



Published in final edited form as:

*Exp Aging Res.* 2014 ; 40(3): 375–393. doi:10.1080/0361073X.2014.897151.

## The Attractiveness Halo Effect and the Babyface Stereotype in Older and Younger Adults: Similarities, Own-Age Accentuation, and OA Positivity Effects

Leslie A Zebrowitz and Robert G. Franklin Jr.

Brandeis University

### Abstract

**Background**—Two well-documented phenomena in person perception are the attractiveness halo effect (more positive impressions of more attractive people), and the babyface stereotype (more childlike impressions of more babyfaced people), shown by children, young adults (YA) and people from diverse cultures. This is the first study to systematically investigate these face stereotypes in older adults (OA) and to compare effects for younger and older adult faces.

**Method**—YA and OA judges rated competence, health, hostility, untrustworthiness, attractiveness, and babyfacedness of older and younger neutral expression faces. Multilevel modeling assessed effects of rater age and face age on appearance stereotypes.

**Results**—Like YA, OA showed both the attractiveness halo effect and the babyface stereotype. However, OA showed weaker effects of attractiveness on impressions of untrustworthiness, and only OA associated higher babyfacedness with greater competence. There also was own-age accentuation, with both OA and YA showing stronger face stereotypes for faces closer to their own age. Age differences in the strength of the stereotypes reflected an OA positivity effect shown in more influence of positive facial qualities on impressions or less influence of negative ones, rather than vice versa.

**Conclusions**—OA own-age biases, previously shown in emotion, age, and identity recognition, and OA positivity effects, previously revealed in attention, memory, and social judgments, also influence age differences in the strength and content of appearance stereotypes. Future research should assess implications of these results for age-related differences in susceptibility to appearance biases that YA have shown in socially significant domains, such as judicial and personnel decisions.

### Keywords

Aging; attractiveness halo; babyface stereotype; face perception

---

Two well-documented phenomena in person perception are the attractiveness halo effect, whereby more positive traits are ascribed to more attractive individuals of many ages and races (Eagly, Ashmore, Makhijani, & Longo, 1991; Langlois et al., 2000; Dion, 2002), and

the babyface overgeneralization effect, whereby more childlike traits are ascribed to more babyfaced individuals of many ages and races (Zebrowitz & Montepare, 1992; Zebrowitz, Montepare & Lee, 1993; Zebrowitz et al., 2012). People from diverse cultures also show the attractiveness halo effect and the babyface stereotype (Dion, 2002; Langlois et al., 2000; Zebrowitz et al., 1993; Zebrowitz et al., 2012). In addition, the halo effect and the babyface stereotype are evident at a young age (Langlois et al., 2000; Montepare & Zebrowitz-McArthur, 1989). The early manifestation of face stereotypes together with their cultural generality suggest universal attunements to variations in attractiveness and babyfacedness. However, there is a paucity of research investigating the robustness of these face stereotypes at the other end of the lifespan, a gap in the literature that we address in the present study. Not only are there reasons to expect age differences, but also examining face stereotypes in older adults (OA) has particular importance in light of the abundant evidence that facial appearance can bias socially significant outcomes in the judicial, occupational, and political arenas, to name a few (Zebrowitz, 1997). Given the ‘greying’ of American society (US Census Bureau, 2000), differences between OA and younger adults’ (YA) facial stereotypes may augur for changes in those social outcomes.

No previous research has investigated the babyface stereotype in OA and only one study has investigated the halo effect in OA (Larose & Standing, 1998). Although that study showed an OA halo effect, it did not control for facial expression or face sex, which makes it unclear whether the effects reflected variations in attractiveness per se. For example, if women were overrepresented among the more attractive faces, those faces may be perceived as more trustworthy due to sex stereotypes rather than attractiveness stereotypes. That study also did not determine whether there were age differences in the strength of the halo effect, because it did not include YA judges. Yet, there is some reason to expect age-related variations.

OA may show a weaker attractiveness halo effect because age brings the wisdom to eschew attractiveness as a basis for forming impressions (cf. Baltes & Smith, 2008). OA and YA also may show differences in the magnitude of the attractiveness halo effect due to an OA ‘positivity effect.’ This positivity effect has previously been demonstrated in OA tendency to be less responsive to negative stimuli and/or more responsive to positive stimuli (Mather & Carstensen, 2005; Murphy & Isaacowitz, 2008). OA lower responsiveness to negative stimuli may translate to weaker halo effects due to a weaker influence of low attractiveness on their impressions, whereas their higher responsiveness to positive stimuli may translate to a stronger halo effect due to a stronger influence of high attractiveness. Age differences in responsiveness to positively and negatively valenced stimuli also have implications for the babyface stereotype, since high babyfacedness is a positive attribute and low babyfacedness resembles anger, eliciting trait impressions paralleling those associated with angry expressions (Marsh, Adams, & Kleck, 2005; Zebrowitz, Kikuchi, & Fellous, 2007).

Previous research on face stereotypes has ignored not only older raters, but also older faces. The present study filled this gap in the literature by examining the moderating effects of face age. OA show better performance for older than younger faces on face recognition (Anastasi & Rhodes, 2005; Fulton & Bartlett, 1991), age recognition (Voelkle, Ebner, Lindenberger, & Riediger, 2011), emotion recognition (Malatesta, Izard, Culver, & Nicolich, 1987), and recognition of criminals in lineups (Wright & Stroud, 2002). This ‘own-age bias’ may also

be shown in the domain of face stereotypes, in which case OA will be more sensitive to variations in the attractiveness or babyfacedness of older than younger faces, with the reverse true for YA.

We tested the following hypotheses:

1. OA, like YA, will show the attractiveness halo and the babyface stereotype.
2. Consistent with previous evidence for own-age biases, there will be an own-age accentuation effect in face stereotypes. OA will show stronger face stereotypes than YA for older faces, with the reverse age difference for younger faces.
3. Consistent with an OA positivity effect, when OA show stronger face stereotypes than YA, this will reflect a greater influence of high attractiveness or high babyfacedness on their impressions rather than a greater influence of low attractiveness or low babyfacedness. When OA show weaker face stereotypes than YA, this will reflect a weaker influence of low attractiveness or low babyfacedness on their impressions rather than a weaker influence of high attractiveness or high babyfacedness.

