Do Older Adults Differ From Younger Adults in Their Trait Impressions and Appearance Stereotypes?

Masters Thesis

Presented to

The Faculty of the Graduate School of Arts and Sciences
Brandeis University
Department of Psychology
Leslie Zebrowitz, Advisor

In Partial Fulfillment
of the Requirements for

Master's Degree

by

Suzanne Hillman

August, 2012
ABSTRACT

Do Older Adults Differ From Younger Adults in Their Trait Impressions and Appearance Stereotypes?

A thesis presented to the Department of Psychology

Graduate School of Arts and Sciences
Brandeis University
Waltham, Massachusetts

By Suzanne Hillman

First impressions from faces have been shown to be similar across cultures, and by both children and young adults (YAs). This study compared older adult (OA) raters with YA raters, and included both older and younger adult faces. Both age groups showed similar levels of within-age agreement in their impressions of competence, health, hostility, and trustworthiness. Consistent with evidence for age-related increases in positivity, OA showed more positive impressions of health, hostility, trustworthiness, and attractiveness of faces, as well as an own-age bias for impressions of competence. These effects were not consistently stronger in OAs for faces with more negatively-valenced trait impressions. The attractiveness halo effect was found for both OA and YA raters. The babyface stereotype was also found for both rater ages, although the effects were weaker for OAs.
# Table of Contents

Do Older Adults Differ From Younger Adults In Their Trait Impressions and Appearance Stereotypes? ........................................................................................................................................... 1

Theoretical explanation for trait ratings from first impressions of faces................................. 1
Consensus and cue utilization in first impressions from faces.................................................. 3
Effects of Rater Age on trait impressions from faces.............................................................. 5
Aims of the proposed study......................................................................................................... 8

Method................................................................................................................................... 10
Facial Stimuli............................................................................................................................. 10
Trait Ratings............................................................................................................................. 11
Control Measures..................................................................................................................... 11
Procedure................................................................................................................................ 12

Results.................................................................................................................................... 14
Inter-rater Agreement in Impressions......................................................................................... 14
Positivity of Trait Impressions.................................................................................................. 14

Rater Age Effects...................................................................................................................... 15
Moderation of Rater Age Effects by Face Age........................................................................ 16
Moderation of Rater Age Effects by Face Negativity............................................................ 17
Moderating Effects of Face Set............................................................................................... 18

Facial Appearance Stereotypes................................................................................................ 19

Attractiveness Halo.................................................................................................................. 20
Moderation of Attractiveness Halo by Negativity................................................................. 20
Babyface Stereotype ................................................................. 21

Control Data Analyses ............................................................ 22

Discussion .............................................................................. 23

Positivity ................................................................................. 23

Halo Effect and Babyfaced Stereotype ..................................... 25

Conclusions ........................................................................... 27

References .............................................................................. 28
Index of Tables

Table 1   Mean attractiveness ratings and numbers for facial stimuli..........................41
Table 2   Cronbach's alphas within sets by both rater and face age..........................42
Table 3   Standardized betas and planned comparisons between OA and YA for
          attractiveness halo, babyfaced stereotype effects, and interactions with face age group. .43
Table 4   Standardized betas for the negativity variable effects on the attractiveness halo
          ............................................................................................................................................44
Table 5   Statistics of control measures.................................................................45
Index of Figures

Figure 1  Older adult (OA) and younger adult (YA) mean trait ratings for all faces and for negatively, neutrally, and positively valenced faces........................................................................46

Figure 2  Comparison of strength of attractiveness halo on untrustworthiness ratings as they interact with negativity, based on rater age.................................................................47

Figure 3  Comparison of the strength of the babyfaced overgeneralization on health ratings as it interacts with face age group, based on rater age.........................................................48

Figure 4  Comparison of the strength of the babyface overgeneralization on hostility ratings as it interacts with face age group, based on rater age.........................................................49
Do Older Adults Differ From Younger Adults In Their Trait Impressions and Appearance Stereotypes?

Young adults (YAs) demonstrate remarkably rapid, consensual, trait ratings in their first impressions from faces (Willis & Todorov, 2006; Zebrowitz & Collins, 1997; Bar, Neta, & Linz, 2006; Hassin & Trope, 2000). Their trait impressions are modified by attractiveness and babyfacedness, as per the attractiveness halo, whereby attractive people are perceived more positivity on multiple dimensions (Eagly, Ashmore, Makhijani, & Longo, 1991) and babyfacedness overgeneralization stereotype, whereby people with facial features which are baby-like are perceived to have more child-like traits (Montepare & Zebrowitz, 1998). It is not clear how these rating abilities are affected by aging, nor how or if the stereotype effects are affected. As such, it is not known if older adults (OAs) show consensual trait impressions, have stereotype effects in their trait ratings, nor if any consensual impressions or stereotype effects agree with YAs. These questions will be addressed in this study.

*Theoretical explanation for trait ratings from first impressions of faces*

Consensus when interpreting traits from faces during very short time periods suggests an adaptive purpose in decoding these traits, implying that this ability has important survival value. This impression is supported by Bar et al. (2006), where they demonstrated that some traits such as threat (or hostility) are detected within the first 39 ms of exposure to a face. Similarly, Willis & Todorov (2006) compared trait detection speeds for five different traits (attractiveness, likability, trustworthiness, competence, and aggressiveness), and found all were detected within a 100ms viewing time, with results
comparable to those for no time limit. Support for a possible evolutionary origin for trait impressions is also available through the cross-cultural consensus between US and indigenous people from the Bolivian rainforest judges found by Zebrowitz et al. (2012), as well as analogous responses to dominance impressions between young children and adults, and attractiveness between infants and adults (Keating & Bai, 1986; Langlois, Roggman, & Rieser-Danner, 1990; Montepare & Zebrowitz-McArthur, 1989). In addition, the possibility of an evolutionary origin combines well with the ecological theory of social perception.

