Published in final edited form as:

Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2017 September ; 24(5): 555–574. doi: 10.1080/13825585.2016.1227423.

The effects of emotion on younger and older adults' monitoring of learning

Sarah K. Tauber^a, John Dunlosky^b, Heather L. Urry^c, and Philipp C. Opitz^d

^aDepartment of Psychology, Texas Christian University, Fort Worth, TX, USA

^bDepartment of Psychological Sciences, Kent State University, Kent, OH, USA

^cDepartment of Psychology, Tufts University, Medford, MA, USA

^dDavis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Abstract

Age-related differences in memory monitoring appear when people learn emotional words. Namely, younger adults' judgments of learning (JOLs) are higher for positive than neutral words, whereas older adults' JOLs do not discriminate between positive versus neutral words. In two experiments, we evaluated whether this age-related difference extends to learning positive versus neutral pictures. We also evaluated the contribution of two dimensions of emotion that may impact younger and older adults' JOLs: valence and arousal. Younger and older adults studied pictures that were positive or neutral and either high or low in arousal. Participants made immediate JOLs and completed memory tests. In both experiments, the magnitude of older adults' JOLs was influenced by emotion, and both younger and older adults demonstrated an emotional salience effect on JOLs. As important, the magnitude of participants' JOLs was influenced by valence, and not arousal. Emotional salience effects were also evident on participants' free recall, and older adults recalled as many pictures as did younger adults. Taken together, these data suggest that older adults do not have a monitoring deficit when learning positive (vs. neutral) pictures and that emotional salience effects on younger and older adults' JOLs are produced more by valence than by arousal.

Keywords

Aging; metamemory; judgments of learning; valence; arousal

The study of metacognitive aging has become an important subdiscipline of cognitive-aging research. First, any age-related declines in metacognitive processes can potentially explain age-related deficits in learning and memory (for reviews, Hertzog & Hultsch, 2000; Tauber & Witherby, 2016). For instance, if older adults show deficits in monitoring their on-going learning, such monitoring deficits could subsequently limit the degree to which they could effectively regulate their learning (e.g., Dunlosky & Rawson, 2012; Thiede, 1999). Second,

CONTACT: Sarah K. Uma Tauber, uma.tauber@tcu.edu.

Disclosure statement

No potential conflict of interest was reported by the authors.

metacognitive processes (such as monitoring and allocation of study time) are malleable, so even if age-related deficits do arise, older adults could potentially improve these metacognitive processes (e.g., through training) and in turn show improvements in cognition (see Dunlosky, Kubat-Silman, & Hertzog, 2003; West, Bagwell, & Dark-Freudeman, 2008). Thus, in the present research, we are interested in the influence of healthy aging on how people monitor their learning.

Although the modal outcome of research on metacognitive aging indicates that monitoring of learning stays intact well into old age (for reviews, see Castel, Middlebrooks, & McGillivray, 2016; Hertzog & Dunlosky, 2011; Hertzog & Hultsch, 2000), some exceptions suggest that older adults' monitoring may be vulnerable (e.g., Daniels, Toth, & Hertzog, 2009; Morson, Moulin, & Souchay, 2015; Toth, Daniels, & Solinger, 2011). Most relevant to the present research, Tauber and Dunlosky (2012) found age differences in monitoring of learning for emotional versus neutral words and an age-related monitoring deficit when adults learn positive versus neutral words. The primary goal of the current research was to evaluate whether this age difference arises when younger and older adults learn positive and neutral scenes depicted in pictures. Our secondary goal was to explore potential dimensions of emotion, specifically valence and arousal that may be responsible for effects on younger and older adults' monitoring of learning. In the next two sections, we review the literature pertaining to each of these goals in turn.

Younger and older adults' monitoring of emotional information

One of the most common measures of monitoring during study are judgments of learning (JOLs; for a recent review, see Rhodes, 2016). In Tauber and Dunlosky (2012), younger and older adults studied negative, positive, and neutral words. Immediately after each word, they made a JOL in which they predicted the likelihood of remembering that word on a future test. JOLs were made on a scale from 0% (not at all likely to remember the word) to 100% (absolutely likely to remember the word). Finally, participants attempted to recall the words. Results from two experiments revealed a consistent emotional salience effect for younger adults: JOLs were higher for emotional words (positive and negative) relative to neutral. Though not investigated with older adults' JOLs, emotional salience effects have also been documented on younger adults' JOLs for learning word pairs (Zimmerman & Kelley, 2010), emotional faces (Nomi, Rhodes, & Cleary, 2013), and emotional pictures (Hourihan & Bursey, 2015). Most important, older adults' JOLs were higher for negative relative to neutral words, and an age-related monitoring difference occurred for positive words. Specifically, older adults' JOLs did not differ between positive and neutral words (Tauber & Dunlosky, 2012). In contrast, younger adults' JOLs were consistently higher for positive than neutral words. This age-related difference was evident when the manipulation involved three valence categories (positive, negative, and neutral words) and two valence categories (positive and neutral words).

It is possible that the age-related difference in the magnitude of JOLs for positive versus neutral information reflects a general lack of monitoring positive aspects of stimuli. Prior research does not address this possibility because the age-related difference has only been demonstrated for words, and single words may not be sufficiently provocative to highlight

positive aspects of the stimuli. If this is the case, then the age-related monitoring difference for positive information may be eliminated with pictures, which are rich with emotional detail and are potentially more emotionally provocative. Indeed, neural responses tend to be more pronounced after viewing pictures than single words (Hinojosa, Carretie, Valcarcel, Mendez-Bertolo, & Pozo, 2009; Kensinger & Schacter, 2006; but see Bayer & Schacht, 2014). One explanation for such outcomes is that pictures are more complex and provide more spatial information relative to words (e.g., Schlochtermeier et al., 2013; Tempel et al., 2013). Given that emotional pictures are more provocative; when older adults study to-be-learned pictures, their JOLs may be higher for positive relative to neutral pictures. To evaluate this possibility, we used pictures of positive and neutral scenes that were selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008), which is a large database of emotional pictures. The IAPS provides normative ratings of valence and arousal for each picture, and Grühn and Scheibe (2008) provide normative ratings from younger and older adults.

Emotion effects on judgments of learning: the contributions of valence and arousal

By capitalizing on the normative valence and arousal ratings for the IAPS pictures, we can also begin exploring two dimensions of emotion that may be responsible for effects on JOLs: valence and arousal. Valence pertains to the degree to which an item is unpleasant to pleasant. It is often divided into positive, negative, and neutral categories. Arousal pertains to the degree to which an item activates one's emotions, ranging on a continuum from calm to exciting. Currently, research has supported the arousal account of emotional influences on JOLs (Nomi et al., 2013; Zimmerman & Kelley, 2010). Zimmerman and Kelley (2010) reasoned that if emotional effects on JOLs are driven by valence, then JOLs should differ between items that differ in valence (positive or negative), but are equivalent in arousal. However, their results demonstrated that younger adults' JOLs did not discriminate between positive and negative words (or word pairs), but did discriminate between emotional and neutral words. That is, participants gave higher JOLs to the words that were more arousing (i.e., positive and negative items) than to words that were less arousing (i.e., neutral items), which supports an arousal account. Nomi et al. (2013) applied the same logic to their research, which evaluated younger adults' JOLs when learning faces with happy, angry, or neutral expressions. Younger adults' JOLs did not differ between happy and angry expressions, but were higher for the emotional expressions (happy and angry) than for neutral expressions, which also supports an arousal account.

