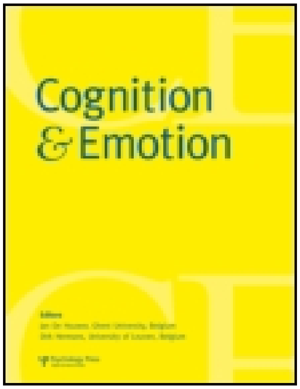

Cognition & Emotion
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pcem20

Test-retest reliability of an emotion maintenance task

Rose Broome [a], David E. Gard [a] & Joseph A. Mikels [b]

[a] Department of Psychology, San Francisco State University, San Francisco, CA, USA

[b] Psychology Department, DePaul University, Chicago, IL, USA

Version of record first published: 14 Nov 2011.

To cite this article: Rose Broome, David E. Gard & Joseph A. Mikels (2012): Test-retest reliability of an emotion maintenance task, Cognition & Emotion, 26:4, 737-747

To link to this article: http://dx.doi.org/10.1080/02699931.2011.613916

COGNITION AND EMOTION
2012, 26 (4), 737–747

Psychology Press
Taylor & Francis Group

BRIEF REPORT

Test–retest reliability of an emotion maintenance task

Rose Broome[1], David E. Gard[1], and Joseph A. Mikels[2]

[1]Department of Psychology, San Francisco State University, San Francisco, CA, USA
[2]Psychology Department, DePaul University, Chicago, IL, USA

Research on working memory has suggested domain-specific components for visual, verbal, and spatial information, and more recently for emotion. Affective working memory has been proposed as the set of processes involved in the maintenance of emotions to guide behaviour. The current study examined the reliability of an emotion maintenance/affective working memory task over two experimental sessions separated by one week. Subjective accuracy based on individual ratings was found to correlate over time and was highest for negatively valenced pictures. Results suggest that this paradigm is a reliable measure of emotion maintenance, underscoring the utility of this measure as an assessment tool for normative and clinical populations.

Keywords: Affect; Working memory; Emotion; Cognition; Emotion maintenance.

Working memory is the cognitive system that integrates memory, attention and perception (e.g., Baddeley, 2003). Defined as the short-term maintenance and manipulation of information used to guide goal-directed behaviour (Goldman-Rakic, 1987), working memory is involved in reasoning, problem solving, and comprehension (Just & Carpenter, 1992). Literature on working memory has focused primarily on the maintenance of visual, verbal, and spatial information, and various tests have been developed to measure each specific subcomponent (e.g., Baddeley, 1998; Engle, Tuholski, Laughlin, & Conway, 1999; Luck & Vogel, 1997).

Recent work suggests that working memory—commonly studied through the maintenance of visual, verbal, and spatial information—may also be involved in other experiences, including the process of holding emotions online (Davidson & Irwin, 1999; Mikels, Larkin, Reuter-Lorenz, & Carstensen, 2005; Mikels, Reuter-Lorenz, Beyer, & Fredrickson, 2008). As defined by Davidson and Irwin (1999), affective working memory is a means of representing an emotion in the absence of immediate elicitors to help guide action and

Correspondence should be addressed to: David E. Gard, Department of Psychology, San Francisco State University, San Francisco, CA 94132, USA. E-mail: dgard@sfsu.edu

Special thanks to Amy Harrison, Hana Kubkova, Jen Chou, Jon Meyer-Shen, Lee Kaplowitz, Ivonne Melgar, and Jorie Pollak for their work on this project.

http://www.psypress.com/cogemotion http://dx.doi.org/10.1080/02699931.2011.613916

organise behaviour around motivationally salient goals.

Multiple behavioural studies have included emotional stimuli *in* working memory tasks, often using methodologies based on emotionally valenced face pictures and delayed match-to-sample working memory tasks (e.g., Gooding & Tallent, 2003; Kensinger & Corkin, 2003; Luciana, Burgundy, Berman, & Hanson, 2001). In other words, these studies have investigated the effect of emotion on working memory, and have generally found effects of image valence and arousal on general working memory performance. Although these methodologies contribute to our understanding of how emotionally charged stimuli influence working memory processes, they do not distinguish between the maintenance of visual and verbal information versus the maintenance of actual subjective emotional information. Although research examining the influence of emotion on working memory is important, understanding the processes by which actual subjective emotions are maintained is an equally important endeavour and has provided insight into how emotion maintenance differs in different populations (see, e.g., Gard et al., 2011; Mikels et al., 2005).