Ecological theory of social perception, as introduced by McArthur and Baron (1983) and grounded in Gibson's theory of object perception (1979), assumes that perception is most attuned to information that is useful in the attainment of a perceiver's goals and desires, as well as those of the species. Detection of some attributes could have enough evolutionary relevance that overestimating the presence of those traits is more strongly adaptive than failing to detect those traits in cases where they are not present. The resultant face overgeneralization effects mean that those with facial features resembling babies, emotions, or those who are ill or unfit are likely to be perceived as having traits relating to babies, specific emotions, or poor health (Zebrowitz & Collins, 1997). Some relevant examples of these overgeneralization effects include the attractiveness halo effect and the babyface stereotype. The attractiveness halo effect is a well-documented effect wherein more attractive individuals are assumed to possess more positive traits, including lower aggressiveness, better health, and higher trustworthiness and competence than their less attractive peers (Eagly, Ashmore, Makhijani, & Longo,
Such impressions may be viewed as overgeneralizations of adaptive responses to very unattractive people whose appearance is correlated with social, cognitive, or physical problems. The babyface overgeneralization effect describes the effects of evolutionarily adaptive responses appropriate for a baby being overgeneralized to individuals whose faces resemble babies (Zebrowitz et al., 2007). The traits associated with resembling a baby's face include low dominance, high trustworthiness, and low competence (Marsh, Adams, & Kleck, 2005; Zebrowitz et al., 2007; Poutvaara et al., 2009; Olivola et al., 2010; Montepare & Zebrowitz, 1998).

**Consensus and cue utilization in first impressions from faces**

We looked at impressions of health, competence, aggression, and untrustworthiness in this study, because there is existing research on consensus for these traits in YAs. There is some variation in the terminology used for each of these in the literature, due to the varying ways of interpreting, and therefore measuring, personality traits. When we refer to literature on “aggression”, we will also be including literature on dominance, threat, and danger. The idea of “competence” is similarly broad, and includes not only general impressions of competence, but also those of intelligence (intellectual competence). References to literature on “trustworthiness” also include research on honesty and conscientiousness.

Research has documented consensus among both YAs and OAs in their impressions of aggressiveness (Carré & McCormick, 2008; Sacco & Hugenberg, 2009; Said, Sebe, and Todorov, 2009; Zebrowitz, Kikuchi, & Fellous, 2007; Zebrowitz, Kikuchi, & Fellous, 2010; Zebrowitz, Wang, Bronstad, Eisenberg, Undurraga, & Reyes-
García, 2011; Zebrowitz, Franklin, Hillman, & Boc, unpublished research), competence (Ballew & Todorov, 2007; Goren, Hall, Mandisodza, Todorov, 2005; Olivola & Todorov, 2010; Poutvaara, Jordahl, & Berggren, 2009; Verhulst, Lodge, & Lavine, 2010; Zebrowitz et al., 2011; Zebrowitz et al., unpublished research), health (Zebrowitz & Rhodes, 2004; Zebrowitz et al, 2011; Zebrowitz et al., unpublished research), and trustworthiness (Willis & Todorov, 2006; Zebrowitz et al., unpublished research), although the consensus in impressions of aggressiveness was significantly lower for OAs as compared to YAs (Zebrowitz et al., unpublished research).

Some of the consensus research involving YA judges has identified specific facial cues to trait impressions. Facial cues to aggressiveness include reduced aggression impressions due to a strong resemblance to a baby's face (Said, Sebe, and Todorov, 2009; Zebrowitz et al., 2007; Sacco & Hugenberg, 2009), and increased aggression impressions due to a greater resemblance to anger expressions (Zebrowitz, Kikuchi, & Fellous, 2010). Facial cues to competence include lowered competence impressions due to a strong resemblance to a baby's face (Poutvaara, Jordahl, & Berggren, 2009; Olivola & Todorov, 2010); and increased competence impressions due to high attractiveness (Verhulst, Lodge, & Lavine, 2010; Olivola et al., 2010). Facial cues to health include increased health impressions due to high facial attractiveness, or its component qualities (Zebrowitz et al., 2002; Zebrowitz & Rhodes, 2004; Zebrowitz et al, 2011; Zebrowitz et al., unpublished research). Facial cues to trustworthiness include increased trustworthiness impressions due to resemblance to happiness expressions, reduced trustworthiness impressions due to resemblance to anger impressions (Said et al. 2009; Oosterhof &
Todorov, 2008; Zebrowitz et al., 2010), and increased trustworthiness impressions due to resemblance to a baby's face (Montepare & Zebrowitz, 1998).

**Effects of Rater Age on trait impressions from faces**

*Consensus in impressions.* OAs have not been investigated for consensus in trait impressions to the same extent as YA. Zebrowitz et al. (unpublished research) found that both OA and YA showed significant consensus in their facial impressions of aggressiveness, competence, health, and untrustworthiness, as well as their impressions of attractiveness and babyfaceness. Aggressiveness was the only trait for which OA impressions showed significantly lower consensus than YA. In addition, both YA and OA generally showed greater within- than between-age agreement, except for OA impressions of aggressiveness and untrustworthiness. The tendency toward greater within-age agreement suggests that some criteria for forming first impressions from faces are age specific.

*Positivity of impressions.* Research provides little evidence of differences in the amount of agreement among OAs as compared to YAs, but there is reason to believe that there may be effects of rater age on the positivity of impressions. For example, Carstensen's socioemotional selectivity theory suggests that perceived time limitations trigger a shift in older adulthood toward emotionally meaningful goals over information seeking goals (Carstensen, 1995), increasing motivation to focus more on the emotional aspects of an experience in order to improve their affect. Consistent with this hypothesis, research shows that OAs, but not YAs, demonstrate biased attention for neutral faces over negative ones and biased memory for positive over neutral faces (Charles, Mather,
& Carstensen, 2003; Mather & Carstensen, 2003; Allard & Isaacowitz, 2008). OA also tend to detect less negativity in ambiguously emotional faces, tend to label negative expressions less negatively than their YA counterparts (Sullivan & Ruffman, 2004; Calder et al., 2003), and do not perform as well as YAs at recognizing negatively-valenced emotions like fear, sadness, and sometimes anger (Ruffman, Henry, Livingstone, & Phillip, 2008; Isaacowitz & Stanley, 2011).

While research on emotion perception has supported the hypothesis that OAs are less sensitive to negatively-valenced information in faces, research on trait impressions is more mixed. For example, Ruffman, Sullivan, & Edge (2006) found that OAs rated highly dangerous faces as less dangerous than YAs, with no difference for medium and low danger in faces. This effect was specific to faces, as OAs and YAs showed no difference in rating dangerous situations. Similarly, Stanley & Blanchard-Fields (2008) found that YAs were better than OAs at detecting deceit in a video about a crime when there was a visual component, with the OA difficulty with detection of fear and shame accounting for the differences between the two. In contrast, Zebrowitz et al. (unpublished manuscript) found that OAs had higher ratings of aggression in all faces, lower health ratings in older or male faces, and higher untrustworthiness ratings in older male faces than YAs.

