Published in final edited form as:

Psychol Aging. 2013 March ; 28(1): 202–212. doi:10.1037/a0030927.

Older and Younger Adults' First Impressions From Faces: Similar in Agreement but Different in Positivity

Leslie A. Zebrowitz, Robert G. Franklin Jr, Suzanne Hillman, and Henry Boc

Department of Psychology, Brandeis University

Abstract

People readily form first impressions from faces, with consensual judgments that have significant social consequences. Similar impressions are shown by children, young adults (YA), and people from diverse cultures. However, this is the first study to systematically investigate older adults' (OA) impressions. OA and YA showed similar levels of within-age agreement in their impressions of competence, health, hostility, and trustworthiness. Both groups also showed stronger within- than between age agreement. Consistent with other evidence for age-related increases in positivity, OA showed more positive impressions of the health, hostility, and trustworthiness of faces. These effects tended to be strongest for the most negatively valenced faces, suggesting that they derive from OA lesser processing of negative cues rather than greater processing of positive cues. An own-age bias in impressions was limited to greater OA positivity in impressions of the hostility of older faces, but not younger ones. Although OA and YA differed in vision and executive function, only OA slower processing speed contributed to age differences in impression positivity. Positivity effects in OA have not been previously linked to processing speed, and research investigating possible explanations for this effect would be worthwhile.

Keywords

aging; face perception; trait impressions; agreement; positivity; own-age bias

Cultural wisdom enjoins us not to “judge a book by its cover,” an admonition suggesting that our natural proclivity is in fact to judge people by their appearance. Considerable research has shown that this is so. Not only do we make judgments about people from their faces, but also there is a remarkable consensus in impressions of a variety of traits with functional significance for perceivers, including aggressiveness, competence, health, and trustworthiness (Carr, McCormick, & Mondloch, 2009; Kalick, Zebrowitz, Langlois, & Johnson, 1998; Olivola & Todorov, 2010; Willis & Todorov, 2006; Zebrowitz, Bronstad, & Lee, 2007). Moreover, first impressions have significant social consequences, predicting, for example, criminal justice decisions (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie Vaughns, & Johnson, 2006; Stewart, 1980; Zebrowitz & McDonald, 1991) and voter preferences (Olivola & Todorov, 2010; Rule et al., 2010). Such “face reading” is manifested in diverse cultures with a cross cultural consensus in trait impressions from faces that extends to indigenous people from the remote Bolivian rainforest (Dion, 2002; Zebrowitz, Montepare, & Lee, 1993; Zebrowitz et al., 2012). Even infants and young children show responses analogous to those of adults (Keating & Bai, 1986; Langlois et al., 2000; Langlois, Roggman, & Rieser-Danner, 1990; Langlois & Styczynski, 1979; Montepare &

Zebrowitz-McArthur, 1989). Surprisingly, there is a paucity of research investigating the robustness of first impressions from faces at the other end of the life span a gap in the literature that we address in the present study.

Although research has not systematically investigated face impressions in older adults (OA), there is some reason to expect age related variations. One possibility is that OA have acquired the wisdom to avoid judging people by their appearance (cf. Baltes & Smith, 2008). Indeed, it is wise to refrain from judging a book by its cover, as trait impression accuracy from faces is a small effect (e.g., Zebrowitz & Rhodes, 2004), and is often offset by inaccurate overgeneralizations (e.g., Zebrowitz, Andreoletti, Collins, Lee, & Blumenthal, 1998). Another possibility is that OA greater experience with first impressions leads them to use more diagnostic facial cues. Consistent with this possibility is evidence that OA show more sensitivity to trait-diagnostic behavioral information than YA (e.g., Hess & Kotter-Grohn, 2011). Still another possibility is that OA first impressions may differ from those of YA due to declines in responsiveness to negatively valenced information. Several studies have found age differences in responses to negative stimuli, consistent with the hypothesis that OA show a “positivity bias” (Mather & Carstensen, 2005). For example, OA give more positive ratings than do YA to pictures depicting people, animals, scenery, or inanimate objects taken from the International Affective Picture System (IAPS) database (Lang, Bradley, & Cuthbert, 1999; Mather & Knight, 2005). Also, whereas YA tend to show preferential attention to and recall of negatively valenced faces and other stimuli, OA do not show this bias (Grady, Hongwanishkul, Keightley, Lee, & Hasher, 2007; Grohn, Scheibe, & Baltes, 2007; Isaacowitz, Goren, Wadlinger, & Wilson, 2006; Skosberg, 2012), although a meta analysis of relevant studies reveals that the stronger negativity biases for YA than OA are small effects (Murphy & Isaacowitz, 2008). In addition, OA show larger deficits in labeling negative than positive facial expressions (Ruffman, Henry, Livingstone, & Phillips, 2008).

Age differences in the wisdom to eschew facial appearance as a basis for judging people or in sensitivity to diagnostic or negative facial cues may have implications for agreement in first impressions as well as for the positivity of impressions. If OA are less likely to judge people by their appearance, their first impressions from faces may show less consistency across raters OA also may show lower within age agreement than YA if they show less consistent use of negative facial cues and if other consensually shared cues do not replace these. On the other hand, if OA and YA each show consensual use of different cues, with OA using more diagnostic or more positive cues, then there should be equal within age agreement among OA and YA, but greater within than between age agreement for each group.

Age differences in cue utilization may also yield mean differences in impressions in addition to age differences in interrater agreement. In particular, if OA show weaker responsiveness to negatively valenced facial cues, this may yield age differences in the positivity of trait impressions from faces. Consistent with this possibility, research has shown that OA perceive less danger than YA in faces at the high end of a dangerousness continuum as identified by YA or OA, with no age difference for less dangerous looking faces (Ruffman et al., 2006). This result suggests that age differences in the positivity of impressions will be most pronounced for the most negatively valenced faces. Although age differences in the positivity of first impressions from faces would be consistent with other evidence that OA show a positivity bias in their responsiveness to various stimuli, the question remains as to the mechanism for this effect. We explored several possibilities, including age differences in mood, vision, executive control, and processing speed.

Another consideration in examining effects of perceiver age on trait impressions from faces is the age of the faces. Hess and colleagues (Hess, Germain, Swaim, & Osowski, 2009; Hess, Leclerc, Swaim, & Weatherhee, 2009) have shown that the personal relevance of tasks affects the performance of OA more than YA. For example, personal accountability for a personnel decision had a more beneficial effect on OA than YA memory for information about the applicants and greater interest in a criminal case scenario augmented the use of complex qualitative evidence to determine guilt more for OA than YA judges. Personal relevance of the person being judged, as manipulated by age, also affects the performance of OA on face recognition tasks, with the advantage of younger raters attenuated or reversed for older faces. This has been shown for face recognition (Anastasi & Rhodes, 2005; Fulton & Bartlett, 1991), age recognition (Voelkle, Ebner, Lindenberger, & Riediger, 2012), emotion recognition (Malatesta, Izard, Culver, & Nicolich, 1987), and picking out perpetrators of crimes when viewing lineups (Perfect & Harris, 2003; Wright & Stroud, 2002).

It is possible that an own age bias will also be shown in first impressions from faces, in which case, OA first impressions will be more sensitive to appearance variations in older faces and YA first impressions will be more sensitive to variations in younger faces, although there is reason to expect the former effect to be stronger. If this is the case, we would predict that each age group would show greater agreement for own age than other age faces. Greater sensitivity to variations in the appearance of own age faces also may yield an attenuation of the OA positivity bias when rating older faces insofar as they show greater sensitivity to negative cues in own age faces. On the other hand, in-group favoritism may simply yield more positive impressions of own age faces.

