a Ph.D., M.D, or J.D. Sixteen percent of participants did not report their highest degree of education.

Design. Participants were randomly assigned to one of four conditions: $IR\ (n = 33)$, $OE\ (n = 35)$, $suppression\ (n = 36)$, and $control\ (n = 34)$.

Participants in all conditions were instructed that the material might be somewhat boring. Participants in the IR, OE, and suppression conditions were given instructions for regulating their boredom while they learned. Participants in the control condition were instructed to try their best even when they experienced boredom. The exact wording of instructions delivered in each condition are in Appendix B.

The IR condition was instructed to use the same IR strategy used in Experiment 1. The OE condition received the same instructions as the IR condition about the definition and use of reappraisal. However, rather than being given instructions to use an explicit reappraisal strategy,
participants in the OE condition were instructed to generate and use their own reappraisal strategy. Participants in the suppression condition were given instructions about how to avoid expressing their affective states throughout the learning task. Instructions for this task were adapted from Gross & Levenson’s (1997) suppression instructions. The control condition was given the same instructions as in Experiment 1, with the exception that they also were instructed that the task might be boring.

Learning Session Interface, Manipulation Check, and Affect Measures. The interface used in Experiment 2 was identical to Experiment 1 with the exception of the changes mentioned above. The same manipulation check and affect measures used in Experiment 1 were used in Experiment 2.

Learning Measures. The learning measures used in Experiment 2 were similar to Experiment 1 but with two important differences. First, Experiment 2 included an identical pretest and posttest so that learning gains (changes from pre to post) could be assessed. Both measures consisted of 10 questions (5 text based, 5 inference) about the U.S. Constitution. Second, we included a 10-item short essay posttest to assess participants’ ability to recall what they learned during the session. All ten questions were inference questions. Participants responded to these questions by typing their response in a textbox presented below each question and were given as much time as needed to compose their answers.

Procedure. The procedure for Experiment 2 was similar to Experiment 1 with the exception of the key differences that were described above.

Scoring of essay posttest. Three ideal answer units were identified for each of the 10 questions on the essay posttest. These answer units represented three key concepts that were necessary in order to adequately answer the question. Participants received one point for each
answer unit they provided in an answer; thus, each participant could receive a score of 0-3 for each question. Participants’ total score was derived by summing scores on each of the 10 questions, yielding a possible total score of 30 points.

Responses were coded independently by two trained coders. In order to obtain inter-rater reliability, the two coders initially coded a subset of 20 responses and achieved a Pearson correlation of .75. All disagreements were resolved through discussions between the two coders. Once disagreements were settled, the coders each coded another 20 essays and obtained a Pearson correlation of .89. Once this acceptable inter-rater reliability score was achieved, the remaining essays were randomly divided between the two coders and were coded independently.

Results

Our analyses examined the effect of condition on affective engagement (measured by valence, arousal, and discrete affect self-reports), other affective experiences, and performance on the learning measures. Similar to Experiment 1, we focused on three comparisons in which we compared the IR condition to every other condition. We predicted that the IR condition would report significantly more affective engagement than the other conditions, and would also achieve significantly higher learning outcomes. Descriptive statistics for each measure across condition can be found in Table 2.

Manipulation Check. A chi-square test revealed that the IR, OE, suppression, and control conditions differed significantly in their use of the three affect regulation strategies, $\chi^2 (6, N = 116) = 21.8, p < .01, \phi = .434$. In the IR condition, 66.7% used reappraisal, 18.5% used suppression, and 14.8% used no strategy. In the OE condition, 35.7% used reappraisal, 28.6% used suppression, and 35.7% used no strategy. In the suppression condition, 17.9% used...
reappraisal, 46.4% used suppression, and 35.7% used no strategy. Finally, in the control condition, 22.6% used reappraisal, 25.8% used suppression, and 51.6% used no strategy.

**Dimensional Affect.** Similar to Experiment 1, we first measured affective engagement by measuring participants’ self-reports of valence and arousal on the affect grid. A one-way MANOVA to explore the joint effect of condition on arousal and valence was significant, Wilks’ $\lambda = .883$, $F(6, 246) = 2.64$, partial $\eta^2 = .061$. Subsequent ANOVAs examining arousal and valence separately revealed that there was no significant main effect of condition on arousal, $F(3, 124) = 2.40$, $MSE = 2.06$, $p = .07$, partial $\eta^2 = .055$. This finding is likely due to the fact that there was very low arousal at the start of the session. To account for this, we removed the first block (trials 1 and 2) from all conditions and repeated the ANOVA. By doing so, we found that the IR condition ($M = 5.69$, $SD = 1.33$) reported significantly more arousal than the OE ($M = 4.97$, $SD = 1.61$), suppression ($M = 4.89$, $SD = 1.45$), and control ($M = 4.65$, $SD = 1.50$) conditions, $F(3, 128) = 3.01$, $MSE = 6.54$, partial $\eta^2 = .066$. This finding suggests that IR did lead to increased arousal, and that this effect was somewhat slower to emerge in Experiment 2.