We believe that variable interpretations of the term 'aggression' may explain the reduced OA agreement and the rater age-based differences in mean ratings of aggression. Specifically, OAs may have had less agreement in their interpretation of the term, with many understanding it to mean 'assertiveness' rather than 'hostility', thereby reducing their
mean ratings and consensus as compared to YAs. To investigate this possibility, the present study substituted the word “hostility” for “aggressiveness” to determine whether this less ambiguous term eliminates the age differences in inter-rater agreement and mean ratings.

One possible explanation for the Zebrowitz et al. (unpublished manuscript) finding that OAs had more negative trait impressions for the health and trustworthiness ratings in older faces is an “own age bias” (Backman, 1991; Lamont et al., 2005), wherein both children and OAs are better at recognizing faces of their own age (Anastasi & Rhodes, 2005; Fulton & Bartlett, 1991). A recent meta-analysis of the own-age bias in facial recognition found that recognition within own's own age group had more hits and fewer misses than for other age groups (Rhodes & Anastasi, 2011). OA show increased selectivity in use of cognitive resources over YA (Hess, 2006), and tasks with higher personal relevance or social accountability enhance the performance of OA more than YA (Hess, Leclerc, et al., 2009; Hess, Germain, et al., 2009). Therefore, due to the increased personal relevance of own age-faces, OAs may be expected to have better agreement in their detection of emotional expressions, and the associated trait impressions, of faces of their own age. This could mean that the positivity effect is specific to OAs looking at younger adult faces, with the own-age bias reducing or removing this effect for older adult faces. Consistent with this hypothesis, one study found a non-significant reversal in the tendency for YA to surpass OA in recognizing emotions when both groups are judging older faces (Malatesta et al., 1987), and others have found that OA are better than YA at recognizing emotions on older adult faces,
although both YA and OA are better at recognizing emotions on younger adult faces (Ebner et al., 2009; Hummert et al., 2002; Isaacowitz & Stanley, 2011)

Attractiveness Halo Effect and Babyface Stereotype. Investigation of the attractiveness halo effect has shown that both YA and OA show consensus in ratings of attractiveness (Langlois & Roggman, 1990; Langlois, Roggman, & Rieser-Danner, 1990; Zebrowitz et al., 2011; Zebrowitz et al., unpublished research), and the attractiveness halo effect has been demonstrated for OA as well as YA judges (Dion, Berscheid, & Walster, 1972; Larose & Standing, 1998; Zebrowitz et al., unpublished research). In addition, the study by Zebrowitz et al. (unpublished research) found that OAs had higher attractiveness ratings than YAs for both older and younger faces.

YAs have been shown to demonstrate the babyface stereotype (Marsh et al., 2005; Montepare & Zebrowitz, 1998; Zebrowitz et al., 2007; Poutvaara et al., 2009; Olivola et al., 2010), while a recent study in this lab (Zebrowitz et al., unpublished research) found that OAs showed no babyface stereotype in their impressions of young adult faces. However, they did associate higher aggressiveness and untrustworthiness with very low babyfacedness in older faces.

Aims of the proposed study

The present investigation extended previous research documenting consensus of first impressions of faces to include OA participants and older and younger faces. We expected OAs to show similar consensus within-group to that of YAs. We also expected that OA would show more positive trait impressions than YA for both older and younger faces. We expected that using a larger set of faces than that used in Zebrowitz et al.
(unpublished manuscript) would successfully demonstrate the extensibility of OA positivity effects to trait impressions, as found in other research (Sullivan & Ruffman, 2004; Calder et al., 2003; Ruffman, Henry, Livingstone, & Phillip, 2008; Isaacowitz & Stanley, 2011). We also expected that the own age bias would increase positivity in OA for older faces. However, another possibility based on Zebrowitz et al. (unpublished manuscript) was that OAs would have reduced positivity for older faces, perhaps due to more in-depth processing of these faces. We expected that greater OA positivity would be particularly marked for the most negatively-valenced faces. We expected a similar attractiveness halo effect in OAs and YAs, such that more attractive faces were rated lower in hostility and untrustworthiness, and higher in health and competence. We expected a similar babyfacedness stereotype in OAs and YAs, with more babyfacedness resulting in lower hostility, untrustworthiness, and competence ratings. A second possibility, based on Zebrowitz et al. (unpublished manuscript), was that OAs would show no babyface stereotype for younger faces, again perhaps due to more in-depth processing of own-age faces.
Method

Participants

32 YA participants (16 females) were recruited from the undergraduate population taking Introduction to Psychology at Brandeis University, and 32 OAs (16 females) from the Greater Boston area, were recruited through fliers in senior centers, ads in newsletters, and existing databases of previous participants. All OA were screened using the Mini-Mental State Examination (Folstein, M, Folstein, S, & McHugh, 1975) with a minimum score of 26 required for inclusion. All participants were Caucasian. The YAs were compensated with 2 class credits, and the OAs with $25. All participants were given a consent form to read and sign, and were debriefed at the end of the session.

Facial Stimuli

Grayscale images of 148 people (74 men) drawn from a sample of men and women from the Intergenerational Studies (IGS) archive were used in this study. The IGS was 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).

Two images of each person were used, one collected when the individuals were 17-18, and one collected when they were between 52 and 62. In order to limit the amount of time spent rating faces to an hour, we removed faces of poor image quality from our
initial set of age-matched faces, and then split the entire set of 296 faces (148 men) into two equivalent sets matched on previously collected attractiveness ratings\(^1\).

The faces naturally separated into groups based on age and sex, which resulted in 4 groups of 74 faces each for each age and sex pairing. We then split each group into two sets of equivalent attractiveness ratings, resulting in a total of 8 groups each containing 37 faces. Each set, A and B, contained 148 faces spread equally across age and sex (Table 1). Because individuals do not necessarily have the same attractiveness ratings at various ages, some had their two faces in different sets.

**Trait Ratings**

Four traits (healthy, hostile, competent, untrustworthy) and two appearance qualities (attractive, babyfaced) were rated on 7-point bipolar scales, with endpoints labeled (1) *not at all* (healthy, competent, untrustworthy, hostile, attractive, babyfaced) and (7) *very* (healthy, competent, untrustworthy, hostile, attractive, babyfaced).

