

Personality at Face Value: Facial Appearance Predicts Self and Other Personality Judgments among Strangers and Spouses

Raluca Petrican · Alexander Todorov · Cheryl Grady

Published online: 11 January 2014
© Springer Science+Business Media New York 2014

Abstract Character judgments, based on facial appearance, impact both perceivers' and targets' interpersonal decisions and behaviors. Nonetheless, the resilience of such effects in the face of longer acquaintanceship duration is yet to be determined. To address this question, we had 51 elderly long-term married couples complete self and informant versions of a Big Five Inventory. Participants were also photographed, while they were requested to maintain an emotionally neutral expression. A subset of the initial sample completed a shortened version of the Big Five Inventory in response to the pictures of other opposite sex participants (with whom they were unacquainted). Oosterhof and Todorov's (in Proc Natl Acad Sci USA 105:11087–11092, 2008) computer-based model of face evaluation was used to generate facial trait scores on trustworthiness, dominance, and attractiveness, based on participants' photographs. Results revealed that structural facial characteristics, suggestive of greater trustworthiness, predicted positively biased, global informant evaluations of a target's personality, among both spouses and strangers. Among spouses, this effect was impervious to marriage length. There was also evidence suggestive of a Dorian Gray effect on personality, since facial trustworthiness predicted not only spousal and stranger, but also self-ratings of extraversion. Unexpectedly, though, follow-up analyses revealed that (low) facial dominance, rather than (high) trustworthiness, was the strongest predictor of self-rated extraversion. Our present findings suggest that subtle emotional cues, embedded in the structure of emotionally neutral faces, exert long-lasting effects on personality judgments even among very well-acquainted targets and perceivers.

Keywords Big five · Facial appearance · Trustworthiness · Dominance · Married couples

R. Petrican (✉) · C. Grady
Rotman Research Institute, University of Toronto, 3560 Bathurst Street, Toronto, ON M6A 2E1,
Canada
e-mail: raluca.petrican@gmail.com

A. Todorov (✉)
Department of Psychology, Princeton University, Princeton, NJ 08540, USA
e-mail: atodorov@princeton.edu

Introduction

Because evaluating the interpersonal danger potential of another individual is crucial to successful functioning and well-being, humans have evolved rapid, intuitive, unreflective mechanisms for making such decisions, based solely on the physical attributes of their conspecifics (Todorov et al. 2008; Zebrowitz et al. 1996). Indeed, a growing body of literature is attesting to the importance of physical, mainly facial, appearance in driving perceivers' evaluation of strangers. For example, Willis and Todorov (2006) demonstrated that perceivers draw on facial appearance alone to make snap judgments on a variety of traits, of varying complexity, and that the outcome of such rapid evaluations remains relatively intact even when perceivers are given unlimited time. The pervasiveness of such effects is intriguing in light of the mixed evidence for the validity of trait judgments based solely on facial appearance (Berry and Brownlow 1989; Berry 1991; Bond et al. 1994; Olivola and Todorov 2010a, b; Zebrowitz et al. 1996).

Interestingly though, some consistent links between facial appearance and personality have indeed been documented. For example, there is some evidence that facial appearance earlier in life predicts subsequent personality development. Nevertheless, the pattern of results is rather complex, because it encompasses both self-fulfilling and self-defeating prophecy effects (especially for males, Zebrowitz et al. 1996, 1998a, b). With respect to the former, Zebrowitz et al. (1996, 1998b) reported that, among men, attractiveness and honest-looking appearance, respectively, earlier in life led to the development of congruent personality profiles in later years. Nonetheless, compatible with a self-defeating prophecy effect, Zebrowitz et al. (1998a) documented that among adolescent boys, childlike features were associated with a host of both positive (i.e., higher academic achievement) and negative (i.e., higher delinquency) behaviors, which were at odds with the personality stereotype of babyfaced individuals.

Complementing these reports on the effect of facial appearance on subsequent personality development, there are studies documenting the impact of earlier personality on facial appearance later in life (i.e., a Dorian Gray effect). For example, Zebrowitz et al. (1998b) documented that, among women, an early personality, congruent with the stereotypical profile of a physically attractive individual, predicted greater physical attractiveness later in life. Further empirical evidence consistent with a Dorian Gray effect comes from studies investigating emotion-relevant personality traits. Thus, Malatesta et al. (1987) documented that for older adults, a lifelong predisposition to experience certain emotional states remained imprinted on their faces, thereby biasing the perceivers' attribution of relevant emotional states (e.g., older adults with a hostile predisposition were judged to be angry looking even when posing with a voluntarily neutral stance).