## Method

### Participants

Forty-eight YA participants (23 men), 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. Forty-eight OA participants (24 men), 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 (Folstein et al., 1975) all scoring above 26 out of 30 ( $M = 28.9$ ,  $SD = 1.2$ ).

### Facial Stimuli

Three sets of facial images were used. One was from the Boston Study of Management Processes, a Boston-area sub-sample of participants from the study of Midlife Development in the United States (MIDUS) (Lachman, 1997). The images were color frontal facial photographs displaying neutral expressions, with a grey cape to mask clothing. Sixty nine younger adults (41 men) between the ages of 25 and 39 ( $M = 32.7$ ,  $SD = 4.2$ ), and 68 older adults (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  $7.5 \times 9.5$  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) who were photographed between 17–18 years of age and between 52 and 62 years of age, for a total of 296 images. It should be noted that the younger faces in the MIDUS set were older than our YA raters. Also, the older faces in the IGS set as well as some in the MIDUS set were younger than our OA raters. Indeed, for MIDUS faces the within-group age range exceeded the between-group age range. Thus is it more appropriate to view the younger and older faces as differing in age-similarity to raters

rather than being ‘own-age vs. other-age,’ which may weaken the predicted own-age accentuation effect.

To reduce the rating task for IGS faces to a reasonable length, the images were divided into two sets, with equal numbers of older and younger men and women included in each set. 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 babyfaced by OA,  $t(30) = 1.76, p = .09$ , with no differences shown in babyface ratings by YA,  $t < 1$ . Images were displayed at an approximate facial size of  $9 \times 11$  cm.

### Dependent measures

**Face impressions**—Participants rated four traits (competent, healthy, untrustworthy, and aggressive or hostile) and two appearance qualities (attractive, babyfaced) using 7-point scales. These traits were selected to represent multiple domains (intellectual, physical, moral) on which previous research with YA had shown reliable attractiveness and/or babyfacedness stereotype effects. Raters rated the faces on scales with endpoints labeled (1) *not at all* (healthy, competent, untrustworthy, aggressive, hostile, attractive, or babyfaced) and (7) *very* (healthy, competent, untrustworthy, aggressive, hostile, attractive, or babyfaced). Competence ratings by one OA man were dropped, because all faces were rated identically. Preliminary analyses of aggressiveness ratings using the Boston Study faces revealed significantly lower agreement among OA than YA. Since this might reflect less agreement regarding the meaning of “aggressive” (i.e. assertiveness versus hostility), we used the more unambiguous term “hostile” for ratings of IGS faces, and we report only the hostile ratings.

**Control measures**—Measures of vision, affect, and cognitive function were administered to assess age differences. Results are consistent with previous studies of community-dwelling OA, demonstrating the representativeness of our sample (see Table 1).

### Procedure

After obtaining informed consent, participants completed a computerized version of the PANAS. Next, MediaLab software (Empirisoft, New York City, NY) was used to present the trait rating task. Participants were randomly assigned to one of two rating orders: healthy, untrustworthy, competent, aggressive/hostile, attractive, and babyfaced, or aggressive/hostile, competent, untrustworthy, healthy, babyfaced, and attractive. They rated one of the three sets of faces, which were shown in one of four orders, counterbalancing age and sex of face. Participants were asked to rate each face in comparison to the other faces of that age/sex grouping so that variations in trait impressions would less likely reflect sex or age stereotypes, as we were interested in whether the face stereotypes of OA were responsive to variations in facial appearance within demographic groups, as has been previously shown for YA. Faces were shown for either 4s (Boston Study faces) or 3s (IGS faces) after which the face disappeared and the rating scale appeared on the screen,

remaining until the rating was made. Participants rated all faces on one trait scale before the next one was introduced. An instruction screen appeared prior to each scale indicating what the trait rating would be. Following the face ratings, a demographic and health questionnaire and the remaining control measures were administered.

## Results

### Overview of Analyses

We used a multi-level model approach examining effects of attractiveness and babyfacedness separately for each trait impression. We created two models for each trait rating, one examining attractiveness and the other examining babyfacedness. To account for both subject-level and face-level variance, participant and face were included as random effects in each model, with face-level factors nested within subject-level factors. More specifically, each model included the appearance quality of interest (attractiveness or babyfacedness), and face age group (younger faces coded as 0, older faces coded as 1) as face-level factors, as well as participant age (YA coded as 0, OA coded as 1) as a subject-level factor together with all two-way and three-way interactions between these three variables. Each model also included the other appearance quality (e.g., babyfacedness in the attractiveness model) as well as face sex (female faces coded as 0, male coded as 1) as face-level control variables, but the interactions involving these variables were not included in the model. Each participant's trait rating for each face served as the dependent variable.

To visualize significant interaction effects, we plotted expected values based on the regression equations derived from the models. These plots were conducted using values +1 and -1 standard deviations from the mean for attractiveness and babyfacedness for each of the four traits to represent high and low values for each of those appearance qualities. To further explicate interactions, we conducted simple slopes tests for OA and YA separately for older and younger faces (see Aiken & West, 1991) to determine whether effects of attractiveness and babyfacedness were significant for each age group rating older and younger faces separately. Finally, we used slope difference tests to determine whether the slopes differed significantly (see Dawson & Richter, 2006).

### Attractiveness Halo

**Overall effects of attractiveness**—Consistent with previous research on the attractiveness halo effect, more attractive faces were rated as less hostile ( $\beta = -.13$ ,  $SE = .016$ ,  $p < .001$ ) and less untrustworthy ( $\beta = -.24$ ,  $SE = .015$ ,  $p < .001$ ), as well as more competent ( $\beta = .28$ ,  $SE = .013$ ,  $p < .001$ ) and more healthy ( $\beta = .28$ ,  $SE = .013$ ,  $p < .001$ ).