**Control Measures**

We administered several control measures to assess age differences in emotional state, visual abilities, executive function, and cognitive function. These included measures assessing visual acuity (Snellen Eye Chart); color vision (Ishihara’s Tests for Color Deficiency, Ishihara, 2010); contrast sensitivity (Mars Letter Contrast Sensitivity Test, Mars Perceptrix, Chappaqua, NY), face perception, (Benton Facial Recognition Test, Benton et al., 1983) executive function (Short-form Wisconsin Card sorting task, E. A. Berg., 1948), general cognitive abilities (Mini–mental state examination, Folstein, M, Folstein, S, & McHugh, 1975; Shipley Vocabulary Test, Shipley, 1946; Pattern
Comparison Task; Salthouse, 1993), affect (Positive And Negative Affect Scale, or PANAS; Watson, Clark, & Tellegen, 1988), and emotion recognition (Mind in the Eyes; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001).

**Procedure**

After obtaining informed consent, participants were seated in front of a computer screen. They first completed a computerized version of the Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988). Next, MediaLab software (Empirisoft, New York City, NY) was used to present the trait rating task.

Participants were randomly assigned to either face set A or B, and one of two scale orders: healthy, untrustworthy, competent, aggressive, attractive, and babyfaced, or aggressive, competent, untrustworthy, healthy, babyfaced, and attractive. Within each scale, faces were shown in one of four orders, counterbalancing age and sex of face: (1) younger men, younger women, older men, older women (2) younger women, younger men, older women, older men (3) older men, older women, younger men, younger women, or (4) older women, older men, younger women, younger men. Participants were asked to rate each face in comparison to the other faces of that age/sex grouping so that variation in ratings would not merely reflect different responses to faces of different ages or sexes. During the scale rating, participants viewed each face for three seconds and then rated it on the 7 point scale after the face disappeared. The scale was on the screen until the rating was made, at which time it disappeared and was replaced with a new face. There was a five minute break halfway through the ratings. The entire rating portion
lasted about an hour, after which a demographic and health questionnaire and the control measures were administered.
Results

Inter-rater Agreement in Impressions

We found that a younger male participant rated all the older faces as 1 on babyfaceness, and dropped his older adult faces babyfacedness ratings from our analysis. We also found that an older male participant rated all faces, regardless of age, as 3 on babyfaceness, and his competence ratings showed a negative correlation with the other raters. We therefore dropped his babyfacedness and competence ratings from our analysis.

With one exception, Cronbach alphas showed acceptable inter-rater reliabilities for the individual trait and appearance ratings of both OA and YA judges, suggesting that OAs, like YAs, show consensual first impressions from faces (Table 2). These high inter-rater reliabilities justified data analyses using mean ratings for each face across judges as the unit of analysis. The exception to acceptable reliability was YAs ratings of untrustworthiness in set A, which was slightly under the convention of .70 (α = .679), and which could not be raised by dropping any deviant participants. Nevertheless, we included this set of YAs untrustworthiness ratings in our analyses for purposes of comparing with OAs.

Positivity of Trait Impressions

To test the hypothesis that OAs would rate faces more positively than YAs, we conducted a 2 (rater age: OA vs. YA) by 2 (face age: older vs. younger) X 2 (face sex:
female vs. male) X 2 (face set: A or B) ANOVA for each of the 4 trait ratings, as well as for babyfaceness and attractiveness. Face was the unit of analysis, with rater age a within-face variable, and face age, face sex, and face set between-face variables. We included face sex in the analysis to account for related variance, but did not report any face sex effects because we had no predictions for face sex. Face set had no theoretical meaning, and therefore effects are only reported when they qualified rater age or face age effects that held true when collapsing across face set.

Because there was reason to expect a greater age difference between mean ratings for negatively-valenced faces, we performed the same ANOVA design above with the negativity variable for each trait as an additional factor. This variable was obtained by identifying faces rated as above the median in the specific trait by both OAs and YAs, below the median in the same trait by both rater ages, and the remaining faces. We report negativity effects only when we found a significant interaction between rater age and the negativity factor.

**Rater Age Effects.**

Contrary to prediction, OA did not rate the competence of faces more positively \( (M = 4.11, SD = .03) \) than YA \( (M = 4.12, SD = .036) \), \( F(1,288) = 1.256, p = .263, \eta^2 = .004 \). As predicted, OAs showed greater positivity \( (M = 4.57, SD = .04) \) in their ratings of health, \( F(1,288) = 338.38, p < .001, \eta^2 = .54 \), as compared to YAs \( (M = 4.12, SD = .04) \). Also as predicted, OAs showed greater positivity \( (M = 2.94, SD = .04) \) in their ratings of hostility, \( F(1,288) = 312.95, p < .001, \eta^2 = .52 \), as compared to YAs \( (M = 3.41, SD = .05) \). As predicted, OAs showed greater positivity \( (M = 3.53, SD = .03) \) in their ratings of
untrustworthiness, \( F(1,288) = 158.08, p < .001, \eta^2 = .35 \), as compared to YAs (\( M = 3.89, SD = .04 \)). As predicted, OAs showed greater positivity (\( M = 3.55, SD = .04 \)) in ratings of attractiveness, \( F(1,288) = 323.53, p < .001, \eta^2 = .53 \), than YAs (\( M = 2.99, SD = .05 \)). A main effect of rater age, \( F(1,228) = 22.05, p < .001, \eta^2 = .07 \), showed that OAs rated faces as less babyfaced (\( M = 3.03, SD = .04 \)) than YAs (\( M = 3.17, SD = .04 \)).

**Moderation of Rater Age Effects by Face Age.**

The lack of a rater age main effect for competence was qualified by an interaction between face age and rater age, \( F(1,288) = 25.37, p < .001, \eta^2 = .08 \). Older faces were rated as more competent by OAs (\( M = 4.32, SD = .05 \)) than by YAs (\( M = 4.22, SD = .05 \)), \( p = .006 \) and younger faces were rated as more competent by YAs (\( M = 4.06, SD = .05 \)) than by OAs (\( M = 3.89, SD = .05 \)), \( p < .001 \). The greater positivity in ratings of attractiveness by OAs was qualified by an interaction between face age and rater age, \( F(1,288) = 24.03, p < .001, \eta^2 = .08 \). The difference between ratings by OAs and YAs was greater for older faces than younger faces, although the rater age effect was significant in both cases (older faces: OA \( M = 3.49, SD = .06 \); YA \( M = 2.77, SD = .07 \); \( p < .001 \). Younger faces: OA \( M = 3.61, SD = .06 \); YA \( M = 3.20, SD = .07 \); \( p < .001 \)). In addition, an overall tendency for older faces to be rated as less attractive (\( M = 3.13, SD = .06 \)) than younger faces (\( M = 3.41, SD = .06 \)), \( F(1,288) = 9.79, p = .002, \eta^2 = .033 \) was significant only for YA raters, \( p < .001 \). The greater positivity in ratings of hostility by OAs was qualified by an interaction between face age and rater age, \( F(1,288) = 19.900, p < .001, \eta^2 = .065 \). The difference between ratings by OAs and YAs was greater for older faces than for younger faces, although the rater age effect was significant in both cases.
(older faces: OA $M = 2.61$, $SD = .06$; YA $M = 3.19$, $SD = .07$; $p < .001$. Younger faces: OA $M = 3.27$, $SD = .06$; YA $M = 3.62$, $SD = .06$; $p < .001$). In addition, there was an overall tendency for older faces to be rated as less hostile ($M = 2.90$, $SD = .06$) than younger faces ($M = 3.45$, $SD = .06$), $F(1,288) = 39.20$, $p < .001$, $\eta^2 = .12$.