There was a significant main effect of condition on valence, $F(3, 124) = 3.56$, $MSE = 1.64$, partial $\eta^2 = .079$, in that the IR condition reported significantly higher valence than the suppression condition ($d = .39$) and control condition ($d = .81$), but not the OE condition.

Taken together with arousal, there are several important implications from these results. First, patterns of affective engagement on the affect grid between the IR and control conditions in Experiment 1 replicated in Experiment 2, which offers further validation for our findings from Experiment 1. Second, patterns between the IR and OE conditions indicate that although arousal differed significantly between conditions, valence did not. Thus, we cannot conclude that the use of IR was associated with more affective engagement on the affect grid than the use of a self-
generated strategy. However, the findings do suggest that our IR strategy and self-generated reappraisals are equally useful for regulating valence, but that our IR strategy is perhaps more useful than self-generated reappraisals for regulating arousal. Lastly, the IR condition reported more affective engagement on the affect grid than the suppression condition.

**Discrete Affective States.** A MANOVA to examine the effect of condition on the combined proportional reports of *engagement, boredom, frustration, confusion, happiness,* and *neutral* was significant, Wilks’ $\lambda = .717, F(15, 232) = 1.98$, partial $\eta^2 = .105$. Subsequent ANOVAs revealed a significant main effect of condition for engagement, $F(3, 88) = 4.19$, $MSE = .078$, partial $\eta^2 = .125$. We found that the IR condition reported significantly more engagement than the OE ($d = .50$), suppression ($d = .65$), and control ($d = .40$) conditions. There was also a main effect of condition for neutral, $F(3, 88) = 2.90$, $MSE = .072$, partial $\eta^2 = .090$. The IR condition reported significantly less neutral affect than the suppression ($d = .73$) and control ($d = .74$) conditions, but not the OE condition. There were no significant differences in self-reported boredom, frustration, confusion, or happiness. Therefore, the pattern for engagement found in Experiment 1 replicated in Experiment 2, but the pattern for confusion and frustration did not.

By considering affective engagement on the affect grid and the discrete self-report measure of engagement together, an interesting picture emerges. First, participants in the IR condition reported significantly higher affective engagement on the discrete self-report measure than participants in all other conditions. Second, the use of IR enhanced affective engagement on the affect grid and on the discrete self-report measure to a greater extent than the use of no strategy or the use of suppression. Perhaps more importantly, the use of IR was associated with more arousal and more engagement on the discrete affect self-report measure than the use of a self-generated reappraisal strategy (in the OE condition). These findings suggest that IR
strategies that incorporate features of utility value and reappraisal interventions may be a potential tool for enhancing affective engagement during learning.

**Reading-Time.** Similar to Experiment 1, there were no condition differences for reading time, $F(3, 131) = .811, \text{MSE} = 1.73, p > .05$, partial $\eta^2 = .038$.

**Learning.** We first conducted an ANOVA using pretest scores as the dependent variable and condition as the independent variable to determine if there were significant differences in prior knowledge across conditions. The model was not significant, $F(3, 134) = 1.00, \text{MSE} = .022, p > .05$, partial $\eta^2 = .023$.

We conducted a $2 \times 4$ repeated measures ANCOVA using test as a two-level within subjects factor (embedded test and posttest), condition as a four level between subjects factor (IR, OE, suppression, and control), and pretest scores as a covariate. There was a significant main effect of test, $F(1, 113) = 31.88, \text{MSE} = .027$, partial $\eta^2 = .220$. Participants performed significantly better on the embedded test questions ($M = 0.64, SD = 0.17$) than the posttest ($M = 0.43, SD = 0.19$), which is similar to our finding in Experiment 1. There was also a significant main effect of condition, $F(3, 113) = 2.78, \text{MSE} = .044$, partial $\eta^2 = .069$. Participants in the IR condition ($M = 0.60, SD = 0.16$) achieved significantly higher scores than participants in the OE ($M = 0.53, SD = 0.15$), suppression ($M = 0.50, SD = 0.15$), and control conditions ($M = 0.51, SD = 0.14$). Unlike Experiment 1, there was no significant test $\times$ condition interaction.

Proportional learning gains were computed as: (posttest – pretest) / (1 – pretest). This measure allowed us to examine the extent to which each participant learned from pretest to posttest. An ANOVA yielded a significant main effect of condition on proportional learning gains, $F(3, 102) = 2.73, \text{MSE} = .180$, partial $\eta^2 = .074$. The IR condition ($M = .30, SD = .17$)
achieved significantly higher scores than the OE ($M = .14, SD = .22$), suppression ($M = .15, SD = .29$), and control conditions ($M = .12, SD = .24$).

The analysis of learning gains provides two key insights. First, we found that although there were no significant differences between the IR and control condition in Experiment 1, the IR condition achieved significantly higher learning gains than the control condition in Experiment 2. Second, we found that the IR condition also achieved significantly higher learning than all other experimental conditions (the OE and suppression conditions) in Experiment 2. This finding suggests that IR is an effective means for improving learning outcomes, even when compared to suppression strategies and self-generated reappraisals.