Although the arousal account for emotional effects on JOLs is compelling, the current evidence in support of it comes from studies in which valence and arousal varied together across the categories of items (e.g., positive vs. neutral or negative vs. neutral). In particular, valence and arousal for emotional items (either positive or negative) were more pleasant or unpleasant (for positive and negative, respectively) *and* more arousing than neutral items (Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). Ideally, support for the arousal account would come from studies that dissociate valence from arousal to examine the degree to which each independently contributes to JOLs. To address this issue, we manipulated

valence and arousal in both experiments using normative ratings provided by Grühn and Scheibe (2008).

Emotion effects on memory

Finally, although the effects of emotional stimuli on people's JOLs were focal to the present research, we also report analyses of memory performance for emotional information. A long-standing issue in metamemory research pertains to whether the effects of a given cue are the same on metamemory judgments and on memory. Prior research has demonstrated a variety of dissociations (for a review, see Dunlosky & Metcalfe, 2009) in which a cue impacts JOLs and not memory (and vice versa), suggesting that JOLs do not directly tap underlying memory strength (e.g., Koriat, 1997; Mazzoni & Nelson, 1995; Susser, Mulligan, & Besken, 2013). Given the paucity of research on emotion and JOLs, exploring whether emotion has a different impact on JOLs and memory is critical (and we revisit this issue in the General Discussion section).

Previous research has established that in some instances, age-related differences occur in memory for emotional information (for reviews see Carstensen & Mikels, 2005; Reed, Chan, & Mikels, 2014). Older adults tend to focus on positive information at the expense of the negative information, whereas younger adults either tend to focus on the negative information more so than positive or they focus equally on positive and negative information (but see Denburg, Buchanan, Tranel, & Adolphs, 2003; Kensinger, Brierley, Medford, Growdon, & Corkin, 2002). Note that "the positivity effect concerns the relative difference between older and younger people in attention to and memory for *positive* as opposed to *negative* material" (Reed & Carstensen, 2012, p. 1, italics in original). For the current experiments, the primary focus was on evaluating older adults' monitoring when learning positive relative to neutral pictures, and hence we could not examine positivity (or negativity) biases on memory. Age-related similarities also occur in memory for emotional information (for a review see Murphy & Isaacowitz, 2008). For instance, younger and older adults both demonstrate an emotional salience effect on memory: They remember more emotional information (positive and negative) relative to neutral. Based on prior research, we predicted that younger and older adults will demonstrate an emotional salience effect on memory in the context of studying positive versus neutral pictures.

Summary and predictions

The primary goal of the present research was to evaluate younger and older adults' monitoring of their learning of meaningful pictures. To do so, participants studied positive and neutral pictures, made judgments about their learning, and received memory tests. If age-related differences in the magnitude of JOLs represent a general lack of monitoring the positive aspects of stimuli, then age-related differences in the magnitude of JOLs should be observed for positive (vs. neutral) pictures as well. As noted above, however, other evidence suggests that the age-related monitoring difference for positive information will be diminished or eliminated when learning pictures. The present experiments will provide evidence to competitively evaluate these possibilities.

Our secondary goal was to evaluate the degree to which valence and arousal contribute to emotional effects on younger and older adults' JOLs. To do so, pictures were selected from the IAPS (Lang et al., 2008) with normative ratings of valence and arousal that were gathered from younger and older adults (Grühn & Scheibe, 2008). Thus, we evaluated the magnitude of younger and older adults' JOLs to normative ratings of valence (henceforth valence) and to normative ratings of arousal (henceforth arousal). If emotional effects on JOLs are due to valence, then JOLs will be higher for positive than neutral pictures, regardless of level of arousal. In contrast, if emotional effects on JOLs are due to arousal, then JOLs will be higher for more arousing pictures regardless of valence. Finally, valence and arousal may both contribute to participants' JOLs. In that case, JOLs will be higher for high-arousal positive pictures than for low-arousal positive pictures and also higher for low-arousal positive pictures than neutral pictures. To preview, in Experiment 2, we further evaluated the magnitude of participants' JOLs for pictures that were normatively rated as low-arousal neutral (e.g., a rock formation) and high-arousal neutral (e.g., lava).

Experiment 1

Younger and older adult participants in Experiment 1 studied pictures and made an immediate JOL for each one. Pictures were selected from the IAPS (Lang et al., 2008) and were positive and high in arousal (e.g., loving embrace or skydiving), positive and low in arousal (e.g., fishing or puppies), or neutral and low in arousal (e.g., a farmer in a field or a rock formation). Following study, participants recalled as many pictures as possible and then completed a recognition memory test for the pictures. For two reasons, our results highlight recall results (whereas recognition results appear in the Appendix). First, recall performance is relevant to estimating monitoring resolution, which although is not the focus of the present experiments it is relevant to a more complete analysis of metacognitive monitoring. Second, the recall test occurred before the recognition test (and the former would likely influence the latter, e.g., the testing effect); thus, the recognition test will not be discussed further. Interested readers can find the descriptive values for it in Table A1 in the Appendix.

Method

Design and participants—A 2 (age group: younger, older) $\times$ 3 (picture type: high-arousal positive, low-arousal positive, low-arousal neutral) mixed-factor design was used with picture type manipulated within-participant. Forty-two younger adults ($M_{age} = 19.8$, $SE = .7$) from Kent State University participated in exchange for course credit, and 40 older adults ($M = 68.2$, $SE = .9$) recruited from the community received \$25 each for their time. The gender distribution (29 younger and 22 older women) did not differ between the age groups, $\chi^2(1, N = 82) = 1.7$, $p = .19$. As is common in research on cognitive aging, vocabulary scores (Ekstrom, French, Harmon, & Dermen, 1976) were higher for older adults ($M = .69$, $SE = .03$) than for younger adults, $M = .47$, $SE = .02$; $t(80) = 6.5$, $p < .001$, $d = 1.4$. No participants reported vision impairments that prevented them from completing any tasks.

Materials—A total of 180 pictures were selected from the IAPS (Lang et al., 2008). Pictures were selected using normative ratings of valence and arousal collected from

younger and older adult populations (Grühn & Scheibe, 2008). Sixty pictures were positive and high in arousal, 60 were positive and low in arousal, and 60 were neutral and low in arousal. Table 1 provides mean valence and arousal ratings for each picture type, and Table 2 provides inferential statistics for analyses of valence and arousal between each picture type. Half of the pictures (i.e., 90 pictures, 30 of each picture type) comprised Set 1, and the other half comprised Set 2. One picture (a high-arousal positive picture) was not visible during study. All measures associated with this picture were removed from the reported analyses and the programming error was eliminated in Experiment 2. The picture sets were counterbalanced between study and the recognition test so that half of the participants in each age group received Set 1 as the study set and Set 2 as new pictures on the recognition test and the other half received Set 2 as the study set and Set 1 as new pictures on the recognition test. Set 1 and Set 2 did not differ in valence or arousal ($t_s < 1$).

