Emotion effects on memory from childhood through adulthood: Consistent enhancement and adult gender differences

Anais F Stenson, Emory University
Jacqueline S Leventon, Towson University
Patricia Bauer, Emory University

Journal Title: Journal of Experimental Child Psychology
Volume: Volume 178
Publisher: Elsevier | 2019-02-01, Pages 121-136
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1016/j.jecp.2018.09.016
Permanent URL: https://pid.emory.edu/ark:/25593/vk0wm

Final published version: http://dx.doi.org/10.1016/j.jecp.2018.09.016

Copyright information:
© 2018 Elsevier Inc.

This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Accessed June 8, 2024 10:36 AM EDT
Emotion effects on memory from childhood through adulthood: Consistent enhancement and adult gender differences

Anaïs F. Stenson\textsuperscript{a}, Jacqueline S. Leventon\textsuperscript{b}, and Patricia J. Bauer\textsuperscript{a}

\textsuperscript{a}Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30222 USA

\textsuperscript{b}Department of Psychology, Towson University, 8000 York Road, Towson, MD 21252 USA

Abstract

Emotion typically enhances memory. This ‘canonical’ emotional memory enhancement (EME) effect has been extensively studied in adults, but its developmental trajectory is unclear. The handful of developmental studies that have manipulated emotion at encoding and then tested subsequent memory have yielded mixed results. To identify whether development change in EME occurs across middle childhood, adolescence, and early adulthood, we examined EME in 206 8- to 30-year-olds, using the same stimuli, paradigm, and analyses for or all participants. At encoding, participants saw negative, neutral, and positive pictures while completing an incidental task. Two weeks later, participants completed a recognition memory test. We calculated negative-neutral and positive-neutral memory difference scores for each participant and then tested whether EME were predicted by age gender. Negative pictures were remembered better than neutral; the magnitude of this difference diminished in older males, but not older females. Positive pictures were also remembered better than neutral, but this EME effect was small and did not change significantly with age or by gender. We also examined whether subjective ratings of stimulus emotion changed with age or between genders, and report small differences. These results suggest that emotion effects on recognition memory are apparent by middle childhood and remain consistent across adolescence and early adulthood for females, whereas for males emotion elicitation and EME effects diminish slightly with age. These findings enrich both the EME literature specifically, and, more generally, what is known about emotion-cognition interactions across middle childhood, adolescence, and early adulthood.

Keywords

memory; development; emotion; gender; middle childhood; adolescence

Corresponding author: Anaïs F. Stenson, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, 49 Jesse Hill Jr. Dr. SE, Atlanta, GA 30303-3049, afstens@emory.edu, Tel. +1 (404) 712-0252.

Anaïs F. Stenson is now at the Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine.

Publisher’s Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Introduction

Emotion has powerful effects on memory (see, e.g., Banich et al., 2009; Kensinger & Schacter, 2008; Roozendaal & McGaugh, 2011, for reviews). Although emotion is ubiquitous throughout development, there are gaps in the account of how emotion impacts memory across the lifespan (see, e.g., Carver, 2014; Hamann & Stevens, 2014, for reviews). Many adult studies systematically manipulate emotion during memory encoding and then test memory performance; typically, memory is enhanced for emotional versus neutral events (see, e.g., Kensinger & Schacter, 2008; Talmi, 2013, for reviews). These studies have helped identify cognitive and neural mechanisms that underpin this “canonical” emotional memory enhancement (EME) effect. In contrast, most developmental studies have examined how children remember naturally occurring emotional events. Few developmental EME studies parallel the approach used to study EME in adults, and thus there have been limited opportunities to examine the canonical emotion effect across development. There are thus two gaps between the developmental and adult accounts of EME. First, studies with adults have utilized laboratory-based paradigms to probe how emotion impacts memory, and this approach has not been widely used across development. Second, the few studies that have used this approach to examine EME in children have produced varied results. Thus, it is not yet clear if this canonical EME effect is consistent across development. The present study addressed this gap by examining EME effects on recognition memory across middle childhood, adolescence, and early adulthood—a window of immense physiological, cognitive, and social-emotional changes (e.g., Berenbaum, Beltz, & Corley, 2015; Casey, 2015; Kilford, Garrett, & Blakemore, 2016).

Studies of the Emotional Memory Enhancement (EME) Effect

Studies of EME in adults typically assess how emotion at encoding impacts subsequent memory by presenting emotional and neutral stimuli and then comparing subsequent memory for those stimuli. These studies often use stimuli from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) or the Affective Norms for English Words (ANEW; Bradley & Lang, 1999) to elicit emotion (e.g., Kensinger & Corkin, 2003; Kensinger & Schacter, 2006). These large, standardized stimulus sets facilitate direct comparisons of EME effects both within- and between-subjects, as well as across studies. In contrast, developmental studies have primarily examined how emotion modulates children’s recollection of personally relevant events, and their results indicate that emotional events are recalled in greater detail than neutral events (e.g., Ackil, Van Abbema, & Bauer, 2003; Bauer et al., 2017; Fivush, Hazzard, McDermott Sales, Sarfati, & Brown, 2003; reviewed by, e.g., Goodman, Quas, & Ogle, 2010). These studies provide insight into how emotion impacts children’s ability to recall events from their own lives. However, because they examine retrospective reports of personal experiences, their results are not directly comparable with those from the adult EME literature. Moreover, the nature of these studies precludes both systematic manipulation of emotion at the time of memory encoding and direct comparison of emotion effects between subjects.

