

Aging and Emotion Recognition: Not Just a Losing Matter

Jocelyn A. Sze

San Francisco VA Medical Center, San Francisco, California

Madeleine S. Goodkind

VA Palo Alto Health Care System, Palo Alto, California and
Stanford University School of Medicine

Anett Gyurak

Stanford University and Stanford University School of Medicine

Robert W. Levenson

University of California, Berkeley

Past studies on emotion recognition and aging have found evidence of age-related decline when emotion recognition was assessed by having participants detect single emotions depicted in static images of full or partial (e.g., eye region) faces. These tests afford good experimental control but do not capture the dynamic nature of real-world emotion recognition, which is often characterized by continuous emotional judgments and dynamic multimodal stimuli. Research suggests that older adults often perform better under conditions that better mimic real-world social contexts. We assessed emotion recognition in young, middle-aged, and older adults using two traditional methods (single emotion judgments of static images of faces and eyes) and an additional method in which participants made continuous emotion judgments of dynamic, multimodal stimuli (videotaped interactions between young, middle-aged, and older couples). Results revealed an Age \times Test interaction. Largely consistent with prior research, we found some evidence that older adults performed worse than young adults when judging single emotions from images of faces (for sad and disgust faces only) and eyes (for older eyes only), with middle-aged adults falling in between. In contrast, older adults did better than young adults on the test involving continuous emotion judgments of dyadic interactions, with middle-aged adults falling in between. In tests in which target stimuli differed in age, emotion recognition was not facilitated by an age match between participant and target. These findings are discussed in terms of theoretical and methodological implications for the study of aging and emotional processing.

Keywords: aging, emotion recognition, empathic accuracy, social interaction

Emotion recognition is critical to social functioning. Our capacity to recognize the emotions of others helps us interpret and predict people's actions, experience shared feelings, and interact effectively. Difficulties in this domain are associated with poor interpersonal functioning and communication, reduced social interest, and lower life satisfaction (Csarton, Kessler, & Pape, 1999; Ciarrochi, Chan, & Caputi, 2000; Feldman, Philippot, & Custrini, 1991; Shimokawa et al., 2001). Given the considerable importance

that older adults afford to social and emotionally meaningful goals (Carstensen, Fung, & Charles, 2003; Carstensen, Isaacowitz, & Charles, 1999; Lawton, 2001; Magai, 2008) and their heightened susceptibility to negative physical and mental health consequences of social isolation and loneliness (Bath & Deeg, 2005), emotion recognition may be among the most important ingredients for successful aging.

Emotion Recognition and Aging: Traditional Tests

A large literature exists that assesses how aging influences emotion recognition ability. In these studies, emotion recognition has almost always been assessed by instructing participants to identify the single emotion depicted in static, posed, socially decontextualized, single-modality stimuli (e.g., a photograph of an emotional facial expression). In studies of this sort, participants typically view stimuli for 10 to 15 s or at a self-paced rate and are asked to choose the emotion depicted from a list of emotion labels.

Although there have been several exceptions (Calder et al., 2003; Williams et al., 2006), studies using these methods have generally found evidence for age-related declines in emotion recognition. A meta-analysis of 28 such studies, incorporating results from 705 older and 962 young adults, found that older adults exhibited mild reductions in recognition of at least some emotions (anger, disgust, fear, happiness, sadness, surprise) in each of four modalities (faces, voices, bodies/contexts, and matching of faces to

This article was published Online First July 23, 2012.

Jocelyn A. Sze, Department of Veterans Affairs, San Francisco VA Medical Center, San Francisco, California; Madeleine S. Goodkind, Sierra-Pacific Mental Illness Research, Education, and Clinical Center, VA Palo Alto Health Care System, Palo Alto, California, and Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine; Anett Gyurak, Department of Psychology, Stanford University and Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine; Robert W. Levenson, Department of Psychology, University of California, Berkeley.

This research was supported by a grant from the National Institute of Aging (R37-AG017766).

Correspondence concerning this article should be addressed to Robert W. Levenson, Institute of Personality and Social Research, Department of Psychology, University of California, 4143 Tolman Hall MC 5050, Berkeley, CA 94720-5050. E-mail: boblev@socrates.berkeley.edu

voices), with certain negative emotions (sadness and anger) proving most difficult (Ruffman, Henry, Livingstone, & Phillips, 2008). Falling outside of the purview of this meta-analysis are several additional studies (Phillips, MacLean, & Allen, 2002; Slessor, Phillips, & Bull, 2007) that show similar age-related declines in performance on the Reading the Mind in the Eyes Test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), in which respondents infer mental states from static images of the eye region of faces.¹

The emotion recognition tests used in these studies typically require individuals to process a single channel of sensory information (e.g., visual or vocal) and match that information to an internal template for what different emotions look like or sound like. These kinds of match-to-prototype tasks are similar to other cognitive tasks for which age-related performance declines are common (Salthouse, 1991). Current models account for these age-related declines in terms of age-related neurodegeneration, particularly in frontal and temporal brain regions (Bartzokis et al., 2001; Calder et al., 2003; Isaacowitz et al., 2007; Phillips et al., 2002; Raz et al., 2005; Ruffman et al., 2008).

Evidence for Preserved Social Functioning in Late Life

Extrapolating from the findings of age-related declines in emotion recognition reviewed thus far, and considering the centrality of emotion recognition for social functioning, we might expect older individuals to be profoundly impaired in the social domain. But this has generally not been observed in the everyday functioning of healthy older adults (Salthouse, 1990). Consistent with this, we have found that older individuals evidence reduced potential for conflict and higher levels of satisfaction in intimate relationships compared with middle-aged individuals (Levenson, Carstensen, & Gottman, 1993), which suggests older individuals have intact skills in this important realm of social functioning.

Such contradictions call into question whether the laboratory methods traditionally used for assessing emotion recognition are ecologically valid, especially when used with older adults. Contextual (Berg & Sternberg, 1985; Dixon, 1992) and life-span theories of human intelligence (Baltes, Dittmann-Kohli, & Dixon, 1984) argue that older adults' abilities must be considered in the context of adults' knowledge of the pragmatics of everyday real-world situations and not just in terms of highly controlled, decontextualized laboratory tests.

Emotion Recognition and Aging: More Ecologically Valid Tests

Recognizing the emotions of others in everyday life is a complex process. Emotions are often short-lived, can change rapidly (Ekman, 1992), are manifest in multiple modalities (face, voice, posture), and occur in social contexts. For these reasons, real-world emotion recognition involves many subprocesses, including tracking multiple actors, tracking multiple channels of information, considering contextual moderators, and continuously updating judgments based on changing information. In addition, because emotional information competes with many other kinds of information in real-world contexts, the motivation to attend to and track emotion in others is quite important (Ickes, Gesn, & Graham, 2000).

The typical laboratory methods for assessing emotion described earlier afford excellent experimental control but do not capture many of these processes that are critical for real-world emotion recognition. In contrast, we, and others (Ickes, 1993; Levenson & Ruef, 1992), have argued that ecologically valid measures of emotion recognition should assess the ability to track emotions as they unfold spontaneously in dynamically changing, socially embedded contexts, using multimodal emotional information.

There is reason to hypothesize that older individuals would do relatively well on emotion recognition tasks that mirror real-world processing. This would particularly be the case when the recognition task relied on processes where age-related declines are not common and where experience and acquired stores of social knowledge are important. Moreover, motivation to attend to social and emotional information is thought to increase with age (Carstensen et al., 1999; Fung & Carstensen, 2003), thus creating optimal motivational conditions for engaging in these more demanding emotional recognition tasks.

