

7-2-2020

Altering age and gender stereotypes by creating the Halo and Horns Effects with facial expressions

Mary Katherine Radeke

Anthony John Stahelski

Follow this and additional works at: <https://digitalcommons.cwu.edu/cotsfac>



Part of the [Cognition and Perception Commons](#), and the [Sociology Commons](#)



ARTICLE



<https://doi.org/10.1057/s41599-020-0504-6>

OPEN

Altering age and gender stereotypes by creating the Halo and Horns Effects with facial expressions

Mary Katherine Radeke¹   & Anthony John Stahelski¹

This study examined the impact of a variable, facial expression, on the social perception and personality trait stereotypic inferences made to age and gender. Twelve facial photographs of young and old female and male models posing with either smiling, scowling, or neutral facial expressions were presented to participants who judged various social perceptions and personality traits. Results indicated that facial expression is strongly associated with two very different inference groupings. Smiling induced positive inferences, creating a Halo Effect, scowling induced negative inferences, creating a Horns Effect. Smiling influenced the age and gender inferences in a positive direction, and scowling did the opposite. The age and gender stereotypical inferences made to the neutral facial expression were in-between smiling and scowling. In all model configurations, the impact of smiling or scowling on the inference process was much stronger than either age or gender. However, significant age and gender inference differences were found in all three facial expression conditions, indicating that facial expressions did not completely subdue the use of these variables as inference inducers. The results are discussed in terms of how specific facial expressions can be used to positively or negatively influence age and gender stereotypes.

¹Psychology Department, Central Washington University, Ellensburg, WA, USA. ✉email: radekem@cwu.edu

Introduction

A stereotype can be defined as a belief that certain attributes are characteristic of members of a particular group (Gilgovich et al., 2019). Stereotyping occurs when a perceiver infers a preconceived set of traits based on visible characteristics of a person, and this may occur quickly and unconsciously, based on limited knowledge of the individual (Greenwald and Banaji, 1995). The use of stereotypes appears to be universal, and stereotype formation starts early in life. Children form stereotypes in the family context (Bryan et al., 1986), and biologically based stereotypes, like age and gender, are formed earlier and remain stronger than non-biologically based stereotypes (Hoffman and Hurst, 1990).

Age and gender are broad social categories that are generally the first aspects that perceivers notice when meeting a person for the first time (Johnson et al., 2015). These categories are used to make judgments about the perceived person, and the judgments are often stereotypical (Ebner et al., 2018; Ellemers, 2018; Lamont et al., 2015; Macrae and Bodenhausen, 2000). Research on age and gender perceptions indicates that these stereotypes are a mixture of positive and negative; for example, people over 65 are perceived as more Agreeable and less impulsive (positive), and less active and competent (negative; Chan et al., 2012; Hack, 2014). Females are perceived as more Agreeable, Conscientious and Open (positive), and sad (negative); and males are perceived as more Extraverted (positive), and angry and threatening (negative; Löckenhoff et al., 2014; Parmley and Cunningham, 2014). Age and gender stereotypes can be conceived as “baseline perceptions”, that appear to be lifelong and are resistant to change (Silberstang, 2011). However, they may be at least temporarily influenced by personal knowledge of, or experience with, the perceived individual, by the use of other social category stereotypes such as race or social class, or by noticeable facial structure features on the perceived person (Todorov et al., 2015).

Additionally, there is a somewhat different type of stereotype that can be influential in the perception process. These stereotypes do not always align with the usual demographic social groupings of age, gender, race, and social class. Edward Thorndike first used the term “Halo Effect” to describe how the use of a perceived visible characteristic of an unknown person led to an overall positive perception of that person (Thorndike, 1920). The Halo Effect is classified as a stereotype; in fact, an alternative phrase for the Halo Effect is the “what is beautiful is good” stereotype (Dion et al., 1972). Early research on the Halo Effect focused on perceived physical attractiveness as an inducer of the effect. High physical attractiveness generally leads to positive inferences (the Halo Effect), while low physical attractiveness generally leads to negative trait inferences (the Horns Effect, the opposite of the Halo Effect; Dion and Berscheid, 1974). More recent research continues to demonstrate that physical attractiveness is an initiator of the Halo Effect (Andreoni and Petrie, 2008; Eagly et al., 1991; Little et al., 2006; Thiruchselvam et al., 2016; Zebrowitz and Franklin, 2014) and the Horns Effect (Cook et al., 2003). A metanalytic review conducted by Langolis et al. (2000), demonstrated that these effects occur cross-culturally.

However, the original Halo Effect definition allows for the possibility of visible characteristics other than attractiveness to function as inference inducers. Zebrowitz and colleagues reported that, independent of perceived facial attractiveness, faces perceived as “baby-faced” (a younger, more baby-like facial appearance) induce Halo Effects, while more mature faces (a chronologically older facial appearance) induce Horns Effects (Zebrowitz and Franklin, 2014; Zebrowitz et al., 1996).

Despite the existence of an entire industry devoted to regaining physical attractiveness and youthfulness by reducing external markers of aging, nothing short of cosmetic surgery can come

close to completely erasing external aging markers. Moreover, although individuals can somewhat attenuate others’ perceptions of their facial attractiveness and age through their clothing, grooming and hairstyle choices, their faces still show signs of aging. Given all of this, neither facial attractiveness nor facial age are completely controlled by the perceived person. Therefore, these inference inducers are relatively difficult to control, meaning that little can be done about them at the immediate moment of perception.

Therefore, the stereotyping caused by attractiveness and babyfacedness may be similar to the attribute stereotyping caused by other difficult to control inference inducers, such as the perceived person’s age and gender. The stereotyping caused by varying attractiveness and babyfacedness, like the stereotyping caused by age and gender, can be either positive or negative, and equally resistant to change. Additionally, stereotyping is not just a perceptual phenomenon. Stereotypes influence the behavior and intentions of both perceivers and the perceived. Research on age, gender, attractiveness and babyfacedness stereotypes indicates the debilitating behavioral and motivational effects of negative stereotypes on perceived persons (Bhanot and Jovanovic, 2005; Dunning and Sherman, 1997; Kwong See and Heller, 2004; Snyder et al., 1977; Sparko and Zebrowitz, 2011).

Could these judgmental stereotypes be shifted by more controllable inference inducers, such as facial expressions? Our research examines facial expressions as (1) possible inducers of the Halo and Horns Effects, and (2) as stimuli that can positively or negatively shift age and gender stereotypes. Unlike the age, gender, attractiveness and babyfacedness inference inducers discussed above, facial expressions are more controllable by the perceived person. The particular facial expression inference inducers in this study are the smile, scowl, and neutral expressions. A review of the research on smiling by LaFrance (2011) describes the powerful impact of smiling in a variety of interpersonal domains. LaFrance’s review indicates that genuine smiling has positive effects in important social interactions such as, the mother-infant pair bond, romance, friendships, and workplace relationships.

Clues regarding smiling as a Halo Effect inducer and scowling as a Horns Effect inducer were first provided by Nisbett and Wilson (1977) and Weitzel et al. (1981). The authors had participants rate either a “warm and friendly” instructor or a “cold and distant” instructor seen in a filmed interview. Results indicated that the warm/friendly instructor was rated much more positively. Although not stated explicitly by the authors, presumably the warm/friendly instructor smiled more, and the cold/distant instructor scowled more, or at least smiled less. More directly relevant results come from a recent study by Senft et al. (2016), which indicate the positive effects of smiling on inferences. Using personality traits from the Big-Five factor structure (Goldberg, 1992), the authors compared three personality trait ratings (Agreeableness, Conscientiousness, and Extraversion) made by participants while looking at either a neutral facial expression or a smiling facial expression. They also varied the gender and the race (Asian or Caucasian) of the stimulus face. Participant personality ratings of the neutral expression varied by both gender and race. However, there were no gender and race rating variations in response to the smiling expression; instead there was universal agreement that the smiling individual was Agreeable, Extraverted, and somewhat Conscientious, regardless of gender or race. Regarding scowling, Tidball et al. (2006) found that a scowling face was rated more Neurotic than a smiling face, and less Extraverted, Conscientious, Open, and Agreeable.