**Essay Posttest.** The ANOVA on essay posttest scores was not significant, $F (3, 108) = .809, MSE = .021, p = .491, \eta^2 = .017$. There were, however, small- to medium-sized effects when the IR condition was compared to both the OE ($d = .33$) and control ($d = .32$) conditions. Differences between IR and the suppression conditions were minimal ($d = .06$. This non-significant finding suggests that although the use of IR led to enhanced performance on multiple choice questions, the facilitating effect of IR did not transfer to more the more challenging essay questions.

**The Mediating Role of Affect in Experiment 2**

Similar to Experiment 1, we conducted a series of mediation analyses to explore the mediating effect of affective engagement (measured by valence, arousal, and engagement on the discrete self-report measure) between condition (IR, OE, Suppression, Control) and learning (Embedded Scores, Gain Scores). Similar to Experiment 1, the IR condition was the reference group and was coded as 0. All other groups were coded as 1. Following the procedure from
Experiment 1, a one-tailed alpha level of $p < .10$ was used for significance testing since all predicted effects are in the positive direction. Essay posttest was not considered because the ANOVAs did not yield any condition differences on this measure.

**Embedded scores.** First, we tested the relationship between condition, valence, and embedded scores (Condition $\rightarrow$ Valence $\rightarrow$ Embedded). The effect of condition on embedded scores ($b = 0.12, p < .05$) and valence ($b = 1.05 p < .05$) was significant, thereby satisfying Conditions 1 and 2. The effect of valence on embedded scores (when controlling for condition) was significant, $b = 0.03, p < .10$ (Condition 3a). However, condition was still a significant predictor of embedded scores ($b = 0.08, p = .06$) after controlling for valence (so Condition 3b was not satisfied). Therefore, there was evidence of partial mediation of condition on learning through valence.

We found no evidence of mediation for arousal (Condition $\rightarrow$ Arousal $\rightarrow$ Embedded), or engagement on the discrete self-report measure (Condition $\rightarrow$ Engagement $\rightarrow$ Embedded). As such, our findings suggest that affective engagement did not mediate the relationship between condition and learning on the embedded questions. However, valence was found to partially mediate this relationship. This finding indicates that there was some evidence of mediation between condition and embedded scores, but only for valence.

**Learning Gains.** Unlike our findings for embedded scores, we found that valence did not mediate the relationship between condition and learning gains. Valence appears to be a partial mediator of condition on learning during the learning activity, but that effect may not be sufficiently strong to influence overall learning gains.

We next tested for mediation effects of arousal (Condition $\rightarrow$ Arousal $\rightarrow$ Learning Gains). The effect of condition on learning gains, $B = 0.18, p = .01$, and arousal, $B = 0.59, p <$
.10, was significant (satisfying Conditions 1 and 2). The effect of arousal on learning gains (when controlling for condition) was significant, $B = 0.06, p < .05$ (Condition 3a). When controlling for arousal, condition, was still a significant predictor of learning gains, $B = 0.15, p < .05$ (Condition 3b), revealing that there was evidence of partial mediation between condition and learning through arousal.

Finally, we tested for mediation effects of engagement on the discrete self-report measure (Condition $\rightarrow$ Engagement $\rightarrow$ Learning Gains). The effect of condition on learning gains, $B = 0.19, p < .01$, and engagement, $B = 0.15, p < .10$, was significant (satisfying Conditions 1 and 2). The effect of engagement on learning gains (when controlling for condition) was significant, $B = 0.34, p < .01$ (Condition 3a), and condition, when controlling for engagement, was still a significant predictor of learning gains, $B = 0.14, p < .05$ (Condition 3b). Hence, although there was no evidence of mediation between condition and learning for valence, there was evidence of partial mediation for arousal and discrete self-reports of engagement.

**Discussion**

In addition to replicating several of the findings from Experiment 1, Experiment 2 provides further insights into the role of IR during learning. When compared to the control condition, we found that the use of the IR strategy was associated with more affective engagement on the affect grid than the use of suppression or no strategy (when accounting for similar levels of arousal between conditions at the beginning of the learning session) and increased engagement on the discrete self-report measure. Indeed, the use of IR was associated with increased engagement on the discrete self-report measure compared to all conditions,
including the OE condition. Unlike Experiment 1, we did not find significant differences in confusion or frustration between conditions.

Similar to Experiment 1, we found that participants in the IR condition achieved significantly higher scores on embedded questions than the control condition. In addition, in Experiment 2 the IR condition achieved significantly higher scores on the posttest and achieved significantly higher proportional learning gains than the control condition. Furthermore, the IR condition achieved higher performance on the embedded questions and posttest than the open-ended and suppression conditions. However, there was no significant effect of IR when participants were asked to write essays in response to questions.

At first blush, this finding suggests that although the use of IR seems to facilitate performance on questions that are accompanied by multiple choice options, the intervention may not be adequately robust to improve open-ended essay responses. There was, however, a small non-significant effect ($d = .21$) showing that the IR condition performed marginally better on the essay posttest than the control condition. Our sample did not have the requisite statistical power to detect this small effect, so replication with a larger sample is warranted.