Personality at Face Value

At zero acquaintance, perceivers can only draw on physical appearance to infer personality—and the validity of such information has been repeatedly questioned (e.g., Olivola and Todorov 2010b; Zebrowitz et al. 1996). Nevertheless, there is evidence that even under such circumstances, some trait judgments are still fairly accurate (e.g., Borke and Liebler 1992, 1993), despite the fact that both perceiver and target characteristics play a decisive role in the success of such enterprises (e.g., Ambady et al. 1995). Fuelled by these earlier findings, more recent investigations on social target legibility attempted to elucidate the differentially predictive value of physical appearance variables for specific personality

factors (e.g., Penton-Voak et al. 2006). Most of this research has employed the Big Five taxonomy, the canonical theoretical framework for assessing variations in personality for the past two decades (e.g., Costa and McCrae 1992; Goldberg 1993). Unsurprisingly, of all the Big Five factors, the two that consistently elicit the greatest self-other agreement (for a recent meta-analysis, see Connolly et al. 2007), conscientiousness and extraversion, have also been shown to possess the most valid physical indicators (Borkenau et al. 2009; Kenny et al. 1992; Naumann et al. 2009). Specifically, personal grooming-related variables, such as style of dress (e.g., formal attire), have been shown to be valid cues of conscientiousness (Albright et al. 1988; Borkenau and Liebler 1992). Likewise, facial appearance variables, such as (higher) attractiveness, (greater) symmetry and spontaneous, positive emotionally expressive behaviors (i.e., smiling) are reportedly accurate indicators of extraversion (Borkenau et al. 2009; Fink et al. 2005; Penton-Voak et al. 2006; Pound et al. 2007).

Nevertheless, complementing these investigations on the predictive value of facial appearance for substantive variations in Big Five ratings, there are studies documenting erroneous uses of physical cues to infer the personality of strangers. Interestingly, spontaneous smiling tops the list of the most “misinterpreted” observable cues, as perceivers seem to interpret it as being indicative of a variety of socially desirable traits. For example, Kenny et al. (1992) documented that observers used smiling as a cue for greater extraversion, openness to experience, and agreeableness. Similarly, Naumann et al. (2009) reported that smiling targets were (erroneously) attributed an overall positive personality profile, as reflected in greater scores on all socially desirable traits (i.e., extraversion, agreeableness, openness to experience, conscientiousness, and emotional stability, i.e., reverse coded neuroticism).

Aging and Personality Judgments

Extant research suggests that personality judgments, both in reference to the self and others, are largely impervious to age-related effects (Keightley et al. 2006; Ruby et al. 2009; Zebrowitz et al. 2013). If anything, relative to their younger counterparts, older adults seem to be more lenient judges not only of their own and close others’ personality (Grady et al. 2012), but also of strangers’ personality (Zebrowitz et al. 2013), an effect that is consistent with previous reports of an age-related positivity bias in information processing (cf. Mather and Carstensen 2005).

Unfortunately, research into potential age differences in the use of facial appearance cues to infer personality is rather scarce. Nevertheless, existing evidence suggests that younger and older adults use facial appearance cues similarly to infer personality. For example, consistent with the attractiveness stereotype (for a review, see Berscheid 1981), judges from both age groups have been found to evaluate the more physically attractive elderly targets (irrespective of their perceived age) as possessing a more positive personality profile, and having enjoyed more success in their personal and professional lives (Johnson and Pittenger 1984). Likewise, null age effects have been reported with respect to the use of subtle emotional cues to infer personality. For example, both younger and older perceivers have been found to use to a similar degree subtle cues of anger and surprise to make judgments of the targets’ danger and naivete, respectively (Franklin and Zebrowitz 2013).

In sum, current literature suggests that younger and older perceivers may be rather similar in their evaluations of others’ personality. More importantly for the present

investigation, both age groups appear to rely upon similar physical appearance cues to judge others, at least in zero acquaintance situations.

Current Research

The present studies examined the relationship between facial appearance—assessed under standardized, emotionally neutral conditions—and both self and other judgments of personality not only in a sample of very well-acquainted targets and perceivers (Study 1: long-term elderly married¹ individuals), but also at zero acquaintance (Study 2). Our aim was to test whether facial appearance exerts a long-lasting effect on personality judgments, which remains salient even for perceivers who are extremely well-acquainted with the target. To assess variations in facial appearance, we used the computational model of face evaluation developed by Oosterhof and Todorov (2008). Capitalizing on the fact that trait judgments from faces are highly intercorrelated (Todorov et al. 2008), these authors used a data-driven approach and demonstrated that most of the variance in social evaluation outcomes could be described along two dimensions (Oosterhof and Todorov 2008). One dimension, dubbed valence/trustworthiness, was shown to reflect variations in perceived positive/negative intentions and determine the valence of interpersonal judgments, as well as subsequent decisions to approach or avoid a target (Todorov et al. 2008).² Scores on this dimension were found to be most sensitive to structural variations, which increase facial resemblance to positive versus negative emotional expressions (Oosterhof and Todorov 2008; Said et al. 2009). The second evaluative dimension, dubbed power/dominance, was shown to capture variations in the perceived ability to implement one's intentions. Scores on this dimension were found to be most sensitive to variations in facial maturity and masculinity, that is, variations in the physical ability required to implement one's intentions.