The purpose of our study was to explore the effect of smiling, neutral, and scowling facial expressions on age and gender stereotypical inferences. In this study, participants were shown facial

photographs of old and young, female and male models who were either smiling, neutrally expressive, or scowling. They were asked to make inferences about the models' social perception characteristics; attractiveness, honesty, facial maturity, pleasing to look at, and threat. Additionally, participants made inferences with regard to the Big-Five personality traits (Agreeableness, Conscientiousness, Emotional Stability (opposite of Neuroticism), Extraversion, and Openness). We hypothesized that (1(a)) the smiling facial expression will trigger a Halo Effect compared to neutral and scowling, and (1(b)) the scowling facial expression will trigger a Horns Effect compared to neutral and smiling facial expressions. Additionally, we hypothesize that, (2) smiling will alter all age and gender inferences positively and scowling will alter all age and gender inferences negatively, compared to age and gender inferences made in the neutral expression condition.

Method

Design. Twelve facial photographs, head and shoulders of the models only (Ebner et al., 2010), were presented to study participants. Participants viewed six photographs of either old male and female models with scowling, neutral, and smiling expressions, or six photographs of young male and female models with scowling, neutral, and smiling expressions. Photographs and questions were randomized to control for demand and expectation confounds. Personality trait and social perception assessment data were collected from participants while they viewed each of the facial photographs.

The decision to present old and young faces in two separate questionnaires was made for two reasons; the length of a combined questionnaire and the original design of the study. First, requiring participants to view old and young male and female models would have resulted in 540 questions total, while having participants view either young male and female models or old male and female models resulted in a much shorter, less fatigue-inducing survey. Second, the selection of age as a variable was made based on our original intention of only examining one age group (older male and female models). After data collection using only older models, we determined that the inclusion of a comparison age group was needed. This resulted in a second round of data collection, using young male and female models exhibiting the same facial expressions.

Participants. Four hundred-seventy male ($n = 212$) and female ($n = 258$) participants were recruited using Amazon's Mechanical Turk (MTurk). Participant age data was collected using age ranges (e.g., 26–30 years) rather than specific ages. Participants ranged from 18 years to 51 or more years of age, with the most common age range between 26 and 30 years (21%). Participants identified as white (White-European 75%), black or African American (9%), Asian (7%), Hispanic or Latino (6%), African Indian or Canadian First Peoples (1%), and other (not indicated 2%). While an American sample was requested, specific nationality was not obtained. Of the 470 participants, 199 participants viewed the old, male and female, scowling, neutral, and smiling facial photographs and 271 participants viewed the young, male and female, scowling, neutral, and smiling facial photographs.

Prior to participating in the online survey, participants were asked to review a consent document and "agree" to the contents of the document. The consent document described the purpose of the study, directions for completing the study, a statement regarding privacy and confidentiality of their information and survey responses, a statement regarding approval of the study by committee, and researcher contact information. Informed consent was obtained from all participants included in the study. This study was approved by the Central Washington University

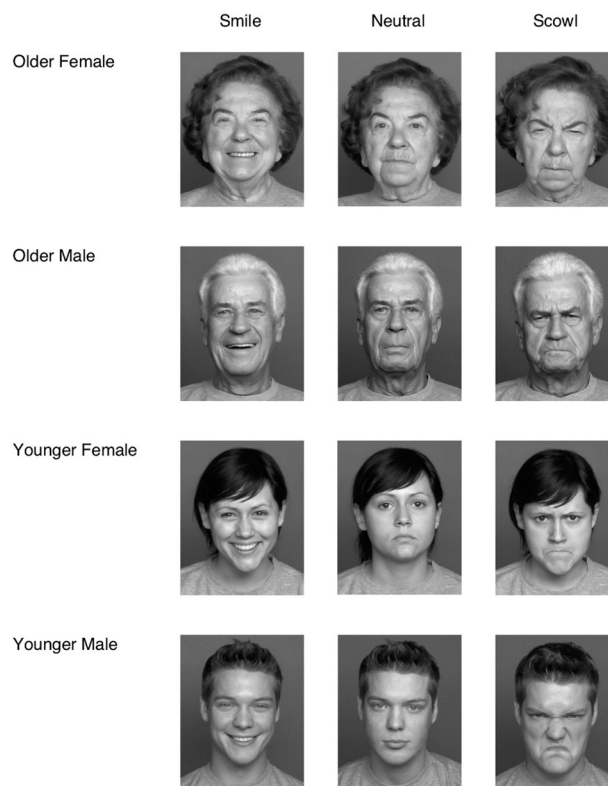


Fig. 1 Facial photographs. Facial photographs of young and old males and females exhibiting smiling, neutral, and scowling facial expressions. Photographs (Ebner et al., 2010) used by permission from the Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany.

Institutional Review Board, Human Subjects Review Council (H17127) and was conducted according to the principles expressed in the Declaration of Helsinki.

Materials

Face stimuli. The 12 face photographs used in this study are shown in Fig. 1. The face photographs were selected from the FACES database created by Ebner et al. (2010) at the Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany and were used with permission (<https://faces.mpdl.mpg.de/imeji/>). For comparison, Ebner et al.'s (2010) validation of the FACES database reported the accuracy of two datasets by three groups of raters ($N = 154$). All raters were White-European, male and female native German speakers, nationality was not reported. Ebner et al. divided raters into three rating groups by age; young (range 21 to 31 years, $M = 25.7$), middle-aged (range 44 to 55 years, $M = 49$), and older (70 to 81 years, $M = 73.6$). Raters rated all faces in a given image set, with a relatively equal number of males and females rating young (range 19–31 years, $M = 24.2$) middle-aged (39–55, $M = 49$), and old face photographs (range 69–80 years, $M = 73.2$). The male and female faces exhibited angry, disgusted, fearful, happy, neutral, and sad expressions. Ebner et al. (2010) reported overall accuracy of 81% for angry, 68% for disgusted, 81% for fearful, 96% for happy, 87% for neutral, and 73% for sad facial expressions (see Ebner et al., 2010 for a full explanation of the validation procedure and results). The photographs for the current study were selected based on the most obvious facial expression, as well as the most obvious age of the model as determined by the authors.

Table 1 Big-Five traits, Cronbach's alpha coefficients for comparison, and Minimarker traits.

Traits	Cronbach's alpha coefficient: Goldberg (1992)/Saucier (1994, 2002)/this study ^a	Adjectives ^b
Extraversion/introversion	0.90/0.83/0.75	Bold, energetic, extraverted, talkative, (bashful, quiet, shy, withdrawn)
Agreeableness/disagreeableness	0.88/0.81/0.85	Cooperative, kind, sympathetic, warm, (cold, harsh, rude, unsympathetic)
Conscientiousness/unconscientiousness	0.90/0.83/0.89	Efficient, organized, practical, systematic, (careless, disorganized, inefficient, sloppy)
Emotionally stable/unstable	0.84/0.78/0.83	Relaxed, unenvious, (envious, fretful, jealous, moody, temperamental, touchy)
Openness/closedness	0.85/0.78/0.89	Complex, creative, deep, imaginative, intellectual, philosophical, (uncreative, unintellectual)

^aCronbach's alpha for this data set are presented for comparison.^bAdjectives in parentheses indicate reversed scored items (Saucier, 1994, 2002).

Social perceptions. Using a Likert-type scale, participants were asked to answer five questions addressing the following social perceptions: *attractiveness* (1 = extremely unattractive to 7 = extremely attractive); *facial maturity* (1 = extremely baby-faced to 5 = extremely mature-faced); *honesty* (1 = extremely dishonest to 7 = extremely honest); *pleasing to look at* (1 = strongly disagree to 7 = strongly agree); and *threatening* (1 = extremely threatening to 7 = extremely non-threatening). These questions were presented for each photograph.

The Big-Five personality traits: mini-markers (MM). Goldberg (1992) developed a set of 100-adjective markers for the Big-Five personality factor structure (Agreeableness, Conscientiousness, Emotional Stability, Extraversion, and Openness) widely used for personality description. As the use of a 100-adjective questionnaire is often not ideal when combined with other assessments, Saucier (1994, 2002) created a validated subset of 40 adjective markers. Eight adjectives represented each of the five factors. For comparison, Cronbach's alpha coefficients for Goldberg's scale, the Mini-markers (Saucier, 1994) and the data set for this study are presented in Table 1.