The rater age effects for health, untrustworthiness, and babyfacedness were not moderated by face age, $F$s < 1.87, $ps > .173$, indicating that OAs and YAs showed equivalent tendencies to rate older faces ($M = 4.24$, $SD = .05$) as less healthy than younger faces ($M = 4.44$, $SD = .05$), $F(1,288) = 7.37$, $p = .007$, $\eta^2 = .03$; less untrustworthy ($M = 3.54$, $SD = .04$) than younger faces ($M = 3.87$, $SD = .04$), $F(1,288) = 29.11$, $p < .001$, $\eta^2 = .09$; and less babyfaced ($M = 2.77$, $SD = .05$) than younger faces ($M = 3.44$, $SD = .05$), $F(1,288) = 81.65$, $p < .001$, $\eta^2 = .22$.

**Moderation of Rater Age Effects by Face Negativity.**

The overall tendency for OAs to rate faces as more healthy was qualified by an interaction with the face negativity factor ($F(2,271)= 7.19$, $p = .001$, $\eta^2 = .05$), with the biggest age difference for more positive-looking faces above the median in health (OA: $M = 5.12$, $SD = .03$; YA: $M = 4.76$, $SD = .04$; $p < .001$), smaller for faces at the median (OA: $M = 4.55$, $SD = .04$; YA: $M = 4.13$, $SD = .05$; $p < .001$), and smallest for more negative-looking faces below the median (OA: $M = 4.00$, $SD = .03$; YA: $M = 3.44$, $SD = .04$; $p < .001$; Figure 1). The overall tendency for OAs to rate faces as less untrustworthy than YAs was qualified by an interaction with the negativity factor ($F(2,271) = 23.56$, $p < .001$, $\eta^2 = .15$), with the biggest age difference for more negative looking faces above the median in untrustworthiness (OA: $M = 3.96$, $SD = .03$; YA: $M = 3.44$, $SD = .04$; $p < .001$, $\eta^2 = .22$).
4.51, SD = .04; p < .001), smaller with faces at the median (OA: M = 3.55, SD = .03; YA: M = 3.94, SD = .05; p < .001), and smallest with faces below the median in untrustworthiness (OA: M = 3.10, SD = .03; YA: M = 3.22, SD = .04; p = .014; Figure 1). The overall tendency for OAs to rate faces as more attractive than YAs was qualified by the negativity factor, F(2,271) = 6.28, p = .002, η² = .04, with the biggest age difference for more negative-looking faces below the median in attractiveness (OA: M = 2.88, SD = .04; YA: M = 2.19, SD = .05; p < .001), smaller for the faces at the median (OA: M = 3.49, SD = .05; YA: M = 2.93, SD = .06; p < .001), and smallest for faces above the median (OA: M = 4.25, SD = .04; YA: M = 3.82, SD = .05; p < .001). The rater age effect for babyfacedness was modified by the negativity factor (F(2,271) = 5.42, p = .005, η² = .038). The overall OA tendency to rate faces as less babyfaced than YAs was significant for faces at the median (OA: M = 3.00, SD = .04; YA: M = 3.16, SD = .05; p = .003) and above the median (OA: M = 3.60, SD = .04; YA: M = 3.85, SD = .05; p < .001), but not below the median (OA: M = 2.45, SD = .03; YA: M = 2.47, SD = .05; p = .608).

The rater age positivity effects for hostility and competence were not qualified by the rated negativity of the faces on those dimensions, respective Fs(2,271) = 1.06 and 1.96, ps > .142.

**Moderating Effects of Face Set.**

The lower hostility ratings by OAs compared to YAs were qualified by face set, F(1,288) = 233.181, p < .001, η² = .447. In set A, OAs and YAs rated faces as equally hostile (OA: M = 3.12, SD = .06; YA: M = 3.18, SD = .07; p = .088). In set B, OAs rated
faces as less hostile than YAs (OA: $M = 2.77, SD = .06$; YA: $M = 3.64, SD = .07, p < .001$). The higher attractiveness ratings by OAs compared to YAs were moderated by face set, $F(1,288) = 17.41, p < .001$, $\eta^2 = .057$. The difference between ratings by OAs and YAs was greater in set A (OA: $M = 3.69, SD = .06$; YA: $M = 3.00, SD = .07$; $p < .001$) than in set B (OA: $M = 3.41, SD = .06$; YA: $M = 2.98, SD = .07$; $p < .001$), with OAs giving higher ratings than YA in both sets. The higher babyfacedness ratings by YAs compared to OAs were qualified by an interaction between rater age and face set, $F(1,228) = 100.04, p < .001$, $\eta^2 = .26$. In set A, OAs rated faces as more babyfaced than YAs (OA: $M = 3.32, SD = .05$; YA: $M = 3.16, SD = .06, p < .001$), whereas in set B, OAs rated faces as less babyfaced than YAs (OA: $M = 2.75, SD = .06$; YA: $M = 3.19, SD = .06, p < .001$).