A few developmental studies have systematically manipulated emotion at encoding and then tested subsequent memory in a manner that parallels adult EME studies. Their results are
mixed. Cordon, Melinder, Goodman, and Edelstein (2013) showed 7- to 9-year-olds and young adults negative and neutral pictures and conducted a recognition memory test one week later. Adults remembered more neutral pictures than children, but memory performance for negative images was equivalent for both age groups. Leventon, Stevens, and Bauer (2014) showed 5- to 8-year-olds negative, neutral, and positive IAPS pictures and tested recognition memory 24 hours later. Children’s behavioral memory performance was statistically equivalent for all emotion conditions across the age range; among 7- to 8-year-olds only, there was evidence of an emotion effect in neural (event-related potential) response to the stimuli. Vasa and colleagues (2012) showed 12- to 17-year-olds and adults negative, neutral, and positive IAPS pictures. Approximately 30 minutes after encoding, participants completed a surprise recall task. EME effects were robust for both age groups. These studies each tested different age ranges and used different paradigms; therefore, it is difficult to offer a robust interpretation of their mixed results.

Potential Sources of Age-related Change in Emotion Processing

There are several potential sources of age-related change in emotion experience and processing across middle childhood, adolescence, and early adulthood (see, e.g., Del Piero, Saxbe, & Margolin, 2016, for a review). First, as children transition into adolescence, they report experiencing more negative emotion during daily life, relative to childhood (e.g., Larson & Lampman-Petratis, 1989; Larson, Moneta, Richards, & Wilson, 2002). Second, changes in emotional reactivity also may contribute to developmental differences, though there are mixed findings regarding changes during adolescence. For instance, Silk and colleagues (2009) observed greater autonomic response, measured via pupil dilation, to emotional versus neutral words in 8- to 17-year-old children whose pubertal development was more advanced than the less-developed children.¹ In contrast, Silvers and colleagues (2012) found that self-reported emotional reactivity to negative and neutral pictures was consistent in participants aged 10–23 years. Third, between early adolescence and adulthood there are changes in both emotion regulation strategies and efficacy (Zimmermann & Iwanski, 2014). This is particularly relevant, as prior research indicates that when children regulate their emotional responses EME effects are diminished (Leventon & Bauer, 2016). Collectively, these results point to developmental changes in the processing of emotional information across middle childhood, adolescence, and early adulthood that might impact EME effects.

The development of brain structure, function, and connectivity also may contribute to changes in processing of emotional information. Neural structures that support emotion processing, including the amygdala, prefrontal cortex, and hippocampus, continue to develop through adolescence (e.g., Goddings et al., 2014; Gogtay et al., 2004; Raznahan et al., 2014; see, e.g., Blakemore, 2012; Casey, Getz, & Galvan, 2008, for reviews). These structures play a central role in EME (reviewed by, e.g., Carver, 2014). As such, their relative developmental immaturity may contribute to age-related differences in EME effects. For example, functional connectivity between the prefrontal cortex and subcortical regions,

---

¹The difference in emotional reactivity as a function of pubertal status remained significant when age was included as a covariate; however, this does not fully disentangle the impact of pubertal status versus older age.
including the amygdala and hippocampus, continues to develop through middle childhood (e.g., Gabard-Durnam et al., 2014) and, in response to emotional stimuli, across adolescence (Guye et al., 2008; Vink, Derks, Hoogendam, Hillegers, & Kahn, 2014). Measures of neural activity during presentation of emotional pictures suggest that both the time course (measured with event-related potentials; e.g., Hajcak & Dennis, 2009) and the relative activation of different brain regions (measured with fMRI; e.g., Monk et al., 2003; Vasa et al., 2012) change between middle childhood and adulthood. This ongoing development of the brain structures that are understood to support EME might be reflected in developmental change in EME effects across middle childhood, adolescence, and early adulthood.

Both adult EME studies and developmental studies of emotion processing suggest that there could be gender differences in EME across development. For instance, some adult studies report amplified EME for females versus males (e.g., Canli, Desmond, Zhao, & Gabrieli, 2002), whereas other studies report no differences (e.g., Spalek et al., 2015). In a sample of 7- to 10-year-olds, Bauer, Stevens, Jackson, and San Souci (2012) observed gender differences in neural activity during emotional autobiographical memory recall. However, because the study examined autobiographical memories, there was not a direct comparison of relative EME effects between genders. In contrast, Cordon and colleagues (2013) documented equivalent emotion ratings and recognition memory for emotional pictures between genders in both children and young adults, but did not measure neural processes. Findings of developmental gender differences in some emotion processing tasks and in adult EME studies also indicate that gender should be also considered in any examination of EME across development. For instance, girls’ emotion recognition performance and impulse control reaches adult-like levels before boys’ (Lawrence et al., 2015; Shulman, Harden, Chein, & Steinberg, 2014). In addition, during adolescence there are gender-specific physiological changes, such as increased levels of circulating sex hormones, which could impact memory performance (for reviews see, e.g., Andreano & Cahill, 2009; Berenbaum, Beltz, & Corley, 2015). Collectively, these findings highlight the need for additional research into the prospect of gender differences in the developmental trajectory of EME.

**Present Study**

To elucidate the developmental trajectory of EME, studies that use consistent stimuli, paradigms, and analyses across a wide age range are required. In addition, because performance on some memory tasks continues to develop through adolescence, it is necessary to examine the impact of emotion on memory using a task that elicits similar performance across the range of participant ages. This approach would enable direct comparison of emotion effects on performance across development and between genders. In the present research, we initiated this effort by examining EME effects on recognition memory from middle childhood to adulthood. Recognition memory performance approaches adult-like levels in middle childhood or early adolescence when the items for which memory is tested are intact and easy to recognize (e.g., Cycowicz, Friedman, Snodgrass, & Duff, 2001; Ghetti & Angelini, 2008; Mandler & Robinson, 1978; but see Ofen, Chai, Schuil, Whitfield-Gabrieli, & Gabrieli, 2012; also see e.g., Bauer, 2007, for a review). We reasoned that because age-related differences in episodic memory for intact neutral stimuli are minimized in recognition paradigms (but see Cordon et al., 2013), observed developmental
differences in EME effects could reasonably be attributed to shifts in emotion processing. Thus, testing EME with a recognition memory paradigm is ideal for examining whether there are developmental changes or gender differences in EME across middle childhood, adolescence, and early adulthood.