It should also be noted that scores on the essay posttest were generally low, with a grand mean of .303. This general low performance might be attributed to: (a) fatigue effects as the essay posttest was the last measurement item, (b) the questions themselves being too difficult, or (c) a combination of both (a) and (b). The low performance suggests that this test might not be sufficiently sensitive to detect subtle effects between treatment and control. Lastly, results from Experiment 2 reveal the possible mediating role of self-reported engagement on the discrete self-report measure, but not affective engagement measured by valence and arousal on the affect grid. Although we found no evidence of mediation in Experiment 1, we found in Experiment 2 that
valence (but not arousal) partially mediated the relationship between condition and embedded scores, and arousal (but not valence) and engagement on the discrete self-report measure partially mediated the relationship between condition and learning gains.

Taken together, findings from Experiment 2 support the control-value theory and theories of affect regulation. Specifically, findings indicate that engagement measured by the discrete self-report measure does partially mediate the effect of IR strategies on learning outcomes, and that the effects of mediation may be more detectable when prior knowledge is taken into account. As a result of these findings, we suggest that the unfolding of IR’s effect on learning outcomes occurred in the following ways: (1) the IR strategy used features of utility value and IR interventions to help participants reappraise the learning task, (2) these reappraisals led to increased engagement, and (3) engagement mediated the effect of IR on learning. We ground this conclusion in the control-value theory, which posits strategies that utility value interventions increase positive affect and engagement, and from theories of affect regulation which posit that giving participants a way to reappraise the learning task can help them experience more positive affect and less negative affect.

**Synthesis of Findings from Experiments 1 and 2**

In order to synthesize the complex pattern of results in both experiments, effect sizes were computed for all key variables. With the exception of valence and arousal (which can be classified as affective engagement on the affect grid, but that need to be evaluated separately), we classified each variable into the larger category of variables it represented. Thus, discrete self-reports of engagement and happiness were grouped in the positive discrete affective states category; boredom, frustration, and confusion were assigned to the negative discrete affective
states category. For Experiment 1, the embedded and posttest measures were classified as learning measures. The learning measures category in Experiment 2 included the embedded measure, posttest proportional learning gains, and the essay posttest. Effect sizes for each category in both experiments were then obtained by computing the average effect size of all variables within that category. Finally, since the IR and control conditions were included in both experiments, the category-level effect sizes were averaged across Experiment 1 and 2, resulting in one set of effect sizes for the instructed reappraisal vs. control condition.

We compared the IR and control condition to address the facilitated reappraisal hypothesis. We used Cohen’s (1992) criteria of d’s of 0.2, 0.5, and 0.8 representing small, medium, and large effects, to interpret these effect sizes. A number of patterns emerge from the effect sizes presented in Table 3. There were medium to large effects for valence and arousal (revealing more affective engagement as reported on the affect grid in the IR condition than the control condition), medium effects for positive discrete affective states, small effects for negative discrete affective states, and a small to medium effect for learning. Taken together, our findings across two experiments suggest that IR is associated with more positive affect and enhanced learning, which provides some support for the facilitated reappraisal hypothesis. Although IR was beneficial for improving positive affect, it was less beneficial for reducing negative affect.

There was a concern that the potential enhancing effects of IR could be due to a different form of processing associated with the task of error searching. To address that concern, we compared effect sizes for affect and learning between the IR and ES conditions. When compared to the ES condition, there were medium effects in favor of the IR condition for valence and arousal, which suggests that the IR condition showed slightly more affective engagement on the affect grid than the ES condition. There were also medium effects in favor of the IR condition for
positive and negative discrete affective states, and a medium to large effect for learning. Overall, participants who used IR experienced more positive affect, less negative affect, and achieved better learning outcomes than learners who engaged only in the error search task. These findings suggest that the enhancing effects of IR are attributable to the reappraisal itself, rather than to incidental processing associated with the error searching task.

We explored the question of whether self-generated reappraisals (used by the OE condition) could afford the same enhancing effects as IRs. We compared effect sizes between the IR and OE conditions and found small effects in favor of the IR condition for valence and arousal and positive discrete affective states. There was no effect for negative discrete affective states, but there was a medium effect in the direction of the IR condition for learning outcomes. The pattern of findings indicates that although the difference between instructed and self-generated reappraisals on learners’ affective states may be less robust, there is a clear advantage of using IR for improving learning outcomes. Therefore, theoretically principled IR strategies that combine features of utility value and reappraisal interventions may be more effective than self-generated reappraisals.

Lastly, there was the question of the efficacy of IR compared to suppression, which is another common affect regulation strategy. Although there was no effect for negative discrete affective states, there were medium effects in favor of the IR condition for valence and arousal, indicating that the IR condition experienced somewhat more engagement on the affect grid than the suppression condition. There were also medium effects in favor of the IR condition for positive discrete affective states and learning. This finding supports previous research in non-learning contexts that demonstrated the negative effects of suppression on affective responding.
and memory (Campbell-Sills et al., 2006; Gross, 2002; Richards & Gross, 2000). As such, suppression appears to be a poor strategy for helping learners regulate their affect and learning.