To isolate working memory for subjectively experienced emotions from other types of working memory, a novel emotion maintenance task was developed (Mikels et al., 2008). With this task Mikels et al. (2008) conducted a series of experiments in which participants viewed emotion-eliciting pictures, maintained their subjectively experienced emotion, and then compared it to an emotion elicited from another picture. Critically, using interference methodologies they found a double dissociation: an emotional-regulation task interfered with emotion maintenance but not cognitive maintenance, while secondary cognitive tasks interfered with cognitive maintenance yet facilitated emotion maintenance. This double dissociation provides strong support for separate domain-specific components of working memory used for the maintenance of affective information. Despite these intriguing findings, the utility of this measure of working memory for emotion remains unknown, yet is vitally important when considering using this task in other contexts such as with clinical populations.

Expanding on the research conducted by Mikels et al. (2008), the current study assessed the reliability of this emotion maintenance task completed at two time points separated by one week. The goal of the current study was to establish the test–retest reliability of the emotion maintenance task, which would contribute to our basic understanding of the consistency of affective working memory as a construct, and would provide vital information on the utility of this task in longitudinal and treatment-outcome research.

In the present study participants completed a maintenance task modelled after Mikels et al. (2008). Specifically, they maintained their emotions elicited by a picture over a 3 s delay and then compared the intensity of their maintained subjectively experienced emotion to that elicited by second picture. In order to score performance on the maintenance task, a standard procedure used in prior studies of emotion maintenance (Gard et al., 2011; Mikels et al., 2005) was employed in the present investigation. After all comparisons were made participants gave in-the-moment ratings of the emotional intensity of the pictures using a visual analogue scale. Accuracy ratings for the comparison tasks were calculated for each participant individually using their own subjective ratings. Emotion maintenance accuracy was measured as the percentage correct between self-report rating of picture intensity and the picture comparisons. In an analogous visual working memory task intended to serve as a "non-emotional" control condition, the maintenance of brightness intensity of neutral pictures was assessed, as was the in-the-moment ratings of brightness of the pictures (also used to calculate accuracy). In order to assess test–retest reliability, participants returned after one week to complete the emotion maintenance and brightness maintenance tasks using novel pictures.

METHOD

Participants

Fifty-one participants (38 female, mean age 23.23, 33% White, 38% Asian American and Pacific Islander, 13% Latino/Hispanic, 7% African American, and 9% other) were recruited from the San Francisco State University student population and received course extra-credit for their participation. Study procedures were approved by the San Francisco State University Committee for the Protection of Human and Animal Subjects. All participants were given a description of the study and provided written informed consent.

Materials

A PC desktop computer with E-Prime software was used to administer the picture comparison and picture rating tasks. Picture pairs were selected as described by Mikels et al. (2008) and consisted of 42 negatively valenced picture pairs, 44 positively valenced picture pairs, and 40 neutral picture pairs, from the IAPS image database (Lang, Bradley, & Cuthbert, 2005).[1]

Design and procedure

All participants completed both the emotion maintenance task, with positive and negatively valenced pictures grouped together, and a brightness maintenance task, consisting of neutral pictures. One week later participants returned to complete both tasks again (with novel pictures). On this second visit participants also completed a separate rating task in which all pictures from sessions one and two were rated (described below). Order of picture group (either affect or brightness), and presentation of picture pairs within each task was randomised for each participant.[2]

Maintenance tasks

In the emotion maintenance task participants viewed an emotionally charged picture on the computer screen for 5 s, and were asked to experience the emotional intensity of the picture and hold its intensity level in mind over a 3 s delay (see Figure 1A). During the delay, a white fixation cross was displayed on the screen to keep participants' focus on the computer. After the delay a second similarly valenced emotionally charged picture was presented and participants were again asked to experience its emotional intensity. Following the second picture, a green cross was presented to prompt the participant to indicate whether the second picture was higher or lower in emotional intensity than the first picture. If the participant felt that the second picture was higher in emotional intensity he/she was instructed to indicate this by pressing a key labelled "H" on the keyboard. If the second picture was felt to be lower in emotional intensity, the participant was instructed to press a key labelled "L" on the keyboard.