In the present research, we examined EME of recognition memory from middle childhood, through adolescence, and into early adulthood (ages 8–30 years). The lower age bound was set because including participants younger than 8 years would have further restricted the range of stimuli that could be used to elicit emotional responses with participants of all ages. To facilitate comparison of performance across this age range, the same stimuli, paradigm, and analyses were used for all participants. This study design allows for direct comparison of EME effects on recognition performance across a wide age range, while testing for other potential sources of developmental differences, such as differential emotion elicitation. As some prior studies have documented adult gender differences in EME effects, and our participants’ age range spans a period of substantial gender-specific changes, we also conducted an exploratory analysis of gender differences in outcome measures. Thus, we sought to determine whether EME effects on recognition memory change between middle childhood and early adulthood, if gender was related to EME, and if any observed developmental changes would be gender-specific.

**Method**

**Participants**

Data were collected at a large private Southeastern university. Altogether, 151 children (75 females), ages 8 to 16, and 88 young adults (45 females), ages 18 to 30, enrolled in the study ($N = 239$). Twelve children (five females) and four adults (one female) were lost to attrition between Sessions 1 and 2. Sixteen children (five females) and five adults (four females) were excluded from analyses due to technical errors, failure to perform the task, or experimenter error. Ultimately, 127 children (65 females) and 79 adults (40 females) were included in the study (final $N = 206$). Children were recruited from a database of families that had previously expressed interest in study participation through the university’s Child Study Center. Although detailed data on socioeconomic status were not collected, the pool is comprised primarily of families from educated middle- to upper-middle-class SES. Information about highest education was not systematically collected from adult participants, but most were either undergraduate or graduate students at the University. Self- or parent-reported race and ethnicity was collected for all participants. Ethnicity was reported for 200 participants; 12 identified as Hispanic or Latino and 188 identified as not Hispanic or Latino. Race was reported for 202 participants; 26 identified as Asian, 37 identified as Black or African-American, 14 identified as bi- or multi-racial, and 125 identified as White or Caucasian. Prior to testing the children, their guardians provided written informed consent. The children received a gift card to a major retail chain for their participation. Adult participants were recruited through the university psychology subject pool. They provided written informed consent and received course credit for their participation. All procedures were reviewed and approved by the university IRB.
Materials and Procedure

A set of 165 child-appropriate pictures (57 negative, 53 neutral, and 55 positive) was selected from the International Affective Picture System (IAPS; Lang et al. 2008) and a similar lab-collected set of pictures. Of these 165 pictures, 150 pictures (50 per valence) were included in the memory task. This set of 150 pictures was used to create eight presentation orders that were used for both child and adult participants. However, before child participants came to the lab, thumbnail images of these 150 pictures were emailed to the guardian for approval (procedure approved by Lang, personal communication). If the guardian requested that specific pictures be removed, they were replaced with alternate pictures of the same valence (maximum number of replaced images was five); there were no requests to replace neutral or positive images. The seven negative pictures that were not used in the presentation orders were used to replace any picture(s) that the guardian wanted removed. Five positive pictures were presented at the end of picture presentation so that the session ended on a positive note. Three neutral images were used for the practice trials. To control for previously reported biases in affective processing of stimuli with humans (Proverbio, Adorni, Zani, & Trestianu, 2009), within each emotion condition, half of the images included humans and half did not.

The study consisted of two sessions separated by approximately 14 days (M = 14.1 (0.91), Range = 11–20). During Session 1, participants viewed pictures and engaged in a behavioral task to ensure attention to the pictures. During Session 2, participants viewed all pictures from Session 1, along with new pictures, and completed a recognition memory task. At the end of Session 2, participants provided subjective ratings of valence and arousal for a subset of the pictures (45) using the modified Self-Assessment Manikin (SAM; see Figure 1) a widely used ratings system that has previously found high correlations (> .88) in ratings from adults, children, and adolescents (Bradley & Lang, 1994; McManis, Bradley, Berg, Cuthbert, & Lang, 2001).

Session 1.—During the encoding task, participants viewed pictures and used a game controller to indicate whether each picture contained a human. All pictures were presented in full color at 30.5 cm (h) × 23 cm (w) in size, subtending a visual angle of approximately 15.6° (h) × 20.6° (w). Each picture was presented for 3000ms and immediately followed by a decision screen lasting for 3000ms that prompted participants to indicate whether the prior picture contained a human. A 850 to 1250ms inter-stimulus interval followed. Stimulus presentation was controlled using Advanced Neuro Technology evokeTM software. A trial schematic is presented in Figure 2, Panel A. The encoding task lasted approximately 9min of the approximately 90min Session 1.

After orientation to the encoding task, participants completed three practice trials to ensure they understood the task. They then viewed 90 pictures (30 negative, 30 neutral, and 30 positive) presented in one of eight pseudo-randomized orders. No more than two images of the same valence were presented consecutively. The images and presentation order were counterbalanced across participants. Five positive pictures were shown at the end of each

2Of the 165 pictures, 79 (48%) were from the IAPS and 86 (52%) were from the supplemental set. Exactly half of the pictures used in the memory task were from the IAPS.
presentation so that the session ended on a positive note. These final positive trials were not included in analyses.