Some support for these predictions can be derived from existing research that involves more complicated social judgments. For example, we found that older adults were better than young adults in judging levels of marital satisfaction, based on observing thin slices (3 min of videotaped conversation) of marital behavior (Ebling & Levenson, 2003). Similarly, older adults have been found to use strategies that lead to more effective appraisal and more satisfying choices of action on emotionally salient interpersonal tasks compared with young adults (Blanchard-Fields, 1986) and to demonstrate increases in contextual, relativistic thinking, particularly regarding familiar real-world situations (Cohen, 2006; Howard, Howard, Dennis, Yankovich, & Vaidya, 2004). Aging is thought to convey a greater ability to see the self and others from a dynamic perspective in which individuals are perceived in terms of changing experiences within the context of multiple frames (e.g., social, psychological; Labouvie-Vief, 1998).

These theories of social and emotional aging and empirical findings led us to hypothesize that the age-related declines in emotion recognition that have been found when using traditional tests would be reversed if emotion recognition was tested in more ecologically valid ways that tap into processes that are well matched to the interests, abilities, and accumulated knowledge base of older individuals.

Expanding on Existing Research

The present study revisited the issue of age differences in emotion recognition using a design that enabled us both to replicate aspects of existing research on emotion recognition and aging, and to expand it to address several important gaps.

Including Traditional Tests of Emotion Recognition

We thought it was important to include several traditional tests of emotion recognition using single judgments and static stimuli. A

¹ The Eyes Test has also been referred to as a test of theory of mind, or, the ability to infer others' mental states and understand that others have mental states different from one's own (Baron-Cohen, 1991; Blair, 2005). In this study, we are describing it broadly as an emotion recognition test, in that the goal of the test is to identify mental states from representations of emotional expressions.

replication of age-related deficits in emotion recognition with these more traditional tests would help establish that our participant sample was similar to those used in previous studies, which would be particularly important if our participants showed a different pattern of age differences on the dyadic test.

Including a New Test of Emotion Recognition That Assesses Continuous Emotion Judgments Based on Dynamic, Socially Embedded, Multimodal, Emotional Information

The primary goal of our study was to test age differences in emotion recognition using a test that closely resembles the way emotion recognition typically occurs in the real world. Thus, as noted earlier, we sought an assessment in which emotional judgments would be made and continuously updated based on dynamically changing, socially embedded, multimodal information. To accomplish this, we adapted a method we had used previously to study empathic accuracy (Levenson & Ruef, 1992). In this test, participants view a videotaped recording of an unrehearsed interaction between a married couple and use a rating dial to provide continuously updated ratings of the valence of emotion being experienced by either the husband or the wife. Valence represents one of the most fundamental ways that emotional judgments are made (Osgood, Suci, & Tannenbaum, 1957), serves as a major organizer of emotion in dimensional theories (Feldman Barrett & Russell, 1999; Watson, Wiese, Vaidya, & Tellegen, 1999), and captures a great deal of the variance in self-reported emotional response (for a review, see Mauss & Robinson, 2009). The viability of obtaining real-time ratings of emotional valence has been validated in prior studies of empathic accuracy (Ickes, Stinson, Bissonnette, & Garcia, 1990; Levenson & Ruef, 1992).

This test allowed us to assess emotion recognition using information that was embedded in a dynamic, social context. Most human emotions occur in social contexts (Scherer, Matsumoto, Wallbott, & Kudoh, 1988), and research indicates that even small amounts of dynamic and contextualized social information (Ambady, Hallahan, & Conner, 1999; Wehrle, Kaiser, Schmidt, & Scherer, 2000) can improve judgment accuracy. The test also enabled us to provide multimodal information from visual (e.g., facial expression) and auditory (e.g., prosody) channels. Research indicates that providing multimodal information facilitates behavioral reactions to emotional stimuli (Collignon et al., 2008; de Gelder & Vroomen, 2000).

Including Middle-Aged Participants

Although there are exceptions (Isaacowitz et al., 2007; Malatesta, Izard, Culver, & Nicolich, 1987), most studies examining emotion recognition and aging have examined only two age groups: young and old. Theories of life-span development often predict linear changes across adulthood; however, nonlinear change is also possible. Either way, inclusion of middle-aged participants is critical for elucidating age-related trends.

Controlling for Age of Targets

Surprisingly, few studies of aging and emotion recognition have considered the match or mismatch of the age of participants and

the age of people depicted in target stimuli. Typically, studies have used only young and middle-aged adult targets, and have not examined possible interactions between age of participants and age of targets. It is possible that age similarity could increase emotion recognition familiarity or by conveying an “in-group” advantage (Elfenbein & Ambady, 2002) in understanding emotional signals and contexts. In one study that considered this issue, similarity between participants and targets was associated with greater accuracy (Malatesta et al., 1987). However, two subsequent studies found no evidence of an own-age advantage in decoding posed facial expressions (Ebner, He, & Johnson, 2011; Ebner & Johnson, 2009).

Hypotheses

We had two hypotheses regarding age and emotion recognition: (a) there would be age-related deficits in making single-emotion judgments using traditional static, nonsocial, visual images of emotional displays, and (b) there would be age-related increases in emotion recognition using our novel and more ecologically valid test requiring continuous tracking of emotional valence based on dynamic, social, multimodality information. Our rationale for the first hypothesis was based on existing literature indicating that older participants do not fare as well in these tests. Our rationale for the second hypothesis is that this new test makes use of processes that remain strong and information that builds with age. Even if mild age-related neurodegeneration produces deficits that might hamper performance on traditional tests of emotion recognition, older adults should be able to compensate for such deficits and even surpass young adults in real-world social contexts due to a variety of age-related advantages, including (a) increased performance on a variety of socioemotional tasks involving judgments and appraisals within social contexts (Blanchard-Fields, 1986; Ebling & Levenson, 2003; Fung & Carstensen, 2003; Rahhal, Colcombe, & Hasher, 2001); (b) increased contextual thinking within familiar, real-world situations (Cohen, 2006; Howard et al., 2004) and increased tendency to see the self and others from a dynamic, contextualized perspective (Labouvie-Vief, 1998); and (c) enhanced motivation to maximize the meaningfulness of social and emotional input (Carstensen & Lockenhoff, 2003). Finally, although there is some evidence supporting an in-group advantage when ages of raters and targets are matched (Malatesta et al., 1987), this evidence did not seem sufficient to support an a priori hypothesis regarding an in-group age advantage, especially for our dynamic rating test.

Method

Participants

A total of 223 participants were studied. Seventy-six young participants (age range, 20 to 30 years, $M = 22.99$, $SD = 2.62$), 73 middle-aged participants (age range, 40 to 50 years, $M = 44.54$, $SD = 2.90$), and 74 older participants (age range, 60 to 80 years, $M = 66.38$, $SD = 5.27$) were recruited using flyers and online postings in the local community and from a research participant database administered by the University of California, Berkeley. These age ranges were selected to encompass a large portion of adulthood (viz, 60 years), have each age group differ by approx-

imately a generation, and to have the older group extend beyond “young–old” age.² Participants had to be in good general health and be sufficiently mobile to travel to the laboratory. The recruitment was designed to ensure that gender and ethnicity were stratified evenly across the three age groups. In terms of gender, 66.4% of the participants were women and 33.6% were men. In terms of ethnicity, the sample was 67.7% percent Caucasian American, 12.1% Asian American, 7.6% African American, 3.6% Latino/a American, and 5.8% indicated “other” as their racial/ethnic background. As would be expected, the age groups differed in income (older and middle-aged participants reporting higher incomes than young participants) and education (older and middle-aged participants reporting higher education than young participants). Means, standard deviations, effect sizes, and comparisons among age groups for income and education are presented in Table 1.