To summarize, the present study addressed the following questions concerning perceiver age effects on trait impressions from faces:

1. **Agreement.** Is within age agreement weaker for OA, as may be predicted if OA are not using the negative facial cues that YA use to achieve agreement? Is within age agreement stronger than between age agreement, as may be predicted if each age group is using different facial cues?
2. **Positivity.** Do OA show more positive impressions than YA, consistent with evidence for less negativity in other domains? Is the OA positivity bias greater for more negative looking faces, as may be predicted if OA are making less use of negative facial cues used by YA?
3. **Own age bias.** Is within age agreement stronger for own age faces? Is the OA positivity bias attenuated for own age faces, as may be predicted from an own age bias in sensitivity to negative facial cues, or is the positivity bias stronger for own age faces, as may be predicted from in-group favoritism?
4. **Mechanisms.** Do age differences in mood, perceptual abilities, and/or cognitive abilities contribute to age differences in agreement and/or positivity?

Method

Participants

Forty eight YA participants (23 men) and 48 OA participants (24 men) completed the study. YA participants, aged 18–22 ($M = 18.8$, $SD = 1.0$), were recruited from a university and completed the study for course credit or payment of \$15. OA participants, aged 65–85, ($M = 76.3$, $SD = 6.4$), were recruited from the local community and were paid \$25 for completing the study. OA were screened using the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) all scoring above 26 out of 30 ($M = 28.9$, $SD = 1.2$).

Facial Stimuli

Three sets of facial images were used. One set was taken from a Boston sample of individuals who participated in the study of Midlife Development in the United States (MIDUS) funded by the John D. and Catherine T. MacArthur Foundation (Lachman, 1997). The images used were color frontal facial photographs of people displaying neutral expressions and wearing a gray cape to mask their clothing. A total of 69 YA (41 men) between the ages of 25 and 39 ($M = 32.7$, $SD = 4.2$), and 68 OA (44 men) between the ages of 60 and 74 ($M = 66.1$, $SD = 4.1$) were included. Images were displayed at an approximate size of 7.5×9.5 cm.

Two additional sets of grayscale images were taken from the Intergenerational Studies (IGS) archive, a longitudinal study of representative samples of individuals born in Berkeley, California, in the late 1920s or attending school in Oakland, California, in the 1930s (Eichorn, 1981). The images included two photographs of 198 individuals (74 men). One photo was taken when the person was between 17 and 18 years old and a second when the person was between 52 and 62 years old, for a total of 296 images. To reduce the face rating task to a reasonable length, these images were divided into two sets, with equal numbers of older and younger men and women included in each set. In addition, the images in each set were matched within demographic group on previously collected attractiveness ratings by YA (for more information about the photographs and the ratings, see Zebrowitz, Olson, & Hoffman, 1993). The two IGS face sets did not differ in the attractiveness ratings by YA or OA in the current sample. However, one set was rated as marginally more baby faced by OA, $t(30) = 1.76$, $p = .09$, with no differences shown in baby face ratings by YA. $t < 1$. Images were displayed at an approximate facial size of 9×11 cm.

Dependent Measures

Face impressions—Four traits (competent, healthy, untrustworthy, and aggressive or hostile) were rated on 7-point scales. Raters evaluated the extent to which they thought that each trait described the face shown, with endpoints labeled 1. *not at all* (healthy, competent, untrustworthy, aggressive or hostile) and 7. *very* (healthy competent, untrustworthy, aggressive or hostile). Competence ratings by one OA man were dropped from the analyses, because all faces received an identical rating. In preliminary analyses of aggressiveness ratings using the MIDUS faces, we found that OA agreement was significantly lower than that of YA. Because it seemed possible that lower agreement in OA may reflect less agreement regarding the meaning of the term aggressive (i.e., assertiveness vs. hostility), we used the more unambiguous term “hostile” for ratings of the IGS faces. After the four trait ratings, faces also were rated on two 7-point appearance scales (attractive, baby faced), but these will not be discussed in depth, as they are relevant to separate questions regarding age differences in appearance stereotypes.

Control measures—Measures of vision, affect, and cognitive function were administered to determine whether any age differences in trait impressions could be linked to differences in these more general processes. Participants completed a computerized version of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) to determine whether there were age differences in affect. Participants also were tested for differences in visual acuity (Snellen Eye Chart), contrast sensitivity (Mars Letter Contrast Sensitivity Test, Mars Perceptrix, Chappaqua, NY), and color vision (Ishihara’s Tests for Color Blindness; Ishihara, 1917/2010). and they completed the Benton Facial Recognition Test (Benton, Van Allen, Hamsher, & Levin, 1983), which assesses the ability to match faces that are the same identity when presented with different orientations, expressions, and lighting conditions. We also included measures of general cognitive abilities. All participants completed the Shipley Vocabulary test (Shipley, 1946) to assess crystallized

intelligence, and a timed Pattern Comparison Task (Salthouse, 1993) to assess Processing speed. To assess executive function, participants who rated the MIDUS faces completed the Letter Number Sequencing Task (Wechsler, 1997), and those who rated the IGS faces completed a computerized version of the Wisconsin Card Sort Task (WCST) (the Berg Card Sort Task, or BCST: downloaded from <http://pebl.sourceforge.net/battery.html/> and validated by Piper et al., 2012).

Procedure

Participants rated one of the three sets of images. After obtaining informed consent, participants were seated in front of a computer screen. They first completed a computerized version of the PANAS. Next, MediaLab software (Empirisoft, New York, NY) was used to present the trait rating task. Participants were randomly assigned to one of two trait orders: healthy, untrustworthy, competent, aggressive/hostile; or aggressive/hostile, competent, untrustworthy, healthy. Faces were shown in one of four orders, counterbalancing age and sex of face and participants were asked to rate each face in comparison with the other faces of that age/sex grouping. This procedure was employed so that variations in trait impressions would less likely reflect sex or age stereotypes, as we were interested in whether the first impressions of OA were responsive to variations in facial appearance within demographic groups, as has been previously shown for YA Faces were shown for either 4 s (MIDUS faces) or 3 s (IGS faces), after which the face disappeared and the rating scale appeared on the screen, remaining until the rating was made. Then the next face appeared. Participants rated all faces on one trial slab before the next one was introduced. Prior to each scale, there was an instruction screen indicating what the trial to be rated would be. The entire rating portion lasted about an hour, after which a demographic and health questionnaire and the remaining control measures were administered.

Results

Agreement in Trait Impressions

To examine OA and YA agreement in impressions we correlated each participant's ratings of a specific trait with the mean of all the other individuals in that person's age group, excluding the participant, as well as with the mean of all the individuals in the other age group. This provided indices of within age and between age agreement for each participant. These correlations were repeated for older faces and younger faces. Controlling for face sex and actual face age when computing these correlations did not change any of the effects reported below. The correlations for each participant were Fisher Z transformed to normalize them (see Franklin & Adams, 2009; Judd, Ryan, & Park, 1991) and then used as data in a series of 2 (rater age: older vs. younger) $\times$ 2 (face age: older vs. younger) $\times$ 2 (agreement type: within age vs. between age) $\times$ 3 (face set: MIDUS, IGS Set 1, or IGS Set 2) mixed factor ANOVAs with face age and agreement type as within subject factors and rater age and face set as between subjects factors. Since face set effects have no theoretical meaning, we only report them when an effect involving a variable of theoretical interest (rater age, face age, agreement type) did not hold true across all face sets, indicating a qualification to generalizability. This method of measuring agreement has several advantages over other methods, such as alpha coefficients. Specifically, the statistics generated allowed inferential statistical comparison of rater age, face age, and agreement type effects as well as their interactions.