**Session 2.**—During the recognition task, participants viewed 150 pictures (50 negative, 50 neutral, and 50 positive). These included the 90 pictures from Session 1 and 60 new pictures (20 negative, 20 neutral, and 20 positive), which the participant had not seen before. Like Session 1, the order of picture presentation was pseudo-randomized so that no more than two images of the same valence were presented consecutively. Both picture sets and presentation orders were counterbalanced so that across participants, all pictures were used equally in the old and new conditions. Five positive images were added to the end of each presentation so that the session ended on a positive note. These trials were not included in any analyses.

After each picture was presented, participants indicated whether they thought it was ‘definitely old,’ ‘maybe old,’ or ‘new’ via a button press on a game controller. The position of the response options was counterbalanced across participants. Participants completed three practice trials (same images as Session 1 practice trials) to ensure that they understood the task. Once the participants affirmed they understood the instructions, picture presentation and data recording began. Each picture was presented for 3000ms, followed by the old/maybe old/new decision screen for 3000ms. Trials were separated by a variable 850–1250ms inter-stimulus interval. A trial schematic is presented in Figure 2, Panel B. Stimulus presentation was controlled using Advanced Neuro Technology eevokéTM software. The task lasted approximately 18 minutes for all participants.

Following the recognition task, participants provided subjective valence and arousal ratings for 45 of the pictures (15 from each emotion condition) using the modified SAM scales, shown in Figure 1. The SAM scales were abbreviated from the 9-point versions used with adults to 5-point versions that reduced participant burden, particularly for the children (Leventon et al., 2014). Both scales were illustrated with a cartoon character that depicted different feelings; participants were instructed that they should select the illustration on the scale that was most similar to how the pictures made them feel. For valence ratings, participants selected from illustrations depicting a range of states from very happy or positive (5) to very bad or unpleasant (1). For arousal ratings, participants selected from illustrations depicting a range of states from very relaxed, calm, and not aroused (1) to highly alert and aroused (5). Participants selected one valence and one arousal rating for each picture.

**Analytic Approach**

All analyses were conducted with R version 3.2.2 (R Core Team, 2015). Raw memory performance data (mean number of hit, miss, false alarm, and correct reject trials) are provided in Table 1, Panel A. Following prior research, we evaluated participants’ discrimination between old and new images by calculating $d'$, a discriminability index, and $C$, an index of response bias (e.g., Banks, 1970; Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988; Wixted, 2007). In order to conduct this analysis, ‘maybe old’ and ‘definitely old’ responses were combined and both treated as ‘old.’ We calculated $d'$ by first calculating...
z scores for hit (HR) and false alarm (FA) rates, and then subtracting z(FA) from z(HR). Thus, for each participant \(d' = z(\text{HR}) - z(\text{FA})\). This \(d'\) value indexes how well participants distinguished old items from new items. We calculated \(C\) using \(z(\text{FA})\) and \(z(\text{HR})\) in the following formula: \(C = -0.5 \times (z(\text{HR}) + z(\text{FA}))\). The resulting \(C\) value indexes the tendency to respond \(old\) rather than \(new\). Hit and FA rates of 1 or 0 were corrected in accordance with Macmillan and Kaplan’s (1985) recommendation: rates of 0 were replaced with 0.5 / \(n\) and rates of 1 were replaced with \((n - 0.5) / n\), where \(n\) is the total possible number of hits or false alarms.

We first tested whether overall \(d'\) and \(C\) bias were significantly different than chance. We then used repeated-measures ANOVAs to examine overall \(d'\) and \(C\) bias as a function of emotion condition. Next, we assessed the difference in \(d'\) and \(C\) between the following conditions: a) negative and neutral, and b) positive and neutral. To do this, we calculated difference scores for each emotion condition relative to neutral (i.e., negative \(d'\) – neutral \(d'\), negative \(C\) – neutral \(C\), positive \(d'\) – neutral \(d'\), and positive \(C\) – neutral \(C\)). We then examined the effects of age and gender on the negative-neutral difference and positive-neutral difference by constructing separate linear models for each of these four difference scores. The approach allowed us to treat age as a continuous variable, which maximized statistical power and prevented overlooking differences that might emerge gradually across development.

We first evaluated the results of the SAM ratings to check that the pictures elicited the predicted emotional responses using repeated-measures ANOVAs with the factor emotion. For both valence and arousal ratings, we calculated within-subjects difference scores for each emotion condition relative to neutral (i.e., \(M_{\text{Negative Valence}} - M_{\text{Neutral Valence}}\), \(M_{\text{Negative Arousal}} - M_{\text{Neutral Arousal}}\), \(M_{\text{Positive Valence}} - M_{\text{Neutral Valence}}\), and \(M_{\text{Positive Arousal}} - M_{\text{Neutral Arousal}}\)). These scores were then used as the dependent variables in multiple regression analyses with the predictors age and gender to determine if there were age- and/or gender-related differences in emotion elicitation that might impact EME effects.

Results

Recognition Memory Performance

Participants’ memory performance was significantly greater than chance (\(d' = 0\)) for all pictures, regardless of emotion, (\(M_{d'} = 1.58, SD = 0.55\), \(t(205) = 40.94, p < .0001\). Overall \(C\) was also significantly greater than 0 (\(M_C = 0.26, SD = 0.43\), \(t(205) = 8.55, p < .0001\), indicating that participants were more likely to respond \(old\) than \(new\). This bias likely reflects the relative proportions of old (2/3) and new (1/3) pictures presented during the memory test. Memory performance was significantly above chance for each emotion condition (see Table 1).