In the present studies, we used Oosterhof and Todorov's (2008) face evaluation model to test a number of hypotheses regarding the link between facial appearance and ratings of personality, assessed within a Big Five framework. Although conceptualized initially as orthogonal, subsequent investigations revealed systematic intercorrelations among scores on the five factors, which are thought to reflect partly the perceivers' global positive evaluation bias, specifically their tendency to rate a target low on socially undesirable traits (i.e., neuroticism) and high on socially desirable traits (i.e., extraversion, agreeableness,

¹ There is evidence that romantic partners may be vulnerable to seeing each other through rose-colored glasses. For example, individuals who are committed to their relationship evaluate their partner's personality in idealized manner (i.e., more positively than the partner evaluates himself or herself or is evaluated by others: Murray et al. 1996). This line of work is distinct from ours, however, since our purpose is to test the link between facial appearance variables and partner personality evaluations, irrespective of their level of idealization.

² To date, Oosterhof and Todorov's (2008) model of face evaluation has been solely tested with university-aged samples. Consequently, we conducted a control study in an attempt to replicate the documented effect of facial trustworthiness on approach-avoidance behaviors (i.e., perception of rapport) in an elderly sample, similar to the one who participated in Studies 1 and 2. Fifty-seven elderly participants (15 male)—unacquainted with each other and with the main study participants—rated their perception of anticipated rapport in response to the FaceGen models of the Study 1 participants. The rapport measure was modeled after the one used by Butler et al. (2003) and it required participants to rate their agreement with two sentences, "The conversation with this person would be warm and smooth." and "I would 'click' with this person." on a scale from 1 (*completely disagree*) to 9 (*completely agree*). Participants' responses to the two items were averaged to create an index of anticipated rapport. As expected, based on prior studies with younger samples (Oosterhof and Todorov 2008), raters were more likely to anticipate greater rapport with more (rather than less) trustworthy looking targets, $b = .42$, $SE = .09$, $t(50) = 4.82$, $p < .01$.

openness to experience and conscientiousness) (cf. Anusic et al. 2009). Therefore, our hypotheses concerned the effect of facial appearance not only on judgments on the five personality factors, but also on the higher-order global evaluation bias. Prior studies suggested that positive emotionally expressive behavior is among the most utilized physical indicator of personality (Borkenau et al. 2009; Kenny et al. 1992; Naumann et al. 2009). Consequently, all our hypotheses focused on the trustworthiness dimension of Oosterhof and Todorov's (2008) model, due to its sensitivity to structural variations that increase resemblance to positive versus negative emotions. Attractiveness and facial dominance were included solely for control purposes. Thus, we tested the following set of hypotheses.

First, we investigated whether facial trustworthiness may be a valid personality cue. Recent studies demonstrated that under circumstances in which spontaneous emotionally expressive behaviors are allowed, a more cheerful appearance is uniquely linked to self and observer reports of greater extraversion (Borkenau et al. 2009). Moreover, there is evidence that a lifelong predisposition to experience certain emotional states may remain imprinted on the face, thereby being visible even under emotionally neutral conditions (Malatesta et al. 1987). Based on this evidence, we reasoned that targets, whose neutral expression would contain subtle positive, rather than negative, emotional cues (i.e., higher facial trustworthiness), would rate themselves and be rated both by their spouses (Study 1) and by strangers (Study 2) higher on extraversion.

Second, we investigated whether facial trustworthiness could impact informant evaluations of a target's personality, not only at zero acquaintance, but also among spouses. To this end, we capitalized on previous findings that the predictive breadth of spontaneous, positive emotionally expressive behaviors, such as smiling, tends to be overestimated, such that they are not only correctly used as an indicator of greater extraversion, but they are also erroneously used to predict more positive global evaluations of a target's personality (Naumann et al. 2009). Because variations in facial trustworthiness are indicative of structural resemblance to positive versus negative emotional expressions and determine the overall valence of interpersonal judgments (Fiske et al. 2007; Oosterhof and Todorov 2008), we reasoned that, consistent with a halo error effect (Thorndike 1920), participants with a more trustworthy looking appearance would receive more positive global personality evaluations both from their spouses (Study 1) and from strangers (Study 2). Importantly, we expected this effect to be an instantiation of biased perception, since recent studies provided compelling evidence that at zero acquaintance, people's trustworthiness judgments, based on facial appearance alone, do not differentiate between war criminals and war heroes, CEOs who committed financial fraud versus those who did not, or students who were found cheating on an exam versus those who were not (cf. Rule et al. 2013). We therefore hypothesized that facial trustworthiness would predict global evaluative bias in informant, but not self-evaluations of personality.