**General Discussion**

In two experiments, we examined the effectiveness of IR for increasing affective engagement and enhancing learning. In this section, we consider implications of our findings in the contexts of lifelong learning and other important contexts, and present limitations and future directions. The subsequent discussion focuses on the broader patterns in the data rather than on specific details because these have been addressed in the Discussion sections of the individual experiments.

**Applied Implications**

Instructed reappraisal is applicable to educational technologies like multimedia, hypermedia, intelligent tutoring systems, and serious games. Our research has demonstrated that IR can be implemented and used in a web-based learning environment in an automated way, in the absence of a human trainer or animated pedagogical agent. Thus, IR interventions can be implemented at a very low cost and can be made available to large populations of learners of all ages. Importantly, the implementation of these interventions into publicly available resources like web-based learning environments caters to the needs of traditional learners, as well as lifelong learners who have moved away from traditional learning environments. These learners, who are more likely to use informal environments that lack the structure of classroom settings, could benefit from the prompting of IR strategies to help them regulate their affective experiences while they learn.
Intelligent tutoring systems (ITSs) are becoming increasingly successful at sensing learners’ affect through facial features, posture patterns, keyboard/mouse behaviors, speech contours, and textual and contextual cues (for a recent review, see Calvo & D’Mello, 2010). As these systems become more efficient at detecting learners’ affect, it will become increasingly important for them to help learners manage their affective states as they arise. In addition to the proactive strategy of IR emphasized in this research, IR might also be applicable as a reactive intervention to help learners reappraise negative affect as it arises. For instance, ITSs could be designed to engage in adaptive automation in order to help learners regulate their affective experiences more effectively. Systems that are able to detect when learners have transitioned to a negative affect state (such as boredom or frustration) could prompt learners to engage in IR strategies that will help them manage these feelings and become affectively engaged. Automatic prompting of IR strategies can help reduce the task demand and cognitive load required for learners to regulate their own affective experiences (in the absence of any kind of prompting), without imposing additional cognitive demands (Leroy et al., 2012). Furthermore, it might not always be necessary to prompt the use of IR. Indeed, it may be the case that ITSs should merely help learners be more vigilant of their affective experiences so that they can deploy IR strategies on their own. In such adaptive systems, the interface and the learner could work collaboratively to help achieve successful affect regulation and learning.

Advances in technology and the web have opened doors to information that in previous decades was only accessible in classrooms, libraries, or multi-volume encyclopedias. Currently, learners of all ages have access to web-based learning encyclopedias, multimedia learning environments, educational games, and intelligent tutoring systems, as well as traditional text books, in order to gain the knowledge and skills. Having access to such a wealth of informational
resources has the obvious benefit of allowing individuals to take more control over their learning. However, having access to virtually limitless information can be accompanied by a host of affective experiences that can impair learning. For instance, learners might feel frustration while they use the web to learn a new skill, confusion while they research a topic they are interested in, or boredom while they attempt to educate themselves on something that is not personally relevant to them (for example, a parent researching an uninteresting topic to help his or her child complete a school project).

Students are not the only individuals who could benefit from using IR strategies. Adult workers often have to participate in continuing education for professional and vocational development, which can be a challenging and potentially overwhelming process. We posit that IR could be a useful strategy for helping even adult learners regulate their affective experiences across a wide range of such learning situations. Aside from learning, the use of IR has applications in other real world contexts where boredom is likely to arise, although it is likely that different reappraisal strategies would be needed in different contexts.

**Limitations and Future Directions**

The first limitation was that the learning task might not have been sufficiently challenging. Challenging topics such as math or physics, and more abstract topics such as philosophy and scientific reasoning, require more cognitive resources (Paas & van Merriënboer, 1994; Sweller, van Merriënboer, & Paas, 1998). Therefore, there is a question of whether the beneficial effects of IR would persist during learning of these types of challenging topics. It is also possible that IR could still be beneficial in these challenging domains if it alleviates cognitive load induced by negative affective states like frustration and anxiety. Additional research investigating IR with more challenging topics is needed to answer this question.
Second, an unforeseen limitation of this research is that the instructions given to participants in the IR condition may have caused them to adopt a performance goal, because the IR instructions mentioned a computerized test of their copy-editing skills. If that is the case, then perhaps our findings are attributable to the possibility that participants in the IR condition were working toward achieving a goal that participants in the other conditions did not work toward. Although we did not collect any data on participants’ goals during learning, we contend that the instructions did not influence goals or serve as the driving force of participants’ affective responses and learning outcomes. Given that (a) the instructions about the fictional computerized test pertained to the error searching task, not the primary task of learning about the Constitution, and (b) participants in all conditions were given the primary goal of learning as much as possible about the U.S. Constitution, we feel that we can reasonably argue that the instructions about the fictional computerized test did not lead to increased goal orientation in the IR condition.