Procedure—All participants sat in front of the computer screen and could adjust their chair and screen to their preferred viewing angle. Following informed consent, reports of demographic information and the vocabulary test, participants were given the following instructions:

Your task is to study each picture so that you will be able to remember it on a future memory test. On the test, your task will be to describe details from the pictures that you studied. Your goal will be to describe as many pictures as possible with enough details so that someone else could pick out which picture you are trying to describe.

Next, the set of pictures were presented one-at-a-time in the center of the computer screen. Pictures were presented for 1 sec each, and the order of pictures was randomized anew per participant. Immediately after studying each picture, participants made a JOL predicting the likelihood of remembering that picture on a future test. JOLs were made on a scale from 0% (certain the picture would not be remembered) to 100% (certain the picture would be remembered). Participants were given unlimited time to make their JOLs.

After the study and JOL phase, participants completed unrelated tasks (participants' selection of a Sudoku puzzle, a crossword puzzle, and/or math problems) during a 20 min retention interval. Next, participants received a free recall test with the instructions to type a description of each remembered picture with enough detail so that someone else could identify that picture. To do so, participants responded to the prompt, "I remember a picture that had..." and hit enter after they typed each description. To help participants keep track of their entries, an on-screen counter increased with the addition of each new picture description. Also, each entry was added to a "Remember List" on the screen so that participants could see the descriptions they had already entered. Participants were instructed to continue this process (i.e., answering the prompt and hitting enter) until they had listed the details for every picture they could remember, and they were given unlimited time to do so.

Following the free recall test, participants completed a recognition test. On the recognition test, participants were presented with the pictures from the study set and from the new set (i.e., 180 total pictures, 90 old and 90 new). The order of pictures was randomized anew per participant, and each picture was presented one-at-a-time in the center of the screen.

Immediately following each recognition decision, participants made a confidence judgment indicating their level of confidence in their recognition decision (i.e., studied or new). Analyses and descriptive values for the recognition test can be found in the Appendix. After the recognition test, participants were debriefed and compensated for their time.

Results

Our primary interest was in younger and older adult's JOLs; thus, analyses of JOL magnitude are presented first. Next, we briefly present some outcomes relevant to recall performance and monitoring resolution.

JOL magnitude—As evident from Figure 1, for both younger and older adults, JOLs were higher for positive pictures than for neutral pictures. In contrast, participants' JOLs did not differ between positive pictures that were highly arousing and positive pictures that were less arousing. These observations were supported by a 2 (age group) $\times$ 3 (picture type) mixed-factor analysis of variance (ANOVA). A significant main effect of age group, $F(1, 80) = 8.1$, $p = .006$, $\eta_p^2 = .09$, indicated that older adults' JOLs were significantly higher than younger adults' JOLs. A significant main effect of picture type occurred $F(2, 160) = 54.3$, $p < .001$, $\eta_p^2 = .40$, which was qualified by a significant picture type by age group interaction, $F(1, 80) = 13.64$, $p < .001$, $\eta_p^2 = .14$.

To interpret the significant interaction, we conducted paired t -tests between picture types within the two age groups. In doing so, we found that JOLs were significantly higher for high-arousal positive pictures than for low-arousal neutral pictures both for younger adults, $t(41) = 7.4$, $p < .001$, $d = .69$, and for older adults, $t(39) = 3.5$, $p = .001$, $d = .22$. JOLs were also significantly higher for low-arousal positive pictures than low-arousal neutral pictures for younger adults, $t(41) = 7.8$, $p < .001$, $d = .64$, and for older adults, $t(39) = 4.2$, $p < .001$, $d = .26$. Finally, JOLs did not differ between high-arousal positive pictures and low-arousal positive pictures for younger, $t(41) = 1.4$, $p = .18$, or older adults ($t < 1$).

Recall performance—Free recall responses were scored by two independent raters who were blind to participants' ages and to picture type. Participants provided 1,600 total responses (45% from older adults). Of those responses 6.6% were excluded (3.0% from older adults) because the description was too vague to discriminate between pictures (e.g., "person," "an animal"). An additional 6.1% of responses (2.6% from older adults) were excluded because they were not descriptions of pictures that had been studied. Raters' scores were in agreement for 87.8% of the remaining responses (1,397 responses, 623 from older adults). Agreement was high for responses from each picture set (Set 1: 87.5% agreement; Set 2: 88.0% agreement) and from each age group (younger responses: 87.7% agreement; older responses 87.8% agreement). For 12.2% (i.e., 171) of these responses (5.4% from older adults), the two raters' scores were in conflict. For 55.6% (i.e., 95) of these conflicts, the conflict was resolved by a third rater. The third rater could not resolve the remaining conflicts (76 responses, 29 from older adults) by identifying a specific picture for each. However, the picture type (i.e., high-arousal positive, low-arousal positive, or low-arousal neutral) for each response was identified. As such, these responses were included in the free recall analyses. A final 2.9% (36 responses, 14 from older adults) of the responses were

removed because the description was of a picture that the participant had already described in an earlier response. The mean percent of pictures correctly recalled by younger and older adults are provided in Table 3.

As evident from Table 3, both age groups demonstrated an emotional salience effect on recall, and younger adults also demonstrated superior memory for high-arousal pictures relative to pictures that were less arousing. These observations were supported by a 2 (age group) $\times$ 3 (picture type) mixed-factor ANOVA. The main effect of age group was not significant, $F(1, 79) = 2.6$, $p = .11$, $\eta_p^2 = .03$. There was a significant main effect of picture type, $F(2, 158) = 24.5$, $p < .001$, $\eta_p^2 = .24$, which was qualified by a significant picture type by age group interaction, $F(2, 158) = 3.7$, $p = .03$, $\eta_p^2 = .05$.

Follow-up tests revealed that recall was significantly greater for high-arousal positive pictures than for low-arousal neutral pictures for younger adults, $t(41) = 5.4$, $p < .001$, $d = .82$, and for older adults, $t(38) = 3.2$, $p = .003$, $d = .40$. Recall was significantly greater for low-arousal positive pictures than for low-arousal neutral pictures for younger adults, $t(41) = 3.5$, $p = .001$, $d = .46$, and older adults, $t(38) = 4.1$, $p < .001$, $d = .48$. Recall was also significantly greater for high-arousal positive pictures than for low-arousal positive pictures for younger adults, $t(41) = 2.9$, $p = .006$, $d = .40$. Recall did not differ between high-arousal positive pictures and low-arousal positive pictures for older adults, $t < 1$.

Monitoring resolution—Though not relevant to our goals, for completeness we assessed monitoring resolution via within-participant gamma correlations (Nelson, 1984).

Correlations were computed between JOLs and recall performance and averaged separately for younger and older adults. Four correlations were calculated, one including all pictures and one for each picture type (i.e., high-arousal positive, low-arousal positive, low-arousal neutral). All correlations were significantly greater than zero (overall, $M = .12$, $SE = .04$; high-arousal positive, $M = .23$, $SE = .05$; low-arousal positive, $M = .24$, $SE = .05$; $t_s > 2.8$), except for correlations for low-arousal neutral pictures ($M = .11$, $SE = .06$) $t(67) = 2.0$, $p = .06$. A 2 (age group) $\times$ 3 (picture type) mixed-factor ANOVA revealed no significant effects, $F_s < 3.9$.