All of the overall within and between age agreement means were significantly greater than zero, as indicated by significant intercept effects in the ANOVAs. all $F_s > 25$. $p_s < .001$ (see Table 1). All marginal means reported below also were significantly greater than zero at $p < .001$. It should be noted that these means are averaged across many independent Z

transformed correlations, and thus should be regarded as mean effect sizes in a meta analysis, rather than a single correlation effect.

Rater-age effects—Contrary to expectation, main effects of perceiver age revealed higher agreement in ratings of competence among OA than YA, $F(1, 89) = 6.94, p = .010, \eta^2 = .072$, as well as higher agreement in ratings of “hostile” among OA than YA, $F(1, 60) = 5.43, p = .023, \eta^2 = .083$. In contrast, a main effect of perceiver age for ratings of “aggressive” revealed lower agreement among OA than YA, $F(1, 30) = 5.89, p = .021, \eta^2 = .164$. There were no significant perceiver age differences in agreement for impressions of healthy or untrustworthy, both $F_s < 1.5$, all $p_s > .25$ (see Table 1).

Agreement-type effects—As predicted, within age agreement (i.e., the correlation between a rater’s own ratings and the mean ratings for the rest of the rater’s own age group) was significantly greater than between age agreement (i.e., the correlation between a rater’s own ratings and the mean ratings for participants in the other age group) for all impressions: Competent $F(1, 89) = 68.2, p < .001, \eta^2 = .43$; healthy, $F(1, 90) = 34.7, p < .001, \eta^2 = .278$; untrustworthy, $F(1, 90) = 5.43, p = .023, \eta^2 = .22$; hostile, $F(1, 60) = 32.9, p < .001, \eta^2 = .354$; and aggressive $F(1, 30) = 19.27, p < .001, \eta^2 = .391$ (see Table 1).

Moderation of rater-age effects by face age—Contrary to the expectation that there would be stronger agreement in impressions of own age than other age faces, the Rater Age $\times$ Face Age interaction was not significant for any of the trait impressions, all $F_s < 1.5$, all $p_s > .26$. Rather, OA and YA were equivalent in showing greater agreement in impressions of the competence of younger than older faces, $F(1, 89) = 4.54, p = .036, \eta^2 = .048$, the health of older than younger faces, $F(1, 90) = 12.57, p = .001, \eta^2 = .123$, the hostility of younger than older faces, $F(1, 60) = 25.2, p < .001, \eta^2 = .296$, and the aggressiveness of older than younger faces, $F(1, 30) = 5.91, p = .021, \eta^2 = .17$. OA and YA also showed equal agreement in impressions of the untrustworthiness of younger and older faces, which did not differ, $F(1, 90) = 1.30, p = .26, \eta^2 = .014$ (see Table 1).

Moderation of rater-age effects by agreement type—Rater Age $\times$ Agreement Type interactions for impressions of hostile, $F(1, 60) = 9.91, p = .003, \eta^2 = .142$, and aggressive, $F(1, 30) = 10.6, p = .003, \eta^2 = .26$, qualified the rater-age effects in agreement. In the case of hostile ratings, the tendency for OA to show stronger agreement than YA was greater for within age agreement, $t(63) = 3.84, p < .001$, than for between age agreement, $t(63) = 2.10, p = .038$. In the case of aggressive ratings, the tendency for OA to show weaker agreement than YA was greater for within age agreement, $t(31) = 3.96, p < .001$, than for between age agreement, $t(31) = 2.57, p = .014$. (See Figure 1).

Moderation of agreement-type effects by face age—A significant Face Age $\times$ Agreement Type interaction for impressions of hostile, $F(1, 60) = 7.50, p = .008, \eta^2 = .111$, revealed that the tendency for within age agreement to be greater than between age agreement was greater for older faces (within age: $M = .59, SD = .25$, between age: $M = .51, SD = .24, t(63) = 5.33, p < .001$, than for younger faces (within age: $M = .70, SD = .24$, between age: $M = .67, SD = .23, t(63) = 2.40, p = .019$). Similarly, a significant Face Age $\times$ Agreement Type interaction for ratings of aggressive, $F(1, 30) = 6.59, p = .013, \eta^2 = .19$, revealed that the tendency for within age agreement to be greater than between age agreement was significant for older faces, within age agreement, $M = .57, SD = .33$; between age agreement $M = .48, SD = .32, t(61) = 5.12, p < .001$, but not younger faces, within age $M = .47, SD = .29$; between age $M = .44, SD = .28, t(31) = 1.50, p = .14$.

Moderating effects of face set—Face set did not moderate any of the rater age effects, all $F_s < 2$, $p_s > .15$, while it qualified one face age effect and one agreement type effect. An interaction between face age and face set, $F(2, 90) = 20.4$, $p < .001$, $\eta^2 = .311$, revealed that the tendency for agreement in impressions of health to be higher for older faces than younger faces was significant for both MIDUS faces, older: $M = .68$, $SD = .18$; younger: $M = .55$, $SD = .20$, $t(31) = 5.41$, $p < .001$, and IGS Set 2 faces, older: $M = .60$, $SD = .16$; younger: $M = .45$, $SD = .16$, $t(31) = 5.65$, $p < .001$, with the reverse for IGS Set 1 faces, older: $M = .55$, $SD = .19$; younger: $M = .45$, $SD = .15$, $t(31) = 5.01$, $p < .001$. An interaction between agreement type and face set, $F(2, 90) = 17.5$, $p < .001$, $\eta^2 = .28$, revealed that tendency for within age agreement in health to be greater than between age agreement was significant for MIDUS faces, within age: $M = .68$, $SD = .20$; between age: $M = .50$, $SD = .18$, $t(31) = 8.27$, $p < .001$, and IGS Set 1 faces, within age: $M = .52$, $SD = .16$; between age: $M = .48$, $SD = .19$, $t(31) = 2.27$, $p = .026$, but not for; IGS Set 2 faces, within age: $M = .52$, $SD = .18$, between age: $M = .53$, $SD = .17$, $t(31) = .08$.

Positivity of Trait Impressions

To test the hypothesis that OA would rate faces more positively than YA, we conducted 2 (rater age: OA vs. YA) $\times$ 2 (face age: older vs. younger) $\times$ 2 (face sex: female vs. male) $\times$ 3 (face set: MIDUS, IGS Set 1, or IGS Set 2) ANOVAs on each of the four trait ratings. Rater was the unit of analysis and rater age and face set were between subjects variables, and face age and face sex were within subject variables. Although we included face sex in the analysis to account for any variance it might contribute, none of the effects of rater age or face age reported below were moderated by face sex. We do not report any other face sex effects, as we had no predictions regarding this variable. In addition, since face set effects have no theoretical meaning, we only report them when an effect involving rater age or face age was not obtained across all face sets, indicating a qualification to the generalizability of that effect.