Third, the assumption in both experiments was that participants’ self-reported affective experiences were indicative of their feelings toward the learning task as a result of their use (or lack of use) of affect regulation strategies. However, it is possible that participants’ self-reports of affect might have been a reflection of their perceived success or failure, their relief of completing the learning task in the previous trial, or their hopefulness that the next trial would be easier or more interesting. In essence, participants’ self-reported affective experiences might have been the result of some unmeasured affective or cognitive experience, which makes the findings somewhat difficult to interpret. On the other hand, if affect was only attributable to hopefulness or success, the instructed reappraisal condition would not have been expected to differ significantly from the other conditions because the other conditions would have also experienced hopefulness and success since the learning task was identical across conditions. The
significant differences in affective responding between the IR condition and the other conditions in both experiments suggests that these differences were likely due to the use of IR strategies, rather than an unmeasured experience like hopefulness or success.

An assumption of this research was that the use of IR would provide a context in which the task might be more personally valuable, which in turn would be associated with more engagement and better learning outcomes. However, we did not measure the extent to which the IR strategy fostered reappraisals of value because our goal was to develop and test an IR strategy that fostered affective engagement by combining features of utility value interventions and reappraisal strategies. As such, we tracked participants’ affective engagement on the affect grid and on the discrete self-report measure (along with other affective states) throughout the session, rather than their appraisals of value. None the less, the absence of this data undermines our ability to determine whether there were associative links among IR, appraisals of value, affective engagement, and learning. Future work should include a more rigorous manipulation check to measure participants’ appraisals of value, interest, motivation, and learning goals prior to and immediately following the learning activity.

Lastly, the absence of the ES condition in Experiment 2 is a potential limitation. The ES condition in Experiment 1 was used to ensure that any positive outcomes in the IR condition were attributable to the use of reappraisal and not to the task of error searching. Having fulfilled its purpose, we felt that it was defensible to eliminate the ES condition in Experiment 2 and focus on testing new conditions instead. Nevertheless, a replication of Experiment 2 with the ES condition may be warranted.

**Concluding Remarks**
In summary, findings from the current experiments offer initial evidence for the usefulness of IR strategies to increase affective engagement and to improve learning outcomes. The findings also support Pekrun’s (2006) control-value theory and theories of affect regulation, which emphasize the importance of appraisals of utility value and reappraisals of affect, respectively. This research serves as an initial step toward understanding the applied relationships among reappraisal, affective engagement, and learning. It is clear that there is still much work to be done in order to determine the most appropriate types of IR across various types of domains. As future research begins to explore innovative ways of instructing learners to use IR during classroom and computerized learning, improved interventions can be put into place that help learners be more capable of regulating their affective states and increasing learning outcomes.
References


Table 1

*Means and standard deviations (in parentheses) by condition for Experiment 1*

<table>
<thead>
<tr>
<th>Measure</th>
<th>IR</th>
<th>ES</th>
<th>Control</th>
<th>Pattern of Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affect Grid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arousal</td>
<td>5.51 (1.22)</td>
<td>5.12 (1.66)</td>
<td>4.52 (1.75)</td>
<td>IR &gt; C</td>
</tr>
<tr>
<td>Valence</td>
<td>5.68 (1.18)</td>
<td>4.78 (1.72)</td>
<td>4.76 (1.46)</td>
<td>IR &gt; C &amp; ES</td>
</tr>
<tr>
<td><strong>Discrete Affect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engagement</td>
<td>.309 (.356)</td>
<td>.201 (.266)</td>
<td>.135 (.160)</td>
<td>IR &gt; C</td>
</tr>
<tr>
<td>Confusion</td>
<td>.011 (.042)</td>
<td>.095 (.131)</td>
<td>.069 (.113)</td>
<td>IR &lt; C &amp; ES</td>
</tr>
<tr>
<td>Frustration</td>
<td>.058 (.117)</td>
<td>.154 (.221)</td>
<td>.152 (.196)</td>
<td>IR &lt; C &amp; ES</td>
</tr>
<tr>
<td>Happiness</td>
<td>.130 (.206)</td>
<td>.052 (.157)</td>
<td>.068 (.152)</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>.258 (.282)</td>
<td>.236 (.309)</td>
<td>.344 (.333)</td>
<td></td>
</tr>
<tr>
<td>Boredom</td>
<td>.162 (.233)</td>
<td>.179 (.248)</td>
<td>.205 (.199)</td>
<td></td>
</tr>
<tr>
<td><strong>Reading Time</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.06 (.617)</td>
<td>1.12 (.651)</td>
<td>1.07 (.686)</td>
<td></td>
</tr>
<tr>
<td><strong>Learning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embedded</td>
<td>.791 (.109)</td>
<td>.712 (.105)</td>
<td>.737 (.115)</td>
<td>IR &gt; C &amp; ES</td>
</tr>
<tr>
<td>Posttest</td>
<td>.715 (.105)</td>
<td>.648 (.109)</td>
<td>.732 (.089)</td>
<td>IR &gt; ES</td>
</tr>
</tbody>
</table>

*Note:* Range of possible valence and arousal scores on the Affect Grid is (1-9). Range of possible scores on the embedded and posttest measures is 0-1 (proportional scores). Only significant patterns are reported.
Table 2