Third, we examined whether the hypothesized effect of facial trustworthiness on spousal global evaluation bias (Study 1), could be attributed to the reflexive, intuitive evaluative mechanisms that underlie first impressions (Todorov et al. 2005), but which may be resilient enough to persist over time. To this end, we tested whether the hypothesized effect of facial trustworthiness on spousal global positivity bias would weaken with longer duration of marriage.

In sum, our current research sought to shed light on the following questions: (a) are targets with a more trustworthy looking appearance perceived to be more extraverted by both their spouses and strangers?; (b) do targets with a more trustworthy looking appearance receive more positive global personality evaluations both from their spouses and strangers, although their own self-judgments are not any more positive than those of targets with a less trustworthy looking appearance ?; (c) in the case of spouses, are the

effects posited at (a) and (b) susceptible to marriage length (i.e., length of time the rater has known the target)?

Study 1

Method

Participants

Participants were both members of fifty-two elderly couples (women's age: $M = 69.79$ years [$SD = 6.22$]; men's age: $M = 71.43$ years [$SD = 6.42$]).³ They were recruited from an adult volunteer participant pool associated with the University of Toronto and by posting flyers in the greater Toronto area. Prior to their laboratory visit, potential participants underwent a phone screening interview. Specifically, they were asked whether (a) they ever had a stroke, tumor, neurological disease, concussion, depression, seizure, head injury, aneurysm, learning disability, psychiatric illness, epilepsy; (b) they had ever been in a serious car accident and/or hit their head badly and/or been unconscious; and (c) what medication (if any) they take on a regular basis. Potential participants who responded "yes" to any of the questions at points "a" and "b" or reported that they were taking psychotropic medication on regular basis were excluded from participating.

Participants had been married between 15 and 60 years ($M = 42.14$, $SD = 9.48$); no outliers on marriage length were observed. All were native English speakers or had used English as the primary language for at least 30 years. The majority (85.7 %) of participants identified themselves as "Caucasian", with the remaining self-identifying as "Asian/American" (9.2 %), "Black/African American" (3.1 %) or of mixed race (2 %).

Measures

Facial Stimuli Photographs of each participant were taken at the beginning of the study session using a Canon PowerShot SD870 1S digital camera. In order to encourage a neutral expression, participants were instructed to imagine that they were posing for a Canadian passport-type photograph in which no smile or any other emotional display is permitted. More than half of the participants were wearing glasses (all had clear frames); because we were concerned that their removal might adversely impact ability to maintain a neutral expression (e.g., by squinting their eyes to be able to see the photographer more clearly), we allowed participants to keep their glasses on (with posing adjusted as needed to minimize glare). As an additional control, we introduced "glasses" as a dummy coded (0 = *self did not wear*; 1 = *self did wear*) variable in all the reported analyses.

Models of participants' faces were generated in Facegen using the Photofit option on one forward facing picture for each person. To obtain facial trait scores, we used the computer face evaluation model, developed by Oosterhof and Todorov (2008) and implemented in Facegen Modeller program (<http://facegen.com>) version 3.2. Thus, once the participants' faces are imported in Facegen, scores on facial trustworthiness, dominance, and attractiveness are automatically generated as standard deviations from zero (i.e., the average of the database of faces on which the face evaluation model is based).

³ Eleven participants were non-Caucasian. The results are unchanged if these participants are eliminated from the analyses.

Big Five Inventory Participants completed an adapted version of the Big Five Inventory (Schimmack et al. 2004) both in reference to themselves and in reference to their spouse. All items started with the stem “I tend to...” (self-report version) or “S/he tends to...” (informant version). The trait subscales demonstrated reasonable reliability, both for the self-report (alphas from .72 to .83) and informant (alphas from .78 to .88, except for openness to experience, alpha .54) versions. To create self and informant report indices of each of the Big Five traits, responses to the relevant trait subscale items were averaged separately for the two versions of the Big Five Inventory.

Procedure

The study period was comprised of two 1.5 h-long sessions separated by a half-hour lunch break. Upon their arrival at the lab, the spouses were taken to separate testing rooms, where they remained for the duration of the two study sessions. (They were reunited during the lunch break.) At the beginning of the first session, participants’ photographs were taken and, subsequently, during the same session, they were asked to fill out a larger questionnaire package that included the Big Five Inventory.

Data Reduction

One participant (male) only completed part of the questionnaire package, which did not include the Big Five measure, which is why we eliminated both members of this couple from all the analyses reported next. Thus, we focused on the 51 couples, who provided data on all the variables of interest.

Data Analytic Strategy

Due to the dependency in our couple data, we used hierarchical linear regression models (HLM 7.01, Raudenbush et al. 2013) to test our hypotheses. Hierarchical linear regression produces essentially the same parameter estimates as simple linear regression, but accounts for the dependency inherent in nested data, such as couple data, and hence uses more appropriate estimates of standard errors to test statistical significance. The model contained two levels, wherein individuals (level-1) were nested within couples (level-2). Our statistical hypotheses focused exclusively on the relationships among the level-1 variables. Following the recommendations of Campbell and Kashy (2002) for analysis of dyadic data in HLM, we tested our hypotheses by running fixed slopes regression models. To obtain standardized coefficients, we conducted our analyses with standardized variables. Because the personality data (see below) departed from normality, we reported the robust standard error estimates for all analyses below (Hox 2002).