We evaluated the effects of emotion on \(d'\) and \(C\) by conducting separate one-way ANOVAs with the factor emotion (negative, neutral, or positive). For cases that did not satisfy the assumption of sphericity, Greenhouse-Geisser corrections were applied. For all post-hoc analyses with multiple comparisons, Holm-Bonferroni corrections were applied to the \(p\)-values. There were significant differences in both \(d'\), \(R(2, 615) = 45.22, p < .0001, \eta^2 = .13\),
and \( C, R(2, 615) = 20.33, p < .0001, \eta^2_p = .06 \), between emotion conditions. Participants’ memory (\( d' \)) was best for negative pictures, followed by positive and then neutral pictures, all \( p s < .02 \). Participants’ response bias (\( C \)) was lowest for negative and positive pictures, which did not differ statistically, \( p = .54 \), and significantly higher for neutral versus negative or positive pictures, both \( p s < .001 \). See Table 1 for descriptive statistics.

We assessed whether EME effects changed with age by deriving negative-neutral and positive-neutral difference scores for both \( d' \) and \( C \). A positive difference score for \( d' \) indicates greater accuracy for the emotional (negative or positive) versus neutral pictures. Conversely, greater accuracy for neutral versus emotional pictures would produce a negative \( d' \) difference score value. We conducted separate multiple regression analyses for the negative-neutral and positive-neutral \( d' \) and \( C \) difference scores, entering age and gender as predictors in the models. Neither age, \( p = .81 \), gender, \( p = .46 \), nor their interaction, \( p = .86 \), predicted negative-neutral memory bias (\( C \)) differences, total \( R^2 = .004, R(3, 202) = 0.29, p = .83, 95\% \text{ confidence interval } [\text{CI}] [-0.01, 0.02] \). The magnitude of memory enhancement (\( d' \)) for negative versus neutral pictures was not predicted by age, \( p = .92 \), or gender, \( p = .33 \), however, their interaction approached statistical significance, \( p = .06 \), and the model explained a small but significant proportion of the variance, total \( R^2 = .04, R(3, 202) = 2.86, p = .04, 95\% \text{ CI } [-0.01, 0.09] \). Therefore, we conducted an exploratory analysis in which participants were stratified by gender and then linear regressions were used to assess gender differences in the relation between age and the negative-neutral \( d' \) difference.

This exploratory analysis of gender differences indicated that, for females, age did not significantly predict negative-neutral EME, \( p = .93 \)—memory for negative pictures was enhanced relative to neutral across the tested age range. In contrast, for males, the magnitude of the negative-neutral EME effect significantly diminished with age, \( \beta = -.29(SE = .09), p = .003, 95\% \text{ CI } [-0.44, -0.09] \). As shown in Figure 3, Panel A, the regression line slopes differed by gender, yet their 95\% confidence intervals overlapped across the age range. Thus, both genders demonstrated enhanced memory for negative versus neutral pictures across the age range, but the magnitude of this enhancement shifted with age only for males.

We again used multiple regression analyses to evaluate the effects of age and gender on the positive-neutral difference in \( d' \) and \( C \) values. Neither age, \( p = .83 \), gender, \( p = .90 \), nor their interaction, \( p = .14 \), explained a significant proportion of the variance in the positive-neutral \( d' \) difference, total \( R^2 = .02, R(3, 202) = 1.76, p = .16, 95\% \text{ CI } [-0.02, 0.07] \). As shown in Figure 3, Panel B, memory was slightly better for positive versus neutral pictures across the age range for females and males alike. The positive-neutral \( C \) difference also was not moderated by age, \( p = .31 \), gender, \( p = .33 \), or their interaction, \( p = .95 \). Together, these predictors did not account for a significant proportion of the variance in \( C \), total \( R^2 = .001, R(3, 202) = 1.04, p = .38, 95\% \text{ CI } [-0.01, 0.01] \). These results indicate that the memory advantage for positive relative to neutral stimuli did not change significantly across the age range or between genders, and that there were not age or gender differences in participants’ response bias for positive versus neutral pictures.
Subjective Ratings of Picture Valence and Arousal

To evaluate the success of the emotion manipulation for participants of all ages and of both genders, participants (n = 204) provided subjective ratings of picture valence and arousal using the SAM. Two participants did not provide complete ratings due to time constraints. Summary statistics for the ratings are provided in Table 2.

Valence ratings differed significantly according to picture emotion, $R^2(2, 609) = 1193, p < .0001, \eta^2 = .79$. Participants rated negative pictures as more unpleasant than either neutral or positive pictures, and positive pictures as more pleasant than either neutral or negative pictures, all $p < .001$. Arousal ratings also differed significantly according to picture emotion, $R^2(2, 609) = 183.3, p < .0001, \eta^2 = .38$. Negative pictures were rated as more arousing than both positive and neutral pictures, both $p < .001$. Positive pictures were rated as more arousing than neutral, $p < .001$. All statistical comparisons are reported in Table 3.

The average difference in valence ratings for negative versus neutral pictures was larger for females than males, $\beta = .26(SE = .07), p < .001, 95\% \text{ CI [0.13, 0.41]}$ but was not predicted by age, $p = .75$, or the interaction of age and gender, $p = .95$, total $R^2 = .07, R(3,200) = 5.15, p = .002, 95\% \text{ CI [0.00, 0.14]}$. On average, females rated the negative pictures as more aversive than neutral pictures ($M_{\text{Difference}} = -1.26, SD = .46)$, relative to males ($M_{\text{Difference}} = -0.98, SD = .54$). This gender difference was consistent across the age range. The difference between arousal ratings for negative and neutral pictures was significantly predicted by age, $\beta = .25(.07), p = .01, 95\% \text{ CI [0.05, 0.35]}$, gender, $\beta = -.14(.11), p = .04, 95\% \text{ CI [−0.44, −0.01]}$, and their interaction, $\beta = -.29(.11), p = .04, 95\% \text{ CI [−0.44, 0.02]}$, total $R^2 = .06, R(3,200) = 03.91, p = .01, 95\% \text{ CI [0.00, 0.11]}$. Females rated the negative pictures as more arousing than neutral pictures ($M_{\text{Difference}} = 1.49, SD = .81$), relative to males ($M_{\text{Difference}} = 1.26, SD = .74$), and the average negative-neutral difference increased slightly with age for females, but not for males.