Experiment 2

Experiment 1 revealed that the magnitude of younger and older adults' JOLs was influenced by valence because they were higher for positive than for neutral pictures. As important, the magnitude of younger and older adults' JOLs was not influenced by arousal. These results support a valence account of JOLs, although they might also reflect a kind of isolation effect (Hunt, 1995). Specifically, the items that likely stood out the most in Experiment 1 were the low-arousal neutral items, given that there were fewer neutral items than positive, and all neutral items were low in arousal. In other words, the to-be-learned pictures included twice as many positive pictures than neutral ones, which may have made valence stand out to participants. By contrast, there were twice as many low-arousal pictures than high-arousal pictures, and hence this stimulus dimension may not have been as salient. The primary goal of Experiment 2 was to replicate and extend the results from Experiment 1 when eliminating these differences in picture set composition. To do so, valence and arousal were factorially

manipulated in Experiment 2, which resulted in four types of items: high-arousal positive, low-arousal positive, high-arousal neutral, and low-arousal neutral.

Method

Design and participants—A 2 (age group: younger, older) $\times$ 2 (valence: positive, neutral) $\times$ 2 (arousal: high, low) mixed-factor design was used with valence and arousal manipulated within-participant. Thirty-eight younger adults ($M_{age} = 20.2$, $SE = .4$) from Kent State University participated in exchange for course credit, and 42 older adults ($M_{age} = 71.0$, $SE = .9$) recruited from the community received \$25 each for their time. The gender distribution (29 younger and 27 older women) did not differ between the age groups, $\chi^2(1, N = 80) = 1.4$, $p = .24$, and vocabulary scores (Ekstrom et al., 1976) were higher for older adults ($M = .70$, $SE = .03$) than for younger adults, $M = .44$, $SE = .02$; $t(73) = 7.6$, $p < .001$, $d = 1.7$. The vocabulary test was administered incorrectly for five older adult participants, and their vocabulary scores were removed from analyses. No participants reported vision impairments that prevented them from completing any tasks.

Materials—Materials were identical to those used in Experiment 1, with the addition of neutral pictures that were high in arousal. High-arousal neutral pictures were selected from the IAPS (Lang et al., 2008) using the same criteria as in Experiment 1 (i.e., normative ratings from Grühn & Scheibe, 2008). High-arousal neutral pictures depicted scenes that were not rated as particularly positive or negative, but that were normatively rated as more arousing relative to low-arousal neutral pictures (e.g., a doctor in a white coat, a pilot in a cockpit, lava). Two picture sets were created (Set 1 and Set 2) each with 120 pictures, 30 from each picture type (high-arousal positive, low-arousal positive, high-arousal neutral, and low-arousal neutral; see Tables 1 and 2). The picture sets were counterbalanced across participants as in Experiment 1, and they did not differ in valence or arousal ($t_s < 1$).

Procedure—After providing written informed consent and prior to study, participants in Experiment 2 were given the following instructions:

Your task is to study each picture so that you will be able to remember it on a future memory test. On the memory test, you will be presented with images. Some of the images will be those that you originally studied, and others will be new. New images will either be slightly altered versions from those you originally studied or will be entirely new images. The altered (new) images will be horizontally reversed from the original image. Here is an example: [example images provided each labeled as 'original' or 'new']. On the test, each image will be presented individually, and your task will be to determine whether you studied the image earlier or whether the image is new.

Participants were provided with these instructions to anticipate the nature of the recognition test, which was slightly different from Experiment 1. The test included horizontally reversed versions of studied pictures to increase the rate of false alarms and decrease the rate of hits, relative to those from Experiment 1. Although these updates were meant to increase the sensitivity of the recognition test, these data were less relevant to the current report and

hence will not be discussed further, but the outcomes are presented in the Appendix. Otherwise, the procedure for study, JOLs, and free recall was identical to Experiment 1.

Results

As in Experiment 1, results are ordered so that analyses of JOL magnitude are presented first followed by recall performance and monitoring resolution.

JOL magnitude—As evident from Figure 2, participants' JOLs replicated the effects established in Experiment 1; JOLs were higher for positive pictures than for neutral pictures. In contrast, the magnitude of participants' JOLs was less influenced by arousal. These observations were supported by a 2 (age group) $\times$ 2 (valence) $\times$ 2 (arousal) mixed-factor ANOVA. There was a significant main effect of valence, $F(1, 78) = 54.2, p < .001, \eta_p^2 = .41$, which was qualified by a significant valence by age group interaction, $F(1, 78) = 5.3, p = .02, \eta_p^2 = .06$, a marginal valence by arousal interaction, $F(1, 78) = 3.7, p = .06, \eta_p^2 = .05$, and a significant three-way interaction between age group, valence, and arousal, $F(1, 78) = 7.7, p = .007, \eta_p^2 = .09$.

To interpret the significant three-way interaction, we conducted paired *t*-tests between picture types within the two age groups. In doing so, we found that JOLs were significantly higher for high-arousal positive pictures than for high-arousal neutral pictures for younger adults, $t(37) = 5.1, p < .001, d = .31$ and older adults, $t(41) = 6.3, p < .001, d = .26$. JOLs were significantly higher for low-arousal positive pictures than for low-arousal neutral pictures for younger adults, $t(37) = 5.2, p < .001, d = .34$ and were marginally higher for low-arousal positive pictures than for low-arousal neutral pictures for older adults, $t(41) = 2.0, p = .06, d = .09$. Notably, JOLs did not differ between high-arousal positive and low-arousal positive pictures for younger adults and older adults, $t < 1$. JOLs did not differ between high-arousal neutral and low-arousal neutral pictures for younger adults ($t < 1$), and were significantly higher for low-arousal neutral pictures than for high-arousal neutral pictures for older adults JOLs, $t(41) = 3.5, p = .001, d = .15$. No other effects were significant $F < 2.1$.

Recall performance—Free recall responses were scored with the same procedures as Experiment 1. Participants provided 1,330 total responses (59.2% from older adults). Of those responses 6.8% were excluded (4.7% from older adults) because the description was too vague to discriminate between pictures (e.g., “light,” “a baby,” “scenery”). An additional 4.6% of responses (3.2% from older adults) were excluded because they were not descriptions of pictures that had been studied. Raters' scores were in agreement for 85.1% of the remaining responses (1,179 responses, 681 from older adults). Agreement was high for responses from each picture set (Set 1: 85.8% agreement; Set 2: 84.2% agreement) and from each age group (younger responses: 91.0% agreement; older responses 80.8% agreement). For 14.9% (i.e., 176) of these responses (11% from older adults), the two raters' scores were in conflict. A third rater resolved 61.4% (i.e., 108) of these conflicts. The third rater could not resolve the remaining conflicts (68 responses, 53 from older adults), but as in Experiment 1, the picture type (i.e., high-arousal positive, low-arousal positive, high-arousal neutral, low-arousal neutral) for each response was identified. Finally, 3.5% (35 responses,

28 from older adults) of the responses were removed because the description was of a picture that the participant had already described in an earlier response.