**Own-age accentuation and OA positivity**—Rater age did not moderate the attractiveness halo effect for ratings of hostility ( $\beta = -.014$ ,  $SE = .022$ ,  $p = .62$ ). However it significantly moderated the effect for ratings of untrustworthiness ( $\beta = -.076$ ,  $SE = .021$ ,  $p < .001$ ), with this interaction reflecting a weaker attractiveness halo effect for OA (simple slope =  $-.18$ ,  $p < .001$ ) than YA (simple slope =  $-.26$ ,  $p < .001$ ), although both effects were significant. Consistent with an OA positivity effect, Figure 1, left panel, shows that the

weaker halo effect for OA was due to a weaker influence of low attractiveness on OA than YA impressions of untrustworthiness.

Rater age also moderated the halo effect for ratings of competence ( $\beta = -.053$ ,  $SE = .019$ ,  $p = .005$ ) and health ( $\beta = -.035$ ,  $SE = .021$ ,  $p < .001$ ), with these significant interactions reflecting a tendency for OA to show a weaker halo effect for competence ratings (simple slope = .25,  $p < .001$ ) than did YA (simple slope = .27,  $p < .001$ ), and a stronger halo effect for health ratings (simple slope = .31,  $p < .001$ ) than did YA (simple slope = .26,  $p < .001$ ), although all effects were significant. Age differences in the strength of the halo effect for competence and health ratings were further moderated by triple order interactions with face age ( $\beta = .060$ ,  $SE = .025$ ,  $p = .019$  for competence;  $\beta = -.035$ ,  $SE = .021$ ,  $p < .001$  for health). The pattern of results shown in these interactions provided support for the predicted own-age accentuation and mixed support for an OA positivity effect, as described below.

The triple order interaction for competence ratings was consistent with own-age accentuation in the strength of the halo effect. The overall tendency for OA to show a weaker halo effect for competence ratings than YA held true for younger faces (OA: simple slope = .23,  $p < .001$ ; YA simple slope = .28,  $p < .001$ ; difference,  $p = .005$ ), but not for older faces (OA: simple slope = .27,  $p < .001$ ; YA simple slope = .26,  $p < .001$ ; difference,  $p = .71$ ). Figure 1, middle panel, shows that, contrary to the OA positivity effect shown in the halo effect for impressions of untrustworthiness, the weaker OA halo effect for younger faces was due to a weaker influence of high attractiveness on OA than YA impressions of competence.

As was true for competence ratings, the triple order interaction for health ratings showed the predicted own-age accentuation. OA showed a stronger halo effect than YA for older faces (OA: simple slope = .40,  $p < .001$ ; YA simple slope = .27,  $p < .001$ ; difference,  $p < .001$ ), whereas the opposite pattern was shown for younger faces (OA simple slope = .22,  $p < .001$ , YA simple slope = .26,  $p < .001$ ; difference,  $p = .053$ ). Health ratings also supported the predicted OA positivity effect. Figure 1, right panel, shows that the stronger halo effect for OA than YA when rating the health of older faces was due to a stronger influence of high attractiveness on OA than YA impressions, and the weaker halo effect for OA than YA rating the health of younger faces was due a weaker influence of low attractiveness on OA than YA impressions.<sup>1</sup>

### Babyface Stereotype

**Overall effects of babyfacedness**—Consistent with previous research on the babyface stereotype, higher babyfacedness was significantly associated with lower ratings of hostility ( $\beta = -.088$ ,  $SE = .014$ ,  $p < .001$ ), and untrustworthiness ( $\beta = -.075$ ,  $SE = .013$ ,  $p < .001$ ). There were no overall effects of babyfacedness on ratings of competence ( $\beta = -.01$ ,  $SE = .011$ ,  $p = .63$ ) or health ( $\beta = .005$ ,  $SE = .011$ ,  $p = .69$ ).

<sup>1</sup>Face age significantly moderated the effect of attractiveness on ratings of hostile ( $\beta = .043$ ,  $SE = .022$ ,  $p = .051$ ) and untrustworthy, ( $\beta = -.038$ ,  $SE = .020$ ,  $p = .066$ ) revealing a weaker halo effect for younger faces that did not vary with rater age.



**Own-age accentuation and OA positivity**—Rater age significantly moderated the babyface stereotype for ratings of competence ( $\beta = .045$ ,  $SE = .017$ ,  $p = .007$ ), with this significant interaction revealing that OA showed a marginally significant positive effect of babyfacedness on competence ratings (simple slope =  $.019$ ,  $p = .056$ ), while YA showed a non-significant negative effect (simple slope =  $-.011$ ,  $p = .21$ ). Rater age also moderated the babyface stereotype for impressions of untrustworthiness ( $\beta = -.047$ ,  $SE = .019$ ,  $p = .013$ ), with this significant interaction revealing a weaker effect for OA (OA: simple slope =  $-.029$ ,  $p = .010$ , YA simple slope =  $-.044$ ). Although rater age alone did not moderate the babyface stereotype for ratings of health and hostility, triple order interactions with face age qualified the effects for these two impressions as well as for impressions of untrustworthiness (untrustworthy  $\beta = -.066$ ,  $SE = .026$ ,  $p = .010$ ; hostile  $\beta = -.062$ ,  $SE = .029$ ,  $p = .031$ ; health  $\beta = .048$ ,  $SE = .022$ ,  $p = .030$ ). The pattern of results shown in these interactions provided support both for the predicted own-age accentuation and the OA positivity effect, as described below.

The triple order interaction for untrustworthy ratings was consistent with own-age accentuation in the strength of the babyface stereotype. YA showed a stronger babyface effect for ratings of the untrustworthiness of younger faces than did OA (YA simple slope =  $-.075$ ,  $p < .001$ ; OA: simple slope =  $-.028$ ,  $p = .062$ , age difference,  $p = .013$ ). The opposite pattern was shown for older faces, with OA showing a significant babyface effect (simple slope =  $-.031$ ,  $p = .040$ ), while YA did not (simple slope =  $-.012$ ,  $p = .37$ ). Although this pattern is also consistent with own-age accentuation in the strength of face stereotypes, the rater age difference was not significant ( $p = .34$ ). Consistent with an OA positivity effect, Figure 2, bottom left panel, shows that the significantly weaker OA babyface stereotype for younger faces was due to a weaker influence of low babyfacedness on OA than YA impressions of untrustworthiness..