The brightness maintenance task was identical to the emotion maintenance task, but served as a control condition measuring visual working memory. In this task only neutral pictures were used, and participants were instructed to maintain the level of brightness intensity over the delay period.

[1]Additional high-arousal picture pairs were added including, erotic, action adventure and threat pictures.

[2]This finding was in contrast to the Mikels et al. (2008) study, where performance on the brightness maintenance task was higher than the emotion maintenance task. One key difference is that in the current study accuracy scores were computed using participant-specific subjective ratings of the pictures, while in the Mikels et al. (2008) study concordance was used (i.e., "accuracy" was based on normed ratings of the pictures collected from a large sample of research participants), which may explain the difference.

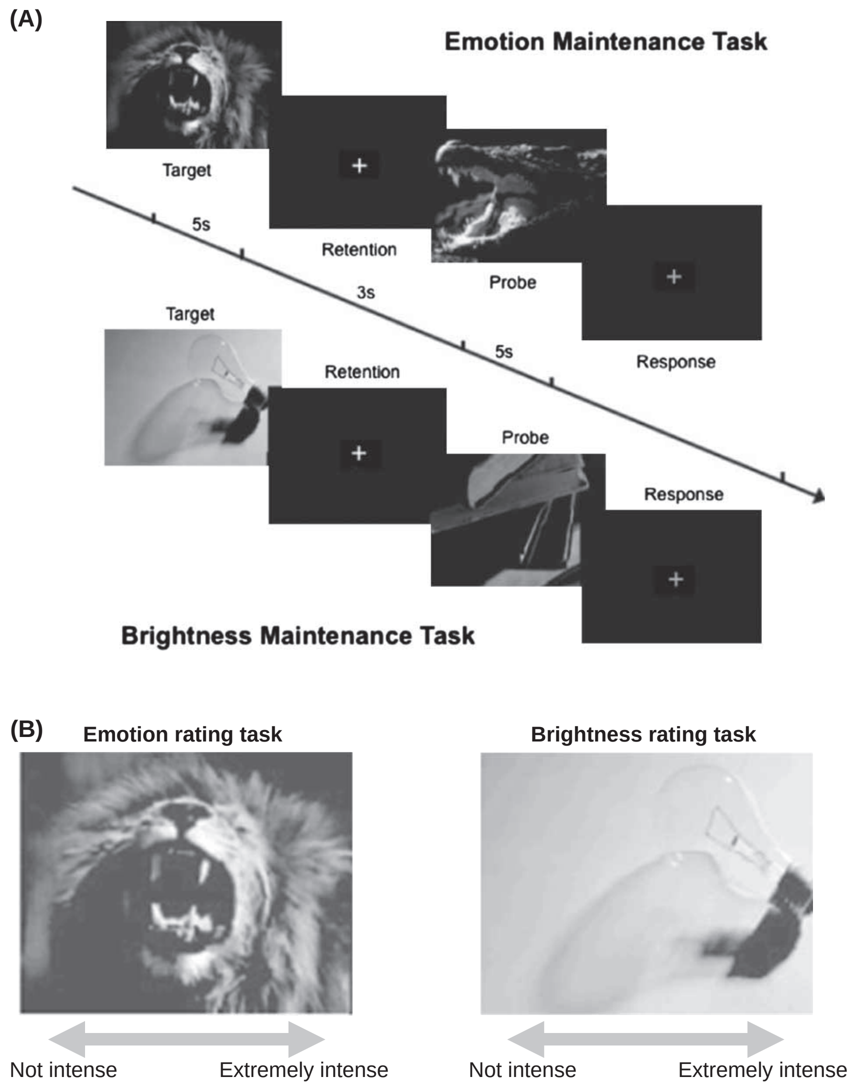

Figure 1. *(A) Emotion maintenance and brightness maintenance task. During the emotion maintenance task a positive or negative target picture was viewed for 5 s and emotional intensity was maintained over a 3 s delay. Then a separate probe picture of the same valence was viewed for 5 s. Following the pictures, participants indicated whether the probe was more or less intense than the target. Participants completed an analogous brightness maintenance task, which was identical except neutral pictures were viewed and brightness intensity was maintained. Figure originally printed in Mikels et al. (2008). Emotion and working memory. Emotion, 8(2), 256–266. Copyright © 2008 American Psychological Association. Reprinted with permission. (B) Emotion and brightness rating task. After both maintenance tasks were completed participants rated their in-the-moment experience of emotion intensity to emotion pictures, and their in-the-moment brightness experience to the brightness (neutral) pictures, all using a visual analogue scale. Using the computer mouse, participants clicked along the continuum for their experience rating.*

Rating task

Following the emotion maintenance task and brightness maintenance task in the second session, participants completed a rating task of all pictures viewed in both sessions. Pictures from the emotion maintenance task were rated on emotional intensity only, and pictures from the brightness maintenance task were rated on brightness intensity only. Pictures were viewed one at a time and participants indicated the intensity level experienced in reaction to each picture using a visual analogue scale. The scale was labelled "Very intense" at one end and "Not intense" at the other end (see Figure 1B). Participants indicated their response by clicking anywhere on the scale with the computer mouse.

Data analysis plan

Ratings for each picture were matched with comparison responses for the corresponding picture pairs, yielding a subjectively determined percent accuracy score based on each participant's subjective ratings of the pictures. In other words, participants' ratings of the pictures were used to score the picture comparison task trials. For example, if a participant indicated picture A was more intense than picture B in the comparison maintenance task, and also rated picture A as having a higher intensity than picture B during the individual