The difference in valence ratings for positive versus neutral pictures was predicted by gender, $\beta = -.19(SE = .05), p = .005, 95\% \text{ CI [−0.25, −0.04]}$, but not age, $p = .39$, or their interaction, $p = .15$, total $R^2 = .05, R(3,200) = 3.33, p = .02, 95\% \text{ CI [0.00, 0.10]}$. Females’ ratings indicated that they perceived the positive pictures to be more pleasant relative to the neutral pictures ($M_{\text{Difference}} = 0.95, SD = .38$), than males did ($M_{\text{Difference}} = 0.80, SD = 0.37$) across the age range. In contrast, the difference in arousal ratings for positive versus neutral pictures was not predicted by gender, age, or their interaction, total $R^2 = .01, R(3,200) = 0.85, p = .47, 95\% \text{ CI [−0.02, 0.04]}$. Participants across the age range reported similar subjective assessment of the valence and arousal of positive versus neutral pictures. There was a significant gender difference in participants’ valence ratings, but not in their arousal ratings.

---

3To see if the decrease in males’ arousal and valence ratings mediated their age-related decrease in negative EME we tested for significant relations between age and the hypothesized mediators (negative-neutral valence and arousal ratings difference scores). Neither relation was significant, $p > .69$, indicating that the necessary condition for mediation was not met (see, e.g., Baron & Kenny, 1986).
Discussion

Emotion has powerful effects on memory. Although these effects have been studied extensively, particularly in adults, there are gaps in our understanding of how emotion impacts memory across development (reviewed by, e.g., Carver, 2014; Hamann & Stevens, 2014; Kensinger & Schacter, 2008; Roozendaal & McGaugh, 2011). We addressed this gap by examining EME effects across middle childhood, adolescence, and early adulthood, a developmental window in which individuals undergo major physiological, cognitive, and social-emotional changes (for reviews see, e.g., Berenbaum, Beltz, & Corley, 2015; Casey, 2015; Kilford, Garrett, & Blakemore, 2016; Steinberg & Morris, 2001). In contrast, recognition memory performance is largely consistent across this age range (e.g., Cycowicz et al., 2001; Ghetti & Angelini, 2008). Accordingly, we utilized a recognition memory task in order to isolate the effects of emotion on memory performance across the tested age range. To facilitate comparisons with prior research, we used both a common paradigm from adult EME studies and a set of developmentally appropriate IAPS and IAPS-like pictures. Specifically, 8- to 30-year-old participants saw negative, neutral, and positive pictures during an encoding session and then completed a recognition memory test two weeks later. This design enabled us to compare memory performance between emotion conditions within-subjects and treat age as a continuous variable, thereby affording greater power to detect differences in EME linked to age and gender.

Emotional Memory Enhancement Effects

In keeping with prior research, we found a robust main effect of emotion on memory performance: participants’ memory for negative and positive pictures was significantly better than for neutral pictures. Overall, the negative-neutral memory difference was larger than the positive-neutral difference; this result also mirrors prior research (reviewed by, e.g., Kensinger & Schacter, 2008). For both negative and positive pictures, EME relative to neutral was largely consistent from middle childhood through early adulthood. However, there were some modest age- and gender-specific changes in the magnitude of the EME effect. Figure 3, Panel B shows that although there were not statistically significant age differences in the positive-neutral memory advantage, this small EME effect diminishes between middle childhood and early adulthood. Figure 3, Panel A shows that the negative-neutral EME effect also decreased slightly across the age range, yet the effect remained robust even in the oldest participants. However, results of our exploratory analysis suggested that the negative-neutral memory advantage diminished with age for males, whereas for females it remained stable. This gender-specific age effect was small but significant; as shown in Figure 3, Panel A, the 95% confidence intervals around the regression lines for females and males overlapped substantially. The exploratory nature of this finding means it should be interpreted cautiously. By comparison, the main effect of emotion on memory—particularly for negative pictures—was robust across participants and replicated the canonical EME effect.

These findings extend the EME and developmental memory literatures by providing evidence that emotion enhances recognition memory from middle childhood onwards. Given our large sample size and within-subjects analysis of EME, these results provide compelling
evidence that, at least as tested in the present research, EME effects on recognition memory are largely consistent from middle childhood through early adulthood for both genders. The exception to this account is the small but significant age-related decrease in negative-neutral EME in males, but not females. There are at least two likely explanations for this gender difference. First, the impact of negative emotion on memory might diminish with age for males. Second, the negative pictures might have elicited a muted emotional response in older males, relative to females and younger males. For reasons discussed next, we favor the second explanation.