Because we anticipated that age differences in mean ratings of faces might be most pronounced for negatively valenced faces, we performed additional analyses on each trait impression using the same ANOVA design, adding a 3 group face negativity factor (negatively valenced, neutral, positively valenced). More specifically, for each set of faces, we computed a median split on each trait rating separately for OA and YA. Medians were computed separately for older and younger male and female faces to ensure that the median splits were not confounded with face age or sex. Faces were designated as “negative” in competence and health if both OA and YA rated them below the median on these trait scales, and they were designated as “negative” in untrustworthiness, hostility, or aggressiveness if both groups rated them above the median. Faces were designated as “positive” in competence and health if both OA and YA rated them above the median, and they were designated as “positive” in untrustworthiness, hostility, or aggressiveness if both age groups rated them below the median. Faces that were not consistently rated above or below the median by both age groups were designated as “neutral.” In the interest of brevity, we report only those face negativity effects that were moderated by rater age and achieved at least marginal significance ($p = .10$), as others are not relevant to our hypotheses.

Rater-age effects—Consistent with the hypothesis that OA would show more positive trait impressions, main effects of rater age revealed higher health ratings by OA than YA, $F(1, 90) = 7.01$, $p = .010$, $\eta^2 = .07$; lower untrustworthiness ratings by OA than YA, $F(1, 90) = 4.53$, $p = .036$, $\eta^2 = .05$; and lower hostility ratings by OA than YA, $F(1, 60) = 5.74$, $p = .020$, $\eta^2 = .087$. Contrary to prediction, a significant main effect of rater age revealed higher aggressiveness ratings by OA than YA, $F(1, 30) = 6.27$, $p = .018$, $\eta^2 = .17$, and there was no significant effect of rater age on competence ratings, $F(1, 89) = .21$ (see Table 2).

Moderation of Rater-Age Effects by Face Negativity

As predicted, rater age differences in impression positivity were moderated by face negativity. For impressions of health, the Rater Age $\times$ Face Negativity interaction was marginally significant, $F(2, 90) = 246, p = .089, \eta^2 = .027$. Planned comparisons within each level of face negativity revealed that, as predicted, the tendency for OA to rate faces as healthier than YA was strongest for the most unhealthy looking faces below the median in rated health, $F(1, 90) = 8.67, p = .004, \eta^2 = .088$, although it was also significant for faces near the median, $F(1, 90) = 5.52, p = .021, \eta^2 = .058$, and for the most healthy looking faces above the median: $F(1, 90) = 4.40, p = .039, \eta^2 = .047$. The Rater Age $\times$ Face Negativity interaction was also significant for impressions of untrustworthiness, $F(2, 90) = 4.18, p = .018, \eta^2 = .086$. Planned comparisons within each level of face negativity revealed that the tendency for OA to rate faces as less untrustworthy than YA was significant for highly untrustworthy looking faces above the median in untrustworthy appearance $F(1, 90) = 7.02, p = .010, \eta^2 = .072$, as well as for neutral faces, $F(1, 90) = 6.74, p = .011, \eta^2 = .070$, but not for trustworthy-looking faces that were below median in untrustworthy appearance, $F(1, 90) = .44$. The Rater Age $\times$ Face Negativity interactions were not significant for impressions of aggressiveness, competence, or hostility, all F s < 1.4 (see Figure 2).

Moderation of rater-age effects by face age—The only significant Rater Age $\times$ Face Age effect was for hostility ratings, $F(1, 60) = 4.41, p = .040, \eta^2 = .068$. The tendency for OA to rate faces as less hostile than YA was significant for older faces, $t(62) = 2.90, p = .005$ but not for younger faces, $t(62) = 1.63, p = .11$, consistent with an in group favoritism effect (see Figure 3). Nevertheless, OA, like YA, rated younger faces as more hostile than older faces, OA: $t(31) = 6.33, p < .001$; YA: $t(31) = 6.83, p < .001$. All other Rater Age $\times$ Face Age interactions were not significant, F s $< 2.5, p$ s $> .15$, indicating that OA and YA showed equivalent tendencies to rate older faces as more competent than younger faces, $F(1, 89) = 19.6, p < .001, \eta^2 = .179$; less healthy than younger faces, $F(1, 90) = 10.4, p = .002, \eta^2 = .10$; less untrustworthy than younger faces, $F(1, 90) = 24.6, p < .001, \eta^2 = .214$; and less aggressive than younger faces, $F(1, 30) = 5.01, p = .03, \eta^2 = .143$ (see Table 2).

Moderating effects of face set—Face set did not moderate any of the foregoing effects for impressions of competent, healthy, untrustworthy, or aggressive (which was rated only for the MIDUS face set). However, it did moderate impressions of hostility (which was rated only for the two IGS face sets). A Rater Age $\times$ Face Set interaction, $F(1, 60) = 4.28, p = .043, \eta^2 = .067$, revealed that the tendency for OA to rate faces as less hostile than YA was significant for IGS Face Set 2, OA, $M = 2.77, SD = .94$; YA: $M = 3.63, SD = .40, t(30) = 3.40, p = .003$, but not IGS Face Set 1, OA, $M = 3.12, SD = .66$; YA: $M = 3.18, SD = .96; t(30) = .22$. There also was a Rater Age $\times$ Face Age $\times$ Face Set effect, $F(1, 60) = 4.48, p = .039, \eta^2 = .069$. The tendency for OA to differ more from YA in their ratings of older faces than younger faces was significant in Face Set 2, $F(1, 30) = 6.06, p = .020, \eta^2 = .168$, but not in Face Set 1, $F(1, 30) = .054$.¹

¹We also examined the rater-age and face-age effects using mixed-effects models (see Judd, Westfall, & Kenny, 2012), finding that, with one exception, the main effects of rater age were also significant when accounting for subject-level and stimulus-level variance. The exception was that the main effect for hostility was not significant with this analysis ($p = .11$). However, the Rater Age $\times$ Face Age effect for hostility was replicated in the mixed-effect smodel ($p = .02$), with OA rating older faces as less hostile than YA ($p < .001$), and no significant rater-age difference for younger faces ($p = .41$). Face-age effects in the mixed-effects analysis replicated the original findings that older faces were rated as less hostile ($p = .005$), less healthy ($p = .04$), and less untrustworthy ($p < .001$), than younger faces. However, the original findings that older faces were rated as more competent and less aggressive were not significant in the cross-random effects analysis, suggesting that these face-age differences may not generalize to other faces.

Control Data Analyses

Age differences—We performed t tests on each of the control measures to assess age differences in affect (PANAS), visual perception (acuity, contrast sensitivity, color vision, face recognition), and cognitive abilities (executive function, processing speed, and crystallized intelligence; see Table 3). OA and YA did not significantly differ on positive or negative affect in the PANAS. OA showed worse visual acuity (Snellen Eye Chart) and contrast sensitivity than YA, but in the normal range for their age, and OA performed worse than YA on the Benton Facial Recognition Test (Benton, Eslinger, & Damasio, 1981), albeit in the normal range for normative samples of OA (e.g., Eslinger, Damasio, Benton, & Van Allen, 1985). There were no age differences in color vision (Ishihara, 1917/2010). OA scored higher than YA on the Shipley Vocabulary test (Shipley, 1946), consistent with their higher education level and the maintenance of crystallized intelligence in older adulthood (Horn & Caltell, 1967). OA scored lower than YA on a timed pattern comparison task (Salthouse, 1993), consistent with decreases in processing speed in older adulthood (Salthouse, 1996). OA who rated the MIDI'S faces performed worse than YA on letter–number sequencing task (Wechsler 1997 indicating a decline in executive function, and OA who rated IGS faces performed worse on the WCST, also indicating decreases in executive function, although the age differences were not shown on preservative errors, which are the most diagnostic of executive dysfunction (Greve, Stickler, Love, Bianchini, & Stanford, 2005).