Participants were asked how accurately each adjective described the model in the photograph on an 8-point Likert scale, from 1 (extremely inaccurate) to 8 (extremely accurate), Table 1 shows how the 40 adjectives align with the personality factors. After data collection, the 40 adjectives were collapsed into the five factors for data analysis. Each "negative" adjective was reverse scored. For example, "bashful", which is indicative of Introversion, was reverse scored (Saucier, 1994).

Procedure. Using a Qualtrics online survey format, accessed via MTurk, participants were told they would be looking at a series of facial photographs and answering questions about each photograph. Participants were then shown a photograph and responded to the questions listed above, while the photograph was still visible. The presentation of the photographs (one group viewed only six young male and female models, and the second group viewed only six old male and female models) was randomized. Additionally, the social perception question set, and the Mini-marker adjective set were randomized.

Results

All statistical tests were performed using SPSS v. 24 software and alpha level of 0.05. Two separate, mixed multivariate analyses of variance (MANOVAs) were performed to examine the effects of facial expression (within-subjects variable with three levels; smiling, scowling, and neutral), age (between-subjects variable with two levels; old, young), and gender (within-subjects variable

with two levels; male, female). The first MANOVA was done on the social perception question set and the second MANOVA on the Minimarker Big-Five personality question set. Univariate and post-hoc tests were performed to examine the effects of facial expression, age, and gender on each of the dependent variables. A Bonferroni adjustment was used for all post-hoc analyses.

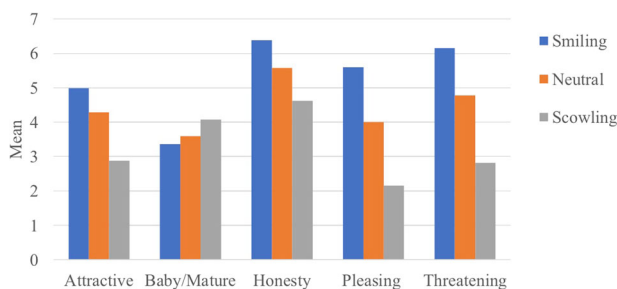
Not all participants completed all survey questions. The Minimarker Big-Five Personality Trait question set suffered the most attrition; 122 participants in the old photograph group and 146 participants in the young photograph group completed the Minimarker question set ($n = 268$), while all 470 participants completed the social perception question set. We concluded that this attrition was likely due to the length of the Minimarker question set and not the overall length of the survey because the presentation of the survey questions, as well as photographs, were randomized for each participant, resulting in the Minimarker question set frequently presented before the social perception question set.

Social perceptions. The MANOVA revealed significant main effects of facial expression (Pillai's Trace = 0.85, $F(10, 459) = 269.21$, $p < 0.001$), age (Pillai's Trace = 0.67, $F(5, 464) = 191.93$, $p < 0.001$), and gender (Pillai's Trace = 0.13, $F(5, 464) = 11.82$, $p < 0.001$). The interaction of facial expression x age was significant (Pillai's Trace = 0.15, $F(10, 459) = 8.35$, $p < 0.001$), as well as the interaction of facial expression x gender (Pillai's Trace = 0.22, $F(10, 459) = 13.04$, $p < 0.001$) and the interaction of age x gender (Pillai's Trace = 0.08, $F(5, 464) = 7.60$, $p < 0.001$). The 3-way interaction of facial expression x age x gender was also significant (Pillai's Trace = 0.31, $F(10, 459) = 20.83$, $p < 0.001$).

Facial expressions and social perceptions. Univariate testing of the individual social perception questions revealed a significant main effect of facial expression for "Attractiveness", $F(2, 936) = 844.84$, $p < 0.001$, $\eta^2 = 0.64$, "Baby-faced/Mature-faced", $F(2, 936) = 172.53$, $p < 0.001$, $\eta^2 = 0.27$, "Honesty", $F(2, 936) = 554.65$, $p < 0.001$, $\eta^2 = 0.54$, "Pleasing to look at", $F(2, 936) = 1589.32$, $p < 0.001$, $\eta^2 = 0.77$, and "Threat", $F(2, 936) = 1238.32$, $p < 0.001$, $\eta^2 = 0.73$. As Table 2 and Fig. 2 shows, participants rated the smiling facial expression as more pleasing to look at, more attractive, more honest, less facially mature, and less threatening, than the neutral and scowling facial expressions ($p < 0.001$). The neutral facial expression was always rated in between smiling and scowling; specifically, participants rated the neutral expression as less pleasing, less attractive, and less honest than the smiling facial expression ($p < 0.001$) and more mature and threatening than the smiling facial expression ($p < 0.001$). The neutral facial expression was rated as more pleasing, more attractive, and more honest than the scowling

Table 2 Facial expressions and social perceptions, age and gender combined (M, SE, and mean difference).

	Smiling expression M (SE)	Mean difference (smiling - neutral)	Neutral expression M (SE)	Mean difference (neutral - scowling)	Scowling expression M (SE)	Mean difference (scowling - smiling)
Attractive	4.99 (0.05)	0.71	4.28 (0.05)	1.40	2.88 (0.06)	-2.11
Baby/mature	3.36 (0.04)	-0.23	3.59 (0.04)	-0.48	4.07 (0.04)	0.71
Honesty	6.38 (0.04)	0.80	5.58 (0.04)	0.96	4.62 (0.05)	-1.76
Pleasing	5.60 (0.05)	1.6	4.00 (0.05)	1.85	2.15 (0.05)	-3.45
Threatening ^a	6.15 (0.04)	1.37	4.78 (0.05)	1.96	2.82 (0.06)	-3.33
Total MD ^b		4.71		6.65		11.36

All mean differences $p < 0.001$.^aScoring is reversed for Threat: 1 (Threat), 7 (Non-threat). All other variables scored as 1 (lowest), 7 (highest).^bTotal mean difference summed without consideration of sign.**Fig. 2 Social perceptions.** Means of the five social perceptions (Attractive, Baby/mature, Honesty, Pleasing to look at, and Threatening) for the smiling, neutral, and scowling facial expressions, age, and gender combined.

facial expression and less mature and less threatening than the scowling expression ($p < 0.001$).

As the next two sections demonstrate, facial expression differences did not completely subdue the inference influence of age or gender. In order to ascertain how facial expression influences age and gender, the following sections present the results for the univariate age and gender main effects and the interaction effects.

Age and social perceptions. Univariate testing revealed a significant main effect of age for “Attractiveness”, $F(1, 468) = 91.24$, $p < .001$, $\eta^2 = 0.16$, “Baby-faced/mature-faced”, $F(1, 468) = 827.01$, $p < 0.001$, $\eta^2 = 0.64$, “Honesty”, $F(1, 468) = 32.03$, $p < 0.001$, $\eta^2 = 0.06$, and “Pleasing to look at”, $F(1, 468) = 23.98$, $p < .001$, $\eta^2 = 0.05$. “Threat” was not significant. As Table 3 shows, participants rated the young models as more attractive, less mature, more pleasing to look at, and less honest than the old models ($p < 0.001$).

The interaction of facial expression \times age was significant for “Attractiveness”, $F(2, 936) = 18.45$, $p < 0.001$, $\eta^2 = 0.038$, “Baby-faced/mature-faced”, $F(2, 936) = 16.77$, $p < 0.001$, $\eta^2 = 0.035$, “Honesty”, $F(2, 936) = 10.26$, $p < 0.001$, $\eta^2 = 0.021$, more “Pleasing to look at” $F(2, 936) = 8.22$, $p < 0.001$, $\eta^2 = 0.017$, and “Threat”, $F(2, 936) = 5.09$, $p = 0.025$, $\eta^2 = 0.011$. As Table 3 shows, pairwise comparisons for facial expression by age revealed that smiling young models were judged as more attractive, less mature, and more pleasing to look at than smiling old models ($p < 0.05$). Neutral young models were judged as more attractive, less mature, less honest, more pleasing to look at, and more threatening than neutral old models ($p < 0.05$). Scowling young models were judged as more attractive, less mature, and less honest than scowling old models ($p < 0.05$).