The triple order interaction for hostility ratings also was consistent with own-age accentuation. OA showed a stronger babyface stereotype for ratings of the hostility of older faces than did YA (OA simple slope =  $-.066$ ,  $p < .001$ ; YA simple slope =  $-.030$ ,  $p = .043$ ), a rater age difference that was marginally significant ( $p = .099$ ). The reverse pattern held true for younger faces (YA simple slope =  $-.088$ ,  $p < .001$ ; OA simple slope =  $-.062$ ,  $p < .001$ ), although the rater age difference was not significant ( $p = .22$ ). The strength of the babyface stereotype for hostility ratings also was consistent with an OA positivity effect. As shown in Figure 2, top panel, the marginally significant tendency for OA to show a stronger babyface stereotype than YA when rating the hostility of older faces was due to a stronger influence of high babyfacedness on OA than YA impressions.

Own-age accentuation was also shown in the triple order interaction for health ratings. OA showed a stronger effect for older faces than did YA (OA simple slope =  $.11$ ,  $p < .001$ ; YA simple slope =  $.039$ ,  $p = .001$ , age difference  $p < .001$ ). In the case of younger faces, there was no own-age accentuation. OA showed a marginally significant babyface effect (simple slope =  $.022$ ,  $p = .083$ ), while the effect for YA was not significant (simple slope =  $.005$ ,  $p = .69$ ), and neither was the rater age difference, ( $p = .27$ ). Consistent with an OA positivity effect, Figure 2, bottom right panel, shows that, the significant age difference in the strength

of the babyface stereotype for older faces was due to a stronger influence of high babyfacedness on OA than YA impressions of health.

## Discussion

The results revealed that stereotypes associated with variations in attractiveness and babyfacedness that have been well-documented in YA are preserved in OA, supporting our first hypothesis. At the same time there were interesting age differences. These included babyface stereotypes not previously documented, as well as support for our second hypothesis, that people would show stronger face stereotypes when forming impressions of people closer to their own age, and support for our third hypothesis that age differences in the strength of face stereotypes would reflect a positivity effect for OA.

### Do OA show face stereotypes?

OA are just as vulnerable to face stereotypes as YA. They showed an attractiveness halo effect, judging more attractive people as more competent and healthy as well as less hostile and untrustworthy, consistent with a large body of research on YA (Eagly et al., 1991; Langlois, 2000). OA also showed a babyface stereotype, judging more babyfaced people as less hostile and untrustworthy, consistent with a large body of research on YA (Montepare & Zebrowitz, 1998; Zebrowitz, Bronstad, & Lee, 2007; Olivola & Todorov, 2010; Sparko & Zebrowitz, 2011). It bears noting that the equivalence of face stereotypes across rater age held true despite poorer performance of OA on many control measures. Also, effects of attractiveness and babyfacedness held with the other appearance variable and face sex controlled, extending a previous study that showed a halo effect for OA without controlling these variables (Larose & Standing, 1998). That research also failed to directly compare the strength of the halo effect shown by OA to YA or to control face age, which the present results reveal is an important moderating variable.

In addition to showing previously documented face stereotype effects, participants, perceived more babyfaced older people as healthier, a component of the babyface stereotype not previously documented in research focusing on younger faces. Moreover, this effect was stronger for OA. OA further associated higher babyfacedness with greater competence, whereas there was no significant relationship for YA. The tendency for OA to associate higher babyfacedness with greater competence and to show a stronger tendency than YA to associate it with greater health suggests that OA are more attuned to the positive implications of a more youthful appearance for health and competence in older individuals. Indeed, there is evidence that people who look younger for their age are healthier (Gunn et al., 2013). Although the null effect for YA is inconsistent with some research documenting a negative relationship between babyfacedness and YA competence ratings (Olivola & Todorov, 2010), most research on the babyface stereotype has documented a positive relationship between babyfacedness and YA ratings of *naivete* (see Montepare & Zebrowitz, 1998, for a comprehensive review). Also, some research has failed to find a relationship between babyfacedness and *perceived intelligence* (McArthur & Apatow, 1983). Competence is a broad term, and it may be that the interpretation of participants in the present study was closer to intelligence than to naivete. Additional research investigating impressions of the



naivete of individuals who vary in babyfacedness would be useful to ascertain whether OA show this well-documented YA stereotype.

### Is there 'own-age' accentuation in face stereotypes?

The hypothesis that stereotypes would be stronger when judging faces closer to one's own age received support in 5 of the 8 stereotypes examined. Consistent with predictions, there was own-age accentuation in the halo effect on impressions of competence and health and in the babyface effect on impressions of health, untrustworthiness, and hostility. More specifically, OA showed a weaker halo effect than YA for impressions of the competence and health of younger faces, and a stronger effect for impressions of the health of older faces. Similarly, babyfacedness had a stronger effect on YA than OA impressions of the untrustworthiness of younger faces and a stronger effect on OA than YA impressions of the hostility and health of older faces. On the other hand, there was no significant own-age accentuation in the halo effect for impressions of trustworthiness and hostility or in the babyface stereotype for impressions of competence.

Whereas some face stereotypes did not show own-age accentuation, in no case did attractiveness or babyfacedness have a stronger effect when faces differed markedly in age from the raters than when they were similar in age. Although some of the 'younger' faces were older than our YA raters and some of the 'older' faces were younger than our OA raters, the effects we found are nevertheless consistent with other evidence for an 'own-age' bias. Just as people show better age, emotion, and identity recognition for own-age faces (Malatesta et al., 1987; Voelkle et al., 2011; Wright & Stroud, 2002), so do they tend to show more sensitivity to attractiveness and babyfacedness when judging similar-age faces. These effects reflect the implications drawn from different levels of attractiveness and babyfacedness rather than own-age accentuation in the appearance ratings (see Footnote <sup>2</sup>).