Materials

Faces Test. This test was modeled on traditional tests of emotion recognition that use forced-choice emotion identification (labeling) of static facial stimuli (Calder et al., 2003; MacPherson, Phillips, & Della Sala, 2006; Moreno, Borod, Welkowitz, & Alpert, 1993; Phillips et al., 2002; Stanley & Blanchard-Fields, 2008). Participants were presented with five sequences of color photographs representing a specific emotion (anger, happiness, fear, disgust, sadness) at six increasing levels of intensity.³ After viewing each photograph, participants selected their answers from eight response choices, including three fillers: neutral, embarrassment, and proud.

We created these stimuli using a set of color photographs provided by Paul Ekman (they are referred to as the New Jersey Institute of Technology (NJIT) image set and can be made available to other researchers through Paul Ekman). The photographs were obtained as part of an ongoing project to develop a new set of high-quality visual stimuli depicting prototypical emotional facial expressions. For this project, a number of different individuals were photographed under carefully controlled conditions (e.g., lighting, angle) as they produced neutral facial expressions and prototypical expressions of different emotions (Ekman & Friesen, 1975). To ensure fidelity, each expression was coded with the Facial Action Coding System (FACS; Ekman & Friesen, 1978). Using high-resolution digital versions of these photographs, we created images of varying emotional intensity using morphing software (Abrosoft FantaMorph 3.7.1) to combine differing amounts of the neutral face with each emotional face. Each subsequent image represented 6.25% (1/16th) greater intensity of the previously presented emotional expression. The five emotion sequences were shown in two different counterbalanced orders across participants. The test was self-paced.

For each target emotion, two scores were computed: (a) the total number identified correctly, and (b) and the lowest intensity level for which the participant identified the correct emotional expression for two consecutive photographs. If the participant only identified the most intense photograph correctly, the score was calculated as a 6; if the participant did not identify two consecutive photographs correctly and did not identify the most intense photograph correctly, the score was calculated as a 7.⁴ A composite score was computed by reverse-scoring the latter variable, stan-

dardizing both variables and then taking the average, so that greater numbers of correct responses and recognition at lower intensity indicated greater ability in facial emotion recognition. Mean performance scores for positive (happy) and negative (anger, fear, disgust, and sadness) expressions were also created to allow examination of recognition ability by emotional valence.

Eyes Test. This test was also modeled on traditional tests of emotion recognition using single emotion judgments based on static, visual stimuli. Participants were presented with 12 photographs of eyes expressing different emotionally laden mental states taken from the Mind in the Eyes Test (Baron-Cohen et al., 2001). After viewing each photograph, participants selected which of a set of four words (e.g., “playful, comforting, irritated, bored”) best described the mental states expressed in the picture. From the 36 pictures in this test, we chose a subset of 12 pictures that were reliably rated as young, middle-aged, and older by a group of five independent young and middle-aged female raters. A stimulus was designated as young, middle-aged, or older if there was age agreement among at least four of the five raters. Three male and one female target were selected for each age group. Greater numbers of correct responses (i.e., higher scores) on the Eyes Test indicated greater ability.

Dyads Test. This test reflected our desire to test emotion recognition in a way that involved continuous tracking of emotion based on dynamic, interpersonal, multimodality stimuli. Participants viewed 12 audiovisual interactions of couples discussing important marital topics (each 3.75 min in length). For each trial, participants were asked to continuously rate how they thought the designated target person was feeling during the interaction using a rating dial (Levenson & Gottman, 1983; Ruef & Levenson, 2007) that consisted of a mechanical dial that moved over a 180-degree valence scale divided into nine divisions ranging from *very negative* (1), to *neutral* (5), to *very positive* (9). A computer sampled the dial position every 5 ms and averaged these readings into 1-s measurement periods using a program written by one of the authors (Robert W. Levenson). This same rating dial had been used previously by target spouses to provide continuous ratings of how they were feeling during the interaction as part of their original study protocol.

² Gerontologists often group older adults into different categories, including “young-old” (59 to 69 years old), “middle-old” (70 to 75 years old), and “old-old” (over 75 years old; Gildengers et al., 2002). To examine whether the larger age range used for older adults (i.e., 20 years versus 10 years) was associated with greater within-group variability, Levene’s Test for Equality of Variances was conducted. Results revealed no age differences in within-group variability for performance on the Dyads, Faces, and Eyes tests ($ps > .10$).

³ The most commonly used traditional test of emotion recognition involves posed facial expressions at “maximum intensity” levels of emotional expression only. However, because such methodology poses limitations in terms of ecological validity and ceiling effects (e.g., 100% hit ratio for maximum intensity happy faces across all groups in Montagne, Kessels, Frigerio, de Haan, & Perrett, 2005), some studies have sought to address these issues by using mid-intensity levels of emotional expression (e.g., 60% intensity levels in Stanley & Blanchard-Fields, 2008) or morphing paradigms in which multiple levels of intensity are presented (Calder et al., 2003; Sullivan & Ruffman, 2004). The present study used this latter paradigm.

⁴ On average, in a given emotion sequence, 16.3% of participants did not correctly identify two consecutive photographs or the most intense photograph.

Table 1
Group Means, Standard Deviations, and Pairwise Comparisons for Demographic Variables and Covariates

	Mean (<i>SD</i>)			Age effect		
	Young	Middle-aged	Older	<i>F</i>	<i>p</i> value	η_p^2
Income (1–8)	2.22 _a (2.04)	2.99 _b (1.90)	3.31 _b (1.82)	6.24	<.01	.06
Education (1–6)	3.41 _a (.88)	4.01 _b (.93)	4.37 _b (1.02)	19.42	<.01	.15
Trait perspective taking	3.70 (.69)	3.67 (.74)	3.62 (.74)	<1	.80	.00

Note. Within each row, different subscripts denote significantly different means at $p < .05$ (two-tailed).

To minimize participant fatigue, the 12 trials were organized in two blocks (six trials each) that were counterbalanced across participants. Four of the tapes depicted younger couples, four depicted middle-aged couples, and four depicted older couples (using the same age criteria as were used for participants). Within each block, the trials were ordered such that targets on adjacent trials were always of different ages. One of the members of the target couple was designated as the target person (within each age group, two of the targets were wives and two were husbands). To keep ethnicity constant, all targets were Caucasian American. Drawing from our extensive library of couples' interactions, the target conversations were selected using a procedure developed in previous research using this type of emotion recognition test (Levenson & Ruef, 1992; Soto & Levenson, 2009). This procedure ensured that the target spouse (a) experienced a sufficient variability and range of emotion (i.e., rated him or herself as feeling positive or negative for at least half the time), and (b) rated his or her own emotion in a way that was reasonable and not unduly idiosyncratic (determined by comparing target ratings with those of a panel of four expert raters).

As in our past research, for each trial, the target's own second-by-second ratings of the valence of his or her emotional experience provided the criterion for determining participants' rating scores. For every second during each trial, recognition accuracy was calculated as the absolute difference between the participant's rating and the target's own rating. Thus, a smaller absolute difference score indicated *higher* accuracy. When we used this approach in our previous research (Levenson & Ruef, 1992), we devised a method for considering accuracy separately for negative and positive emotional moments that was based solely on target's own valence ratings. For present purposes, we sought to improve on this by classifying each second of targets' emotion as positive, negative, or neutral using objective behavioral coding conducted by trained raters using the Specific Affect Coding System (SPAFF; Gottman, 1989), which was designed to code emotion during social interaction based on facial expression, voice tone and pitch, and verbal content. To classify each second of the interaction with a single code,⁵ SPAFF codes of affection, humor, interest, validation, joy were considered positive; codes of anger, contempt, disgust, belligerence, domineering, defensiveness, fear/tension/worry, sadness, whining were considered negative; and neutral codes were considered neutral.