Covariance analyses—We performed 2 (rater age) $\times$ 2 (face age) $\times$ 2 (face sex) analyses of covariance, with each of the control variables as separate covariates to determine whether age differences in mood, visual perception, or cognitive function contributed to age differences in each of the impressions. The results we report below focus on whether the significant rater age difference in impressions lost significance when controlling for the covariate. Because different measures of executive function were used for participants who rated MIDI'S and IGS face, we computed z scores for the letter–number sequencing task and the WCST to use as the executive function covariate. It should be noted that these analyses were exploratory, with the aim of identifying possible mediators of age differences in positivity that may then be tested in an experiment that manipulates them (cf. Miller & Chapman 2001, pp. 45–46).

The significant tendency for OA to rate faces as more healthy than YA in the original ANOVA lost significance when controlling for pattern comparison speed, $F(1, 90) = 2.60$, $p = .11$, $\eta^2 = .028$, as did the tendency for OA to rate faces as less untrustworthy, $F(1, 90) = 1.40$, $p = .24$, $\eta^2 = .016$; and as less hostile, $F(1, 60) = 2.14$, $p = .15$, $\eta^2 = .035$. Controlling for pattern comparison speed also eliminated the significant tendency for OA to rate faces as more aggressive than YA, $F(1, 30) = 1.02$, $p = .32$, $\eta^2 = .034$, but so did controlling for visual acuity, $F(1, 30) = 2.65$, $p = .11$, $\eta^2 = .084$; contrast sensitivity, $F(1, 30) = 1.48$, $p = .23$, $\eta^2 = .048$; and performance on the letter–number sequencing task, $F(1, 30) = 2.31$, $p = .14$, $\eta^2 = .074$. Controlling the other covariates did not render the effect of rater age nonsignificant for ratings of health, all $F_s > 4$, $ps < .025$; untrustworthy, all $F_s > 3.5$, $ps < .075$; hostile, all $F_s > 3.5$, $ps < .075$; or aggressive, all $F_s > 3.5$, $ps < .05$.

Discussion

The present findings reveal that the tendency to form consensual trait impressions from faces persists into older adulthood. At the same time, there were reliable differences between OA and YA impressions, most of which reflected greater positivity for OA, as predicted. Contrary to expectation, there was no consistent evidence for an own age bias either in the level of agreement in first impressions or in differences in their positivity.

Agreement

OA, like YA, showed consensual impressions, with significant within age agreement in their first impressions of faces. OA showed weaker within age agreement than YA in impressions of the aggressiveness of MIDUS faces, consistent with the hypothesis that they may underutilize the negative facial cues that YA use to achieve agreement. However, another possible explanation for OA weaker agreement is that there may have been greater variability in how they construed the term “aggressive,” with some emphasizing the connotation of assertiveness rather than the connotation of hostility. To assess this possibility, we used the less ambiguous negative trait “hostility” when assessing impressions of the IGS faces. In contrast to impressions of aggressiveness, impressions of hostility showed higher within age agreement among OA than YA. OA also showed higher within age agreement in impressions of competence across all face sets. Thus, whatever wisdom OA have acquired about judging people from appearance and whatever tendency they have to eschew negative facial cues do not eliminate their ability to achieve a consensus with age peers when asked to form first impressions solely on the basis of facial cues.

OA not only showed significant within age agreement in their first impressions of faces, they also showed significant agreement with YA. Thus, people who look competent, healthy, trustworthy, aggressive, or hostile to YA will tend to create similar impressions in OA. Although this suggests that facial appearance will lead to similar social consequences in OA, like election outcomes (Olivola & Todorov, 2010; Rule et al., 2010), there also may be some differences, since both age groups showed stronger within than between age agreement in their impressions of competence, health, and untrustworthiness. This suggests that each age group has some unique standards for judging these qualities from faces in addition to shared standards used across age groups. The greater within than between age agreement in many impressions is consistent with previous evidence that shared perceptual experiences augment agreement in face perception (Bronstad & Russell, 2007; Elfenbein & Ambady, 2003; Jones & Hill, 1993; Zebrowitz, Wang, et al., 2012). This result also highlights the value of research investigating whether the reduced between age agreement translates into age differences in the accuracy of impressions.

Positivity

Consistent with other evidence for greater positivity in OA than YA (Grohn et al., 2007; Mather & Carstensen, 2005; Murphy & Isaacowitz, 2008; Skosberg, 2012), OA showed greater positivity in their trait impressions of health, untrustworthiness, and hostility, although the effect for hostility was significant for only one of two face sets. The absence of age differences in positivity for impressions of competence might be related to evidence that OA showed lower valuation of competence related than morality related traits (Hess & Auman, 2001). Perhaps OA greater positivity is most pronounced for impressions that they care more about.

Our results also provided some support for the prediction that the OA positivity effects in trait impressions would be more marked for more negative looking faces, as would be expected if the positivity bias derives from less processing of negative cues rather than greater processing of positive cues. OA greater positivity in ratings of health was strongest for the most unhealthy-looking faces, although it remained significant for faces average or high in healthy appearance. In addition, OA greater positivity in impressions of untrustworthiness was limited to faces high or average in the untrustworthiness of appearance, with no significant effect for the more positive looking faces below the median in untrustworthiness. This effect is consistent with previous evidence that, compared with YA, OA perceived less danger in faces that were categorized as highly dangerous by independent groups of YA or OA judges, whereas there was no age difference for less

dangerous looking faces (Ruffman et al., 2006). Whereas the Ruffman et al., (2006) study did not equate the age or sex of faces across the high and low danger categories, our findings indicate that OA are less responsive than YA to facial cues to danger in very untrustworthy looking faces, that are matched in age and sex to the more trustworthy looking ones.

Although the greater positivity of OA impressions of hostility was comparable across all three levels of hostile appearance based on the median split for all IGS faces, it is noteworthy that these rater age effects were limited to one set of faces, which, as it turned out, were more threatening looking than those in the other set. Specifically, we found that the set that showed greater OA positivity in impressions of hostility, was also rated as less baby-faced than the other set by OA. Since low baby faceness is associated with higher perceived hostility (Zebrowitz, 1997; Zebrowitz, Kikuchi, & Fellous, 2010), these results can be construed as consistent with the hypothesis that the OA positivity; bias would be greater for more negative looking faces.

An exception to the greater positivity of OA impressions was that they rated faces as more aggressive than did YA. One possible explanation for this surprising effect is that, as noted above, some OA may have emphasized the connotation of assertiveness rather than the connotation of hostility when rating aggressiveness which is supported by the lower interrater agreement in aggressive ratings shown by OA than YA. On this account, the tendency for OA to rate faces as more aggressive than YA actually could be construed as a positivity effect insofar as it reflects a more positive spin on the meaning of the ambiguous word “aggressive.” The fact that OA rated another set of faces as less hostile, a less ambiguous descriptor, is consistent with this suggestion.

Own-Age Bias

Although level of agreement showed no own age bias, it was sensitive to face age, with both OA and YA showing higher agreement in their impressions of the competence and hostility of younger than older faces, and in their impressions of the health and aggressiveness of older than younger faces, with no effects of face age on agreement in impressions of untrustworthiness. These un-predicted main effects of face age are intriguing and they suggest that variations in apparent competence and hostility are more clearly shown in younger faces, while variations in apparent health and aggressiveness (which may be assertiveness) are more clearly shown in older face. It is noteworthy that these effects held true even when statistically controlling variations in the actual age of faces within each age group.