Gender and social perceptions. Univariate testing revealed a significant main effect of gender for “Baby-faced/Mature-faced”,

$F(1, 468) = 25.19$, $p < 0.001$, $\eta^2 = 0.05$, “Honesty”, $F(1, 468) = 10.24$, $p < 0.01$, $\eta^2 = 0.021$, and “Threat”, $F(1, 468) = 32.06$, $p < 0.001$, $\eta^2 = 0.06$. As Table 4 shows, males were judged as less facially mature, less honest, and more threatening than females ($p < 0.05$).

The interaction effect of facial expression \times gender (male, female) was significant for “Attractiveness”, $F(2, 936) = 13.58$, $p < 0.001$, $\eta^2 = 0.028$, “Baby-faced/mature-faced”, $F(2, 936) = 23.07$, $p < .001$, $\eta^2 = 0.047$, “Honesty”, $F(2, 936) = 7.19$, $p < 0.01$, $\eta^2 = 0.015$, “Pleasing to look at” $F(2, 936) = 16.23$, $p < 0.001$, $\eta^2 = 0.034$, and “Threat”, $F(2, 936) = 24.93$, $p = 0.001$, $\eta^2 = 0.051$. As Table 4 shows, pairwise comparisons for facial expression by gender revealed that smiling male models were judged as more attractive and less mature than smiling female models ($p < 0.05$). Neutral male models were judged as less mature than neutral female models ($p < 0.05$), and scowling male models were judged as less pleasing to look at and more threatening than scowling female models ($p < 0.05$). See Figs. 3–7 for graphic representations of the interaction between the facial expressions, age, and gender for the social perception questions.

Big-Five personality traits. The MANOVA revealed significant main effects of facial expression (Pillai’s Trace = 0.76, $F(10, 257) = 82.94$, $p < 0.001$), age (Pillai’s Trace = .17, $F(5, 262) = 10.42$, $p < 0.001$), and gender (Pillai’s Trace = 0.20, $F(5, 262) = 11.57$, $p < 0.001$). The interaction of facial expression \times age was significant (Pillai’s Trace = 0.26, $F(10, 257) = 8.91$, $p < 0.001$), as was the interaction of facial expression \times gender (Pillai’s Trace = 0.11, $F(10, 257) = 3.16$, $p < 0.01$) and age \times gender (Pillai’s Trace = 0.12, $F(5, 262) = 7.03$, $p < 0.001$). The 3-way interaction of facial expression \times age \times gender was also significant (Pillai’s Trace = 0.27, $F(10, 257) = 9.35$, $p < 0.001$).

Facial expression and Big-Five personality traits. Univariate testing of the individual Big-Five personality traits revealed a significant main effect of facial expression for Agreeableness, $F(2, 532) = 563.08$, $p < 0.001$, $\eta^2 = 0.68$, Conscientiousness, $F(2, 532) = 187.91$, $p < 0.001$, $\eta^2 = 0.41$, Emotional Stability, $F(2, 532) = 482.49$, $p < 0.001$, $\eta^2 = 0.65$, Extraversion, $F(2, 532) = 319.43$, $p < 0.001$, $\eta^2 = 0.55$, and Openness, $F(2, 532) = 243.62$, $p < 0.001$, $\eta^2 = 0.48$. As Table 5 and Fig. 8 show, the smiling models were perceived as more Agreeable, Conscientious, Emotionally Stable, Extraverted, and Open than the scowling and neutral models ($p < 0.001$). The neutral models were perceived as more Agreeable, Conscientious, Emotionally Stable and Open than the scowling models ($p < 0.001$). The neutral models were also less Extraverted than the scowling models ($p < 0.001$).

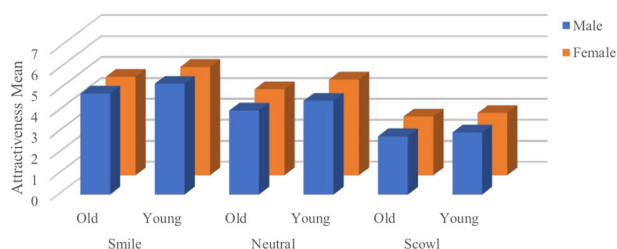
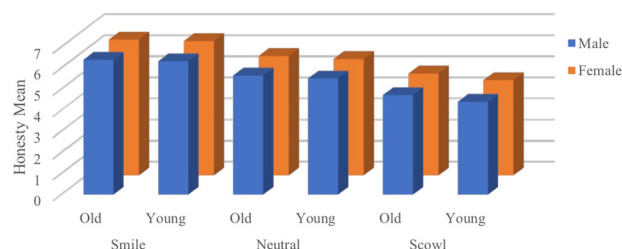
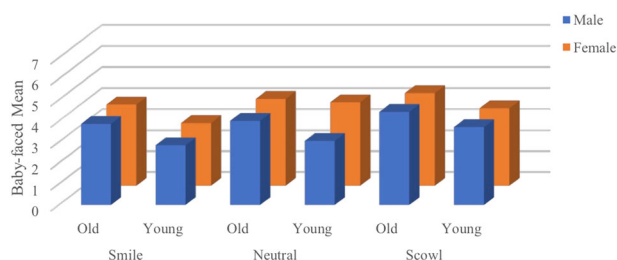
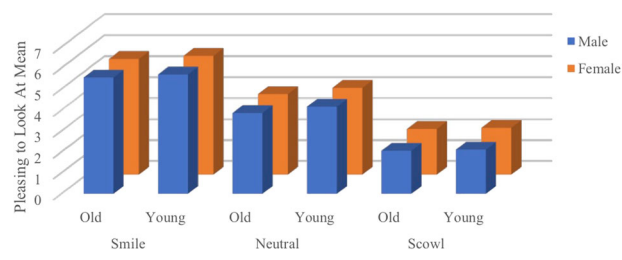
Similar to the main effect sizes for facial expression and the social perception questions, the main effect sizes for facial expression and the Big-Five personality trait questions were substantially large; nonetheless the MANOVA main effects for

Table 3 Facial expression x age, social perception questions (M, SE, mean difference).

	Smile			Neutral			Scowl		
	Young M (SE)	Old M (SE)	MD ^a	Young M (SE)	Old M (SE)	MD ^b	Young M (SE)	Old M (SE)	MD ^c
Attractiveness	5.46 (0.06)	4.52 (0.07)	0.94*	4.75 (0.06)	3.82 (0.07)	0.93*	3.05 (0.07)	2.70 (0.09)	0.35*
Baby/mature	2.48 (0.05)	4.23 (0.06)	-1.75*	2.64 (0.05)	4.55 (0.05)	-1.91*	3.34 (0.05)	4.78 (0.06)	-1.44*
Honesty	6.31 (0.06)	6.45 (0.06)	-0.14	5.44 (0.05)	5.73 (0.06)	-0.29*	4.30 (0.07)	4.93 (0.08)	-0.63*
Pleasing	5.75 (0.06)	5.46 (0.07)	0.29*	4.30 (0.07)	3.70 (0.08)	0.60*	2.20 (0.06)	2.09 (0.08)	0.11
Threat ^d	6.09 (0.06)	6.20 (0.07)	-0.11	4.89 (0.07)	4.66 (0.08)	0.23*	2.73 (0.08)	2.91 (0.09)	-0.18

* $p < 0.05$.^aMean difference between smiling young and old.^bMean difference between neutral young and old.^cMean difference between scowling young and old.^dScoring is reversed for Threat: 1 (Threat), 7 (Non-threat). All other variables scored as 1 (lowest), 7 (highest).**Table 4 Facial expression x gender, social perception questions (M, SE, mean difference).**