Results

Preliminary Analyses

Gender Effects Because preliminary analyses revealed statistically significant gender differences on self and informant ratings of conscientiousness, as well as on facial dominance, we controlled for gender in all the reported analyses.

Facial Trait Variables In line with prior reports and with the previously documented halo effect (see Fiske et al. 2007), trustworthiness tended to be correlated positively with attractiveness, $b = .24$, $SE = .11$, $t(48) = 2.25$, $p = .03$, but correlated negatively with dominance scores, $b = -.31$, $SE = .09$, $t(48) = -3.59$, $p < .01$.

The Big Five Table 1 presents the means, standard deviations, and intercorrelations among the self and informant Big Five ratings. In line with previous findings (for self-informant agreement within a married sample, see Watson et al. 2000), self and spousal ratings of the Big Five traits were moderately to strongly correlated. Moreover, replicating previous findings (e.g., Borkenau and Liebler 1993), the highest self-other agreement was found for extraversion, $b = .61$, $SE = .06$, $t(50) = 9.52$, $p < .01$, and conscientiousness, $b = .51$, $SE = .10$, $t(50) = 5.30$, $p < .01$.

Positive Evaluation Bias

Self-Report Data To extract the global positive self-evaluation bias factor, we conducted a principal-components analysis of participants' self-reported neuroticism, extraversion, openness to experience, agreeableness and conscientiousness aggregate scores, in which we constrained factor extraction to one factor. The resulting factor (that had an eigenvalue greater than 1) accounted for 41.78 % of the variance and was theoretically consistent with the previously documented positive self-evaluation bias (see Anusic et al. 2009), since it had moderate to high loadings (Hair et al. 1998) on all Big Five scores: $-.73$ (neuroticism), $.55$ (extraversion), $.51$ (openness to experience), $.79$ (agreeableness), and $.61$ (conscientiousness).

Spousal Data Similarly to the self-report data, we ran a principal-components analysis of participants' reports of their spouse's Big Five traits. The resulting factor (that had an eigenvalue greater than 1), which accounted for 47.24 % of the variance, was theoretically consistent with a spouse's global positive evaluation bias, since it had moderate to high loadings on all Big Five spousal reported scores: $-.73$ (neuroticism), $.72$ (extraversion), $.63$ (openness to experience), $.65$ (agreeableness), and $.71$ (conscientiousness).

Hypothesis Testing

Spousal Ratings Positive evaluation bias. A regression analysis, predicting spousal global evaluation bias from facial trustworthiness scores, provided support to our hypothesis that targets with a more trustworthy looking appearance would receive more positive global personality evaluations from their spouses, $b = .27$, $SE = .10$, $t(48) = 2.72$, $p < .01$ (see Fig. 1; Table 1). This effect remained statistically significant, $b = .29$, $SE = .09$, $t(46) = 3.28$, $p < .01$, after controlling for attractiveness and dominance, thus suggesting that it is the unique variance in facial trustworthiness scores that predicts stronger spousal positivity bias scores.

Extraversion. Subsequently, we tested our hypothesis that facial trustworthiness would predict higher spousal ratings of extraversion. Results of the analysis, where we regressed spousal ratings of extraversion on facial trustworthiness scores, supported our prediction, $b = .30$, $SE = .08$, $t(48) = 3.59$, $p < .01$. Controlling for attractiveness and dominance did not alter the statistical significance of this effect, $b = .28$, $SE = .09$, $t(46) = 3.22$,

Table 1 Intercorrelation among self- and spousal Big Five judgments and facial ratings of trustworthiness, dominance and attractiveness

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Self-reported neuroticism	–													
2. Self-reported extraversion	–.33**	–												
3. Self-reported openness to experience	–.11	.26*	–											
4. Self-reported agreeableness	–.53**	.14	.34*	–										
5. Self-reported conscientiousness	–.28**	.20	.10	.40**	–									
6. Informant reports of Neuroticism	.31**	–.11	–.02	–.12	–.15	–								
7. Informant reports of extraversion	–.19	.61**	.20	.16	.19	–.46**	–							
8. Informant reports of openness to experience	.13	.10	.30**	–.11	–.02	–.27*	.40**	–						
9. Informant reports of agreeableness	–.19	–.04	–.05	.26**	.19	–.46**	.21*	.23*	–					
10. Informant reports of conscientiousness	.02	.08	–.05	–.01	.50**	–.37**	.37**	.33**	.37**	–				
11. Self positivity bias	–.73**	.54**	.51**	.80**	.61**	–.23*	.38**	.01	.20	.14	–			
12. Informant positivity bias	–.16	.26**	.12	.13	.29**	–.73**	.71**	.63**	.65**	.71**	.28**	–		
13. Facial ratings of trustworthiness	.06	.16	–.04	–.11	.04	–.17	.30**	.16	.03	.23	–.02	.27**	–	
14. Facial ratings of dominance	–.07	–.37**	.14	.20	.02	.07	–.19	–.04	.12	–.08	.03	–.08	–.33**	–
15. Facial ratings of attractiveness	.11	.15	–.05	–.03	.20	.07	.05	.02	–.06	.07	.06	.01	.26*	.02