*Means and standard deviations (in parentheses) by condition for Experiment 2*

<table>
<thead>
<tr>
<th>Measure</th>
<th>IR</th>
<th>O.E</th>
<th>Suppression</th>
<th>Control</th>
<th>Pattern of Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affect Grid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arousal</td>
<td>5.59 (.24)</td>
<td>5.05 (.58)</td>
<td>5.07 (.44)</td>
<td>4.63 (.45)</td>
<td>IR &gt; OE, S &amp; C*</td>
</tr>
<tr>
<td>Valence</td>
<td>5.92 (.14)</td>
<td>5.58 (.34)</td>
<td>5.33 (.11)</td>
<td>4.91 (.47)</td>
<td>IR &gt; S &amp; C</td>
</tr>
<tr>
<td><strong>Discrete Affect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engagement</td>
<td>.453 (.252)</td>
<td>.262 (.297)</td>
<td>.214 (.248)</td>
<td>.222 (.319)</td>
<td>IR &gt; OE, S &amp; C</td>
</tr>
<tr>
<td>Boredom</td>
<td>.121 (.179)</td>
<td>.210 (.191)</td>
<td>.152 (.150)</td>
<td>.231 (.221)</td>
<td></td>
</tr>
<tr>
<td>Confusion</td>
<td>.043 (.063)</td>
<td>.033 (.058)</td>
<td>.052 (.108)</td>
<td>.045 (.085)</td>
<td></td>
</tr>
<tr>
<td>Frustration</td>
<td>.110 (.137)</td>
<td>.081 (.136)</td>
<td>.104 (.124)</td>
<td>.027 (.088)</td>
<td></td>
</tr>
<tr>
<td>Happiness</td>
<td>.046 (.092)</td>
<td>.081 (.112)</td>
<td>.071 (.119)</td>
<td>.027 (.088)</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>.225 (.237)</td>
<td>.332 (.272)</td>
<td>.404 (.255)</td>
<td>.231 (.221)</td>
<td>IR &lt; S &amp; C</td>
</tr>
<tr>
<td><strong>Reading Time</strong></td>
<td>1.14 (.650)</td>
<td>1.15 (.571)</td>
<td>1.05 (.750)</td>
<td>1.31 (.677)</td>
<td></td>
</tr>
<tr>
<td><strong>Learning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>.263 (.139)</td>
<td>.318 (.148)</td>
<td>.282 (.151)</td>
<td>.315 (.144)</td>
<td></td>
</tr>
<tr>
<td>Embedded</td>
<td>.703 (.138)</td>
<td>.606 (.161)</td>
<td>.591 (.163)</td>
<td>.590 (.158)</td>
<td>IR &gt; OE, S &amp; C</td>
</tr>
<tr>
<td>Posttest</td>
<td>.500 (.190)</td>
<td>.425 (.171)</td>
<td>.428 (.231)</td>
<td>.436 (.164)</td>
<td>IR &gt; OE, S &amp; C</td>
</tr>
<tr>
<td>Gain Score</td>
<td>.301 (.275)</td>
<td>.139 (.216)</td>
<td>.152 (.292)</td>
<td>.118 (.240)</td>
<td>IR &gt; OE, S &amp; C</td>
</tr>
<tr>
<td>Essay Posttest</td>
<td>.330 (.170)</td>
<td>.282 (.124)</td>
<td>.320 (.163)</td>
<td>.282 (.125)</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Range of possible valence and arousal scores is (1-9). Range of possible scores on discrete affective states embedded and essay posttest measures is (0-1). Only significant patterns are reported.*

*Note: * indicates that this pattern only emerged when the first block of trials was removed.*
Table 3

*Effect sizes (Cohen’s $d$) for key variables*

<table>
<thead>
<tr>
<th>Measure</th>
<th>IR vs. C</th>
<th>IR vs. ES</th>
<th>IR vs. OE</th>
<th>IR vs. S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>.73</td>
<td>.70</td>
<td>.24</td>
<td>.55</td>
</tr>
<tr>
<td>Arousal</td>
<td>.63</td>
<td>.66</td>
<td>.25</td>
<td>.34</td>
</tr>
<tr>
<td>Pos. Discrete Affective States</td>
<td>.44</td>
<td>.38</td>
<td>.29</td>
<td>.48</td>
</tr>
<tr>
<td>Neg. Discrete Affective States</td>
<td>-.18</td>
<td>-.49</td>
<td>.12</td>
<td>.04</td>
</tr>
<tr>
<td>Reading Time</td>
<td>-.14</td>
<td>-.09</td>
<td>-.02</td>
<td>.13</td>
</tr>
<tr>
<td>Learning</td>
<td>.35</td>
<td>.68</td>
<td>.52</td>
<td>.44</td>
</tr>
</tbody>
</table>

*Note:* IR = Instructed Reappraisal, C = Control, ES = Error Search, OE = Open Ended, S = Suppression. Effect sizes larger than 0.2 sigma are in bold. According to Cohen (1992), effective sizes of 0.2, 0.5, and 0.8 denote small, medium, and large effects, respectively.
Figure Captions

Figure 1. Flow chart of participants’ progress through Experiment 1.