Participants’ subjective valence and arousal ratings of the pictures differed between genders and, for negative pictures, across the age range. Females rated negative pictures as more aversive and arousing, relative to neutral, than males did. For females, there was also a slight age-related increase in the negative-neutral arousal rating difference. Females also rated positive pictures as more pleasant, relative to neutral, than did males; arousal ratings were not significantly different between genders. There were not age-related changes in the positive-neutral valence or arousal ratings. Prior research has also documented gender differences in valence and arousal ratings of emotional pictures; specifically, using the SAM, females rate both negative and positive IAPS pictures as more intensely arousing and extreme in valence than males do (McManis et al., 2001). We observed a similar pattern of gender differences in SAM ratings in the present study. In addition, we report that the gender difference increases with age for arousal ratings of negative pictures, relative to neutral pictures. This suggests that diminished negative-neutral memory enhancement in males versus females likely reflects that the negative pictures did not elicit the same degree of emotional response across the age range in both genders. In addition, recent, well-powered adult studies report equivalent emotion effects on recognition memory performance for females and males when the stimuli are selected to be emotionally arousing for adults (e.g., Spalek et al., 2015). Thus, we hypothesize that the age-related decrease in EME for negative versus neutral pictures for males reflects the fact that our developmentally-appropriate pictures elicited somewhat less emotional arousal in males versus females; this account is corroborated by the picture ratings, which showed both that females rated the negative pictures as more aversive than males and that this gender difference increased across the age range.

The present findings generally align with Vasa and colleagues (2012) finding of consistent EME for both adults and adolescents. However, results from other studies of EME in middle childhood have varied. Cordon and colleagues (2013) reported that 7- to 9-year-olds and adults had equivalent memory for aversive stimuli, but adults also remembered more neutral items than the children. Thus, the EME effect was larger for children than for adults. Leventon and colleagues (2014) did not find evidence of EME effects on recognition memory for 5- to 8-year-olds. This result merits special consideration, as the present study and Leventon and colleagues (2014) employed similar stimuli and laboratory environments. However, there was little overlap in participants’ age: the top of the Leventon and colleagues (2014) range was 8 years, which was the bottom of our range. This raises the possibility that EME effects on recognition memory emerge during middle childhood, and do not appear in younger children.
The difference between Leventon and colleagues (2014) findings and the present results could reflect a transition to adult-like EME effects that occurs in middle childhood, around 8 years of age. Indeed, Leventon and colleagues found age-specific changes in the neural response to previously viewed (‘old’) versus new negative, but not neutral or positive, pictures. Specifically, for five- to 7.5-year-old participants, there were not significant differences between the event-related potentials (ERPs) triggered by old and new pictures, regardless of emotion. In contrast, there were significant differences between the ERPs elicited by old versus new negative, but not neutral or positive, pictures in participants older than 7.5 years—even though there were not significant differences in their behavioral memory performance for emotional and neutral pictures. These age differences in the neural response to emotional stimuli could foreshadow the emergence of behavioral EME effects in middle childhood.

An additional difference between Leventon and colleagues (2014) and the present study was the delay between memory encoding and test: 24 hours versus 2 weeks, respectively. The duration of study-test delay is critical for EME, because emotion is believed to impact memory in two ways: attention mediation and preferential consolidation. The former is a relatively fast-acting process, whereas the latter requires time for consolidation to occur before its effects are apparent (e.g., LaBar & Cabeza, 2006; Talmi & McGarry, 2012; Talmi, 2013). It is possible that in the age range tested by Leventon and colleagues (2014) the effects of emotion on attention were not adult-like, and the 24-hour delay was not long enough for preferential consolidation of the emotional stimuli to unfold. Thus, one explanation for the mixed results from studies of EME in children could be that the effects of emotion on attention during encoding are different from adults, and that preferential consolidation of emotional information requires longer periods (i.e., >24 hours) to unfold in children. These possibilities await future empirical tests.

We tested EME in a manner that should eliminate most sources of developmental difference: encoding was incidental, mnemonic strategies were not necessary, and participants did not need to report details about their memory or provide judgments of memory strength. This contribution is important, but leaves open questions about the development of EME. There remains a need to examine developmental change in EME for other types of memory, including recollection and recall. Ideally, future studies of EME across development will use multiple measures of emotion elicitation. Specifically, behavioral outcome measures can be bolstered with measures of physiological response and overt attention that directly index emotion processing and attention modulation, respectively (e.g., Leventon et al., 2014). Finally, there is a need for developmental studies that systematically probe the mechanisms thought to generate EME effects, to determine whether the contributions of these mechanisms to EME effects are developmentally continuous. For instance, the relative magnitude of emotion effects on attention could be compared between children and adults.

Limitations and Directions for Future Research

Our results should be interpreted in the context of three limitations of the present study. First, we tested recognition memory for IAPS pictures. Whereas IAPS provide a well-controlled stimulus set, it is unlikely that they elicit emotion either of the same kind or
magnitude as that experienced in contexts outside the laboratory. This limitation could be addressed in future studies by using dynamic stimuli, such as film clips or story passages. Second, our stimuli were not appropriate for, and our paradigm was too demanding to include, participants younger than 8 years old. Ideally, future studies of EME across development will be designed so younger participants can be included, as it is plausible that age differences in EME effects on recognition memory occur prior to middle childhood (e.g., Leventon et al., 2014). Conversely, using pictures that are appropriate for children as young as 8 years old precluded the use of many pictures that adults consider to be more intensely emotionally arousing, such as erotica and violent imagery. Thus, it is possible that some of the age-related change we observed in EME in males reflects an attenuated emotional response in the older participants. Third, we utilized a cross-sectional design that cannot address whether EME effects change within an individual as they develop. In short, the present work should be complemented by studies that utilize a variety of stimuli and paradigms, examine multiple measures of emotion processing, include younger participants, and examine whether EME effects shift within individuals across development.