	Smiling			Neutral			Scowling		
	Male M (SE)	Female M (SE)	MD ^a	Male M (SE)	Female M (SE)	MD ^b	Male M (SE)	Female M (SE)	MD ^c
Attractiveness	5.11 (0.05)	4.87 (0.06)	0.24*	4.19 (0.05)	4.38 (0.06)	-0.19	2.86 (0.07)	2.89 (0.06)	-0.03
Baby/mature	3.22 (0.05)	3.50 (0.05)	-0.28*	3.48 (0.04)	3.71 (0.04)	-0.23*	4.10 (0.05)	4.03 (0.04)	0.07
Honesty	6.35 (0.05)	6.42 (0.05)	-0.07	5.58 (0.05)	5.59 (0.05)	-0.01	4.52 (0.06)	4.72 (0.06)	-0.20
Pleasing	5.64 (0.05)	5.57 (0.05)	0.07	4.03 (0.06)	3.98 (0.06)	0.05	2.03 (0.05)	2.27 (0.06)	-0.24*
Threat ^d	6.05 (0.05)	6.24 (0.05)	-0.19	4.83 (0.06)	4.72 (0.06)	0.11	2.56 (0.07)	3.07 (0.07)	-0.51*

* $p < 0.05$.^aMean difference between smiling male and female.^bMean difference between neutral male and female.^cMean difference between scowling male and female.^dScoring is reversed for Threat: 1 (Threat), 7 (Non-threat). All other variables scored as 1 (lowest), 7 (highest).**Fig. 3 Facial expression and ratings of Attractiveness.** This graph presents the interaction of facial expression, age, and gender on mean ratings of Attractiveness.**Fig. 5 Facial expression and Honesty.** This graph presents interaction of facial expression, age, and gender on mean ratings of Honesty.**Fig. 4 Facial expression and Baby-faced/mature-faced.** This graph presents interaction of facial expression, age, and gender on mean ratings of Baby-faced/mature-faced.**Fig. 6 Facial expression and Pleasing to Look At.** This graph presents interaction of facial expression, age, and gender on mean ratings of Pleasing to Look At.

age and gender were significant. The following sections present the results for the univariate main effects of age and gender, and the interaction effects of facial expression, age, and gender for the perceptions of the Big-Five personality traits.

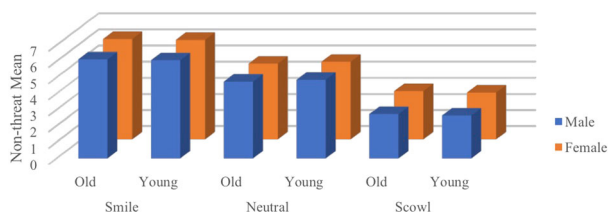
Age and Big-Five personality traits. Univariate results revealed a significant main effect of age for Conscientiousness only, $F(1, 266) = 9.37$, $p < 0.001$, $\eta^2 = 0.03$, participants rated the young models as less Conscientious than the old models. Results

revealed a significant interaction of facial expression x age with regard to Agreeableness, $F(2, 532) = 5.03$, $p < 0.01$, $\eta^2 = 0.019$, Conscientiousness, $F(2, 532) = 9.37$, $p < 0.001$, $\eta^2 = 0.034$, Emotional Stability, $F(2, 532) = 4.72$, $p = 0.009$, $\eta^2 = 0.017$, and Extraversion $F(2, 532) = 10.56$, $p < 0.001$, $\eta^2 = 0.038$, Openness was not significant.

As Table 6 shows, pairwise comparisons for facial expression by age revealed that smiling young models were judged as less Agreeable, less Conscientious, and less Emotionally Stable than smiling old models ($p < 0.05$). Neutral young models were judged as more Agreeable, less Conscientious, and less Extraverted than neutral old models ($p < 0.05$). Scowling young models were judged as less Conscientious and more Emotionally Stable than scowling old models ($p < 0.05$).

Gender and Big-Five personality traits. Univariate results revealed a significant main effect of gender for Agreeableness, $F(1, 266) = 6.50$, $p = 0.01$, $\eta^2 = 0.02$ and Extraversion, $F(1, 266) = 51.50$, $p < 0.001$, $\eta^2 = 0.16$. Conscientiousness, Emotional Stability, and Openness were not significant. Males were perceived as less Agreeable ($p < 0.02$), and more Extraverted ($p < 0.001$) than females.

Results revealed a significant interaction of facial expression x gender for Agreeableness $F(2, 532) = 9.30$ $p < 0.001$, $\eta^2 = 0.03$, and Emotional Stability, $F(2, 532) = 5.88$, $p = 0.001$, $\eta^2 = 0.03$. Conscientiousness, Extraversion, and Openness were not significant. As Table 7 shows, pairwise comparisons for facial expression and gender revealed that Scowling male models were judged as less Agreeable than scowling female models ($p < 0.02$). See Figs. 9–13 for graphic representations of the interaction between the facial expressions, age, and gender for the Big-Five personality traits.



* Note here that threat has been reversed scored to indicate levels of non-threat.

Fig. 7 Facial expression and Non-threat. This graph presents interaction of facial expression, age, and gender on mean ratings of Non-threat. Note that the ratings of Threat presented previously have been reverse scored here for better visual comparison.

Discussion

Results from this study support hypotheses 1a, 1b, and 2, and also indicate the practical usefulness of the Halo and Horns concepts. This study used facial photographs of four models who varied in facial expression, age and gender. When the models in the photographs were smiling, the Halo Effect occurred regardless of the particular model, age of the model or gender of the model. Participants generally rated the smiling faces as pleasing to look at, non-threatening, honest, Agreeable, Conscientious, Emotionally Stable, Extraverted, and Open. The scowling face results supported the Horns Effect; scowling faces were more likely to be perceived as not pleasing to look at, threatening, less honest, less Conscientious, less Agreeable, Emotionally Unstable, less Extraverted, and less Open, again regardless of the particular model, or the age or gender of the model. Additionally, all smiling faces were rated as more attractive and baby-faced when compared to all scowling faces, which were rated as more unattractive and mature-faced. These results are not surprising, given the research on physical attractiveness and baby/mature-facedness (Andreoni and Petrie, 2008; Zebrowitz and Franklin, 2014). The results of this study complement previous research by demonstrating an association between facial expressions and the perceptions of attractiveness and baby-facedness.

Facial expressions had a much greater impact on the facial inference stereotyping process, as measured by effect sizes, than either age or gender. The facial expression effect sizes were substantial, many of them a half standard deviation or greater. The facial expression main effect sizes for the social perception and personality trait questions were larger than the main effect sizes for either age or gender, with the exception of the baby/mature-faced question, where age had a large effect size. The same conclusion is reached by examining the mean differences between the two levels of each independent variable: smiling minus scowling for facial expression, ignoring neutral; young minus old for age; and male minus female for gender. Again, the facial expression mean differences are larger than either the age or gender mean differences, with the exception of the baby/mature-faced question for age.

As noted in the results, as well as in Tables 3, 4, 6 and 7, a comparison of the means provides further explanation of the facial expression inferences versus the age and gender inferences. The tables present the inference means from all ten dependent variables for the smiling, neutral and scowling expressions, as well as for age and gender. In nine of out of ten variables, the neutral means fall between the smiling and scowling means. Conceivably, the neutral expression means reflect 'pure' age and gender perceptions, made without the distracting influence of an obvious emotionally laden facial expression. In all dependent variables, regardless of the specific age and gender configuration, smiling made inferences more positive, compared to neutral, and scowling, with the exception of extraversion, made inferences more

Table 5 Facial expressions and the Big-Five personality traits (M, SE, Mean difference).

	Smiling M (SE)	Mean difference (smiling - neutral)	Neutral M (SE)	Mean difference (neutral - scowling)	Scowling M (SE)	Mean difference (scowling - smiling)
Agreeableness	6.31 (0.06)	1.76	4.54 (0.07)	1.43	3.12 (0.07)	-3.19
Conscientiousness	5.69 (0.06)	0.45	5.24 (0.06)	0.83	4.41 (0.06)	-1.29
Emotional Stability ^a	3.05 (0.06)	-1.26	4.32 (0.06)	-1.36	5.67 (0.06)	2.62
Extraverted	5.90 (0.05)	1.57	4.33 (0.05)	-0.45	4.78 (0.05)	-1.12
Openness	5.26 (0.06)	0.61	4.65 (0.06)	0.73	3.92 (0.06)	-1.34
Total MD ^b		5.65		4.80		9.56

All mean differences $p < 0.001$.

^aScaling for Emotional Stability: 1 (highest), 8 (lowest). Scaling for the other four traits: 1 (lowest), 8 (highest).