picture rating task, then the comparison trial was scored as correct (reflected in a higher accuracy score). If, however, the participant indicated picture A was more intense than picture B in the comparison task, but instead rated picture *B* higher in intensity than picture A in the picture rating task, the comparison trial would be scored as incorrect, reflecting a discrepancy or *inaccuracy* in maintenance. Test–retest reliability was computed as a Pearson correlation between accuracy at session one and session two for both affective maintenance and brightness maintenance.

RESULTS

See Table 1 for accuracy scores broken down by session. Affective maintenance between session one and session two was correlated at $r(49) = .415$, $p = .002$, indicating reliability of the affective working memory task (see Figure 2). Broken down by domain, negative emotional accuracy had the highest reliability, $r(49) = .301$, $p = .032$, followed by positive emotional accuracy, $r(49) = .281$, $p = .046$. Brightness accuracy, which served as a visual working memory control condition, was not found to significantly correlate between session one and session two, $r(49) = .22$, $p = .13$. One possible reason for this lack of reliability may have been that the brightness comparisons were more difficult than the emotion maintenance comparisons (reflected in lower accuracy), and, indeed, percentage accuracy for the brightness maintenance task was significantly lower overall than for the emotion maintenance task, $t(50) = 2.55$, $p = .014$. To investigate whether difficulty of the task influenced the reliability of the brightness maintenance task, the brightness pictures were split based on Mikels et al.'s (2008) separation between easy and difficult comparisons (see appendix for brightness task photo details). For the easy comparisons, i.e., comparison pairs with the largest difference in brightness intensity between the pairs, there was a significant correlation in accuracy between sessions one and two, $r(49) = .317$, $p = .023$. See discussion below for more on this topic.

In addition to the reliability of the task over two sessions, accuracy scores were further analysed using a repeated-measures analysis of variance (ANOVA) with a Bonferroni correction. A main effect was found for task domain, $F(1,\ 50) = 4.97$, $p = .03$, indicating participants performed significantly *better* in the emotion maintenance task compared to the brightness maintenance task (this is in addition to the emotion maintenance task showing higher *reliability* than the brightness maintenance task between session one and session two). Within the

Table 1. *Accuracy means and standard deviations by task domain and session*[2]

	Session 1	*Session 2*	*Overall*
Domain	*M% (SD)*	*M% (SD)*	*M% (SD)*
Negative	74.08 (13.30)	79.67 (11.61)	77.13 (10.31)
Positive	71.48 (12.40)	74.30 (12.00)	72.49 (9.34)
Affect (overall)	72.64 (10.74)	76.92 (10.31)	74.73 (8.86)
Brightness	69.84 (16.04)	73.59 (11.02)	71.11 (11.39)

emotion maintenance task, participants performed significantly better on negatively valenced picture pairs than positively valenced picture pairs, $F(1, 50) = 10.38$, $p = .002$, suggesting a negativity bias. Lastly, a main effect was found for Session, $F(1, 50) = 7.86$, $p = .007$, with the second session having significantly higher accuracy.