As evident from Table 3, younger and older adults recalled more emotional pictures than neutral pictures; which was true for pictures that were high in arousal and for pictures that were low in arousal. These observations were supported by a 2 (age group) $\times$ 2 (valence) $\times$ 2 (arousal) mixed-factor ANOVA. There was no significant main effect of age group, $F(1, 78) = 2.0$, $p = .16$, $\eta_p^2 = .03$. The main effect of arousal was also not significant, and it did not interact with age group or valence, $F_s < 1.9$. There was, however, a main effect of valence, $F(1, 78) = 84.9$, $p < .001$, $\eta_p^2 = .52$, which did not interact with age group, $F(1, 78) = 2.6$, $p = .11$, but was qualified by a significant three-way interaction between age group, valence, and arousal, $F(1, 78) = 5.7$, $p = .02$, $\eta_p^2 = .07$.

Follow-up paired t -tests to interpret the three-way interaction revealed that recall was significantly greater for high-arousal positive pictures than for high-arousal neutral pictures for younger adults, $t(37) = 3.7$, $p = .001$, $d = .46$, and older adults, $t(41) = 7.2$, $p < .001$, $d = 1.09$. Recall was significantly greater for high-arousal positive pictures than for low-arousal neutral pictures for younger adults, $t(37) = 3.5$, $p = .001$, $d = .48$, and for older adults, $t(41) = 5.3$, $p < .001$, $d = .73$. Recall was significantly greater for low-arousal positive pictures than for low-arousal neutral pictures for younger adults, $t(37) = 4.9$, $p < .001$, $d = .65$, and for older adults, $t(41) = 3.6$, $p = .001$, $d = .49$. Recall was also significantly greater for low-arousal positive pictures than for high-arousal neutral pictures for younger adults, $t(37) = 4.5$, $p < .001$, $d = .62$, and for older adults, $t(41) = 5.4$, $p < .001$, $d = .80$. Recall did not differ between high-arousal positive and low-arousal positive pictures for younger adults, $t(37) = 1.1$, $p = .27$, or older adults, $t(41) = 1.3$, $p = .21$. Recall did not significantly differ between high-arousal neutral and low-arousal neutral pictures for younger ($t < 1$), and recall was marginally greater for low-arousal neutral pictures than for high-arousal neutral pictures for older adults, $t(41) = 2.0$, $p = .06$.

Monitoring resolution—As in Experiment 1, within-participant correlations were computed between JOLs and memory performance. Five correlations were calculated, one including all pictures and one for each picture type (i.e., high-arousal positive, low-arousal positive, high-arousal neutral, low-arousal neutral). All correlations were significantly greater than zero (overall, $M = .29$, $SE = .03$; high-arousal positive, $M = .24$, $SE = .05$; low-arousal positive, $M = .16$, $SE = .05$; high-arousal neutral, $M = .25$, $SE = .07$; low-arousal neutral, $M = .27$, $SE = .06$; $t_s > 3.0$). A 2 (age group) $\times$ 2 (valence) $\times$ 2 (arousal) mixed-factor ANOVA revealed no significant effects, $F_s < 3.3$.

General discussion

In the current experiments, we evaluated older adults' monitoring of learning and memory for emotional pictures. Results revealed that the magnitude of older adults' JOLs was influenced by valence when learning positive and neutral pictures. This emotional salience effect was also evident on younger adults' JOLs (cf. Hourihan & Bursey, 2015; Nomi et al., 2013; Zimmerman & Kelley, 2010). This outcome contrasts with previous results using words in which older adults' JOLs did not differ between positive and neutral words,

whereas younger adults' JOLs were higher for positive than neutral words (Tauber & Dunlosky, 2012). In contrast with single words, emotional pictures likely provide participants with richer emotional cues on which to base their JOLs (cf. Hinojosa et al., 2009; Kensinger & Schacter, 2006). For instance, normatively positive pictures (e.g., a smiling baby) and neutral pictures (e.g., a towel) can be starkly different, whereas the differences seem smaller between normatively positive words (e.g., bunny) and neutral words (e.g., fabric). The stark differences in emotional cues for pictures with different emotional content may be necessary for valence to impact older adults' JOLs. Whereas for words, older adults may not consistently detect the smaller differences, or they may not believe that the differences would influence recall. Evaluating these possibilities will provide an important challenge for future research aimed at understanding when age differences arise in monitoring of emotional content.

Picture type did not impact monitoring resolution, and no age-related differences in monitoring resolution were evident in either experiment. Also, given that no meaningful age-related differences occurred in JOL magnitude, the present outcomes fit within the larger literature that has established nominal differences between younger and older adults' monitoring of learning (e.g., Connor, Dunlosky, & Hertzog, 1997; Hertzog, Sinclair, & Dunlosky, 2010; Rabinowitz, Ackerman, Craik, & Hinchley, 1982; Tauber & Rhodes, 2012) and support theoretical perspectives on metacognitive aging that maintain that monitoring of learning remains largely intact with age (for reviews see Castel et al., 2016; Hertzog & Dunlosky, 2011).

In addition to our first goal of examining whether age-related differences would arise in monitoring the learning of emotional pictures, our second goal was to assess the contributions of valence and arousal to younger and older adults' JOLs by factorially manipulating arousal and valence of the to-be-learned pictures. The magnitude of both older and younger adults' JOLs was influenced by valence (positive vs. neutral) and was not influenced by arousal (high vs. low arousal). These outcomes were found when arousal was manipulated for positive pictures only (Experiment 1), which could have been explained as a type of isolation effect on participants' JOLs. However, the effects held when arousal was manipulated for both positive and neutral pictures (Experiment 2). Thus, the structure of the lists used in Experiment 1 cannot account for the effect of valence and no effect of arousal on participants' JOLs. Instead, these data support the hypothesis that valence (and not arousal) is the affective dimension more responsible for emotional salience effects on JOLs in the context of positive emotion. Previously, researchers have favored an arousal explanation for emotional effects on JOLs (Nomi et al., 2013; Zimmerman & Kelley, 2010); however, the evidence in support of this explanation was indirect. In particular, researchers found that JOLs did not differ based on valence (negative vs. positive), but this lack of difference does not directly implicate arousal as the factor producing emotional salience effects on JOLs. Moreover, an explanation based on valence could also account for the emotional salience effects on JOLs in prior research using pictures, given that both valence and arousal differed between emotional items (positive or negative) and neutral. Another possibility is that the contributions of valence and arousal differ depending on material type, with valence being influential for pictures and arousal being influential for words. Currently, however, the most parsimonious explanation for emotion effects on JOLs, which has

garnered relatively direct support in the present research, implicates valence as a critical factor.

The potency of valence does not rule out the possibility that arousal could impact JOLs under different task conditions or with different materials. One possibility is that the magnitude of younger and older adults' JOLs was not influenced by arousal because the discrepancy between low-arousal and high-arousal items was difficult to detect. Specifically, the mean arousal rating for low-arousal items was 2.74 and the mean arousal rating for high-arousal items was 3.43. Even though arousal ratings significantly differed (see Table 2), given that they were made on a nine-point scale, it is possible that a larger disparity between low-arousal and high-arousal items is necessary for participants to use this cue as a basis for their JOLs. Even so, it is noteworthy that the disparity between valence ratings for positive and neutral items was also small on the nine-point scale: the mean valence rating for positive items was 6.55 and the mean valence rating for neutral items was 5.58. Thus, in the current experiments, the strength of the arousal manipulation ($d = 1.29$) and valence manipulation ($d = 1.44$) were similar. Even so, an important direction for future research will be to use items with more expansive normative ratings of arousal to determine whether the magnitude of younger and older adults' JOLs can be influenced by it in other contexts.