^bTotal mean difference summed without consideration of sign.

negative, compared to neutral. The smiling results are supportive of findings by Hack (2014), which showed that participants rated smiling female faces more warmly than neutral female faces. These results are supportive of those reported by Senft et al. (2016) study mentioned previously, which showed that smiling negated age and gender as inference inducers.

Figure 2 through 13 graphically confirm the results of the hypothesis tests. Figure 2 displays the results for all of the five social perception variables, for smiling, neutral, and scowling, respectively. The most obvious differences are the heights of the bars, with the bars in the smiling condition the highest (most positive), and the bars in scowling the lowest (least positive), with the bar heights in the neutral condition in between. In Figures 3 through 7 (graphically examining each of the social perception variables) there are two obvious age differences. In all three facial expression conditions the old faces are perceived as more mature-

faced, and in the smiling and neutral conditions the old faces are perceived as less attractive than the young faces. Scowling reduced the age attractiveness difference and the gender differences are small, as indicated by the small effect sizes compared to facial expression effect sizes.

Figure 8 displays the results for all of the Big-Five factors, for smiling, neutral and scowling, respectively. A comparison of the three facial expression conditions shows the most positive results for all five factors in the smiling condition, the least positive results in scowling, with the in the neutral condition falling in between. In Figs. 9 through 13, which examine each of the Big-Five traits separately, the age and gender differences are small, a reflection of the small effect sizes.

Although the effect sizes in our data show that facial expressions are powerful inducers of the Halo and Horns effects, age and gender differences continue to influence the inference process even when unambiguous facial expressions of either positive or negative valence are present. We found that models with younger faces were perceived as more attractive, pleasing to look at, baby-faced, and less honest than older faced individuals. These results are supportive of previous findings on the effects of age on facial inferences (Chan et al., 2012). Additionally, females were perceived as less threatening, more mature and honest than males, results that are supportive of previous research on gender stereotypes (Lockenhoff et al., 2014). These results are not surprising given the results of the research cited in the introduction section on stereotype formation and maintenance. Stereotypes are cognitive shortcuts that are very useful in social interaction. Results from the Martin et al. study (2014) indicates that stereotypes aid in the passage of social information from one person to another.

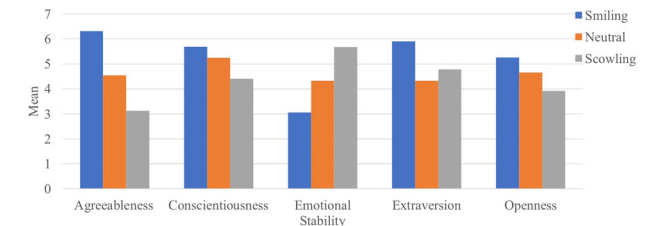


Fig. 8 Big-five personality traits. Means of the Big-five personality traits (Agreeableness, Conscientiousness, Emotional Stability, Extraversion, and Openness) for the smiling, neutral, and scowling facial expressions, age, and gender combined.

Table 6 Facial expression x age, Big-Five personality traits (M, SE, Mean difference).									
	Smile			Neutral			Scowl		
	Young M (SE)	Old M (SE)	MD ^a	Young M (SE)	Old M (SE)	MD ^b	Young M (SE)	Old M (SE)	MD ^c
Agreeableness	6.15 (0.08)	6.46 (0.09)	-0.31*	4.69 (0.10)	4.40 (0.12)	0.29*	3.15 (0.09)	3.09 (0.10)	0.06
Conscientiousness	5.60 (0.07)	5.80 (0.08)	-0.20*	5.06 (0.08)	5.42 (0.09)	-0.36*	4.02 (0.08)	4.80 (0.09)	-0.78*
Emotionally Stable ^d	3.20 (0.08)	2.91 (0.09)	0.29*	4.23 (0.09)	4.40 (0.09)	-0.17	5.80 (0.08)	5.54 (0.09)	0.26*
Extraverted	5.87 (0.07)	5.94 (0.07)	-0.07	4.13 (0.07)	4.53 (0.07)	-0.40*	4.87 (0.06)	4.69 (0.07)	0.18
Openness	5.27 (0.08)	5.25 (0.08)	0.02	4.74 (0.07)	4.56 (0.08)	0.18	3.89 (0.08)	3.96 (0.09)	-0.07
^a p < 0.05. ^b Mean difference between smiling young and old. ^c Mean difference between neutral young and old. ^d Mean difference between scowling young and old. ^e Scaling for Emotional Stability: 1 (highest), 8 (lowest). Scaling for the other four traits: 1 (lowest), 8 (highest).									

Table 7 Facial expression and gender, Big-Five personality traits (M, SE, mean difference).									
	Smile			Neutral			Scowl		
	Male M (SE)	Female M (SE)	MD ^a	Male M (SE)	Female M (SE)	MD ^b	Male M (SE)	Female M (SE)	MD ^c
Agreeableness	6.26 (0.07)	6.35 (0.06)	-0.09	4.59 (0.08)	4.50 (0.08)	0.09	2.95 (0.07)	3.28 (0.08)	-0.33*
Conscientiousness	5.70 (0.06)	5.70 (0.06)	0.0	5.25 (0.07)	5.23 (0.07)	0.02	4.35 (0.07)	4.46 (0.07)	-0.11
Emotionally Stable ^d	3.01 (0.07)	3.09 (0.07)	-0.08	4.28 (0.07)	4.35 (0.07)	-0.07	5.76 (0.06)	5.58 (0.07)	0.18
Extraverted	6.00 (0.06)	5.81 (0.05)	0.19	4.45 (0.06)	4.22 (0.06)	0.23	4.92 (0.05)	4.64 (0.05)	0.28
Openness	5.23 (0.06)	5.29 (0.06)	-0.06	4.61 (0.06)	4.68 (0.07)	-0.07	3.89 (0.07)	3.96 (0.07)	-0.07
^a p < 0.02. ^b Mean difference between smiling male and female. ^c Mean difference between neutral male and female. ^d Mean difference between scowling male and female. ^e Scaling for Emotional Stability: 1 (highest), 8 (lowest). Scaling for the other four traits: 1 (lowest), 8 (highest).									

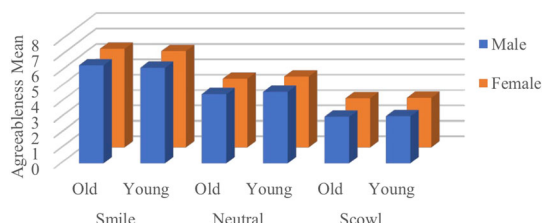


Fig. 9 Facial expression and Agreeableness. This graph presents interaction of facial expression, age, and gender on mean ratings of Agreeableness.

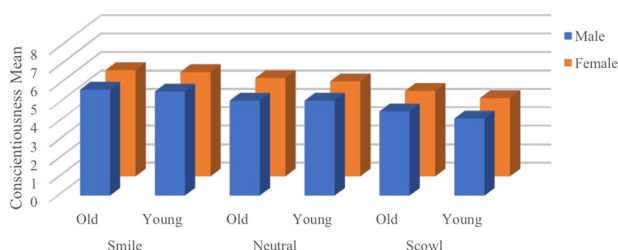


Fig. 10 Facial expression and Conscientiousness. This graph presents interaction of facial expression, age, and gender on mean ratings of Conscientiousness.

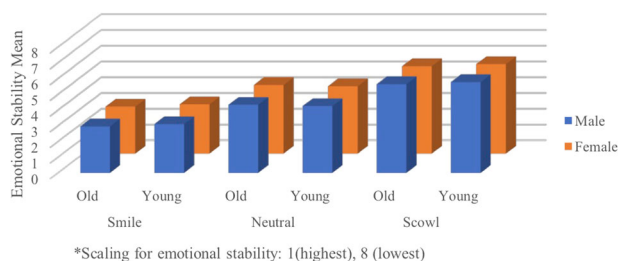


Fig. 11 Facial expression and Emotional Stability. This graph presents interaction of facial expression, age, and gender on mean ratings of Emotional Stability.