**Facial Appearance Stereotypes**

In order to examine the babyface stereotype and attractiveness halo, we ran regression analyses with the mean trait ratings for each face as the dependent variable, and mean ratings of their attractiveness and babyfacedness, face age group, face sex, and the interactions between face age group and both appearance quality mean ratings. These regression analyses were computed separately for OAs and YAs, and the magnitude of the effects for OA and YA were compared using the method described in Rosenthal and Rosnow (1991). In this method, t-values for OA and YA were converted to z-scores, and the significance of the difference between each pair of z-scores was calculated. This was done for attractiveness, babyfacedness, and the interaction of face age with each of those appearance qualities.
Due to the expectation that rater age differences would be greater for more negatively valenced faces, we also performed the multiple regression analysis described above for attractiveness with the negativity variable for attractiveness as an additional factor. This variable was obtained by identifying faces rated as above the median in attractiveness by both OAs and YAs, below the median in attractiveness by both rater ages, and the remaining faces. We report negativity effects only when we found significant interactions between attractiveness and the negativity factor.

**Attractiveness Halo.**

Controlling for babyfacedness ratings, face sex, and face age group, both OA and YA showed the attractiveness halo effect, with significantly more positive impressions of more attractive faces of both ages on all trait dimensions, all $p < .001$ (Table 3). There were no significant effects of rater age on the strength of the halo effect, all betas $< 0.15$, $p > .10$. There were no significant effects of face age on the strength of the halo effect, all $p > .05$. Planned comparisons between the attractiveness effect and interaction of this effect and face age as shown by OAs and YAs revealed no significant rater age differences, all $z < 1.29$, $p > .10$ (Table 3).

**Moderation of Attractiveness Halo by Negativity.**

When looking at the effect of negativity on ratings of faces (Table 4), YAs showed a marginally significant effect of face negativity for the halo effect on impressions of untrustworthiness (Figure 2, Panel 1), $p < .10$. Faces above the median in attractiveness showed significance of $p < .001$, below the median in attractiveness showed significance of $p = .004$, and at the medium showed significance of $p = .045$. In
addition, the effect of attractiveness negativity on the halo effect for untrustworthiness was significantly greater for YAs than OAs, \( p < .01 \) (Table 4; Figure 2). The strength of the halo effect on other trait impressions was not significantly influenced by whether faces were high or low in attractiveness, all \( ps > .130 \).

**Babyface Stereotype.**

Controlling for attractiveness ratings and the sex and age group of the faces, both YAs and OAs showed the babyface stereotype in their impressions of the untrustworthiness of faces, with those who were more babyfaced judged as less untrustworthy, \( ps < .05 \). YAs also showed the babyface stereotype in their impressions of hostility, with those who were more babyfaced judged as less hostile, \( p < .01 \), whereas this effect was not significant for OA, \( p = .418 \) (Table 3). However, OAs hostility impressions showed an interaction effect between babyfacedness and face age group (Figure 4: Panel 1), \( p < .01 \), wherein younger faces showed significance of \( p < .001 \), and older faces showed no significance, \( p = .427 \).

Planned comparisons between the babyfacedness overgeneralization effects for OA and YA, collapsing across face age, revealed that OAs, \( \beta = .01 \), showed a weaker relationship between babyfacedness and rated competence than did YAs, \( \beta = -.14, p < .05 \). Similarly, OAs showed a weaker relationship between babyfacedness and health (OA: \( \beta = .01 \), YA: \( \beta = -.11 \)), and babyfacedness and hostility (OA: \( \beta = -.06 \), YA: \( \beta = -.29 \)) than did YAs, both \( ps < .05 \). However, OAs, \( \beta = -.24 \), showed a marginally significant stronger relationship between babyfacedness and untrustworthiness than did YAs, \( \beta = -.16, p < .10 \) (Table 3).
In addition, the babyfacedness by face age interaction for impressions of health was significantly stronger for OAs than YAs, $p < .05$, (Figure 3). Conversely, the babyfaceness by face age interaction for impressions of hostility was marginally weaker for OAs than YAs, $p < .10$, (Figure 4).

**Control Data Analyses.**

Control data analyses revealed several differences between OA and YA (see Table 5). Visual acuity (Snellen Eye Chart) and Contrast sensitivity without reading glasses were lower for OAs than for YAs, but in the normal range for both age groups. There were no significant age differences in color vision (Ishihara’s Tests for Color Deficiency, Ishihara, 2010), on the Eyes Test (Baron-Cohen et al., 2001), the Benton Facial Recognition Test (Benton, Van Allen, Hamsher, & Levin, 1983). There were no age differences in positive affect on the PANAS, but OA had marginally lower negative affect, consistent with previous research (Windsor & Anstey, 2010). OAs scored higher than YAs on the Shipley Vocabulary test (Shipley, 1946), consistent with their higher education level and the maintenance of crystallized intelligence in older adulthood (Horn & Cattell, 1967). The lower scores of OAs than YAs on a timed Pattern Comparison Task (Salthouse, 1993) are consistent with decreases in processing speed and executive control in older adulthood (Salthouse, 1996). OAs performed worse on the WCST than YAs, indicating a decline in executive function, and had significantly more perseverative errors, which are the most diagnostic of executive dysfunction (Greve, Stickle, Love, Bianchini, & Stanford, 2005).
Discussion

These findings reveal that the tendency to form consensual trait impressions from faces persists into older adulthood. There also were reliable differences between OA and YA impressions, most of which reflected greater positivity for OA, as predicted. The attractiveness halo effect was found for both OA and YA raters, as was the babyface stereotype, although the babyface effects were weaker for OA. We did not find any consistent tendency for the OA positivity effect in trait impressions to be stronger for more negatively-valenced faces.

Positivity

As expected, OAs were more positive in their impressions of health, hostility, untrustworthiness, and attractiveness. These results are consistent with existing research on age differences in emotion recognition as described by the positivity bias, wherein OA raters demonstrate a tendency to interpret emotional expressions more positively than their YA counterparts (Sullivan & Ruffman, 2004; Calder et al., 2003; Ruffman, Henry, Livingstone, & Phillip, 2008; Isaacowitz & Stanley, 2011). These results also suggest that this bias extends to trait impressions, as Ruffman, Sullivan, & Edge (2006) found for impressions of danger from faces and Stanley & Blanchard-Fields (2008) found for impressions of deceit. These results support the supposition that a larger group of faces than that used in Zebrowitz et al. (unpublished research) would demonstrate the positivity effect.
Contrary to expectations, OAs were not more positive than YAs in their overall impressions of competence. However, both age groups rated faces of their own age group more positively than faces of the other age group, suggesting that competence has a stronger own-age bias than for the other trait impressions, for which there was simply a positivity effect. The research we cited earlier indicated that the own-age bias was found for recognition of faces (Anastasi & Rhodes, 2005; Fulton & Bartlett, 1991; Rhodes & Anastasi, 2011), with some effects for emotion recognition (Malatesta et al., 1987; Ebner et al., 2009; Hummert et al., 2002; Isaacowitz & Stanley, 2011). This suggests that this bias could reasonably be extended into trait impressions. In addition, as we suggested earlier, this bias could be expected to increase the positivity effect for OAs rating older faces, as was found for competence. We also found that the tendency for OAs to rate faces as more attractive and less hostile than YAs was greater for older faces than younger faces, partly confirming our hypothesis that the OA positivity effect would be stronger for older faces. As there were no rater age by face age interactions for health or untrustworthiness, however, our hypothesis was not completely confirmed.