"Concordance" scores were also computed using pilot data collected by Mikels et al. (2008). Specifically, emotion intensity ratings were used for the emotion condition images, and brightness intensity ratings were used for the brightness condition images. These concordance scores were used as a comparison-dependent measure in addition to the accuracy scores calculated in this study with the participant's specific ratings of each image. Accuracy scores were significantly higher than concordance scores for the emotion maintenance condition overall, $t(50) = 3.54$, $p = .001$, as well as broken down by negative affect, $t(50) = 2.74$, $p = .008$, and positive affect $t(50) = 3.11$, $p = .003$. There was no significant difference between accuracy and concordance scores for the brightness condition, $t(50) = 1.26$, $p = .213$.

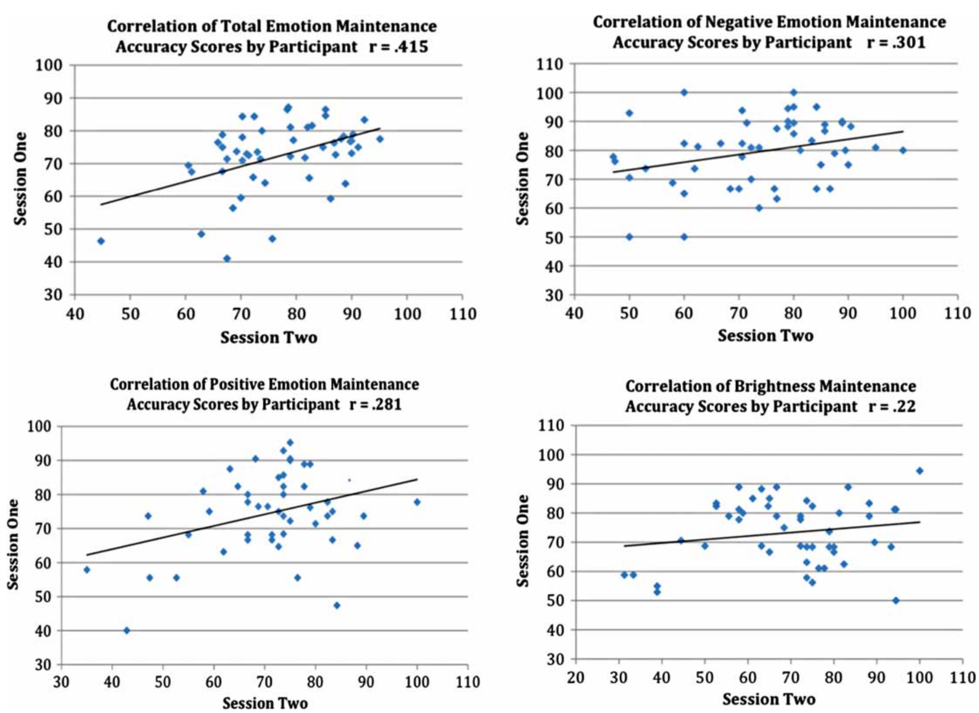

Figure 2. *Correlations in maintenance accuracy scores between session one and session two for each participant separated by condition: total emotion maintenance, negative emotion maintenance, positive emotion maintenance, and brightness maintenance.*

Table 2. *Brightness maintenance task IAPS picture numbers, intensity scores, and intensity difference by picture pair. Picture pairs with a high intensity difference are "easy" (i.e., it is easy to tell which picture is brightener), while those with a low intensity difference are "hard"*