* $p < .05$; ** $p < .01$. $N = 102$ individuals embedded in 51 couples. Correlations were computed in a two-level HLM model, which collapsed across all participants (Level-1), irrespective of gender, but accounted for the interdependence in the data provided by the two spouses (Level-2: couple level). Thus, in these analyses, the pattern of correlations is assumed to be identical for males and females. In all the correlations involving the facial trait variables, we controlled for target gender and whether s/he wore glasses



Fig. 1 Examples of trustworthy (on the left) and untrustworthy (on the right) looking male faces. These faces were created by morphing two of the most trustworthy and untrustworthy, respectively, looking male faces in our sample

$p < .01$. Finally, further supporting our hypothesis regarding the unique effect of facial trustworthiness on spousal ratings of extraversion, the effect of trustworthiness on extraversion remained statistically significant after controlling for spousal ratings of the other Big Five traits, $b = .18$, $SE = .08$, $t(44) = 2.31$, $p = .03$.

*Acquaintanceship (i.e., marriage) Length.*⁴ Results of this set of analyses provided no evidence that marriage length moderated the effect of facial trustworthiness on spousal global evaluation bias ($p > .75$) or on spousal ratings of extraversion ($p > .60$).

Self-Ratings Positive evaluation Bias. As predicted, we found no evidence that facial trustworthiness would exert a statistically significant effect on the global self-evaluation bias ($p > .84$, see Table 1).

Extraversion. To test whether a more trustworthy looking appearance is indicative of a more extraverted personality (from the target's perspective), we regressed self-ratings of extraversion on facial trustworthiness scores and verified that individuals with a more trustworthy looking appearance tended to evaluate themselves as being more extraverted, $b = .16$, $SE = .08$, $t(48) = 1.95$, $p = .06$. We found no evidence of similar effects of facial trustworthiness on self-ratings on the other Big Five traits (all $ps > .26$, see Table 1). Nevertheless, controlling for self-ratings on the other Big Five traits strengthened the aforementioned effect of facial trustworthiness, $b = .17$, $SE = .08$, $t(44) = 2.07$, $p = .05$. However, a regression analysis predicting self-reported extraversion from facial trustworthiness, attractiveness and dominance, revealed only a statistically significant effect of dominance, $b = -.35$, $SE = .09$, $t(46) = -4.14$, $p < .01$ (see Fig. 2), whereas the effect of the other two facial trait variables failed to reach statistical significance, $b = .02$, $SE = .09$, $t(46) = .29$, $p = .78$ (trustworthiness) and $b = .14$, $SE = .07$, $t(46) = 1.94$, $p = .06$ (attractiveness). Thus, it seems that among our elderly participants, a submissive, rather than trustworthy looking, facial appearance was a valid physical cue of (self-reported) extraversion.

In sum, Study 1 provided evidence that subtle positive emotional cues, embedded in the structure of affectively neutral faces, may constitute a valid indicator of extraversion, at least in older adulthood. Moreover, these same emotional cues, suggestive of greater

⁴ There were no relationship length data for four couples, so the analyses involving relationship length are based on data from 47 couples.



Fig. 2 Examples of dominant (on the *left*) and submissive (on the *right*) looking male faces. These faces were created by morphing two of the most dominant and submissive, respectively, looking male faces in our sample

trustworthiness (cf. Oosterhof and Todorov 2008), were found to predict more positive global evaluations of target's personality from the perspective of their spouse. Intriguing as (we hope) these findings may be, they face an important limitation: the raters were the targets' spouses and, thus, had access to a wealth of information beyond the target's facial appearance. Thus, it is plausible that it was not a target's facial trustworthiness, but rather an extraneous variable, which happens to covary with facial trustworthiness, that impacted the spouses' evaluations of the target's personality, broadly, and extraversion, more specifically. To control for this possibility, we conducted Study 2, in which raters were unacquainted with the targets and, thus, their sole source of information for making their personality judgments, was the target's facial appearance. This design thus allowed us to test the hypothesis that perceivers use a target's facial trustworthiness as a cue for inferring a more positive personality profile, broadly, and for inferring greater extraversion, more specifically.