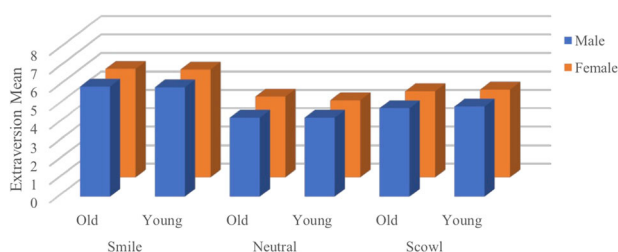


Fig. 12 Facial expression and Extraversion. This graph presents interaction of facial expression, age, and gender on mean ratings of Extraversion.

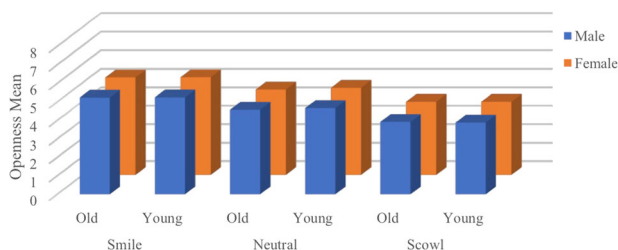


Fig. 13 Facial expression and Openness. This graph presents interaction of facial expression, age, and gender on mean ratings of Openness.

From childhood on, people discuss the actions and attributes of others, and stereotypes are widely used in these discussions (Allport, 1954). Essentially stereotypes are used in person perception all the time, but they can be altered in significant ways, as the results of this study show.

Limitations and future directions. It is possible that the availability bias and demand characteristics are partially responsible for the results of this study. The availability bias states that individuals make decisions based on information that is easily available to them. In this study, the information most easily available to respondents would be the obviously different facial expressions presented on the same faces, which presumably indicate the inference expectations of the researchers. A reasonable conclusion would be that respondents gave positive inferences to the smiling faces and negative inferences to the scowling faces because seeing different expressions of the same faces gave clues as to the researchers' expectations. This would be reasonable if the respondents had no or limited experience with smiling and scowling faces in their lives. However, given that all the respondents were 18 or older, a more reasonable conclusion regarding the results is that the respondents have seen many smiling and scowling faces in their lifetimes, that they have made inferences from the many faces they have seen, and that the inferences they made to the smiling and scowling photographs of real people in the study were drawn from their memories of inferences they had made to the smiling and scowling facial expressions of real people they had seen in the past.

A similar argument can be made regarding the age and gender stereotypes that were seen in all three expression conditions. The participants have been using age and gender stereotypes since early childhood to make attribute inferences, and they remembered those inferences as they viewed the appropriate age and gender configurations in the presented facial photographs. In other words, human faces, however seen, are lifelong sources of information about others, that we automatically use to make perceptions about those others, and many, but not all, of those perceptions are stereotypes. Indeed, calling these inferences stereotypes does not imply that they are always inaccurate. Research on the "kernel of truth" hypothesis indicates that inferences can be accurate when assessing the personality of others from their faces (Berry, 1990). Nonetheless, our results indicate that age, gender, attractiveness and babyfaced-ness stereotypes, regardless of their accuracy or inaccuracy, can be shifted by facial expressions. Creating a positive social interaction opportunity that could overcome a negative inference has positive value, regardless of the accuracy or inaccuracy of the inference.

As with most within-subjects designs, possible demand characteristics may be lessened by the use of randomization, and when possible, counterbalancing. As indicated in the Method section, we randomized the sequence of both photographs and questions and, in addition, we presented one of the independent variables as a between-subjects variable (age of the model). However, it is possible these controls did not entirely eliminate demand characteristics. Future replications of this work should attempt to verify the absence of these characteristics using a completely between-subjects design.

In this study, both the stimuli and the participants are hardly representative of global diversity. Specifically, two limitations of this research is the exclusive use of White-European (in this case, German) faces as stimuli and the lack of participant diversity (age and ethnicity) in the MTurk sample. A related limitation has to do with the limited number of facial models used in the study. It is possible that the results are due to the idiosyncrasies of the faces selected as models. This could be overcome by using a larger

number of diversely different faces, posing the same facial expressions. In addition to the use of a more diverse set of facial stimuli, the use of dynamic versus static stimuli may provide further insight into gender and age stereotypes for the expressions of smiling, neutral, and scowling (Biele and Grabowska, 2006).

As noted in the Results section, the loss of participant responses to the Big-Five adjectives requires further examination. It is possible that a qualitative difference between the social perception questions and the Big-Five adjectives resulted in this loss of responses, rather than the tedious nature of the use of 40 adjectives to assess faces over and over again. The social perception questions were singular and simple, requiring a quick perceptual judgment that respondents had likely made many times in the past based on their own experiences and perceptions of attractiveness, baby/mature-faced appearance, pleasing to look at, honesty, and threat. In contrast, the 40 adjective Minimarker Big-Five scale may have required a deeper level of inference; participants had to think about aspects of a stranger's face that they normally would not consider, such as how bold or bashful an individual is. However, it is worth noting that the results from the social perception questions and the Big-Five questions were similar; respondents made positive inferences to smiling faces and negative inferences to scowling faces with both sets of questions, indicating that the qualitative difference between the question sets likely had little influence on the results.

While we do not perceive the following as a limitation of our study, future analysis of this data will include a comparison of the accurate and inaccurate emotion perception with regard to Ekman's list of emotions (Ekman and Friesen, 1976). One question that arose after analysis of this data was: would participants make different social and personality inferences if they inaccurately identified the emotion of the model? If participants do make different social characteristic and personality trait inferences based on different emotional labels from the same facial expressions, perhaps the emotional label is serving as an appraisal to further inferences. That was the conclusion made by Schacter and Singer (1962) in their seminal study of emotion attribution, a conclusion that has been expanded into emotion appraisal theory (Scherer and Grandjean, 2008; Keltner et al., 2019). This theory postulates that perceivers of facial expressions go through an initial quick appraisal process, leading to a general perception of pleasantness or unpleasantness. This is followed by a secondary appraisal process based on a search for possible causes of the observed facial expression, and the attributional conclusion leads to a specific emotional designation (label). The label then serves as a guide to a wider range of internal inferences about the observed person. Appraisal theory points to an important future direction in facial inferencing research. Data analysis from inferencing studies needs to clearly distinguish the external stimulus, the facial expression, from the internal emotional label, rather than assuming, for example, that scowling is automatically paired with anger.

Conclusion

In this study, the participants exposed to the young female and male models saw the same three facial expressions as did the participants exposed to the old female and male models. It is reasonable to assume that the participants knew that it was the same model who was either smiling, scowling, or neutrally expressive, yet the inferences attributed to each model's different expressions varied greatly. This is indicative of the power of facial expressions. While this seems like an obvious point, it has practical personal implications, especially for those stigmatized on the basis of physical appearance. The Horns Effects in particular, and negative age and gender stereotyping more generally,

may be somewhat countered by genuine smiling (LaFrance, 2011). The negative inferences comprising the Horns Effect reflect prejudice, and potential discrimination. Perhaps smiling is a relatively effortless and inexpensive anti-discrimination tool that may somewhat counteract ageism and sexism. This does not happen because smiling makes negative age and gender stereotypes disappear completely; rather, smiling can create a window of opportunity for positive social interaction between the negatively stereotyped person and the person doing the stereotyping. Smiling may create at least a temporary bridge between two people who otherwise would not engage in interaction.

Permission to use facial stimuli. The facial photographs used in this study were used by permission from FACES database created by the Max Planck Institute (Ebner et al., 2010).

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to ongoing use in current research but are available from the corresponding author upon request.