Picture 1 IAPS number	*Intensity 1*	*Picture 2 IAPS number*	*Intensity 2*	*Intensity difference*
8311	5.25	5920	3.38	1.88
93079	5.15	2580	3.28	1.88
93025	5.15	2385	3.43	1.73
2850	5	5950	3.38	1.63
93049	5	9700	3.38	1.63
5731	5.05	2485	3.5	1.55
2220	4.88	5533	3.38	1.5
7150	4.43	2890	2.93	1.5
93012	4.68	7175	3.18	1.5
2514	4.95	1313	3.55	1.4
7402	4.83	2130	3.63	1.2
2620	4.8	5940	3.63	1.18
2020	4.85	2480	3.7	1.15
5534	4.88	7004	3.73	1.15
7830	4.5	9070	3.38	1.13
60323	4.78	7620	3.65	1.13
1670	4.73	5532	3.63	1.1
2840	4.78	5530	3.68	1.1
7140	4.85	2681	3.75	1.1
7705	4.7	7283	3.63	1.08
1560	4.6	7351	3.58	1.03
5000	4.78	5410	3.78	1
1121	4.6	2575	3.73	0.88
93073	4.98	7920	4.1	0.88
7035	4.55	2600	3.78	0.78
5520	4.53	7030	3.83	0.7
7490	4.85	2383	4.15	0.7
7002	4.58	7234	3.9	0.68
5120	4.43	7320	3.8	0.63
2702	4.5	5510	3.9	0.6
2516	4.58	7034	4.05	0.53
4610	4.68	2487	4.15	0.53
5779	4.58	60060	4.05	0.53
7170	4.55	5740	4.03	0.52
5130	4.45	7829	3.95	0.5
5220	4.48	7190	3.98	0.5
7160	4.55	7090	4.05	0.5
6150	4.2	2810	3.73	0.48
4571	4.13	7080	3.83	0.3
5800	4.38	7560	4.13	0.25

In order to assess reliability due to individual differences in maintenance, we completed a median split of the data into high and low performers (based on average accuracy scores in the affect task). Reliability between session one and session two emotion maintenance accuracy scores was significant for high performers, $r(25) = .45$, $p = .02$, but not significant for low performers, $r(25) = .19$, $p = .34$. For brightness accuracy scores, significant reliability was not found in either high performers, $r(25) = .25$, $p = .22$, or low performers, $r(25) = .12$, $p = .54$.

Finally, we investigated whether there was a relationship between emotion and brightness maintenance accuracy. Thus, correlations were computed for emotion and brightness accuracy for session one, session two, and for both sessions combined. For session one there was a trend towards a relationship between maintenance of emotion and brightness, $r(49) = .25$, $p = .07$. This trend became significant in session two, $r(49) = .40$, $p = .003$ (possibly related to the finding that participants did better in session two than in session one), and there was a significant correlation for both sessions combined, $r(49) = .50$, $p < .0001$.

DISCUSSION

Affective working memory has been defined as the processes involved in representing affect in the absence of immediate elicitors for the purpose of organising behaviour toward salient goals (e.g., Davidson & Irwin, 1999). In the present study, reliability of a delayed-response emotion maintenance task was assessed over two experimental sessions separated by one week. Emotion maintenance accuracy correlated significantly within subjects over both sessions supporting the reliability of this affective working memory task.

One point to consider is the strength of the reliability coefficient for the affective maintenance accuracy scores between session one and session two. Though this reflects a relatively modest correlation, it is in line with previous research suggesting reliability between two sessions for working memory tasks has generally not been found to be high by psychometric standards (Beckmann, Holling, & Kuhn, 2007; Conway et al., 1995). For example, test–retest reliability for reading span working memory tasks has been reported at .41, .47, and .52 in three separate studies (McDonald, Almor, Henderson, Kempler, & Andersen, 2001; Towse, Hitch, Hamilton, Peacock, & Hutton, 2005; Waters & Caplan, 1996). Additionally, an assessment of visuospatial working memory found test–retest reliability scores for *n*-back task accuracy ranging from .302 to .732 ($M = 0.51$, $SD = 0.11$; Hockey & Geffen, 2004). Overall, these findings indicate the test–retest reliability levels of the current study are consistent with those found in other working memory tasks.