The tendency for OAs to rate faces as more healthy than YAs was the most pronounced for more positive-looking faces above the median in rated health, counter to our prediction that OA positivity effects would be greater for the most negative-looking faces. Conversely, the tendency for OAs to rate faces as less untrustworthy than YAs was the most pronounced for more negative-looking faces above the median in untrustworthiness, confirming our hypothesis that OA positivity effects would be greater for the most negative-looking faces. Similarly, the tendency for OAs to rate faces as more
attractive than YAs was the most pronounced for more negative-looking faces below the median in attractiveness, also confirming this hypothesis. Competence and hostility showed no difference in rater age effects between faces above or below the median for the respective trait, counter to our prediction that OA positivity effects would be greater for the most negative-looking faces. It is possible that having performed our analyses on faces rather than raters, thereby collapsing information across raters, may have lost variations between the raters which could have shown negativity effects for competence and hostility.

The overall positivity effect for OAs wherein they rated faces as less hostile than YAs only held true for face set B, in which they also rated faces as less babyfaced than YAs. At the same time, they only rated faces as more attractive than YAs in face set A.

**Halo Effect and Babyfaced Stereotype**

As predicted, the attractiveness halo held true across all four traits, competence, health, hostility, and untrustworthiness, with no significant differences in the effects between OAs and YAs. These results extend existing evidence for a halo effect in OA (Larose & Standing, 1998) by demonstrating that the effect is equivalent in magnitude to that shown by YA. These results also show that the halo effect holds true for both older and younger faces, including when variations in attractiveness are not confounded by variations in babyfaceness or face sex, which were not controlled in the Larose and Standing (1998) study.

Unlike the equally strong halo effect shown by OAs and YAs, OAs showed a weaker babyface stereotype. YAs associated significantly less hostility with more
babyfaced older and younger adult faces, whereas OAs did not. On the other hand, both OAs and YAs associated significantly less untrustworthiness with more babyfaced older and younger faces. These results may shed light on the high rate of fraud on OAs (Mackin, 1994), as they may be just as biased as YAs by babyfacedness for ratings of untrustworthiness, while being more vulnerable due to cognitive declines (Salthouse, 1996). The tendency for greater babyfacedness to reduce impressions of hostility was stronger for younger than for older faces. Neither age group showed the babyface stereotype for competence. However, competence is a concept encompassing many possible interpretations, including intelligence and naivete, perhaps explaining this lack. Research by McArthur and Apatow (1984) found no relationship between perceived intelligence and babyfacedness, but did find that babyfacedness individuals are associated with greater perceived naivete.

OAs had a weaker relationship between babyfacedness and ratings of competence, health, and hostility than YAs, and a marginally significant stronger relationship between babyfacedness and untrustworthiness than YAs. This suggests that OAs are generally less affected by the babyfaced stereotype than YAs, except possibly in the case of untrustworthiness which, as noted above, may increase their vulnerability to fraud. The reduced babyfacedness stereotype in OA impressions of hostility could minimize a YA tendency to believe that mature faced defendants are more likely to be guilty of crimes involving intention than are more babyfaced defendants (Zebrowitz & McDonald, 1991).
Conclusions

Like YAs, OAs form consensual first impressions from faces. However, there were significant age differences in the positivity of impressions. OAs judged faces as healthier, more hostile, less untrustworthy, and more attractive than YAs. OAs judged older faces as more competent, more attractive, and less hostile than younger faces, partially supporting our hypothesis that the own-age effect would increase ratings of older faces by OAs. The greater positivity of OA impressions was stronger for healthy faces with a positive valance, for attractive and untrustworthy faces with a negative valance, and had no negativity effect for hostility and competence. These inconsistent results may have been due to the use of face, not rater, as the unit of analysis. We found similar attractiveness halo effects in OAs and YAs, while OAs tended to have reduced effects of the babyfacedness stereotype as compared to YAs.
References


Footnotes

1 Both facial images of all 148 men and women were previously rated on a 7-point scale of attractiveness and babyfaceness (Zebrowitz et al., 1993).
Table 1

Mean attractiveness ratings and numbers for facial stimuli

<table>
<thead>
<tr>
<th>Set</th>
<th>Male Faces</th>
<th></th>
<th>Female Faces</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Younger Adult</td>
<td>Older Adult Raters (M = 3.41)</td>
<td>Younger Adult</td>
<td>Older Adult Raters (M = 3.06)</td>
</tr>
<tr>
<td></td>
<td>(M = 3.51)</td>
<td></td>
<td>(M = 3.06)</td>
<td></td>
</tr>
<tr>
<td>Set 1</td>
<td>M = 2.98, N = 37</td>
<td>M = 3.67, N = 37</td>
<td>M = 3.01, N = 37</td>
<td>M = 3.70, N = 37</td>
</tr>
<tr>
<td>Set 2</td>
<td>M = 3.06, N = 37</td>
<td>M = 3.28, N = 37</td>
<td>M = 2.89, N = 37</td>
<td>M = 3.54, N = 37</td>
</tr>
</tbody>
</table>
Table 2