Rater age differences in positivity were moderated by face age only in the case of hostility impressions. Specifically, OA greater positivity was significant for older but not younger faces, indicating an own age bias in favorability of impressions, rather than in sensitivity to negative facial cues. Although face age moderated the rater age differences in only this single instance, it did influence impressions, with both groups rating older faces as more competent, less healthy, less untrustworthy, less aggressive, and less hostile. Thus, there was a general tendency to form more positive impressions of the older faces, with the exception of impressions of health. The perception of older adults as less healthy than younger adults is not at all surprising. On the other hand, the perception of older adults as more competent than younger adults contradicts evidence for declining competence in many domains across the adult life span (but see Footnote ¹) However, it should be recalled that the majority of the older faces depicted individuals between 52 and 62 years of age, which is certainly not at the nadir of competence. Moreover, the majority of younger faces depicted individuals between 17 and 18 years of age, which is certainly not at the peak of competence. Finally, the perception of older faces as less hostile and less untrustworthy is consistent with evidence for benign stereotypes of older adults on these dimensions, as well as with crime

statistics indicating that younger adults are more likely to engage in antisocial behaviors (Moffitt, 1993).

A caveat is warranted regarding the finding of little support for the expectation that raters might show an own age bias in their level of agreement in first impressions or in the positivity of those impressions. Own age biases may be stronger when raters are directly comparing older to younger faces, in contrast to the procedure in the present study, which blocked faces by age as well as sex (Biernat, 2003). Blocking faces by age and sex and instructing participants to rate faces compared with others in the same demographic group enabled us to determine whether OA first impressions of faces would be responsive to variations in facial appearance within demographic group, as has Ken previously shown for YA. and OA significant within age and between age agreement showed that they are. Had we not used this procedure, mean ratings of both OA and YA would likely have been highly influenced by age stereotypes with less variation reflecting rater age differences in positivity. Indeed, although our procedure may have underestimated own age biases, it was not sufficient to eliminate effects of face age. As noted above, there were many significant differences in the mean ratings of older and younger faces shown by OA and YA alike. Another consideration in evaluating the lack of own age biases is that, although “own age” faces were closer in age to participants than “other age” faces, they were not exactly matched in age. Specifically, our YA participants were younger than their “own age” MIDUS faces, whereas our OA participants were, on average, older than their “own age” MIDUS and IGS faces. Thus, it remains possible that greater agreement and positivity would be shown in impressions of people more closely matched in age to perceivers.

Mechanisms

The question remains as to the mechanism underlying rater age differences in the positivity of impressions. Consistent with previous research, OA and YA differed on many of the control variables, including measures of visual function, crystallized intelligence, executive control, and processing speed. Although OA poorer vision might impair sensitivity to negative facial cues, such as subtle resemblance to negative emotions influencing first impressions (Zebrowitz et al., 2010), controlling age differences in visual abilities did not eliminate the age differences in impression positivity. Similarly, the tendency for OA to form more positive impressions of hostility, health, and untrustworthiness was unaffected when age differences in executive control scores were controlled. This result is not surprising, given the theoretical account of executive control effects in previous research. Specifically, it has been argued that greater executive control facilitates OA greater motivation to regulate negative emotions, with age differences in positivity strongest when considering OA with higher levels of control (Mather & Knight, 2005). However, such motivation may not have been activated by the neutral expression faces in the present study. Although some of these faces elicited negative first impressions, they lacked the strong negative valence of negative emotion expressions or negative IAPS pictures, which may have activated emotion regulation processes in previous research. There were no age differences in mood in the present study, although such differences would be a prime contender to explain age differences in impression positivity. Specifically, socioemotional selectivity theory (Mather & Carstensen, 2005) predicts that OA will likely be in a more positive mood, and other research indicates that judgments may be biased in a mood-congruent manner (Forgas, 1995; Isen & Shaker, 1982; Schwarz & Clore, 2003). However, controlling for mood did not eliminate the age differences in impression positivity, as would be expected given the absence of differences in mood. The only rater age difference that consistently contributed to differences in impression positivity was OA poorer performance on a timed pattern comparison task (Salthouse, 1993), which assesses speed of judging whether two patterns are the same or different. Positivity effects in OA have not been

previously linked to processing speed, and research manipulating this possible mediator of age differences is currently underway. Although the exploratory covariance analyses did not implicate other control variables as possible mediators, research that manipulates various contenders also would be worthwhile.

Conclusion

Like YA, OA form consensual first impressions from faces. Although these impressions show significant agreement with those of YA, between age agreement was weaker than within-age agreement for both age groups, and there were significant age differences in the positivity of impressions OA judged faces as healthier, less untrustworthy, and less hostile than YA, as well as more aggressive, which may have been due to some OA having a more positive slant on the meaning of this term. The greater positivity of OA impressions tended to be stronger for the most negatively valenced faces, suggesting that this age difference reflects OA lower responsiveness to negative cues more than their greater responsiveness to positive cues. Contrary to expectation, there was no evidence for an own age bias in the level of agreement in first impressions, and minimal evidence for such a bias in impression positivity. Although differences between OA and YA on control measures were consistent with previous research, the only difference that consistently contributed to age differences in first impressions was OA slower processing speed, something that has not previously been linked to OA positivity effects. An important question for future research concerns the relative accuracy of OA and YA trait impressions from faces.

Acknowledgments

This research was supported by United States Department of Health and Human Services, National Institutes of Health, Grant AG38375 to the first author. We thank Margie Lachman for her helpful comments on a draft. We also thank Bosiljka Milosavljevic for her help in recruiting participants and collecting data, as well as Ilana Abramson, Joanne Qiao and Justine Zayhowski.

References

- Anastasi JS, Rhodes MG. An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*. 2005; 12:1043–1047. [PubMed: 16615326]
- Baltes PB, Smith J. The fascination of wisdom: Its nature, ontogeny, and function. [Special issue: From philosophical thinking to psychological empiricism]. *Perspectives on Psychological Science*. 2008; 3:56–64.
- Benton, A.; Van Allen, M.; Hamsher, K.; Levin, H. Test of facial recognition manual. Iowa City, IA: Benton Laboratory of Neuropsychology; 1983.
- Biernat M. Toward a broader view of social stereotyping. *American Psychologist*. 2003; 58:1019–1027. [PubMed: 14664690]
- Blair IV, Judd CM, Chapleau KM. The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*. 2004; 15:674–679. [PubMed: 15447638]
- Bronstad PM, Russell R. Beauty is in the “we” of the beholder: Greater agreement on facial attractiveness among close relations. *Perception*. 2007; 36:1674–1681. [PubMed: 18265847]
- Carr JM, McCormick CM, Mondloch CJ. Facial structure is a reliable cue of aggressive behavior. *Psychological Science*. 2009; 20:1194–1198. [PubMed: 19686297]
- Dion, KK. Cultural perspectives on facial attractiveness. In: Rhodes, C.; Zebrowitz, L.A., editors. *Facial attractiveness: Evolutionary, cognitive, and social perspectives*. Westport, CT: Ablex; 2002. p. 239–259.
- Eberhardt JL, Davies PG, Purdie-Vaughns VJ, Johnson SL. Looking deathworthy: Perceived stereotypicality of black defendants predicts capital sentencing outcomes. *Psychological Science*. 2006; 17:383–386. [PubMed: 16683924]
- Eichorn, DH. *Present and past in middle life*. New York, NY: Academic Press; 1981.