Accuracy and reliability were highest for negatively valenced pictures compared to positively valenced pictures and also compared to the maintenance of visual brightness of neutral pictures. This higher performance (and reliability) with negative valenced stimuli in our young adult sample is consistent with the *negativity bias*, which is well documented in emotion literature and may stem from an evolutionarily adaptive response to danger stimuli (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Mikels et al., 2005).

Higher accuracy rates found in both the affective and brightness conditions in the second session may be a sign of practice effects. Practice effects are commonly found in repeated administrations of memory tests especially when the same version of the test is used during both administrations (Benedict & Zgaljardic, 1998). In the current study, however, different pictures were used for session one and session two in an attempt to minimise this effect. Nevertheless, reliability between the sessions was established and no ceiling effect was found, indicating the task was sufficiently difficult to measure the variation between sessions. One possible explanation for the practice effect is that participants were able to better remember the session two comparisons while completing the individual intensity ratings, which were all collected at the end of session two. For administrations where practice effects are a concern, a greater delay between session two comparisons and the rating task may be beneficial, perhaps having the ratings on a separate administration.

One notable limitation of this study is that reliability was not found in the brightness maintenance task overall, which was meant to be a cognitive (i.e., non-emotional) visual working memory control condition. However, after a median split was performed reliability was found on the easier comparisons, possibly indicating the brightness maintenance task was disproportionately difficult. Thus, for future studies it is important to consider the difficulty of the task in order to achieve reliable measurement of visual working memory.

Similarly, we found individual differences in the reliability of the emotion maintenance task, where high performers were also more reliable than low performers. This, too, may indicate the need for researchers to select image pairs with an appropriate degree of difficulty, and suggests that some individuals are more consistent in this ability. The implications of these individual differences are a possible direction for future studies.

Though we do not have information on the stability of subjective ratings for the images (since images were rated only once at the end of session two), we have additional evidence to support the notion that the variability is indeed due to a deficit in emotion maintenance and not due to variability in evaluation during the rating task. Specifically, these images received normed (in-the-moment) intensity scores collected from a pool of participants in a previously published study (Mikels et al., 2008). We compared the "concordance" of subject's comparisons in the maintenance tasks with these normed ratings to see if their accuracy improved, relative to their own ratings of the images. As expected, participants performed better with their own ratings of the images than when we used the normed pilot data. This suggests that the subjective ratings are stable to a degree, and indicate a meaningful rating by participants.

Another limitation of this study was the restricted age range of the participants, especially since age-related effects of emotion on memory have been reported (Mikels et al., 2005). An affective working memory study on a larger general population, as opposed to a student population, would give us a clearer picture of this construct. Additionally, psychophysiological measures such as affective startle modulation may add additional validity by providing objective measures to compliment participant picture comparisons and ratings (e.g., Bradley, Codispoti, Cuthbert, & Lang, 2001; Gard, Germans Gard, Mehta, Kring, & Patrick, 2007; Lang, 1995).

Finally, there was some evidence that there was a relationship between emotion maintenance and brightness maintenance, even though the specific tasks are quite different (e.g., maintaining an emotion versus maintaining a visual percept). Although this relationship could be argued to reflect an overall maintenance ability, this appears to be the case only in session two, and when sessions one and two were combined. Further, there are data to indicate that these maintenance processes are separable. First, Mikels et al. (2008) found a double dissociation with the emotion and brightness tasks, where including an additional emotion task in the delay disrupted emotion maintenance accuracy but not brightness (i.e., cognitive) maintenance accuracy, and an added cognitive task during the delay disrupted brightness accuracy, but not emotion maintenance. Additionally, these tasks have been very useful in differentiating cognitive versus emotion maintenance abilities in older adults and schizophrenia patients. For instance, although older adults exhibited a deficit in brightness maintenance relative to younger adults, the groups showed equivalent emotion maintenance performance (Mikels et al., 2005). Furthermore, in a recent study in schizophrenia using this task we found evidence that these processes may be separable. Specifically, in a sample of schizophrenia patients and healthy community adults, we found that emotion maintenance was disrupted in schizophrenia, even when we controlled for brightness maintenance accuracy (Gard et al., 2011). Additionally, these maintenance processes were unrelated in both our schizophrenia and healthy community samples.