Our rationale for using behavioral coding to classify emotional valence rather than target self-ratings was that behavioral coding would minimize variations among individual targets in terms of what constitutes a positive, negative, or neutral moment. Rather than relying on the somewhat arbitrary cutoffs in rating dial values

that we had used previously (Levenson & Ruef, 1992), SPAFF coding used both verbal and nonverbal information to identify emotional moments in ways that were consistent across targets. Importantly, this new method also reduced potential confounds by identifying emotional moments based on behavior, which was independent of the way that accuracy was calculated (i.e., based on participant and target rating-dial data).

Self-reported trait perspective taking. Trait perspective taking was assessed using the perspective-taking subscale of the Interpersonal Reactivity Index (Davis, 1980). Sample items included, "When I'm upset at someone, I usually try to 'put myself in his shoes' for a while." Internal consistencies were adequate for this scale ($\alpha = .80$).

General Procedure

At-home questionnaires. Three to 7 days prior to their laboratory visit, participants completed a questionnaire packet that included the Trait Perspective Taking scale.

Laboratory assessment. Participants were greeted by a female experimenter and seated in a chair in a 3 × 6-m experimental room. The experimental protocol (2 hr) consisted of a series of tests designed to assess emotional understanding. The following emotion recognition tests (as previously described) were administered in the following order: (a) Dyads Test Block 1 (six trials, counterbalanced with Block 2; approx. 45 min total); (b) Faces Test (five emotion trials, counterbalanced; approx. 14 min total); (c) Eyes Test (12 photographs; approx. 7 min total); and (d) Dyads Test Block 2 (six trials; approx. 45 min total). All stimuli were presented on a 21-inch LCD computer screen. The participant sat alone in the experimental room and the experimenter sat in an adjacent room where she could view the participant on a monitor and communicate over an intercom system.

Data Analysis

An initial series of analysis of variance analysis (ANOVA) tests were used to evaluate the effects of the counterbalanced orderings of trials within tests. These analyses revealed no significant main effects or interactions involving order for any of our dependent variables. Thus, we collapsed across order and conducted our

⁵ Means and standard deviations for the number of SPAFF behavioral emotion codes per 225-s dyadic interaction (with one code assigned per second) were as follows: for positive codes, $M = 29.25$, $SD = 21.70$, for negative codes, $M = 51.75$, $SD = 31.76$, and for neutral codes, $M = 144.00$, $SD = 29.60$.

primary analysis using a $3 \times 2 \times 3$ ANOVA, with participant age (young, middle-aged, older) and participant gender (male, female) treated as between-subjects factors, and test (Dyads [reverse-scored], Faces, Eyes) treated as a within-subject factor. When appropriate, we included additional within-subject factors for emotional valence (for the Dyads Test and Faces Test), and target age (for the Dyads Test and Eyes Test). When significant main effects or interactions involving participant age were found, we conducted polynomial trend analyses and tested whether a linear pattern captured the effect of age. Because these trend analyses capture overall patterns of difference across groups and not differences between particular groups, we also conducted Bonferroni-adjusted pairwise comparisons to identify differences between age groups. To test our hypotheses of different patterns of age-related differences in performance within the different tests, significant interactions involving participant age and test were followed up within each test separately. Because there were no main effects or interaction effects involving gender in this overall analysis, we present all analyses without stratifying by gender.⁶

Results

Correlations Among Emotion Recognition Tests

The Faces Test and the Eyes Test were only modestly correlated with each other ($r = .21, p < .01$). This is consistent with other research using these kinds of emotion recognition tests (Keightley, Winocur, Burianova, Hongwanishkul, & Grady, 2006; Phillips et al., 2002). The Dyads Test was not significantly correlated with either the Faces or the Eyes tests ($ps > .05$). This is consistent with our assertion that the Dyads Test captures a different aspect of emotion recognition than the more traditional tests.

Age Differences in Emotion Recognition

As predicted, there was a significant interaction between participant age and test, $F(2, 426) = 5.11, p < .01, \eta_p^2 = .05$. Thus, follow-up ANOVAs were conducted to examine age differences separately within each test.

Faces Test: Poorer performance with age. There was a significant main effect for participant age, $F(2, 220) = 4.15, p < .05, \eta_p^2 = .04$. The linear trend for participant age was significant, contrast estimate = $-.18, p < .01$. As hypothesized, and as consistent with the previous literature, young adults performed best, with middle-aged adults intermediary, and older adults performing worst. In addition, pairwise comparisons revealed that young adults performed significantly better than older adults ($p < .05$). The Age \times Emotional Valence interaction was not significant, $F(2, 220) < 1, \eta_p^2 = .00$, indicating that age effects did not differ as a function of positive and negative faces. Means, standard deviations, pairwise comparisons, and effect sizes are presented in Table 2.

Because prior studies have found age differences in the recognition of particular emotions (Ruffman et al., 2008), we conducted an additional Participant Age \times Specific Emotion (anger, disgust, fear, happiness, and sadness) repeated measures ANOVA, with specific emotion as a within-subjects factor. Results revealed a significant Participant Age \times Specific Emotion interaction, $F(8, 880) = 2.41, p < .01, \eta_p^2 = .03$. Follow-up univariate ANOVAs

within specific emotions revealed age differences in recognition of sadness, $F(2, 220) = 8.53, p < .01, \eta_p^2 = .07$, and disgust, $F(2, 220) = 6.30, p < .01, \eta_p^2 = .05$; for both emotions, young adults performed best, middle-aged adults intermediary, and older adults performed worst (contrast estimates = $-.45, -.38$, respectively, $ps < .01$).

Eyes Test: Poorer performance with age. There was a significant Participant Age \times Target Age interaction, $F(4, 440) = 4.10, p < .01, \eta_p^2 = .04$. This interaction was broken down in two ways. To test for age differences in performance, we conducted analyses separately within target age. This revealed that the participant age effect was not significant for young and middle-aged targets, $F(2, 222) = 2.01, p = .14, \eta_p^2 = .02$, and $F(2, 222) < 1, \eta_p^2 = .00$, respectively. However, participant age was significant for older targets, $F(2, 222) = 5.09, p < .01, \eta_p^2 = .04$. The linear relationship for participant age was significant, contrast estimate = $-.34, p < .01$, with young adults performing best, middle-aged intermediary, and older adults performing worst. In addition, pairwise comparisons revealed that young adults performed significantly better than older adults ($p < .01$) and marginally better than middle-aged adults ($p < .10$). While our findings of age-related declines on the Faces Test and Eyes Test are largely consistent with prior literature also finding age-related declines, the differences found in the present study are somewhat less extensive than what has been found in prior studies (age-related declines have also been found for recognition of anger and fear, see Ruffman et al., 2008, for a review). Means, standard deviations, and pairwise comparisons are presented in Table 3.

To test for age-matching effects, we conducted analyses separately within participant age. Rather than revealing an in-group advantage, results revealed an age "mismatch" effect in which young participants performed better at rating *older* targets versus younger targets, $t(75) = 2.84, p < .01$, older participants performed better at rating *young* targets versus older targets, $t(73) = -2.96, p < .01$, and middle-aged participants performed equally at rating older versus younger targets, $t(75) = 1.05, p = .30$. Means, standard deviations, and pairwise comparisons are presented in Table 3.