- Elfenbein HA, Ambady N. When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology*. 2003; 85:276–290. [PubMed: 12916570]
- Eslinger PJ, Damasio AR, Benton AL, Van Allen M. Neuropsychologic detection of abnormal mental decline in older persons. *JAMA: The Journal of the American Medical Association*. 1985; 253:670–674. [PubMed: 3968802]
- Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198. [PubMed: 1202204]
- Forgas JP. Mood and judgment: The affect infusion model. *Psychological Bulletin*. 1995; 117:39–66. [PubMed: 7870863]
- Franklin RG Jr, Adams RB Jr. A dual-process account of female facial attractiveness preferences: Sexual and nonsexual routes. *Journal of Experimental Social Psychology*. 2009; 45:1156–1159.
- Fulton A, Bartlett JC. Young and old faces in young and old beads: The factor of age in face recognition. *Psychology and Aging*. 1991; 6:623–630. [PubMed: 1777151]
- Grady CL, Hongwanisukul D, Keightley M, Lee W, Hasher L. The effect of age on memory for emotional faces. *Neuropsychology*. 2007; 21:371–380. [PubMed: 17484600]
- Greve KW, Stickler TR, Love JM, Bianchini KJ, Stanford MS. Latent structure of the Wisconsin Card Sorting Test: A confirmatory factor analytic study. *Archives of Clinical Neuropsychology*. 2005; 20:355–364. [PubMed: 15797171]
- Grohn D, Scheibe S, Baltes PB. Reduced negativity effect in older adults' memory for emotional pictures: The heterogeneity-homogeneity fist paradigm. *Psychology and Aging*. 2007; 22:644–649. [PubMed: 17874962]
- Hess TM, Auman C. Aging and social expertise: The impact of trait-diagnostic information on impressions of others. *Psychology and Aging*. 2001; 16:497–510. [PubMed: 11554526]
- Hess TM, Germain CM, Swaim EL, Osowski NL. Aging and selective engagement: The moderating impact of motivation on older adults' resource Utilization. *The Journals of Gerontology: Series B. Psychological Sciences and Social Sciences*. 2009; 64:447–456.
- Hess TM, Kotter-Grohn D. Social knowledge and goal-based influences on social information processing in adulthood. *Psychology and Aging*. 2011; 26:792–802. [PubMed: 21604887]
- Hess TM, Leclerc CM, Swaim E, Weatherbee SR. Aging and everyday judgments: The impact of motivational and processing resource factors. *Psychology and Aging*. 2009; 24:735–740. [PubMed: 19739930]
- Horn JL, Cattell RB. Age differences in fluid and crystallized intelligence. *Acta Psychologica*. 1967; 26:107–129. [PubMed: 6037305]
- Isaacowitz DM, Wadlinger HA, Goren D, Wilson HR. Selective preference in visual fixation away from negative images in old age? An eye-tracking study. *Psychology and Aging*. 2006; 21:40–48. [PubMed: 16594790]
- Isen AM, Shalke TE. The effect of feeling state on evaluation of positive, neutral, and negative stimuli: When you "accentuate the positive." do you "eliminate the negative?". *Social Psychology Quarterly*. 1982; 45:58–63.
- Ishihara, S. Tests for color-blindness. Tokyo, Japan: Kanehara Trading Company; 2010. (Original work published 1917).
- Jones D, Hill K. Criteria of facial attractiveness in five populations. *Human Nature*. 1993; 4:271–296. [PubMed: 24214367]
- Judd CM, Ryan CS, Park B. Accuracy in the judgment of in-group and out-group variability. *Journal of Personality and Social Psychology*. 1991; 61:366–379. [PubMed: 1941509]
- Judd CM, Westfall J, Kenny DA. Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*. 2012; 103:54–69. [PubMed: 22612667]
- Kalick SM, Zebrowitz LA, Langlois JH, Johnson RM. Does human facial attractiveness honestly advertise health? Longitudinal data on an evolutionary question. *Psychological Science*. 1998; 9:8–13.

- Keating CF, Bai DL. Children's attributions of social dominance from facial cues. *Child Development*. 1986; 57:1269–1276.
- Lachman, ME. Midlife development in the United States (MIDUS): Boston Study of Management Processes, 1995–1997. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research; 1997. Retrieved from <http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/203/studies/3596?q=MIDUS&searchIn=ALL&paging.startRow=1>
- Lang, PJ.; Bradley, MM.; Cuthbert, BN. International affective picture system (IAPS): Instruction manual and affective ratings. The Center for Research in Psychophysiology, University of Florida; 1999.
- Langlois JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*. 2000; 126:390–423. [PubMed: 10825783]
- Langlois JH, Roggman LA, Rieser-Danner LA. Infants' differential social responses to attractive and unattractive faces. *Developmental Psychology*. 1990; 26:153–159.
- Langlois JH, Styczynski LE. The effects of physical attractiveness on the behavioral attributions and peer preferences of acquainted children. *International Journal of Behavioral Development*. 1979; 2:325–341.
- Malatesta CZ, Izard CE, Culver C, Nicolich M. Emotion communication skills in young, middle-aged, and older women. *Psychology and Aging*. 1987; 2:193–203. [PubMed: 3268208]
- Mather M, Carstensen LL. Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences*. 2005; 9:496–502. [PubMed: 16154382]
- Mather M, Knight M. Goal-directed memory: The role of cognitive control in older adults' emotional memory. *Psychology and Aging*. 2005; 20:554–570. [PubMed: 16420131]
- Miller GA, Chapman JP. Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*. 2001; 110:40–4S. [PubMed: 11261398]
- Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*. 1993; 100:674–701. [PubMed: 8255953]
- Montepare JM, Zebrowitz-McArthur L. Children's perceptions of baby faced adults. *Perceptual and Motor Skills*. 1989; 69:467–472. [PubMed: 2812994]
- Murphy NA, Isaacowitz DM. Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging*. 2008; 23:263–286. [PubMed: 18573002]
- Olivola CY, Todorov A. Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*. 2010; 34(2):83–110.
- Perfect TJ, Harris LJ. Adult age differences in unconscious transference: Source confusion or identity blending? *Memory & Cognition*. 2003; 31:570–580. [PubMed: 12872873]
- Piper B, Li V, Eiwaz M, Kobel Y, Benice T, Chu A, Muller S. Executive function on the Psychology Experiment Building Language tests. *Behavior Research Methods*. 2012; 44:110–123. [PubMed: 21534005]
- Ruffman T, Henry JD, Livingstone V, Phillips LH. A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*. 2008; 32:863–881. [PubMed: 18276008]
- Ruffman T, Sullivan S, Edge N. Differences in the way older and younger adults rate threat in faces but not situations. *The Journals of Gerontology: Series B. Psychological Sciences*. 2006; 61:P187–P194.
- Rule NO, Ambady N, Adams RB Jr, Ozono H, Nakashima S, Yoshikawa S, Watabe M. Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology*. 2010; 98:1–15. [PubMed: 20053027]
- Salthouse TA. Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology*. 1993; 48:29–36.
- Salthouse TA. The processing-speed theory of adult age differences in cognition. *Psychological Review*. 1996; 103:403–425. [PubMed: 8759042]
- Schwarz N, Clore GL. Mood as information: 20 years later. *Psychological Inquiry*. 2003; 14:296–303.