Study 2

Method

Participants

One year after completing the first session, all couples were contacted and invited to participate in a follow-up session in which they completed the personality judgment task (see below). Only 31 couples ($M = 72.88$ years [$SD = 5.44$] for males, and $M = 71.21$ years [$SD = 6.37$] for females) were willing to return to the lab. *T* test analyses revealed no statistically significant differences between couples who participated in the Time 2 session versus those who did not on any of the measures collected, apart from relationship length, $t(50) = -2.28$, $p = .03$ (all other $ps > .10$). That is, couples who participated in the Time 2 session had been married longer at Time 1 ($M = 44.33$ years, $SD = 8.72$ years) relative to those who did not ($M = 38.34$ years, $SD = 9.76$ years). Nevertheless, since the Study 2 tasks did not involve participants' spouses, we saw no need to control for relationship length in the reported analyses.

Tasks and Procedure

As part of a larger study session, participants completed a self-paced personality evaluation task, in which they were required to make Big Five trait judgments in response to the FaceGen models of a subset of opposite sex participants ($N = 90$) from Study 1. Each FaceGen model was evaluated by two participants. As a Big Five measure, we used Rammstedt and John's (2007) short version of the Big Five Inventory, which required participants to rate on a scale from 1 (*completely disagree*) to 7 (*completely agree*) the degree to which they thought that the 10 Big Five adjectives (2/trait) were descriptive of each target face. The two adjectives, relevant to a specific trait, were presented on the same screen and raters made a single judgment regarding how well the two adjectives characterized the respective face. The average correlation between the two raters that evaluated each face was .22 across all Big Five traits, ranging from .10 (for conscientiousness) to .31 (for extraversion). The low interrater agreement for conscientiousness is rather unsurprising, given that personal grooming-related variables, which were eliminated from our target faces, have been shown to be valid cues of conscientiousness that drive interrater agreement at zero acquaintance (Albright et al. 1988; Borkenau and Liebler 1992). In light of these findings, we first conducted all the planned analyses separately for each of the two raters who evaluated each face. Because the pattern of results was virtually identical across the two raters, we opted to conduct all the analyses on the Big Five scores, averaged across the two raters.⁵

⁵ For each target, raters 1 and 2 were defined chronologically, i.e., rater 1 was the individual who evaluated the respective target face before rater 2 did. Although all the faces were evaluated by two judges, raters 1 and 2 evaluated a different number of faces. Specifically, rater 1 evaluated only two opposite sex faces. Rater 2 evaluated more than two opposite sex faces. This occurred because the initial personality evaluation task was designed to involve the whole sample and, thus, in order to obtain two personality ratings per face, we had to compensate for the fact that only 31 of the 52 couples who participated initially were willing to return to the lab.

Importantly, though, the pattern of results was very similar across the two raters. Thus, with respect to the global positivity bias, a principal components analysis (constrained to a single factor extraction) of rater 1's Big Five judgments, yielded a solution (that had an eigenvalue greater than 1), which accounted for 52.42 % of the variance and was theoretically consistent with the postulated rater positive evaluation bias. Specifically, this factor had moderate to high loadings on all Big Five scores: $-.73$ (neuroticism), $.80$ (extraversion), $.66$ (openness to experience), $.77$ (agreeableness), and $.65$ (conscientiousness). A similar analysis of rater 2's Big Five judgments yielded a factor (that had an eigenvalue greater than 1), which accounted for 69.25 % of the variance and was also theoretically consistent with the postulated rater positive evaluation bias. This factor had moderate to high loadings on all Big Five scores: $-.85$ (neuroticism), $.73$ (extraversion), $.87$ (openness to experience), $.89$ (agreeableness), and $.81$ (conscientiousness).

Of note, the association between the two raters' positive global evaluation scores in reference to a given target reached statistical significance, $b = .35$, $SE = .08$, $t(41) = 4.20$, $p < .01$. Since neither rater was acquainted with the target, this finding is compatible with our hypothesis that facial appearance variables may drive perceivers' (largely consensual) global personality evaluations of a target at zero acquaintance. Indeed, as expected, across both raters, targets with a more trustworthy looking appearance received more positive global personality evaluations, $b = .30$, $SE = .10$, $t(41) = 2.94$, $p < .01$ for rater 1, and $b = .33$, $SE = .10$, $t(41) = 3.22$, $p < .01$ for rater 2. This effect remained statistically significant, $b = .34$, $SE = .11$, $t(39) = 3.13$, $p < .01$ for rater 1 and $b = .30$, $SE = .10$, $t(39) = 2.86$, $p < .01$ for rater 2, after controlling for attractiveness and dominance.