Dyads Test: Better performance with age. There was a significant main effect for participant age, $F(2, 212) = 3.23, p < .05, \eta_p^2 = .03$. The linear relationship for participant age was significant, contrast estimate = $.12, p < .01$. As hypothesized, young adults performed worst, middle-aged adults intermediary, and older adults performed best. In addition, pairwise comparisons revealed that older adults performed significantly better than young adults ($p < .05$). No interaction between participant age and emotional valence was found, $F(4, 848) = 1.87, p = .12, \eta_p^2 = .02$, indicating that age effects did not differ as a function of positive, neutral, and negative emotional moments (as identified by SPAFF behavioral coding). The Participant Age \times Target Age interaction was not significant, $F(4, 848) < 1, \eta_p^2 = .02$, indicating no age matching advantage or disadvantage on this test. Means, standard deviations, and pairwise comparisons are presented in Table 3.

⁶ The current lack of gender differences is largely consistent with prior literature showing few overall gender differences when emotion recognition is evaluated against objective criteria (Eisenberg & Lennon, 1983; Russell et al., 2007).

Table 2

Faces Test Scores: Summary of ANOVA Results and Group Means, Standard Deviations, and Pairwise Comparisons by Overall and Specific Response Emotion and Age Group

	Mean (SD)			Age effect		
	Young	Middle-aged	Older	F	p value	η_p^2
Overall positive and negative emotion All intensity levels (standardized)	.15 _a (.51)	-.02 (.48)	-.13 _b (.48)	6.58	<.01	.06
Anger	-.14 (1.07)	.08 (.94)	.08 (.91)	1.34	.26	.012
Disgust	.29 _a (.95)	-.06 (.98)	-.25 _b (.93)	6.30	<.01	.054
Fear	.19 (.95)	-.10 (.99)	-.11 (1.02)	2.17	.12	.019
Happiness	.12 (.99)	-.05 (.94)	-.07 (.98)	<1	.41	.01
Sadness	.31 _a (.92)	.02 (.96)	-.32 _b (.96)	8.53	<.01	.072
	Young	Middle-aged	Older	F	p value	Cramer's V
Max Intensity Only (% correct)						
Anger	75.0%	80.6%	77.3%	.66	.72	.05
Disgust	86.8% _a	68.1% _b	73.3% _b	7.73	.02	.19
Fear	84.2%	75.0%	73.3%	2.97	.13	.12
Happiness	77.6%	76.4%	70.7%	1.10	.58	.07
Sadness	88.2% _a	81.9%	68.0% _b	9.80	<.01	.21

Note. Within each row, different subscripts denote significantly different means at $p < .05$ (two-tailed).

Additional Analyses

High-intensity expressions. Some previous studies have examined emotion recognition only using high-intensity expressions. Thus, we conducted 3×2 chi-square analyses on each emotion to examine whether age differences emerged in the percentage of adults who gave correct answers on the highest intensity facial expressions only. These results largely mirrored the findings when expressions of all intensities were considered. There were significant age differences for sadness, $\chi^2(2, N = 223) = 9.80, p < .01$, and disgust, $\chi^2(2, N = 223) = 7.55, p < .05$. Follow-up 2×2 chi-square analyses revealed that, for sadness, more young adults answered correctly than older adults, $\chi^2(1, N = 148) = 7.55, p < .05$; and for disgust, more young adults answered correctly compared with both middle-aged and older adults, $\chi^2(1, N = 148) = 7.53$ and $4.33, ps < .01$ and $.05$, respectively. Means, standard deviations, percentages, pairwise comparisons, and effect sizes are presented in Table 2.

Table 3

The Eyes Test and Dyads Test Performance: Group Means, Standard Deviations, and Pairwise Comparisons by Participant Age Group

	Mean (SD)		
	Young	Middle-aged	Older
Eyes Test			
Young targets	2.85 (.96)	3.08 (.82)	3.17 (.96)
Middle-aged targets	2.50 (.95)	2.40 (.94)	2.50 (1.10)
Older targets	3.22 _a (.85)	2.92 (.96)	2.77 _b (.95)
Dyads Test (inaccuracy scores)			
All age targets	2.40 _a (.37)	2.29 (.35)	2.25 _b (.31)

Note. The Dyads Test scores refers to difference scores, where higher scores mean greater inaccuracy. Because there was no Participant Age \times Target Age interaction for the Dyads Test, means are presented collapsed across target age. Within each row, different subscripts denote significantly different means at $p < .05$ (two-tailed).

Judging emotional valence. A key difference between the Faces Test and the Dyads Test is that the former required participants to identify discrete emotions, whereas the latter required participants to identify moment-to-moment changes in emotional valence. In exploratory analyses, we attempted to make the tasks more equivalent by rescoring performance on the Faces Test in terms of correct valence rather than correct specific emotion. Results revealed the age-related declines that had been found for recognizing specific emotions on the Faces Test were not found when the task was scored in terms of valence, $F(2, 220) = 1.05, p > .05, \eta_p^2 = .01$.

Controlling for participant characteristics. We examined several participant characteristics that could have influenced emotion recognition: trait perspective taking, education, and income. Trait perspective taking has been associated with better emotion recognition in at least one prior study (Shamay-Tsoory, Harari, Szepeswol, & Levkovitz, 2009). Education and income were both related to age in our sample. Prior research on education and emotion recognition has yielded mixed results, with some studies finding a positive relation (Keightley et al., 2006) and others finding no relation (Orgeta & Phillips, 2008; Phillips et al., 2002). We are not aware of prior work on the association between income and emotion recognition.

We conducted zero-order correlations between these three participant characteristics and performance on each emotion recognition test. As Table 4 indicates, none of the variables were significantly associated with any emotion recognition test, thus ruling out these factors as alternative explanations for the age differences found in emotion recognition.

Discussion

The primary goal of the present study was to examine age differences in emotion recognition using a test that more closely resembles how emotions are perceived in the real world than those used in previous studies. Whereas previous studies had typically used tests that had participants provide single emotion judgments

Table 4
Correlations Between Participant Characteristics and Emotion Recognition Tests

Participant characteristics	Dyads Test performance	Faces Test performance	Eyes Test performance
Education	.13	-.07	-.06
Income	.02	.00	.03
Trait perspective taking	.05	.12	.09

Note. For ease of interpretation, the Dyads Test (i.e., inaccuracy) scores were reverse-scored (multiplied by -1) so that positive correlations would indicate that higher performance on one measure was associated with higher performance on another measure.

* $p < .05$ (two-tailed).

based on static, posed, nonsocial, single-modality stimuli, the present study used a novel test that required participants to provide continuous emotion ratings based on dynamic, socially embedded, multimodal, naturalistic stimuli (i.e., emotionally laden, unrehearsed, dyadic interactions between real married couples). By including two traditional tests of emotion recognition (i.e., single emotion judgments of photographs of faces and eyes), we were able to demonstrate that previous findings of age-related declines were largely replicated in our sample of participants, thus increasing confidence that findings of age-related improvement in emotion recognition using the Dyads Test truly reflected age and not peculiarities of our particular sample.