Received: 2 December 2019; Accepted: 19 May 2020;

Published online: 02 July 2020

References

- Allport G (1954) *The nature of prejudice*. Addison-Wesley Publishing Co, Cambridge, MA
- Andreoni J, Petrie R (2008) Beauty, gender, and stereotypes: evidence from laboratory experiments. *J Econ Psychol* 29(1):73–93. <https://doi.org/10.1016/j.joep.2007.07.008>
- Berry D (1990) Taking people at face value: evidence for the kernel of truth hypothesis. *Soc Cognition* 8:343–361
- Bhanot R, Jovanovic J (2005) Do parents' academic gender stereotypes influence whether they intrude on their children's homework? *Sex Roles* 52:597–607
- Biele C, Grabowska A (2006) Sex differences in perception of emotion intensity in dynamic and static facial expressions. *Exp Brain Res* 171:1–6. <https://doi.org/10.1007/s00221-005-0254-0>
- Bryan L, Coleman M, Ganong L, Bryan S (1986) Person perception: family structure as a cue for stereotyping. *J Marriage Fam* 48:169–174
- Chan W, McCrae RR, De F et al. (2012) Stereotypes of age differences in personality traits: Universal and accurate? *J Pers Soc Psychol* 103:1050–1066
- Cook GI, Marsh RL, Hicks JL (2003) Halo and devil effects demonstrate valenced-based influences on source-monitoring decisions. *Conscious Cogn* 12:257–278
- Dion KK, Berscheid E (1974) Physical attractiveness and peer perception among children. *Sociometry* 37:1–12. <https://doi.org/10.2307/2786463>
- Dion KK, Berscheid E, Walster E (1972) What is beautiful is good. *J Pers Soc Psychol* 24:285–290
- Dunning D, Sherman D (1997) Stereotypes and tacit inference. *J Pers Soc Psychol* 73:459–471
- Eagly AH, Ashmore RD, Mahijani MG, Longo LC (1991) What is beautiful is good, but...: a meta-analytic review of research on the physical attractiveness stereotype. *Psychol Bull* 110:109–128. <https://doi.org/10.1037/0033-2909.110.1.109>
- Ebner NC, Riediger M, Lindenberg U (2010) FACES-A database of facial expressions in young, middle-aged, and older women and men: development and validation. *Behav Res Methods* 42:351–362. <https://doi.org/10.3758/BRM.42.1.351>
- Ebner NC, Joerg L, Voelke MC, Riediger M, Lin T, Lindenberg U (2018) An adult developmental approach to perceived facial attractiveness and distinctiveness. *Front Psychol* 9:561. <https://doi.org/10.3389/fpsyg.2018.00561>
- Ekman P, Friesen WV (1976) *Pictures of facial affect*. Consulting Psychologists Press, CA
- Ellemers N (2018) Gender stereotypes. *Annu Rev Psychol* 69:275–298
- Gilgovich T, Keltner D, Chen S, Nisbett RE (2019) *Social psychology*, 5th edn. W. W. Norton and Co, New York, NY
- Goldberg LR (1992) The development of markers for the Big-Five factor structure. *Psychol Assess* 4:26–42
- Greenwald AG, Banaji MR (1995) Implicit social recognition: attitudes, self-esteem, and stereotypes. *Psychol Rev* 102:4–27. <https://doi.org/10.1037/0033-295X.102.1.4>
- Hack T (2014) Forming impressions: effects of facial expressions and gender stereotypes. *Psychol Rep* 14:557–571

- Hoffman C, Hurst N (1990) Gender stereotypes. *J Pers Soc Psychol* 58:197–208
- Johnson KL, Lick DJ, Carpinella LM (2015) Emergent research in social vision: an integrated approach to the determinants and consequences of social categorization. *Soc Personal Psychol Compass* 9:15–30. <https://doi.org/10.1111/spc3.12147>
- Keltner D, Sauter D, Tracy D, Cowen A (2019) Emotional expression: advances in basic emotion theory. *J Nonverbal Behav* 43:133–160
- Kwong See S, Heller R (2004) Judging older targets' discourse: how do age stereotypes influence evaluations? *Exp Aging Res* 30:63–73
- LaFrance M (2011) *Lip service*. W.W. Norton and Co, New York, NY
- Lamont RA, Swift HJ, Abram D (2015) A review and meta-analysis of age-based stereotype threat: negative stereotypes, not facts, do the damage. *Psychol Aging* 30:180–193
- Langolis JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M (2000) Maxims or Myths of beauty? a meta-analytic and theoretical review. *Psychol Bull* 126:390–423. <https://doi.org/10.1037/0033-2909.126.3.390>
- Little AC, Burt DM, Perrett DI (2006) What is good is beautiful: face preference reflects desired personality. *Pers Ind Diff* 41:1107–1118. <https://doi.org/10.1016/j.paid.2006.04.015>
- Löckenhoff CE, Chan W, McCrae RR et al. (2014) Gender stereotypes of personality: universal and accurate? *J Cross Cult Psychol* 45:675–694. <https://doi.org/10.1177/0022022113520075>
- Macrae CN, Bodenhausen GV (2000) Social cognition: thinking categorically about others. *Annu Rev Psychol* 51:93–120
- Martin D, Hutchison J, Slessor G, Urquhart J, Cunningham S, Smith K (2014) The spontaneous formation of stereotypes via cumulative cultural evolution. *Psychol Sci* 25:1777–1786
- Nisbett RE, Wilson TD (1977) The Halo Effect: evidence for unconscious alteration of judgments. *J Pers Soc Psychol* 35:2150–256. <https://doi.org/10.1037/0022-3514.35.4.250>
- Parmley M, Cunningham JG (2014) She looks sad, but he looks angry: the effects of age, gender and ambiguity on emotion perception. *J Soc Psychol* 154:323–338. <https://doi.org/10.1080/00224545.2014.901287>
- Saucier G (1994) Mini-markers: a brief version of Goldberg's unipolar Big-Five markers. *J Pers Assess* 63:506–516
- Saucier G (2002) Orthogonal markers for orthogonal factors: the case of the Big Five. *J Res Pers* 36:1–31
- Schachter S, Singer JE (1962) Cognitive, social, and physiological determinants of emotional state. *Psychol Rev* 69:379–399
- Scherer KR, Grandjean D (2008) Facial expressions allow inference of both emotions and their components. *Cogn Emot* 22:789–801
- Senft N, Chentsova-Dutton Y, Patten GA (2016) All smiles perceived equally: facial expressions trump target characteristics in impression formation. *Motiv Emot* 40:577–587. <https://doi.org/10.1007/s11031-016-9558-6>
- Snyder M, Tanke E, Berscheid E (1977) Social perception and interpersonal behavior: on the self-fulfilling nature of social stereotypes. *J Pers Soc Psychol* 35:656–666
- Silberstang J (2011) Learning gender: the effects of gender-role stereotypes on women's lifelong learning and career advancement opportunities. In: London M (ed) *The Oxford handbook of lifelong learning*. Oxford University Press, London, pp 370–388
- Sparko A, Zebrowitz L (2011) Moderating effects of facial expression and movement on the babyface stereotype. *J Nonverb Behav* 35:243–257
- Thiruchselvam R, Harper J, Homer A (2016) Beauty is in the belief of the beholder: cognitive influences on the neural response to facial attractiveness. *Soc Cogn Affect Neurosci* 11:1999–2008. <https://doi.org/10.1093/scan/nsw115>
- Thorndike EL (1920) A constant error in psychological ratings. *J Appl Psychol* 4:25–29. <https://doi.org/10.1037/h0071663>
- Tidball B, Prabhala S, Gallimore J (2006) Making faces: exploring perceptions of personality based on emotional expressions. *Proc Hum Factors Erg Soc Annu Meet* 50:885–888. <https://doi.org/10.1177/154193120605000905>
- Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki P (2015) Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annu Rev Psychol* 66:519–545. <https://doi.org/10.1146/annurev-Psych-113011-143831>
- Weitzel CG, Wilson TD, Kort J (1981) The halo effect revisited: forewarned is not forearmed. *J Exp Soc Psychol* 17:427–439. [https://doi.org/10.1016/0022-1031\(81\)90049-4](https://doi.org/10.1016/0022-1031(81)90049-4)
- Zebrowitz LA, Franklin RG (2014) The attractiveness halo effect and the babyface stereotype in older and younger adults: similarities, own-age accentuation, and older adult positivity effects. *Intern J Devoted Sci Aging Proc* 40:375–393. <https://doi.org/10.1080/0361073X.2014.897151>
- Zebrowitz LA, Voinescu L, Collins MA (1996) “Wide-eyed” and “crooked-faced”: determinants of perceived and real honesty across the life span. *Pers Soc Psychol Bull* 22:1258–1269. <https://doi.org/10.1177/01461672962212006>

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.K.R.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020