These results contribute to the small literature that addresses basic emotion processing and emotion effects on memory across development. Our results provide evidence of remarkable consistency in EME effects on recognition memory between middle childhood and adulthood. This provides a baseline from which changes in other types of emotion processing and emotion-cognition interactions can be compared. In particular, identifying the developmental trajectory of emotion processing, and its effects on memory, can augment interpretation of the many recent studies that report adolescent-specific changes in performance on tasks, such as risk and reward decision making paradigms, that often elicit emotional responses (for reviews see, e.g., Del Piero, Saxbe, & Margolin, 2016; Heller & Casey, 2015).

**Conclusion**

In conclusion, the findings of the present research extend and enrich the EME literature. First, we conducted a direct comparison of EME effects for both positive and negative stimuli across a wide range of ages, using the same stimulus set, paradigm, and methodology. Second, our large sample allowed us to evaluate both age and gender differences in EME, as well as possible interactions between these factors. Third, we identified gender differences in developmental trajectory of EME that emerge in early adulthood. We found that EME is present and robust for both genders in middle childhood, but that by early adulthood the magnitude of this enhancement remains stable for females, whereas it decreases for males. We propose that future work should assess the extent to which this finding might have resulted from age and/or gender differences in how effectively our stimuli elicited emotion, in order to determine whether there are in fact gender differences in the mechanisms of EME that emerge across development. Our results provide compelling evidence that emotion consistently enhances recognition memory from middle childhood through adulthood for both genders, but indicate that there could be small but significant gender differences in the developmental trajectory of EME. This result is striking: in spite of the vast cognitive, social, and physiological change that unfolds between
middle childhood and adulthood, EME effects on recognition memory remain largely consistent throughout it all.

Acknowledgments

This research was supported in part by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development awarded to Patricia J. Bauer.

References


Highlights:

- Emotion enhances recognition memory in children, adolescents, and adults.
- For females, emotion consistently enhanced memory in participants aged 8–30 years.
- For males aged 8–30, emotional memory enhancement diminished with age.
- Gender differences in emotion effects on memory may emerge in adulthood.
Table 1. 

<table>
<thead>
<tr>
<th>Valence</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arousal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

**Figure 1.**
Self-Assessment Manikin (SAM) rating scale.
Figure 2.
Trial Schematics.
Panel A. Difference in memory performance ($d'$) for negative versus neutral pictures.

Panel B. Difference in memory performance ($d'$) for positive versus neutral pictures.

Figure 3.
Magnitude of emotional memory enhancement for males and females aged 8 to 30. Shaded areas show 95% confidence intervals around the regression lines.
Table 1.

The number of hits, misses, false, alarms, and correct rejections (rows 1–4), as well as the hit rate, false alarm rate, d’ and C values (rows 5–8) for the three different emotion conditions (N = 206). Standard deviations are presented in parentheses. One-sample t-tests confirmed that memory performance (d’) was significantly better than chance (d’ = 0) for all emotion conditions, as indicated by an asterisk.

<table>
<thead>
<tr>
<th></th>
<th>Negative</th>
<th>Neutral</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hits</td>
<td>22.8(4.6)</td>
<td>17.5(5.8)</td>
<td>20.5(5.2)</td>
</tr>
<tr>
<td>2 Misses</td>
<td>6.8(4.4)</td>
<td>12.1(5.6)</td>
<td>9.1(5.0)</td>
</tr>
<tr>
<td>3 False alarms</td>
<td>3.3(2.9)</td>
<td>3.4(3.1)</td>
<td>4.2(3.1)</td>
</tr>
<tr>
<td>4 Correct rejections</td>
<td>16.4(3.0)</td>
<td>16.3(3.2)</td>
<td>15.5(3.2)</td>
</tr>
<tr>
<td>5 Hit rate</td>
<td>.76(.15)</td>
<td>.58(.19)</td>
<td>.68(.17)</td>
</tr>
<tr>
<td>6 FA rate</td>
<td>.17(.14)</td>
<td>.18(.15)</td>
<td>.21(.15)</td>
</tr>
<tr>
<td>7 d’</td>
<td>1.92(0.68)*</td>
<td>1.33(0.65)*</td>
<td>1.48(0.65)*</td>
</tr>
<tr>
<td>8 C</td>
<td>0.16(0.46)</td>
<td>0.43(0.50)</td>
<td>0.18(0.47)</td>
</tr>
</tbody>
</table>

Table 2.
Self-assessment manikin (SAM) ratings of picture valence and arousal (n = 204).

<table>
<thead>
<tr>
<th>Rating</th>
<th>Picture Valence</th>
<th>Mean(SD)</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence Negative</td>
<td>2.0(0.5)</td>
<td>1.0</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>3.1(0.3)</td>
<td>1.9</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>4.0(0.4)</td>
<td>3.1</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>Arousal Negative</td>
<td>3.1(0.7)</td>
<td>1.3</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>1.8(0.8)</td>
<td>1.0</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>2.8(0.8)</td>
<td>1.0</td>
<td>4.8</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.
Results of Welch’s two sample t-tests comparing participants’ SAM ratings of picture valence and arousal by emotion.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Comparison</th>
<th>Mean Difference</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>Negative-neutral</td>
<td>−1.1</td>
<td>−26.9</td>
<td>344.6</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td></td>
<td>Positive-neutral</td>
<td>0.9</td>
<td>24.4</td>
<td>389.1</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td></td>
<td>Negative-positive</td>
<td>−2.0</td>
<td>−44.5</td>
<td>384.9</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Arousal</td>
<td>Negative-neutral</td>
<td>1.4</td>
<td>19.6</td>
<td>402.0</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td></td>
<td>Positive-neutral</td>
<td>1.0</td>
<td>13.4</td>
<td>389.7</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td></td>
<td>Negative-positive</td>
<td>0.4</td>
<td>4.9</td>
<td>401.4</td>
<td>&lt; .0001</td>
</tr>
</tbody>
</table>