There are diverse future directions for both applied and basic research on emotion maintenance. For example, in clinical work, in addition to the schizophrenia findings mentioned above, there is also evidence that emotion maintenance may be related to crucial predictors of depression, such as rumination, which is a repetitive focus on negative emotions (e.g., Nolen-Hoeksema, 2000). And in neuroscience work, research has noted that the orbitofrontal cortex is crucial in tasks that involve "working memory for value", which may aid in efficient decision-making (e.g., Wallis, 2007).

In conclusion, the findings of the current study provide evidence for the test–retest reliability of an emotion maintenance task, and are important in that they expand the existing literature on emotion maintenance and affective working memory, and indicate the stability of this construct. The reliability of this construct may also provide researchers with a clear target for remediation in some clinical populations.

Manuscript received 7 January 2011
Revised manuscript received 13 July 2011
Manuscript accepted 15 July 2011
First published online 11 November 2011

REFERENCES

Baddeley, A. D. (1998). Recent developments in working memory. *Current Opinion in Neurobiology*, *8*, 234–238.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370.

Beckmann, B., Holling, H., & Kuhn, J. (2007). Reliability of verbal-numerical working memory tasks. *Personality and Individual Differences*, *43*, 703–714.

Benedict, R. H. B., & Zgaljardic, D. J. (1998). Practice effects during repeated administrations of memory tests with and without alternative forms. *Journal of Clinical and Experimental Neuropsychology*, *20*(3), 339–352.

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

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786.

Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, *3*, 11–21.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–311.

Gard, D. E., Cooper, S., Fisher, M., Genevsky, A., Mikels, J. A., & Vinogradov, S. (2011). Evidence for an emotion maintenance deficit in schizophrenia. *Psychiatry Research*, *187*(1–2), 24–29.

Gard, D. E., Germans Gard, M. K., Mehta, N., Kring, A. M., & Patrick, C. P. (2007). Impact of motivational salience on affect modulated startle at early and late probe times. *International Journal of Psychophysiology*, *66*(3), 266–270.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology: The nervous system* (pp. 373–417). Bethesda, MD: American Physiological Society.

Gooding, D. C., & Tallent, K. A. (2003). Spatial, object, affective working memory in social anhedonia: An exploratory study. *Schizophrenia Research*, *63*, 247–260.

Hockey, A., & Geffen, G. (2004). The concurrent validity and test–retest reliability of a visuospatial working memory task. *Intelligence*, *32*(6), 591–605.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149.

Kensinger, E. A., & Corkin, S. (2003). Effect of negative emotional content on working memory and long-term memory. *Emotion*, *3*, 378–393.

Lang, P. J. (1995). The emotion probe. Studies of motivation and attention. *American Psychologist*, *50*(5), 372–385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual* (Technical Report A-6). Gainesville, FL: University of Florida.

Luciana, M., Burgundy, E. D., Berman, M., & Hanson, K. L. (2001). Effects of tryptophan loading on verbal, spatial and affective working memory functions in healthy adults. *Journal of Psychopharmacology*, *15*, 219–230.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D., & Andersen, E. S. (2001). Assessing working memory and language comprehension in Alzheimer's disease. *Brain & Language*, *78*, 17–42.

Mikels, J. A., Larkin, G. R., Reuter-Lorenz, P. A., & Carstensen, L. L. (2005). Divergent trajectories in the aging mind: Changes in working memory for affective versus visual information with age. *Psychology and Aging*, *20*, 542–553.

Mikels, J. A., Reuter-Lorenz, P. A., Beyer, J. A., & Fredrickson, B. L. (2008). Emotion and working memory: Evidence for domain-specific processes for affective maintenance. *Emotion*, *8*, 256–266.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*(3), 504–511.

Towse, J. N., Hitch, G. I., Hamilton, Z., Peacock, K., & Hutton, U. (2005). Working memory period: The endurance of mental representations. *Quarterly Journal of Experimental Psychology*, *58A*, 547–571.

Wallis, J. D. (2007). Orbitofrontal cortex and its contribution to decision-making. *Annual Review of Neuroscience*, *30*, 31–56.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology*, *49A*, 51–74.