Current theories of monitoring of learning suggest that JOLs are inferential and can be based on available cues during learning (Koriat, 1997). The inferential view itself is perhaps best supported by the presence of dissociations among JOLs and memory, and taken together, evidence from emotion research supports the same view. Namely, in the present experiments involving free recall, arousal did influence younger adults' recall for positive words (Table 3), whereas it did not influence their JOLs; similarly, prior evidence has shown that older adults' free recall is greater for positive than neutral words, but their JOLs do not differ (Tauber & Dunlosky, 2012). For cued recall (Zimmerman & Kelley, 2010), younger adults' JOLs have been greater for negative than neutral word pairs, with no influence on recall performance. This double-dissociation across experiments suggests that memory and JOLs tap (at least one) different dimension of emotion (but for discussion of why double-dissociations do not provide definitive evidence against a single-process model, see Dunn & Kirsner, 1988). Even so, demonstrating the same dissociation within a single study using the same test format (e.g., all free recall) would provide even stronger evidence for this conclusion.

Moreover, any cue (e.g., valence) could influence JOLs by two non-exclusive mechanisms: fluency and beliefs (for a review, see Dunlosky, Mueller, & Tauber, 2015). Given that valence is merely a human interpretation about whether a picture is normatively positive or negative, in the present context (where arousal and valence were dissociated via design), we view the valence effects on JOLs as arising from an explicit interpretation of a picture as positive or neutral. In contrast, arousal is partly a physiological response to emotional stimuli and could influence younger and older adults' JOLs through the experience of performing the task. Thus, in the present context in which valence accounted for the impact of emotion on JOLs, it seems reasonable that younger and older adults based their JOLs on beliefs about how emotion influences their memory (cf. Jia et al., 2016; Mueller, Dunlosky, Tauber, & Rhodes, 2014; Mueller, Tauber, & Dunlosky, 2013; Susser, Jin, & Mulligan,

2016) rather than on processing experiences (akin to fluency) during study (cf. Hertzog, Dunlosky, Robinson, & Kidder, 2003; Kelley & Jacoby, 1996; Undorf & Erdfelder, 2011, 2013). To the degree that such beliefs about valence are not responsible for emotion effects on memory (which seems a priori plausible), such beliefs may be responsible for the aforementioned different effects of emotion (i.e., dissociations) on JOLs and memory. To more analytically evaluate such possibilities, a challenge for future research will involve directly measuring people's beliefs about emotions and memory as well as their experiences while processing emotional stimuli, so that the contribution of these two factors to emotional effects on JOLs can be estimated.

Although age-related effects on memory were not the focus of this research, the effects of emotion on older adults' recall is currently of much interest in the field, so we briefly consider these effects in the present experiments. Younger and older adults recalled more positive pictures than neutral pictures (see Table 3). Moreover, no age-related differences were revealed on free recall as evident from the lack of a main effect of age group in both experiments. This is a surprising outcome given prior research demonstrating younger adults' superior episodic memory relative to older adults. Even so, other evidence suggests that older adults' recall of emotional pictures can be equivalent with that of younger adults (Tomaszczyk & Fernandes, 2013). Tomaszczyk and Fernandes (2013) investigated the contributions of valence and arousal to memorability decisions and memory. Younger and older adults studied pictures that were negative, neutral, or positive. Half of each picture type was high in arousal, and half was low in arousal. Following study, participants made picture selections by choosing the pictures that they would best remember on a future test, and they received a free recall test. Older adults selected more positive pictures than neutral or negative ones, whereas younger adults' selections were not influenced by valence. More important, age group did not influence free recall of the emotional pictures. Both age groups recalled more positive pictures than negative or neutral ones, and recalled more highly arousing pictures than pictures low in arousal.

Outcomes from the current experiments are consistent with theoretical perspectives on aging and emotional information, such as the socioemotional selectivity theory. From this perspective, age-related differences occur in memory for emotional information because older adults favor positive information relative to negative; by contrast, younger adults favor negative relative to positive, or show no preference (for reviews see Carstensen & Mikels, 2005; Reed et al., 2014). Although the current outcomes do not evaluate such positivity or negativity effects as strictly defined by a comparison between positive and negative items (Reed & Carstensen, 2012), they are nevertheless broadly consistent with the idea that positive information is a priority for older adults, at least for metamemory monitoring. This observation was evident on the magnitude of older adults' JOLs as well as on recall. Consistent with this literature (Murphy & Isaacowitz, 2008), younger and older adults were also similar in that they both demonstrated emotional salience effects on JOLs and on recall.

In summary, the current experiments demonstrated that the magnitude of younger adults' and older adults' JOLs was influenced by emotion when they study pictures. Younger and older adults expected to and actually did better recall positive pictures relative to neutral

pictures. Most important, older adults did not demonstrate impairments in monitoring of learning relative to younger adults. This outcome suggests that prior work showing age-related differences in JOLs using positive verbal stimuli may arise from a dampened emotional response of older adults to positive words, but this conjecture will require future research to verify.

Acknowledgments

We thank Melissa Bishop, Kianna Blackeney, Lauren Bottoms, Michael Goodrich, Tosha Jones, Kayle McCullough, Steve Rafidi, Messa Soussou, and Kelsey Zempel for their assistance with data collection and scoring.

Funding

This research was supported by the James S. McDonnell Foundation 21st Century Science Initiative in Bridging Brain, Mind, and Behavior Collaborative Award.

References

- Bayer M, Schacht A. Event-related brain responses to emotional words, pictures, and faces – a cross-domain comparison. *Frontiers in Psychology*. 2014; 5:1–10. DOI: 10.3389/fpsyg.2014.01106 [PubMed: 24474945]
- Carstensen LL, Mikels JA. At the intersection of emotion and cognition – Aging and the positivity effect. *Current Directions in Psychological Science*. 2005; 14:117–121. DOI: 10.1111/j.0963-7214.2005.00348.x
- Castel, AD., Middlebrooks, CD., McGillivray, S. Monitoring memory in old age: Impaired, spared, and aware. In: Dunlosky, J., Tauber, SK., editors. *Oxford handbook of metacognition*. New York, NY: Oxford University Press; 2016. p. 519-535.
- Connor LT, Dunlosky J, Hertzog C. Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*. 1997; 12:50–71. DOI: 10.1037/0882-7974.12.1.50 [PubMed: 9100268]
- Daniels KA, Toth JP, Hertzog C. Aging and recollection in the accuracy of judgments of learning. *Psychology and Aging*. 2009; 24:494–500. DOI: 10.1037/a0015269 [PubMed: 19485665]
- Denburg NL, Buchanan TW, Tranel D, Adolphs R. Evidence for preserved emotional memory in normal older persons. *Emotion*. 2003; 3:239–253. DOI: 10.1037/1528-3542.3.3.239 [PubMed: 14498794]
- Dunlosky J, Kubat-Silman AK, Hertzog C. Training monitoring skills improves older adults' self-paced associative learning. *Psychology and Aging*. 2003; 18:340–345. DOI: 10.1037/0882-7974.18.2.340 [PubMed: 12825781]
- Dunlosky, J., Metcalfe, J. *Metacognition*. Thousand Oaks, CA: Sage Publications, Inc; 2009.
- Dunlosky, J., Mueller, MK., Tauber, SK. The contribution of subjective fluency and theories of memory to people's judgments of memory. In: Lindsay, DS, Kelley, CM, Yonelinas, AP., Roediger, HL., III, editors. *Remembering: Attributions, processes, and control in human memory: Papers in honour of Larry L. Jacoby*. New York: Psychology Press; 2015. p. 46-64.
- Dunlosky J, Rawson KA. Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*. 2012; 22:271–280. DOI: 10.1016/j.learninstruc.2011.08.003
- Dunn JC, Kirsner K. Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review*. 1988; 95:91–101. DOI: 10.1037/0033-295X.95.1.91 [PubMed: 3353477]
- Ekstrom, RB., French, JW., Harmon, HH., Dermen, D. *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service; 1976.
- Grühn D, Scheibe S. Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age?