Cronbach's alphas within sets by both rater and face age.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Younger raters, set A</th>
<th>Younger raters, set B</th>
<th>Older raters, set A</th>
<th>Older raters, set B</th>
<th>Older faces, set A</th>
<th>Older faces, set B</th>
<th>Younger faces, set A</th>
<th>Younger faces, set B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hostile</td>
<td>0.866</td>
<td>0.865</td>
<td>0.914</td>
<td>0.914</td>
<td>0.928</td>
<td>0.907</td>
<td>0.946</td>
<td>0.942</td>
</tr>
<tr>
<td>Attractiveness</td>
<td>0.861</td>
<td>0.902</td>
<td>0.844</td>
<td>0.881</td>
<td>0.904</td>
<td>0.95</td>
<td>0.939</td>
<td>0.927</td>
</tr>
<tr>
<td>Babyfacesness</td>
<td>0.805</td>
<td>0.815</td>
<td>0.785</td>
<td>0.784</td>
<td>0.881</td>
<td>0.817</td>
<td>0.869</td>
<td>0.871</td>
</tr>
<tr>
<td>Competence</td>
<td>0.759</td>
<td>0.738</td>
<td>0.783</td>
<td>0.835</td>
<td>0.831</td>
<td>0.88</td>
<td>0.895</td>
<td>0.834</td>
</tr>
<tr>
<td>Health</td>
<td>0.861</td>
<td>0.828</td>
<td>0.775</td>
<td>0.803</td>
<td>0.895</td>
<td>0.932</td>
<td>0.924</td>
<td>0.888</td>
</tr>
<tr>
<td>Untrustworthiness</td>
<td>0.679</td>
<td>0.831</td>
<td>0.71</td>
<td>0.721</td>
<td>0.773</td>
<td>0.872</td>
<td>0.814</td>
<td>0.875</td>
</tr>
</tbody>
</table>
Table 3

Standardized betas and planned comparisons between OA and YA for attractiveness halo, babyfaced stereotype effects, and interactions with face age group

<table>
<thead>
<tr>
<th></th>
<th>Competence</th>
<th>Health</th>
<th>Hostility</th>
<th>Untrustworthiness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Older Raters</td>
<td>Younger Raters</td>
<td>Older Raters</td>
<td>Younger Raters</td>
</tr>
<tr>
<td>Attractiveness</td>
<td>0.737**</td>
<td>0.782**</td>
<td>0.891**</td>
<td>0.856**</td>
</tr>
<tr>
<td>Babyfacedness</td>
<td>0.005&lt;sub&gt;a&lt;/sub&gt;</td>
<td>-0.136&lt;sub&gt;b&lt;/sub&gt;</td>
<td>0.014&lt;sub&gt;a&lt;/sub&gt;</td>
<td>-0.106&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>Attractiveness X Age</td>
<td>-0.015</td>
<td>-0.094</td>
<td>-0.094</td>
<td>-0.033</td>
</tr>
<tr>
<td>Babyfacedness X Age</td>
<td>-0.006</td>
<td>0.044</td>
<td>0.01&lt;sub&gt;a&lt;/sub&gt;</td>
<td>0.099&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

** < .01, * < .05

Note. Betas for contrasting groups with different subscripts differ at p < .01 or better. Betas for contrasting groups with the same subscripts differ at p < .10
Table 4

Standardized betas for the negativity variable effects on the attractiveness halo

<table>
<thead>
<tr>
<th>Competent</th>
<th>Healthy</th>
<th>Hostile</th>
<th>Untrustworthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older Adult</td>
<td>Younger Adult</td>
<td>Older Adult</td>
<td>Younger Adult</td>
</tr>
<tr>
<td>Attractiveness X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negativity Factor</td>
<td>0.021</td>
<td>0.03</td>
<td>0.009</td>
</tr>
</tbody>
</table>

*** <.001, ** <.01, * <.05, †<.10
Betas for contrasting groups with different subscripts differ at p < .01 or better.
Table 5

Statistics of control measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Younger Adults</th>
<th>Older Adults</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snellen Visual Acuity (denominator)</td>
<td>20.25</td>
<td>34.69</td>
<td>-4.665</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Mars Letter Contrast Sensitivity without reading glasses</td>
<td>1.75</td>
<td>1.54</td>
<td>7.014</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Mars Letter Contrast Sensitivity with reading glasses</td>
<td>1.61</td>
<td>1.56</td>
<td>1.151</td>
<td>0.254</td>
</tr>
<tr>
<td>Ishihara's Test for Color Deficiency</td>
<td>13.38</td>
<td>13.31</td>
<td>0.117</td>
<td>0.907</td>
</tr>
<tr>
<td>Benton Facial Recognition Test</td>
<td>47</td>
<td>45.47</td>
<td>1.592</td>
<td>0.117</td>
</tr>
<tr>
<td>Pattern Comparison Test</td>
<td>41.53</td>
<td>28.41</td>
<td>10.159</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Shipley Vocabulary Test</td>
<td>30.69</td>
<td>34.81</td>
<td>-4.405</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>PANAS Positive Affect</td>
<td>3.03</td>
<td>3.25</td>
<td>-1.179</td>
<td>0.243</td>
</tr>
<tr>
<td>PANAS Negative Affect</td>
<td>1.62</td>
<td>1.4</td>
<td>1.695</td>
<td>0.096</td>
</tr>
<tr>
<td>Mind in Eye Test</td>
<td>24.56</td>
<td>24.58</td>
<td>-0.018</td>
<td>0.986</td>
</tr>
<tr>
<td>WCST Correct responses</td>
<td>36.4</td>
<td>29.03</td>
<td>3.737</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>WCST Perseverative errors</td>
<td>5.44</td>
<td>8.37</td>
<td>-2.098</td>
<td>0.043</td>
</tr>
<tr>
<td>WCST Non-perseverative errors</td>
<td>6.16</td>
<td>10.6</td>
<td>-2.203</td>
<td>0.032</td>
</tr>
<tr>
<td>WCST Total errors</td>
<td>11.6</td>
<td>18.97</td>
<td>-3.737</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>WCST Trials to complete first category</td>
<td>10.4</td>
<td>13.7</td>
<td>-1.832</td>
<td>0.074</td>
</tr>
<tr>
<td>Level of Education*</td>
<td>2.59</td>
<td>4.69</td>
<td>-7.006</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

*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 degree
Figure 1

Older adult (OA) and younger adult (YA) mean trait ratings for all faces and for negatively, neutrally, and positively valenced faces.

*** $p < .001$, ** $p < .01$, * $p < .05$
Comparison of strength of attractiveness halo on untrustworthiness ratings as they interact with negativity, based on rater age.

Panel 1: YAs
- below median
- median
- above median

Below: R2 Linear = 0.117
Median: R2 Linear = 0.489
Above: R2 Linear = 0.543

Panel 2: QAs
- below median
- median
- above median

Below: R2 Linear = 0.43
Median: R2 Linear = 0.58
Above: R2 Linear = 0.552
Figure 3

Comparison of the strength of the babyface overgeneralization on health ratings as it interacts with face age group, based on rater age.

Panel 1: QAs

Panel 2: YAs
Figure 4

Comparison of the strength of the babyface overgeneralization on hostility ratings as it interacts with face age group, based on rater age

Panel 1: OA

Panel 2: YA