The stronger halo effect for impressions of the health and competence of similar-age faces may be interpreted within the context of previous research demonstrating stronger halo effects for traits that are more culturally valued (Wheeler & Kim, 1997; Shaffer, Crepaz, & Sun, 2000). Arguably, people may place more value on the health of similar-age peers, which could explain the stronger halo effect for OA than YA impressions of the health of older faces, with a reverse trend for younger faces. It also seems reasonable to suggest that YA college students care more about the competence of younger people than do OA, which could explain the stronger halo effect for YA than OA on impressions of the competence of younger faces. On this account, the absence of own-age accentuation in the halo effect for

<sup>2</sup>Analyses of variance on attractiveness and babyface ratings revealed that OA rated the faces as significantly more attractive (YA:  $M = 2.92$ ,  $SD = .71$ , OA:  $M = 3.49$ ,  $SD = .71$ ,  $F(1,90) = 15.8$ ,  $p < .001$ ), but there was no rater age difference in ratings of babyfacedness (YA:  $M = 3.15$ ,  $SD = .76$ , OA:  $M = 3.01$ ,  $SD = .90$ ;  $F(1,90) = .55$ ). In addition, younger faces were rated as more attractive (younger faces  $M = 3.34$ ,  $SD = .78$ , older faces  $M = 3.07$ ,  $SD = .92$ ;  $F(1,90) = 11.2$ ,  $p = .001$ ) and more babyfaced, (younger faces  $M = 3.33$ ,  $SD = .86$ , older faces  $M = 2.83$ ,  $SD = .95$ ;  $F(1,90) = 49.3$ ,  $p < .001$ ). The face age effect for babyface ratings was not moderated by rater age,  $F(1,90) = 1.08$ , but there was a marginally significant moderation for attractiveness ratings,  $F(1,90) = 3.07$ ,  $p = .083$ . Comparisons within each rater age revealed that OA ratings of the attractiveness of younger and older faces were not significantly different, (younger faces:  $M = 3.55$ ,  $SD = .77$ , older faces:  $M = 3.43$ ,  $SD = .82$ ;  $t(47) = 1.52$ ,  $p = .22$ ), whereas YA rated younger faces as more attractive (younger faces:  $M = 3.12$ ,  $SD = .73$ , older faces:  $M = 2.72$ ,  $SD = .88$ ;  $t(47) = 3.61$ ,  $p = .001$ ). Although YA showed own-age accentuation in attractiveness ratings, they did not show any own-age accentuation in the halo effect. Thus, these results indicate that own-age accentuation found in the strength of the babyface and attractiveness stereotypes cannot be attributed to own age accentuation in mean ratings of attractiveness or babyfacedness.

impressions of hostility and untrustworthiness suggests that people place equal value on these traits in people whether they are similar or dissimilar in age. It is noteworthy that variations in attractiveness predict actual variations in competence and health (Zebrowitz, Hall, Murphy, & Rhodes, 2002; Zebrowitz & Rhodes, 2004). Thus, the stronger effects of attractiveness on impressions of similar-age faces may reflect not only a stronger valuation of health and competence in those faces, but also a greater sensitivity to cues in those faces that are diagnostic of these traits

### **Is there a positivity effect in OA face stereotypes?**

Based on previous evidence for a positivity bias in OA (Castle et al., in press; Mather & Carstensen, 2005; Murphy & Isaacowitz, 2008; Ruffman et al., 2006; Zebrowitz et al., in press), we predicted that when OA showed stronger face stereotypes than YA, this would reflect stronger effects on impressions of attractiveness and babyfacedness at the high, positive end of these appearance continua rather than at the low, negative end, while weaker OA stereotypes would reflect weaker effects of attractiveness and babyfacedness at the negative end of the continua rather than at the positive end. These predictions were supported in all but one of the instances in which we found age differences in the strength of appearance stereotypes. Specifically the weaker OA halo effects for overall impressions of untrustworthiness and impressions of younger faces' health, and the weaker OA babyface effects for impressions of younger faces' untrustworthiness and older faces' hostility all reflected a weaker influence of low attractiveness or low babyfacedness on OA than YA impressions rather than a weaker influence of high attractiveness or high babyfacedness. This can be seen in Figures 1 and 2, where YA and OA ratings diverge more at the low than the high end of the attractiveness continuum. The one exception was the OA weaker halo effect for impressions of younger faces' competence, which reflected a weaker influence of high attractiveness on OA than YA impressions. Complementing the overall support for the predicted OA lower responsiveness to negative facial cues was evidence for their greater responsiveness to positive cues. Specifically, the stronger OA halo effect for impressions of the health of older faces and the stronger OA babyface effects for impressions of the health and hostility of older faces all reflected a stronger influence of high attractiveness or high babyfacedness on OA than YA impressions rather than a stronger influence of low attractiveness or babyfacedness. This can be seen in Figures 1 and 2, where YA and OA ratings diverge more at the high than the low end of the attractiveness continuum. In addition, the tendency for OA to associate high babyfacedness with greater competence, whereas YA did not, can also be construed as an OA positivity effect. In sum, the weaker face stereotypes shown by OA than YA were largely due to OA showing more positive impressions of people with negative facial cues, and the stronger face stereotypes shown by OA were largely due to OA showing more positive impressions of people with positive facial cues.

Future research should examine the mechanisms underlying the documented own-age accentuation and OA positivity effects in face stereotypes. One possible mechanism is variations in attention. Previous research using eye-tracking has demonstrated that OA show greater attention to more positively valenced faces and less attention to negatively valenced faces (Isaacowitz, Wadlinger, Goren, & Wilson, 2008). It may also be that people show

greater attention to similar-age faces. Another related mechanism is variations in differentiation, the degree to which people perceive variation among faces. People's ratings of various traits or appearance qualities may reveal greater differentiation of own-age faces, and OA may show greater differentiation of positively valenced faces and less differentiation of negatively valenced ones. This could be assessed using an index developed by Linville, Salovey, & Fischer (1986) that has been applied to investigate the extent to which people perceive variations in other-race vs. own-race faces (Zebrowitz, Montepare, & Lee, 1993).