- Behavior Research Methods. 2008; 40:512–521. DOI: 10.3758/BRM.40.2.512 [PubMed: 18522062]
- Hertzog C, Dunlosky J. Metacognition in later adulthood: Spared monitoring can benefit older adults' self-regulation. *Current Directions in Psychological Science*. 2011; 20:174–180. DOI: 10.1177/0963721411409026 [PubMed: 25866452]
- Hertzog C, Dunlosky J, Robinson AE, Kidder DP. Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2003; 29:22–34. DOI: 10.1037/0278-7393.29.1.22
- Hertzog, C., Hultsch, DF. Metacognition in adulthood and old age. In: Craik, FIM., Salthouse, TA., editors. *The handbook of aging and cognition*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2000. p. 417-466.
- Hertzog C, Sinclair SM, Dunlosky J. Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. *Developmental Psychology*. 2010; 46:939–948. DOI: 10.1037/a0019812 [PubMed: 20604613]
- Hinojosa JA, Carretie L, Valcarcel MA, Mendez-Bertolo C, Pozo MA. Electrophysiological differences in the processing of affective information in words and pictures. *Cognitive, Affective, & Behavioral Neuroscience*. 2009; 9:173–189. DOI: 10.3758/CABN.9.2.173
- Hourihaan KL, Bursey E. A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory, Advanced Online Publication*. 2015; doi: 10.1080/09658211.2015.1122809
- Hunt RR. The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review*. 1995; 2:105–112. DOI: 10.3758/BF03214414 [PubMed: 24203592]
- Jia X, Li P, Li X, Zhang Y, Cao W, Cao L, Li W. The effect of word frequency on judgments of learning: Contributions of beliefs and processing fluency. *Frontiers in Psychology*. 2016; 6:1–10. DOI: 10.3389/fpsyg.2015.01995
- Kelley CM, Jacoby LL. Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*. 1996; 35:157–175. DOI: 10.1006/jmla.1996.0009
- Kensinger EA, Brierley B, Medford N, Growdon JH, Corkin S. Effects of normal aging and Alzheimer's disease on emotional memory. *Emotion*. 2002; 2:118–134. DOI: 10.1037/1528-3542.2.2.118 [PubMed: 12899186]
- Kensinger EA, Schacter DL. Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*. 2006; 6:110–126. DOI: 10.3758/CABN.6.2.110
- Koriat A. Monitoring one's own knowledge during study: A cue utilization approach to judgments of learning. *Journal of Experimental Psychology: General*. 1997; 126:349–370. DOI: 10.1037/0096-3445.126.4.349
- Lang, PJ., Bradley, MM., Cuthbert, BN. *International affective picture system (IAPS): Affective ratings of pictures and instruction manual (Technical Report A-8)*. Gainesville, FL: University of Florida; 2008.
- Loftus GR, Masson MEJ. Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*. 1994; 1:476–490. DOI: 10.3758/BF03210951 [PubMed: 24203555]
- Mazzoni G, Nelson TO. Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1995; 21:1263–1274. DOI: 10.1037/0278-7393.21.5.1263
- Morson SM, Moulin CJA, Souchay C. Selective deficits in episodic feeling of knowing in ageing: A novel use of the general knowledge task. *Acta Psychologica*. 2015; 157:85–92. DOI: 10.1016/j.actpsy.2015.02.014 [PubMed: 25747574]
- Mueller ML, Dunlosky J, Tauber SK, Rhodes MG. The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language*. 2014; 70:1–12. DOI: 10.1016/j.jml.2013.09.007
- Mueller ML, Tauber SK, Dunlosky J. Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review*. 2013; 20:378–384. DOI: 10.3758/s13423-012-0343-6 [PubMed: 23188739]

- Murphy NA, Isaacowitz DM. Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging*. 2008; 23:263–286. DOI: 10.1037/0882-7974.23.2.263 [PubMed: 18573002]
- Nelson TO. A comparison of current measures of the accuracy of feeling-of knowing predictions. *Psychological Bulletin*. 1984; 95:109–133. DOI: 10.1037/0033-2909.95.1.109 [PubMed: 6544431]
- Nomi JS, Rhodes MG, Cleary AM. Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition & Emotion*. 2013; 27:141–149. DOI: 10.1080/02699931.2012.679917 [PubMed: 22712473]
- Rabinowitz JC, Ackerman BP, Craik FIM, Hinchley JL. Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology*. 1982; 37:688–695. DOI: 10.1093/geronj/37.6.688 [PubMed: 7130642]
- Reed AE, Carstensen LL. The theory behind the age-related positivity effect. *Frontiers in Psychology*. 2012; 3doi: 10.3389/fpsyg.2012.00339
- Reed AE, Chan L, Mikels JA. Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging*. 2014; 29:1–15. DOI: 10.1037/a0035194 [PubMed: 24660792]
- Rhodes, MG. Judgments of learning: Methods, data, and theory. In: Dunlosky, J., Tauber, SK., editors. *Oxford handbook of metacognition*. New York, NY: Oxford University Press; 2016. p. 65-80.
- Schlochtermeier LH, Kuchinke K, Pehrs C, Urton K, Kappelhoff H, Jacobs AM. Emotional picture and word processing: An fMRI study on effects of stimulus complexity. *PLoS One*. 2013; 8:e55619.doi: 10.1371/journal.pone.0055619 [PubMed: 23409009]
- Susser JA, Jin A, Mulligan NW. Identity priming consistently affects perceptual fluency but only affects metamemory when primes are obvious. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2016; 42:657–662. DOI: 10.1037/xlm0000189
- Susser JA, Mulligan NW, Besken M. The effects of list composition and perceptual fluency on judgments of learning (JOLs). *Memory & Cognition*. 2013; 41:1000–1011. DOI: 10.3758/s13421-013-0323-8 [PubMed: 23661189]
- Tauber SK, Dunlosky J. Can older adults accurately judge their learning of emotional information? *Psychology and Aging*. 2012; 27:924–933. DOI: 10.1037/a0028447 [PubMed: 22663156]
- Tauber SK, Rhodes MG. Multiple bases for young and older adults' judgments of learning in multitrial learning. *Psychology and Aging*. 2012; 27:474–483. DOI: 10.1037/a0025246 [PubMed: 21942899]
- Tauber, SK., Witherby, AE. Metacognition in older adulthood. In: Panchana, NA., editor. *Encyclopedia of geropsychology*. New York, NY: Springer; 2016. p. 1-15.
- Tempel K, Kuchinke L, Urton K, Schlochtermeier LH, Kappelhoff H, Jacobs AM. Effects of positive pictograms and words: An emotional word superiority effect? *Journal of Neurolinguistics*. 2013; 26:637–648. DOI: 10.1016/j.jneuroling.2013.05.002
- Thiede KW. The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*. 1999; 6:662–667. DOI: 10.3758/BF03212976 [PubMed: 10682210]
- Tomaszczyk JC, Fernandes MA. A positivity effect in older adults' memorability judgments of pictures. *Experimental Aging Research*. 2013; 39:254–274. DOI: 10.1080/0361073X.2013.779178 [PubMed: 23607397]
- Toth JP, Daniels KA, Solinger LA. What you know can hurt you: Effects of age and prior knowledge on the accuracy of judgments of learning. *Psychology and Aging*. 2011; 26:919–931. DOI: 10.1037/a0023379 [PubMed: 21480715]
- Undorf M, Erdfelder E. Judgments of learning reflect encoding fluency: Conclusive evidence for the ease-of-processing hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2011; 37:1264–1269. DOI: 10.1037/a0023719
- Undorf M, Erdfelder E. Separation of encoding fluency and item difficulty effects on judgements of learning. *The Quarterly Journal of Experimental Psychology*. 2013; 66:2060–2072. DOI: 10.1080/17470218.2013.777751 [PubMed: 23574510]
- West RL, Bagwell DK, Dark-Freudeman A. Self-efficacy and memory aging: The impact of a memory intervention based on self-efficacy. *Aging, Neuropsychology, and Cognition*. 2008; 15:302–329. DOI: 10.1080/13825580701440510