Our findings supported our two primary hypotheses. First, older adults showed evidence of poorer performance than young adults on the two traditional tests of emotion recognition based on viewing faces (for recognizing sadness and disgust) and eyes (for recognizing emotion in older eyes), with middle-aged adults falling in between. This finding replicates many previous studies using these kinds of rating tests and stimuli (Ruffman et al., 2008). It should be noted, however, that some previous studies using facial stimuli have also found age-related declines for other negative emotions that were not found in the present study (e.g., anger and fear in Calder et al., 2003). Second, and most importantly, older adults showed better performance than young adults on our novel, arguably more ecologically valid, Dyads Test, with middle-aged adults performing in between. Findings of age-related improvements in the socioemotional domain are not without precedent (e.g., estimating marital satisfaction based on thin slices of behavior, Ebling & Levenson, 2003; using positive reappraisal strategies to regulate emotion, Shiota & Levenson, 2009; greater prosocial responding to the distress of others, Sze, Gyurak, Goodkind, & Levenson, in press), but this, to our knowledge, is the first such finding using this kind of test of emotion recognition.

Despite hints in the literature of an age-based in-group advantage for emotion recognition, we found no evidence of this on any of our three tests. Rather, we found some evidence that age mismatch improved performance on the Eyes Test (for young and older participants only). Finally, we evaluated the contribution to our findings of several participant characteristics other than age. We found no support that differences in trait perspective taking, education, or income were related to emotion recognition.

Age Advantage on an Ecologically Valid Emotion Recognition Test

Results for the Dyads Test supported our hypothesis that emotion recognition ability, as assessed on this test, would increase with age. Compared with traditional tests of emotion recognition, the Dyads Test has a number of unique features, including (a) participants make ratings of emotional valence that must be continuously updated as emotions unfold, (b) judgments are based on emotional information that is dynamic and naturalistic, (c) emotional information is embedded within a social interaction, and (d) multiple modalities of emotion (i.e., visual and auditory) are presented.

We believe that older adults' abilities may be more optimally tuned to recognize emotion in the context of natural, real-world social situations compared to more static, controlled, decontextualized laboratory tests. In the former context, older adults can draw upon a number of advantages that are thought to occur with healthy aging. As we noted earlier, these include (a) increased performance on a variety of socioemotional tasks involving judgments and appraisals within social and emotional contexts (Blanchard-Fields, 1986; Ebling & Levenson, 2003; Fung & Carstensen, 2003; Rahhal et al., 2001); (b) enhanced ability to engage in contextual, relativistic thinking when considering familiar, real-world situations (Blanchard-Fields, 1986; Ebling & Levenson, 2003; Fung & Carstensen, 2003; Rahhal et al., 2001) and increased tendency to see the self and others from a dynamic, contextualized perspective (Cohen, 2006; Howard et al., 2004); (c) greater motivation in social and emotional domains (Carstensen & Lockenhoff, 2003); and (d) greater experience in social and emotional domains.

Although evaluated using other methods, there are other findings in the literature that are consistent with these, including studies showing that older adults evidence increased attunement to and understanding of intimate relationships and social/emotional contexts (Labouvie-Vief, 1998). Changes in the aging brain may also play an important role. Although mild age-related neurodegeneration in certain frontal and temporal brain regions has been thought to contribute to age-related deficits in traditional tests of emotion recognition (Bartzokis et al., 2001; Calder et al., 2003; Isaacowitz et al., 2007; Phillips et al., 2002; Raz et al., 2005; Ruffman et al., 2008), we have found that dynamic rating tests appear to rely on somewhat different brain regions (e.g., specific to the right lateral orbital frontal cortex in Goodkind et al., in press). Performance on more naturalistic tests may also benefit from other changes that occur in the aging brain, including more complex dendritic branching (Prickaerts, Koopmans, Blokland, & Scheepens, 2004; Segovia, del Arco, & Mora, 2009; Sun & Bartke, 2007) that is thought to reflect accumulated learning and experience over the life span, as well as compensatory changes in brain activity that have been documented in healthy older adults (e.g., a more bilateral pattern of frontal recruitment across right and left hemispheres, Cabeza, Anderson, Locantore, & McIntosh, 2002; Pujol et al., 2002).

Lack of Gender Differences and of Age-Matched Advantage

We found no evidence of differences associated with gender in terms of emotion recognition. While gender differences are reliably found on self-report measures of empathy (Eisenberg & Lennon, 1983), evidence for gender differences when assessing

emotion recognition ability against objectively measured criteria is mixed (for a review, see Hall et al., 1978). While some studies point to a female advantage when speed of recognition is emphasized (Hampson, van Anders, & Mullin, 2006) or when participants are primed with feelings of sympathy (Klein & Hodges, 2001), we, and others, have generally failed to find differences when these conditions are not the focus (Eisenberg & Lennon, 1983; Ickes et al., 1990; Levenson & Ruef, 1992; Russell, Tchanuria, Rahman, & Schmidt, 2007; Soto & Levenson, 2009).

We also found no support for an age-matched in-group advantage in emotion recognition in the Dyads Test. In prior studies using a version of the Dyads Test with individuals from different ethnic groups, we also failed to find an in-group advantage (Soto & Levenson, 2009). Of course, it may be that such advantages are subtle and require greater statistical power to detect. However, we should note that, in the Dyads Test in the present study, the effect size of the relevant Participant Age \times Target Age interaction was essentially zero ($\eta_p^2 = .00$). Finally, we found some evidence of an age-mismatch advantage in the Eyes Test. However, because this result was isolated, not hypothesized, and without much precedent in the literature, we do not feel it wise to offer further interpretation, pending replication. In replication studies, it would be especially important to include stimuli with more objective information about age ranges (we classified Eyes Test stimuli using ratings by young and middle-aged judges).

Limitations

There are several limitations in the present study worthy of note. First, the cross-sectional design makes it impossible to determine whether differences among the three age groups were truly related to aging or were due to cohort or survivorship effects. For example, members of our older cohort grew up during the post-WWII era, and their experiences with widespread suffering and distress might have had an impact on their attunement to social interactions. Second, our sample was drawn from a particular geographic region (the area around the University of California, Berkeley) and thus reflected the high levels of education and income that typify that region. Establishing the generalizability of findings involving the Dyads Test beyond this sample will be important, but our confidence in the generalizability of these findings was increased by finding that performance on the Faces Test and Eyes Test in this sample was similar to that found with other samples in the literature.

We also take note that our Dyads Test was designed to capture as many of the characteristics of real-world emotion recognition as possible. This reflected our desire to maximize ecological validity and to evaluate age differences in emotion recognition under such conditions. However, in doing so, the Dyads Test differed in a number of ways from the traditional Faces and Eyes tests. Having now found the hypothesized age-related advantage with the Dyads Test, additional studies will be needed to determine which factor or factors (i.e., continuously updated ratings, valence ratings, dynamic stimuli, social stimuli, multimodal stimuli) are responsible for the reversal of age effects between the two kinds of tests. Thus, for example, in exploratory analyses using a different scoring method, we found that the age-related declines in performance on the Faces Test were eliminated when answers were scored in terms of correct valence rather than correct emotion. However, there was

no evidence of age-related improvement with this alternative scoring. Thus, valence by itself is unlikely to account for the age-related improvement found in the Dyads Test, but it could be an important contributor. Studying age differences in emotion recognition in ways that enable isolating the contributions of particular subprocesses (valence vs. discrete emotion judgments, dynamic vs. static stimuli, single vs. multiple modality information, etc.) is clearly an important and potentially highly profitable area for follow-up studies.