Figure 2. Interactions between test type and condition in Experiment 1.
Running Head: Affect regulation during learning

Figure 1.
Figure 2.
Appendix A

Experimental instructions for Experiment 1

**Instructed Reappraisal.** Your primary responsibility during this task is to learn all you can about the U.S. Constitution. However, sometimes this material can get really boring. When you’re feeling bored, using the strategy of reappraisal can help you manage that boredom so that you can focus, feel more interested, and achieve better understanding of what you’re reading. While you read the next several pages, try using the following reappraisal strategy: Imagine that you are at a job interview for a position as a copy-editor at one of the most prestigious law firms in your city. You have met Jacob, Lisa, and Allen who will be your supervisors if you get the job, and they all seem to like you. However, before you can be hired, the partners of the firm want to see how good your copy-editing skills are, since your potential new job will require you to be able to identify typographical and grammatical errors with precision. They have developed a computerized test that you have been asked to take. You will be reading the U.S. Constitution and need to show that you have what it takes to be a good copy editor by looking for errors in the document. Do a good job, and you will be moving into your new office in no time.

**Error Search.** Your primary responsibility during this task is to learn all you can about the U.S. Constitution. However, as you read, your secondary task is to identify typographical and grammatical errors with precision throughout the document. Please try to remember these instructions and use them throughout the experiment.

**Control.** Your primary responsibility is to learn all you can about the U.S. Constitution and Bill of Rights.
Appendix B
Experimental instructions for Experiment 2 by condition

Instructed Reappraisal. Your primary responsibility during this task is to learn all you can about the Constitution. However, sometimes this material can get really boring. Research over the last decade has demonstrated that reappraisal, or thinking about something in a different way, can help you alter or regulate your emotions. When you’re feeling bored, using the strategy of reappraisal can help you manage that boredom so that you can focus, feel more interested, and achieve better understanding of what you’re reading.

While you read the next several pages, try using the following reappraisal strategy: Imagine that you are at a job interview for a position as a copy-editor at one of the most prestigious law firms in your city. You have met Jacob, Lisa, and Allen who will be your supervisors if you get the job, and they all seem to like you. However, before you can be hired, the partners of the firm want to see how good your copy-editing skills are, since your potential new job will require you to be able to identify typographical and grammatical errors with precision. They have developed a computerized test that you have been asked to take. You will be reading the U.S. Constitution and need to show that you have what it takes to be a good copy editor by looking for errors in the document. Do a good job, and you will be moving into your new office in now time. Keeping this strategy in mind will help prevent you from experiencing boredom, and may actually help you learn the Constitution better than if you used no strategy at all.

Open Ended Reappraisal. Your primary responsibility during this task is to learn all you can about the Constitution. However, sometimes this material can get really boring. Research over the last decade has demonstrated that reappraisal, or thinking about something in a different way, can help you alter or regulate your emotions. When you are feeling bored, using the strategy of reappraisal can help you manage that boredom so that you can focus, feel more interested, and achieve better understanding of what you are reading.

While you read the next several pages, try using a reappraisal strategy of your own. There are many different ways to reappraise, or think about the situation differently, so you should choose a reappraisal strategy that works for you. It does not matter what you think about, it only matters that you think about the task in a different way. For example, you could imagine that you are learning about the Constitution for a class you are taking. Or, you could imagine that you are trying to get into law school and need to brush up on your legal history. Whatever method you choose, make sure it is relevant to you and helps you think about the task differently. Keeping this strategy in mind will help prevent you from experiencing boredom, and may actually help you learn the Constitution better than if you used no strategy at all.

Suppression. Your primary responsibility during this task is to learn all you can about the Constitution. However, sometimes this material can get really boring. Research over the last decade has demonstrated that emotional suppression, or avoiding expressing your emotions through your facial expressions or bodily movements, can help you alter or regulate your emotions. When you are feeling bored, using the strategy of suppression can help you manage that boredom so that you can focus, feel more interested, and achieve better understanding of what you are reading.

While you read the next several pages, try using the following suppression strategy: People often express boredom through their facial expressions and bodily movements. Boredom is often expressed by yawning, excessive movement, tapping (of the feet or of objects like pencils), and being fidgety. Throughout this experiment, try to suppress these behaviors while you learn. Specifically, please try to avoid: yawning, swiveling in your chair, leaning over in your chair, and propping your head up in your hands. Instead, you should focus on sitting forward and upright in your chair (keeping your back straight and avoiding slouching), remaining as still as you can, and avoiding expressing any kind of emotions on your face. To help you do this, imagine that someone is watching a video of you while you complete this experiment. A complete stranger watching your video should not be able to detect any boredom or any other emotion you experience throughout the experiment. Keeping this strategy in mind will help prevent you from experiencing boredom, and may actually help you learn the Constitution better than if you used no strategy at all.

Control. Your primary responsibility during this task is to learn all you can about the Constitution. However, sometimes this material can get really boring. In order to learn this material, it is important that you try your best to persevere even when you do feel bored. Doing this can help you focus, feel more interested, and achieve better understanding of what you are reading. Keeping this in mind may help you learn the Constitution better than if you did not try to persevere.