- Shipley, WC. Institute of living scale. Los Angeles, CA: Western Psychological Services; 1946.
- Skosberg, N. The choices we make: An examination of situation selection & attentional deployment in young & older adults. Waltham, Massachusetts: Brandeis University; 2012. Manuscript in preparation.
- Stewart JE. Defendant's attractiveness as a factor in the outcome of criminal trials: An observational study. *Journal of Applied Social Psychology*. 1980; 10:348–361.
- Voelkle MC, Ebner NC, Lindenberger U, Riediger M. Let me guess how old you are: Effects of age, gender, and facial expression on perceptions of age. *Psychology and Aging*. 2012; 27:265–277. [PubMed: 21895379]
- Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*. 1988; 54:1063–1070. [PubMed: 3397865]
- Wechsler, D. Wechsler Adult Intelligence Scale-III (WAIS-III) manual. New York, NY: The Psychological Corporation; 1997.
- Willis J, Todorov A. First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science*. 2006; 17:592–598. [PubMed: 16866745]
- Wright DB, Stroud JN. Age differences in lineup identification accuracy: People are better with their own age. *Law and Human Behavior*. 2002; 26:641–654. [PubMed: 12508699]
- Zebrowitz, LA. Reading faces: Window to the soul?. Boulder, CO: Westview Press; 1997.
- Zebrowitz LA, Andreoletti C, Collins MA, Lee SY, Blumenthal J. Bright, bad, babyfaced boys: Appearance stereotypes do not always yield self-fulfilling prophecy effects. *Journal of Personality and Social Psychology*. 1998; 75:1300–1320. [PubMed: 9866189]
- Zebrowitz LA, Bronstad PM, Lee HK. The contribution of face familiarity to ingroup favoritism and stereotyping. *Social Cognition*. 2007; 25:306–338.
- Zebrowitz LA, Kikuchi M, Fellous JM. Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology*. 2010; 98(2):175–189. [PubMed: 20085393]
- Zebrowitz LA, McDonald SM. The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*. 1991; 15:603–623.
- Zebrowitz LA, Montepare JM, Lee HK. They don't all look alike: Individual impressions of other racial groups. *Journal of Personality and Social Psychology*. 1993; 65(1):85–101. [PubMed: 8355144]
- Zebrowitz LA, Olson K, Hoffman K. Stability of babyfacedness and attractiveness across the life span. *Journal of Personality and Social Psychology*. 1993; 64:453–466. [PubMed: 8468672]
- Zebrowitz LA, Rhodes G. Sensitivity to 'bad genes' and the anomalous face overgeneralization effect: Accuracy, cue validity, and cue utilization in judging intelligence and health. *Journal of Nonverbal Behavior*. 2004; 28:167–185.
- Zebrowitz LA, Wang R, Bronstad PM, Eisenberg D, Undurraga E, Reyes-Garcia V. First impressions from faces among U.S. and culturally isolated Tsimane' people in the Bolivian rainforest. *Journal of Cross-Cultural Psychology*. 2012; 43:119–134.

Figure 1.
Rater Age $\times$ Agreement Type interaction for agreement in ratings of hostile and aggressive.

Figure 2. Older adult (OA) and younger adult (YA) mean trait ratings for all faces and for negatively, neutrally, and positively valenced faces. Rater-age differences in mean ratings are indicated above the bars. + $p < .10$, * $p < .05$, ** $p < .01$.

Figure 3.
Rater Age $\times$ Face Age interaction for hostile ratings, ^{pp} $p < .01$.

Table 1

Mean Agreement in Trait Ratings

	<u>Competent</u>		<u>Healthy</u>		<u>Untrustworthy</u>		<u>Hostile</u>		<u>Aggressive</u>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Rater age										
Older adult	.49 ^a	.20	.54 ^a	.21	.37 ^a	.34	.68 ^a	.20	.38 ^a	.37
Younger adult	.41 ^b	.20	.56 ^a	.17	.39 ^a	.28	.56 ^b	.27	.61 ^b	.16
Agreement type										
Within age	.49 ^a	.21	.57 ^a	.20	.41 ^a	.33	.64 ^a	.25	.52 ^a	.31
Between age	.41 ^b	.18	.52 ^b	.18	.35 ^b	.30	.59 ^b	.25	.46 ^b	.30
Face age										
Older faces	.43 ^a	.21	.58 ^a	.19	.36 ^a	.28	.55 ^a	.25	.53 ^a	.32
Younger faces	.47 ^b	.19	.52 ^b	.19	.38 ^a	.33	.68 ^b	.24	.46 ^b	.28

Note. *N* = 48 older adult raters (OA) and *N* = 48 younger adult raters (YA) for competent, healthy, and untrustworthy ratings; *N* = 32 OA and *N* = 32 YA for hostile ratings; and *N* = 16 OA and *N* = 16 YA for aggressive ratings.

Means for contrasting groups with different superscripts (a vs. b) differ at $p < .05$ or better.

Table 2

Mean Trait Ratings

	<u>Competent</u>		<u>Healthy</u>		<u>Untrustworthy</u>		<u>Hostile</u>		<u>Aggressive</u>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Rater age										
Older adults	4.15 ^a	.66	4.46 ^a	.72	3.63 ^a	.73	2.94 ^a	.82	3.91 ^a	.53
Younger adults	4.09 ^a	.57	4.11 ^b	.58	3.91 ^b	.50	3.41 ^b	.76	3.43 ^b	.56
Face age										
Older faces	4.23 ^a	.69	4.16 ^a	.76	3.64 ^a	.73	2.91 ^a	.84	3.58 ^a	.59
Younger faces	4.01 ^b	.63	4.41 ^b	.78	3.91 ^b	.66	3.45 ^b	.86	3.76 ^b	.65

Note. *N* = 48 older adult raters (OA) and *N* = 48 younger adult raters (YA) for competent, healthy, and untrustworthy ratings; *N* = 32 OA and *N* = 32 YA for hostile ratings; and *N* = 16 OA and *N* = 16 YA for aggressive ratings.

Means for contrasting groups with different superscripts (a vs. b) differ at *p* < .05 or better.

Table 3

Older Adult and Younger Adult Scores on Control Measures

Measure	<u>Older adults</u>		<u>Younger adults</u>		<i>t</i> -value	<i>p</i> -value
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
PANAS (PA)	2.90	.70	2.78	.84	.79	.434
PANAS (NA)	1.73	.67	1.92	.66	1.37	.173
Snellen Visual Acuity (denominator)	34.69	14.27	19.75	5.47	6.77	<.001
Mars Letter Contrast Sensitivity	1.56	.19	1.73	.06	8.87	<.001
Ishihara's Test for Color Deficiency	13.46	1.87	13.48	1.65	.06	.954
Benton Facial Recognition Test	41.35	9.31	42.21	8.58	.468	.641
Pattern Comparison Test	28.17	5.16	42.56	5.85	12.79	<.001
Shipley Vocabulary Test	35.48	3.37	31.77	3.89	4.99	<.001
Wisconsin Card Sorting Test	29.03	8.92	36.40	5.56	3.59	.001
Letter-Number Sequencing Test	8.19	1.52	12.56	2.80	5.49	<.001
Education ^a	5	2-7	3	2-3		

Note.

^a Level of Education was coded for highest level attained: 1 = no high school diploma, 2 = high school diploma, 3 = some college, 4 = Bachelor's degree, 5 = some graduate work, 6 = masters degree, 7 = doctorate. Medians and range are reported.