Conclusion

As research on the psychological aspects of aging accumulates, it has become increasingly clear that different aspects of functioning evidence different trajectories of change. Traditionally, research on aging has focused on themes of loss: loss of physical health, loss of loved ones, and loss of cognitive abilities such as memory and executive functioning (Craik & Salthouse, 2007). More recently, a number of studies have documented areas of preserved functioning in the socioemotional realm among older adults and even evidence of areas in which levels of functioning increase with age, including some kinds of emotion regulation (Shiota & Levenson, 2009) and some aspects of emotional empathy and prosocial behavior (Sze et al., in press). The present study confirms prior research suggesting that older adults do not recognize single emotions in static photographic images as well as young adults. However, it adds yet another area of improved functioning, with older adults performing better than younger adults in a seemingly more complex task: continuously tracking the emotional valence of others who are engaged in dynamic, unrehearsed, social interactions. Moreover, the present study may help explain the seeming contradiction between impairments observed in laboratory tests and functioning in the everyday lives of healthy older adults (e.g., Salthouse, 1990). Our findings suggest that with more ecologically valid tests, performance measured in the laboratory can be more similar to that seen in the real world.

In terms of real-world implications, it is possible that older adults' deficits in recognizing discrete emotions from photographs represent a vulnerability that could lead to interpersonal misunderstandings and difficulties in everyday life. If so, they could potentially benefit from training that addresses these deficits. However, because most emotional information in everyday life occurs in dynamic, socially embedded, multimodal contexts, our findings suggest that this is an area of life in which older adults do quite well and might have much to teach others.

References

- Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*, 77, 538–547. doi:10.1037/0022-3514.77.3.538
- Baltes, P. B., Dittmann-Kohli, F., & Dixon, R. A. (1984). New perspectives on the development of intelligence in adulthood: Toward a dual-process conception and a model of selective optimization with compensation. *Life-span development and behavior*, 6, 33–76.
- Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. *Natural theories of mind*, 233–252.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning

- autism. *Journal of Child Psychology and Psychiatry*, 42, 241–251. doi:10.1111/1469-7610.00715
- Bartzokis, G., Beckson, M., Lu, P. H., Nuechterlein, K. H., Edwards, N., & Mintz, J. (2001). Age-related changes in frontal and temporal lobe volumes in men: A magnetic resonance imaging study. *Archives of General Psychiatry*, 58, 461–465.
- Bath, P. A., & Deeg, D. (2005). Social engagement and health outcomes among older people: Introduction to a special section. *European Journal of Ageing*, 2, 24–30. doi:10.1007/s10433-005-0019-4
- Berg, C. A., & Sternberg, R. J. (1985). A triarchic theory of intellectual development during adulthood. *Developmental Review*, 5, 334–370.
- Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition: An International Journal*, 14, 698–718. doi:10.1016/j.concog.2005.06.004
- Blanchard-Fields, F. (1986). Reasoning on social dilemmas varying in emotional saliency: An adult developmental perspective. *Psychology and Aging*, 1, 325–333. doi:10.1037/0882-7974.1.4.325
- Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *NeuroImage*, 17, 1394–1402. doi:10.1006/nimg.2002.1280
- Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia*, 41, 195–202. doi:10.1016/S0028-3932(02)00149-5
- Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*, 27, 103–123. doi:10.1023/A:1024569803230
- Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, 54, 165–181. doi:10.1037/0003-066X.54.3.165
- Carstensen, L. L., & Löckenhoff, C. E. (2003). Aging, emotion, and evolution. *Annals of the New York Academy of Sciences*, 1000, 152–179.
- Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior*, 23, 91–100. doi:10.1023/A:1021339410262
- Ciarrochi, J. V., Chan, A. Y., & Caputi, P. (2000). A critical evaluation of the emotional intelligence concept. *Personality and Individual Differences*, 28, 539–561. doi:10.1016/S0191-8869(99)00119-1
- Cohen, G. D. (2006). *The mature mind: The positive power of the aging brain*. New York, NY: Basic Books.
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, 1242, 126–135. doi:10.1016/j.brainres.2008.04.023
- Craik, F. I. M., & Salthouse, T. A. (2007). *The handbook of aging and cognition* (3rd ed.). New York, NY: Psychology Press.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14, 289–311. doi:10.1080/026999300378824
- Dixon, R. A. (1992). Contextual approaches to adult intellectual development. In R. J. Sternberg & C. A. Berg (Eds.), *Intellectual development* (pp. 350–380). New York: Cambridge University Press.
- Ebling, R., & Levenson, R. W. (2003). Who are the marital experts? *Journal of Marriage and Family*, 65, 130–142. doi:10.1111/j.1741-3737.2003.00130.x
- Ebner, N. C., He, Y., & Johnson, M. K. (2011). Age and emotion affect how we look at a face: Visual scan patterns differ for own-age versus other-age emotional faces. *Cognition and Emotion*, 25, 983–987.
- Ebner, N. C., & Johnson, M. K. (2009). Young and older emotional faces: Are there age group differences in expression identification and memory? *Emotion*, 9, 329–339. doi:10.1037/a0015179
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, 94, 100–131. doi:10.1037/0033-2909.94.1.100
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6, 169. doi:10.1080/02699939208411068
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face. A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.
- Ekman, P., & Friesen, W. (1978). *Facial action coding system: A technique for the measurement of facial movement: Investigator's guide 2 parts*. Palo Alto: Consulting Psychologists Press.
- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological Bulletin*, 128, 203–235.
- Feldman, R. S., Philippot, P., & Custrini, R. J. (1991). In R. S. Feldman (Ed.), *Social competence and nonverbal behavior* (pp. 329–350). Cambridge, England: Cambridge University Press.
- Feldman Barrett, L., & Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science*, 8, 10–14. doi:10.1111/1467-8721.00003
- Fung, H. H., & Carstensen, L. L. (2003). Sending memorable messages to the old: Age differences in preferences and memory for advertisements. *Journal of Personality and Social Psychology*, 85, 163. doi:10.1037/0022-3514.85.1.163
- Gildengers, A. G., Houck, P. R., Mulsant, B. H., Pollock, B. G., Mazumdar, S., Miller, M. D., . . . Reynolds, C. F. (2002). Course and rate of antidepressant response in the very old. *Journal of affective disorders*, 69, 177–184.
- Goodkind, M. S., Sollberger, M., Gyurak, A., Rosen, H. J., Rankin, K. P., Miller, B., & Levenson, R. (2012). Tracking emotional valence: The role of the orbitofrontal cortex. *Human Brain Mapping*, 33, 753–762. doi:10.1002/hbm.21251
- Gottman, J. M., & Krokoff, L. J. (1989). Marital interaction and satisfaction: A longitudinal view. *Journal of consulting and clinical psychology*, 57, 47.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological bulletin*, 85, 845–857.
- Hampson, E., van Anders, S. M., & Mullin, L. I. (2006). A female advantage in the recognition of emotional facial expressions: Test of an evolutionary hypothesis. *Evolution and Human Behavior*, 27, 401–416. doi:10.1016/j.evolhumbehav.2006.05.002
- Howard, J. H., Jr., Howard, D. V., Dennis, N. A., Yankovich, H., & Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology*, 18, 124–134. doi:10.1037/0894-4105.18.1.124
- Ickes, W. (1993). Empathic Accuracy. *Journal of Personality*, 61, 587–610. doi:10.1111/j.1467-6494.1993.tb00783.x
- Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships*, 7, 95–109. doi:10.1111/j.1475-6811.2000.tb00006.x
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 59, 730–742. doi:10.1037/0022-3514.59.4.730
- Isaacowitz, D. M., Lockenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, 22, 147–159. doi:10.1037/0882-7974.22.1.147
- Keightley, M. L., Winocur, G., Burianova, H., Hongwanishkul, D., & Grady, C. L. (2006). Age effects on social cognition: Faces tell a different story. *Psychology and Aging*, 21, 558–572. doi:10.1037/0882-7974.21.3.558
- Klein, K., & Hodges, S. D. (2001). Gender Differences, Motivation, and Empathic Accuracy: When it Pays to Understand. *Personality and Social Psychology Bulletin*, 27, 720–730. doi:10.1177/0146167201276007
- Labouvie-Vief, G. (1998). Cognitive-emotional integration in adulthood.