Zimmerman CA, Kelley CM. I'll remember this! Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*. 2010; 62:240–253. DOI: 10.1016/j.jml.2009.11.004

Appendix

To evaluate recognition memory, the proportion of hits (correct indication that a picture had been studied) and false alarms (incorrect indication that a picture had been studied) were calculated per participant and for each picture type. Hits and false alarms were then used to calculate d' , which is a measure of item discriminability. Descriptive statistics are provided in Table A1.

For Experiment 1, a 2 (age group) $\times$ 3 (picture type) mixed-factor ANOVA revealed a main effect of picture type, $F(2, 92) = 29.1, p < .001, \eta_p^2 = .39$. Follow-up tests indicated that d' scores were significantly greater for low-arousal neutral pictures than for high-arousal positive pictures, $t(57) = 7.0, p < .001, d = .84$, and were significantly greater for low-arousal positive pictures than for high-arousal positive pictures, $t(61) = 5.9, p < .001, d = .57$. d' scores did not differ between low-arousal neutral and low-arousal positive pictures, $t(50) = 1.6, p = .12$. No other effects were significant.

For Experiment 2, a 2 (age group) $\times$ 2 (valence) $\times$ 2 (arousal) mixed-factor ANOVA revealed that no effects were significant, $F_s < 2.5$.

Table A1

Mean hits, false alarms, and d' scores for each picture type in Experiment 1 and Experiment 2.

	Hits				False alarms				d'					Overall
	Positive		Neutral		Positive		Neutral		Positive		Neutral			
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low		
Experiment 1														
Young	.78	.80	–	.83	.18	.12	–	.13	1.9	2.2	–	2.3	2.1	
Older	.73	.80	–	.84	.20	.13	–	.11	1.6	2.2	–	2.4	2.0	
Experiment 2														
Young	.72	.72	.70	.72	.35	.35	.32	.35	1.1	1.0	1.1	1.2	1.0	
Older	.66	.71	.67	.71	.34	.34	.30	.35	.88	1.0	1.0	1.0	1.0	

High: high-arousal; low: low-arousal. For hits and false alarms, standard errors ranged from .01 to .03. For d' , standard errors ranged from .1 to .2. High-arousal neutral pictures were not used in Experiment 1.

Figure 1. Mean judgments of learning as a function of picture type for both age groups in Experiment 1. Error bars are standard errors of the mean, which were calculated using within-participant error (Loftus & Masson, 1994) and separately for each age group.

Figure 2. Mean judgments of learning as a function of picture type for both age groups in Experiment 2. Error bars are standard errors of the mean, which were calculated using within-participant error (Loftus & Masson, 1994) and separately for each age group.

Table 1

Mean normative ratings of valence and arousal for the pictures used in experiments 1 and 2.

	Positive		Neutral	
	High-arousal	Low-arousal	High-arousal	Low-arousal
Arousal	3.5 (.07)	2.7 (.05)	3.4 (.07)	2.8 (.09)
Valence	6.5 (.08)	6.6 (.09)	5.5 (.07)	5.6 (.10)

Means were calculated by averaging normative ratings from younger and older adults (Grühn & Scheibe, 2008). All ratings were made on a nine-point self-assessment manikin (SAM) scale. Higher arousal ratings are indicative of higher arousal, and higher valence ratings are indicative of positive emotion. Standard errors are in parentheses. Inferential statistics provided in Table 2. High-arousal neutral pictures were not used in Experiment 1.

Table 2

Inferential statistics for the normative ratings of the picture stimuli used in both experiments.

Comparison		Arousal	Valence
High-arousal positive	Low-arousal positive	$t = 9.6, d = 1.8^*$	$t = 1.3, p = .20, d = .24$
	Low-arousal neutral	$t = 6.2, d = 1.1^*$	$t = 6.3, d = 1.1^*$
	High-arousal neutral	$t = 1.0, p = .30, d = .19$	$t = 8.8, d = 1.6^*$
Low-arousal positive	Low-arousal neutral	$t = 1.2, p = .25, d = .21$	$t = 7.3, d = 1.3^*$
	High-arousal neutral	$t = 8.4, d = 1.5^*$	$t = 9.8, d = 1.8^*$
High-arousal neutral	Low-arousal neutral	$t = 5.3, d = .97^*$	$t = 1.2, p = .25, d = .21$

For every analysis, $df = 118$. Means and standard errors for each picture type are provided in Table 1. High-arousal neutral pictures were not used in Experiment 1.

* $p < .001$.

Table 3

Mean percent recalled in Experiment 1 and Experiment 2.

	Positive		Neutral	
	High-arousal	Low-arousal	High-arousal	Low-arousal
Experiment 1				
Younger	24.3 (1.9) ^a	19.8 (1.7) ^a	-	15.8 (1.7)
Older	17.7 (2.0) ^a	18.7 (1.8) ^a	-	12.6 (1.8)
Experiment 2				
Younger	12.4 (1.7) ^{ab}	13.6 (1.6) ^{ab}	7.8 (1.2)	7.9 (1.0)
Older	17.8 (1.3) ^{ab}	16.1 (1.6) ^{ab}	8.5 (1.2)	11.0 (1.4)

Standard errors in parentheses. High-arousal neutral pictures were not used in Experiment 1.

^aSignificantly greater than recall of low-arousal neutral pictures, $p < .001$.

^bSignificantly greater than recall of high-arousal neutral pictures, $p < .001$.