Whatever the mechanism proves to be, our results raise the interesting possibility that previously documented appearance biases in the judicial, occupational, and political domains may differ when OA and YA are making the decisions. The fact that OA perceived more babyfaced people as more competent together with their weaker association of attractiveness and babyfacedness with trustworthiness may mitigate judicial biases whereby more attractive defendants and plaintiffs fare better, more babyfaced defendants accused of willful misconduct are more likely to be exonerated, and more babyfaced defendants accused of negligence are more likely to be found at fault (Stewart, 1980; Zebrowitz & McDonald, 1991). The same age differences in facial stereotypes may weaken or even reverse documented effects of attractiveness and babyfacedness on occupational and political outcomes (Collins & Zebrowitz, 1995; Hosoda, Stone-Romero, & Coats, 2003; Olivola, & Todorov, 2010; Rule et al., 2010; Zebrowitz, Tenenbaum, & Goldstein, 1991). In addition, the finding that age differences in the strength of face stereotypes largely reflected OA lesser responsiveness to negative facial cues or greater responsiveness to positive ones rather than vice versa suggests that decisions made by OA in socially significant domains are more likely to favor people who are high in attractiveness rather than penalizing those who are low. The reverse is likely to be true for YA given evidence that the halo effect for them reflects negative responses to unattractive faces more than positive responses to attractive ones (Griffin & Langlois, 2006). Finally, the stronger face stereotypes shown for own-age faces suggests that biasing effects of appearance on significant social outcomes are likely to be stronger when people are making decisions about similar-age peers.

## Acknowledgments

The authors would like to thank Stacey Ng for her helpful comments on an earlier draft. We also thank Bosiljka Milosavljevic for her help in recruiting participants and collecting data as well as Ilana Abramson, Suzanne Hillman, Joanne Qiao, and Justine Zayhowski.

### Funding

This work was supported by the National Institutes of Health (grant number AG38375 to L. A. Z.).

## References

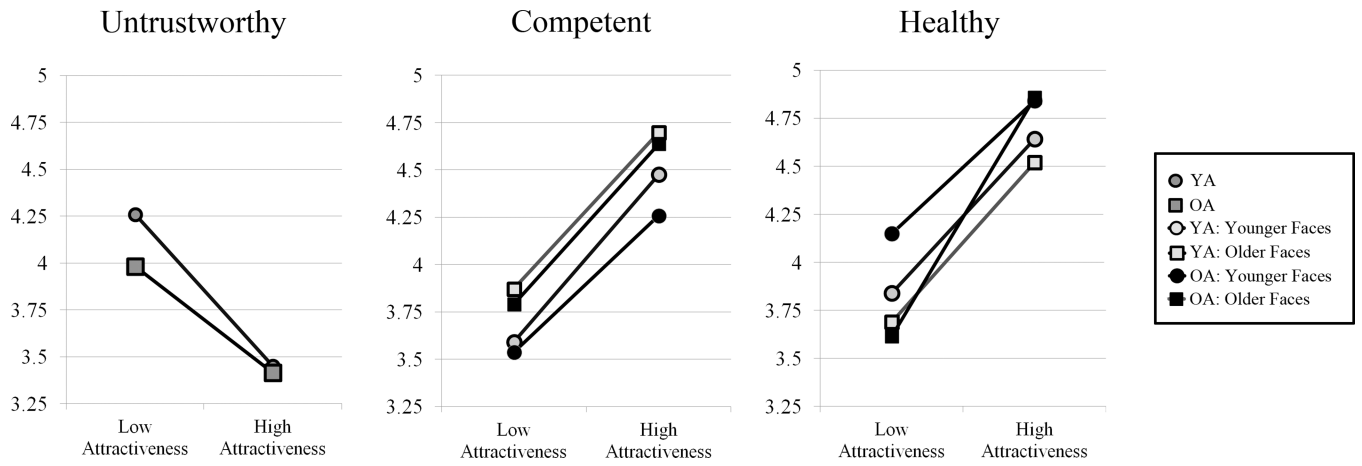
- Aiken, L.S.; West, S.G. *Multiple Regression: Testing and Interpreting Interactions*. SAGE; 1991.
- Anastasi JS, Rhodes MG. An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*. 2005; 12(6):1043–1047. [PubMed: 16615326]
- Baltes PB, Smith J. The fascination of wisdom: Its nature, ontogeny, and function. *Perspectives on Psychological Science*. 2008; 3(1):56–64.