- In K. W. Schaie & M. P. Lawton (Eds.), *Annual Review of Gerontology and Geriatrics* (Vol. 17, pp. 206–237). New York, NY: Springer.
- Lawton, M. P. (2001). Emotion in later life. *Current Directions in Psychological Science*, *10*, 120–123. doi:10.1111/1467-8721.00130
- Levenson, R. W., Carstensen, L. L., & Gottman, J. M. (1993). Long-term marriage: Age, gender, and satisfaction. *Psychology and Aging*, *8*, 301–313. doi:10.1037/0882-7974.8.2.301
- Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: Physiological linkage and affective exchange. *Journal of Personality and Social Psychology*, *45*, 587–597. doi:10.1037/0022-3514.45.3.587
- Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology*, *63*, 234–246. doi:10.1037/0022-3514.63.2.234
- MacPherson, S. E., Phillips, L. H., & Della Sala, S. (2006). Age related decline in the ability to perceive sad facial expressions. *Aging: Clinical and Experimental Research*, *18*, 418–424.
- Magai, C. (2008). Long-lived emotions: A lifecourse perspective on emotional development. In J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 376–392). New York, NY: Guilford Press.
- Malatesta, C. Z., Izard, C. E., Culver, C., & Nicolich, M. (1987). Emotion communication skills in young, middle-aged, and older women. *Psychology and Aging*, *2*, 193–203. doi:10.1037/0882-7974.2.2.193
- Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*, 209–237. doi:10.1080/02699930802204677
- Montagne, B., Kessels, R., Frigerio, E., de Haan, E., & Perrett, D. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? *Cognitive Processing*, *6*, 136–141. doi:10.1007/s10339-005-0050-6
- Moreno, C., Borod, J. C., Welkowitz, J., & Alpert, M. (1993). The perception of facial emotion across the adult life span. *Developmental Neuropsychology*, *9*, 305–314. doi:10.1080/87565649309540559
- Orgeta, V., & Phillips, L. H. (2008). Effects of age and emotional intensity on the recognition of facial emotion. *Experimental Aging Research*, *34*, 63–79. doi:10.1080/03610730701762047
- Osgood, C. E., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Phillips, L. H., MacLean, R. D. J., & Allen, R. (2002). Age and the understanding of emotions neuropsychological and sociocognitive perspectives. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *57B*, P526–530. doi:10.1093/geronb/57.6.P526
- Prickaerts, J., Koopmans, G., Blokland, A., & Scheepens, A. (2004). Learning and adult neurogenesis: Survival with or without proliferation? *Neurobiology of Learning and Memory*, *81*, 1–11. doi:10.1016/j.nlm.2003.09.001
- Pujol, J., López-Sala, A., Deus, J., Cardoner, N., Sebastián-Gallés, N., Conesa, G., & Capdevila, A. (2002). The lateral asymmetry of the human brain studied by volumetric magnetic resonance imaging. *NeuroImage*, *17*, 670–679. doi:10.1006/nimg.2002.1203
- Rahhal, T. A., Colcombe, S. J., & Hasher, L. (2001). Instructional manipulations and age differences in memory: Now you see them, now you don't. *Psychology and Aging*, *16*, 697–706. doi:10.1037/0882-7974.16.4.697
- Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., . . . Acker, J. D. (2005). Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cerebral Cortex*, *15*, 1676–1689.
- Ruef, A. M., & Levenson, R. W. (2007). Continuous measurement of emotion. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 286–297). New York, NY: Oxford University Press.
- Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, *32*, 863–881. doi:10.1016/j.neubiorev.2008.01.001
- Russell, T. A., Tchanturia, K., Rahman, Q., & Schmidt, U. (2007). Sex differences in theory of mind: A male advantage on Happé's "cartoon" task. *Cognition and Emotion*, *21*, 1554–1564. doi:10.1080/02699930601117096
- Salthouse, T. A. (1990). Cognitive competence and expertise in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 310–319). San Diego, CA: Academic Press.
- Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum.
- Scherer, K. R., Matsumoto, D., Wallbott, H. G., & Kudoh, T. (1988). Emotional experience in cultural context: A comparison between Europe, Japan, and the USA. In K. R. Scherer (Ed.), *Facets of emotion: Recent research* (pp. 5–30). Hillsdale, NJ: Erlbaum.
- Segovia, G., del Arco, A., & Mora, F. (2009). Environmental enrichment, prefrontal cortex, stress, and aging of the brain. *Journal of Neural Transmission*, *116*, 1007–1016. doi:10.1007/s00702-009-0214-0
- Shamay-Tsoory, S., Harari, H., Szepeswol, O., & Levkovitz, Y. (2009). Neuropsychological evidence of impaired cognitive empathy in euthymic bipolar disorder. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *21*, 59–67. doi:10.1176/appi.neuropsych.21.1.59
- Shimokawa, A., Yatomi, N., Anamizu, S., Torii, S., Isono, H., Sugai, Y., & Kohno, M. (2001). Influence of deteriorating ability of emotional comprehension on interpersonal behavior in Alzheimer-type dementia. *Brain and Cognition*, *47*, 423–433. doi:10.1006/brcg.2001.1318
- Shiota, M. N., & Levenson, R. W. (2009). Effects of aging on experimentally instructed detached reappraisal, positive reappraisal, and emotional behavior suppression. *Psychology and Aging*, *24*, 890–900.
- Slessor, G., Phillips, L. H., & Bull, R. (2007). Exploring the specificity of age-related differences in theory of mind tasks. *Psychology and Aging*, *22*, 639–643. doi:10.1037/0882-7974.22.3.639
- Soto, J. A., & Levenson, R. W. (2009). Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage. *Emotion*, *9*, 874–884. doi:10.1037/a0017399
- Stanley, J. T., & Blanchard-Fields, F. (2008). Challenges older adults face in detecting deceit: The role of emotion recognition. *Psychology and Aging*, *23*, 24–32. doi:10.1037/0882-7974.23.1.24
- Sullivan, S., & Ruffman, T. (2004). Emotion recognition deficits in the elderly. *International Journal of Neuroscience*, *114*, 403–432. doi:10.1080/00207450490270901
- Sun, L. Y., & Bartke, A. (2007). Adult neurogenesis in the hippocampus of long-lived mice during aging. *The Journals of Gerontology: Series A: Biological and Medical Sciences*, *62*, 117–125. doi:10.1093/gerona/62.2.117
- Sze, J., Gyurak, A., Goodkind, M. S., & Levenson, R. W. (in press). Greater prosocial and empathic responding in late life. *Emotion*.
- Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*, 820–838. doi:10.1037/0022-3514.76.5.820
- Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, *78*, 105–119. doi:10.1037/0022-3514.78.1.105
- Williams, L. M., Brown, K. J., Palmer, D., Liddell, B. J., Kemp, A. H., Olivieri, G., . . . Gordon, E. (2006). The Mellow Years?: Neural Basis of Improving Emotional Stability over Age. *J. Neurosci.*, *26*, 6422–6430. doi:10.1523/JNEUROSCI.0022-06.2006

Received June 27, 2011

Revision received March 26, 2012

Accepted May 14, 2012 ■