- Benton A, Van Allen M, Hamsher K, Levin H. Test of facial recognition manual. Iowa City: Benton Laboratory of Neuropsychology. 1983
- Berry DS, Zebrowitz-McArthur L. What's in a face? Facial maturity and the attribution of legal responsibility. *Personality and Social Psychology Bulletin*. 1988; 14(1):23–33.
- Castle E, Eisenberger NI, Seeman TE, Moon WG, Boggero IA, Grinblatt MS, Taylor SE. Neural and behavioral bases of age differences in perceptions of trust. *Proceedings of the National Academy of Sciences*. (in press).
- Collins MA, Zebrowitz LA. The contributions of appearance to occupational outcomes in civilian and military settings. *Journal of Applied Social Psychology*. 1995; 25(2):129–163.
- Dawson JF, Richter AW. Probing three-way interactions in moderated multiple regression: Development and application of a slope difference test. *Journal of Applied Psychology*. 2006; 91(4):917–926. [PubMed: 16834514]
- Dion, KK. Cultural perspectives on facial attractiveness. In: Rhodes, C.; Zebrowitz, LA., editors. *Facial attractiveness: Evolutionary, cognitive, and social perspectives*. Westport, CT: Ablex; 2002. p. 239-259.
- Eagly AH, Ashmore RD, Makhijani MG, Longo LC. What is beautiful is good, but...: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*. 1991; 110(1): 109–128.
- Eichorn, DH. *Present and past in middle life*. New York: Academic Press; 1981.
- 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]
- Fulton A, Bartlett JC. Young and old faces in young and old heads - The factor of age in face recognition. *Psychology and Aging*. 1991; 6(4):623–630. [PubMed: 1777151]
- Griffin AM, Langlois JH. Stereotype directionality and attractiveness stereotyping: Is beauty good or is ugly bad? *Social Cognition*. 2006; 24(2):187–206. [PubMed: 17016544]
- Gunn DA, de Craen AJM, Dick JL, et al. Facial Appearance Reflects Human Familial Longevity and Cardiovascular Disease Risk in Healthy Individuals. *Journals of Gerontology Series A- Biological Sciences and Medical Sciences*. 2013; 68:145–152.
- Hosoda M, Stone-Romero EF, Coats G. The effects of physical attractiveness on job-related outcomes: a meta-analysis of experimental studies. *Personnel Psychology*. 2003; 56(2):431–462.
- Isaacowitz DM, Wadlinger HA, Goren D, Wilson HR. Is there an age-related positivity effect in visual attention? A comparison of two methodologies. *Emotion*. 2006; 6(3):511. [PubMed: 16938091]
- Ishihara, S. *Test for Colour-blindness*. Tokyo: Kanehara Trading Company; 2010.
- Lachman, ME. Inter-university consortium for Political and Social Research. University of Michigan; 1997. *Midlife Development in the United States (MIDUS): Boston Study of Management Processes, 1995–1997*.
- 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]
- Larose H, Standing L. Does the halo effect occur in the elderly. *Social Behavior and Personality: an International Journal*. 1998; 26(2):147–150.
- Linville, PW.; Salovey, P.; Fischer, GW. Stereotyping and perceived distributions of social characteristics: An application to ingroup-outgroup perception. In: Dovidio, JF.; Gaertner, SL., editors. *Prejudice, discrimination, and racism*. San Diego, CA: Academic Press; 1986. p. 165-208.
- Malatesta CZ, Izard CE, Culver C, Nicolich M. Emotion communication skills in young, middle-aged, and older women. *Psychology and Aging*. 1987; 2(2):193–203. [PubMed: 3268208]
- Marsh AA, Adams RB, Kleck RE. Why do fear and anger look the way they do? Form and social function in facial expressions. *Personality and Social Psychology Bulletin*. 2005; 31(1):73–86. [PubMed: 15574663]
- Mather M, Carstensen LL. Aging and motivated cognition: the positivity effect in attention and memory. *Trends in Cognitive Sciences*. 2005; 9(10):496–502. [PubMed: 16154382]

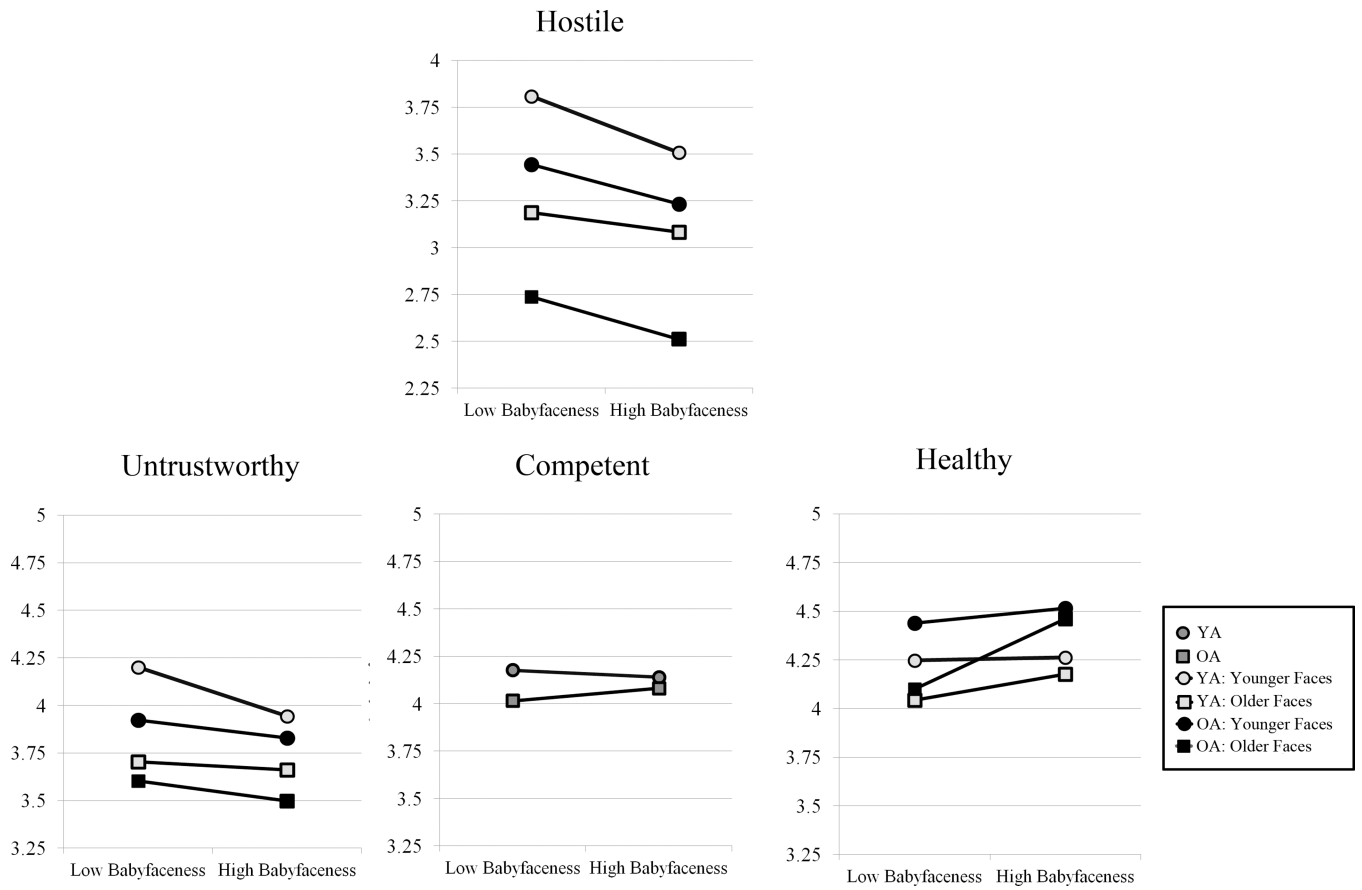
- Montepare, JM.; Zebrowitz, LA. *Advances in Experimental Social Psychology*. Vol. Vol. 30. Academic Press; 1998. Person perception comes of age: The salience and significance of age in social judgments; p. 93-161.
- Murphy NA, Isaacowitz DM. Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging*. 2008; 23(2):263–286. [PubMed: 18573002]
- Olivola CY, Todorov A. Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*. 2010; 46(2):315–324.
- Olivola CY, Todorov A. Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*. 2010; 34:83–110.
- Piper BJ, Li V, Eiwaz MA, Kobel YV, Benice TS, Chu AM, et al. Executive function on the psychology experiment building language tests. *Behavior Research Methods*. 2012; 44:110–123. [PubMed: 21534005]
- Ruffman T, Sullivan S, Edge N. Differences in the way older and younger adults rate threat in faces but not situations. *Journals of Gerontology: Psychological Sciences*. 2006; 61:187–194.
- 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 mediation of adult age differences in cognition. *Developmental Psychology*. 1993; 29(4):722–738.
- Shaffer DR, Crepez N, Sun C-R. Physical attractiveness stereotyping in cross-cultural perspective similarities and differences between Americans and Taiwanese. *Journal of Cross-Cultural Psychology*. 2000; 31(5):557–582.
- Shipley, WC. *Institute of Living Scale*. Los Angeles: Western Psychological Services; 1946.
- Sparko A, Zebrowitz L. Moderating effects of facial expression and movement on the babyface stereotype. *Journal of Nonverbal Behavior*. 2011; 35(3):243–257.
- 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.
- U.S. Census Bureau. [Retrieved April 12, 2013] The next four decades: The older population in the United States: 2010 to 2050. 2010. from <http://www.census.gov/prod/2010pubs/p25-1138.pdf>
- 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*. 2011; 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(6):1063–1070. [PubMed: 3397865]
- Wechsler, D. *Wechsler Adult Intelligence Scale-III (WAIS-III) manual*. New York: The Psychological Corporation; 1997.
- Wheeler L, Kim Y. What is beautiful is culturally good: The physical attractiveness stereotype has different content in collectivistic cultures. *Personality and Social Psychology Bulletin*. 1997; 23(8):795–800.
- Wright DB, Stroud JN. Age differences in lineup identification accuracy: People are better with their own age. *Law and Human Behavior*. 2002; 26(6):641–654. [PubMed: 12508699]
- Zebrowitz, LA. *Reading Faces: Window to the Soul?*. Boulder, CO: Westview Press; 1997.
- Zebrowitz LA, Bronstad PM, Lee HK. The contribution of face familiarity to ingroup favoritism and stereotyping. *Social Cognition*. 2007; 25:306–338.
- Zebrowitz LA, Fellous JM, Mignault A, Andreoletti C. Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review*. 2003; 7:194–215. [PubMed: 12788687]
- Zebrowitz LA, Franklin RGJ, Hillman S, Boc H. Comparing older and younger adults' first impressions from faces. *Psychology and Aging*. (in press).
- Zebrowitz LA, Hall JA, Murphy NA, Rhodes G. Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*. 2002; 28(2):238–249.

- Zebrowitz L, McDonald S. The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*. 1991; 15(6):603–623.
- Zebrowitz LA, Montepare JM. Impressions of babyfaced individuals across the life span. *Developmental Psychology*. 1992; 28(6):1143–1152.
- Zebrowitz LA, Montepare JM, Lee HK. They don't all look alike: Individuated impressions of other racial groups. *Journal of Personality and Social Psychology*. 1993; 65:85–101. [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(3):453–466. [PubMed: 8468672]
- Zebrowitz LA, Rhodes G. Sensitivity to “bad genes” and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior*. 2004; 28(3):167–185.
- Zebrowitz LA, Tenenbaum DR, Goldstein LH. The impact of job applicants' facial maturity, gender, and academic achievement on hiring recommendations. *Journal of Applied Social Psychology*. 1991; 21(7):525–548.
- Zebrowitz LA, Wang R, Bronstad PM, Eisenberg D, Undurraga E, Reyes-García V, Godoy R. First impressions from faces among U.S. and culturally isolated Tsimane' people in the Bolivian rainforest. *Journal of Cross-Cultural Psychology*. 2012; 43(1):119–134.
- Zebrowitz-McArthur L, Montepare JM. Contributions of a babyface and a childlike voice to impressions of moving and talking Faces. *Journal of Nonverbal Behavior*. 1989; 13(3):189–203.





**Figure 1.** Interactions between attractiveness, rater age, and face age as predictors of ratings of untrustworthy, competent, and healthy.



**Figure 2.** Interactions between babyfaceness, rater age, and face age as predictors of ratings of hostile, untrustworthy, competent, and healthy

Table 1

## OA and YA Scores on Control Measures

Measure	Older Adults			Younger Adults			p-value
	M	SD	M	SD	M	t-value	
PANAS (PA) (Watson, Clark, & Tellegen, 1988)	2.90	.70	2.78	.84	.79	.434	
PANAS (NA) (Watson, Clark, & Tellegen, 1988)	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 (Mars Perceptrix, Chappaqua, NY)	1.56	.19	1.73	.06	8.87	< .001	
Ishihara's Test for Color Deficiency (Ishihara, 2010)	13.46	1.87	13.48	1.65	.06	.954	
Benton Facial Recognition Test (Benton, Van Allen, Hamscher, & Levin, 1983)	41.35	9.31	42.21	8.58	.468	.641	
Timed Pattern Comparison Test (Salthouse, 1993)	28.17	5.16	42.56	5.85	12.79	< .001	
Shipley Vocabulary Test (Shipley, 1946)	35.48	3.37	31.77	3.89	4.99	< .001	
Wisconsin Card Sorting Test (the Berg Card Sort Task (BCST validated by Piper et al., 2011; <a href="http://pebl.sourceforge.net/battery.html">http://pebl.sourceforge.net/battery.html</a> ))	29.03	8.92	36.40	5.56	3.59	.001	
Letter-Number Sequencing Test (Wechsler, 1997)	8.19	1.52	12.56	2.80	5.49	< .001	
Education <sup>a</sup>	5	2-7	3	2-3			

<sup>a</sup>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. Medians and range are reported.