Finally, at the level of individual traits, as predicted, both raters judged more trustworthy looking targets as being more extraverted, $b = .20$, $SE = .11$, $t(41) = 1.85$, $p = .07$ for rater 1, and $b = .32$, $SE = .09$, $t(41) = 3.59$, $p < .01$ for rater 2. Controlling for attractiveness and dominance either strengthened or left the effect unchanged, $b = .24$, $SE = .11$, $t(39) = 2.17$, $p = .04$ for rater 1 and $b = .30$, $SE = .10$, $t(39) = 3.00$, $p < .01$ for rater 2 (the association of neither attractiveness, nor dominance with extraversion reached statistical significance, ps ranging from .15 to .53 across the two raters), therefore indicating the

Control Variables

Like in Study 1, we controlled for target gender and whether s/he wore glasses in all the reported analyses involving facial appearance (including the interrater agreement analyses reported above, as well as the analyses involving self-stranger agreement on Big Five judgments, which are reported below).

Data Analytic Strategy

Due to the dependency among our target faces, we used hierarchical linear regression models (HLM 7.01, Raudenbush et al. 2013) to test our hypotheses.

Results

Preliminary Analyses

The Big Five None of the correlations between Study 1 participants' self-ratings and Study 2 participants' informant ratings of the Big Five traits reached significance (all $ps > .09$). Indeed, prior studies documented that at zero acquaintance, under standardized target presentation conditions, such as the ones used in our present research, statistically significant correlations between self and (single) informant ratings—as opposed to ratings done by six or more raters) emerge only for extraversion (cf. Naumann et al. 2009). Since unlike those studies, ours focused exclusively on facial (rather than whole-body) appearance and used computerized versions of participants' faces (to enable computation of facial trustworthiness scores), the absence of statistically significant correlations between self and informant ratings of personality may not be particularly surprising.

Global Evaluation Bias To extract the global positive evaluation bias, we conducted a principal components analysis of the two raters' averaged Big Five scores and constrained factor extraction to one factor. This factor (that had an eigenvalue greater than 1), which accounted for 68.63 % of the variance, was theoretically consistent with the postulated rater positive evaluation bias (see Anusic et al. 2009), since it had moderate to high loadings (cf. Hair et al. 1998) on all Big Five scores: $-.85$ (neuroticism), $.81$ (extraversion), $.83$ (openness to experience), $.90$ (agreeableness), and $.74$ (conscientiousness).

Hypothesis Testing

Extraversion To verify that facial trustworthiness would predict higher informant ratings of extraversion at zero acquaintance, we regressed informant ratings of extraversion on facial trustworthiness scores. Results of this analysis were in line with our hypothesis, $b = .34$, $SE = .08$, $t(41) = 4.06$, $p < .01$. Controlling for attractiveness and dominance left the effect unchanged, $b = .37$, $SE = .10$, $t(39) = 3.78$, $p < .01$ (the association of neither attractiveness, nor dominance with extraversion reached statistical significance, ps of $.24$ and $.06$, respectively), therefore indicating the specificity of the association between facial trustworthiness and informant ratings of extraversion at zero acquaintance.

Footnote 5 continued

specificity of the association between facial trustworthiness and informant ratings of extraversion at zero acquaintance.

Positive Evaluation Bias In line with prior findings that smiling is erroneously used as a cue for inferring a more positive global personality profile (Naumann et al. 2009), we found that greater facial trustworthiness predicted lower informant ratings on neuroticism (i.e., the socially undesirable trait), $b = -.28$, $SE = .10$, $t(41) = -2.85$, $p < .01$, as well as higher informant ratings on all the desirable Big Five traits, $b = .33$, $SE = .08$, $t(41) = 3.86$, $p < .01$ (agreeableness), $b = .23$, $SE = .10$, $t(41) = 2.31$, $p = .03$ (openness to experience), and $b = .35$, $SE = .09$, $t(41) = 3.68$, $p < .01$ (conscientiousness). Consequently, rather unsurprisingly, a regression analysis predicting informant global evaluation bias from facial trustworthiness scores, provided support to our hypothesis that, at zero acquaintance, targets with a more trustworthy looking appearance would receive more positive global personality evaluations, $b = .37$, $SE = .09$, $t(41) = 4.02$, $p < .01$. This effect remained statistically significant, $b = .38$, $SE = .11$, $t(39) = 3.45$, $p < .01$, after controlling for attractiveness and dominance. Taken together with the Study 1 findings, these results suggest that it is the unique variance in facial trustworthiness scores that predicts stronger informant positivity bias scores within both unacquainted and well-acquainted target-perceiver dyads.

Discussion

Our present studies provided a glimpse at the effect of facial appearance on self and other Big Five judgments both at zero-acquaintance, as well as in a sample of well-acquainted targets and informants, specifically among long-term married couples. Previous research documented that positive emotionally expressive behaviors (i.e., smiling) are among the most utilized indicators of personality (Kenny et al. 1992; Naumann et al. 2009), which indeed provide valid information for some traits (i.e., extraversion, Borkenau et al. 2009). Consequently, in the present studies, we used Oosterhof and Todorov's (2008) model of face evaluation and tested whether, under standardized, emotionally neutral conditions, structural facial characteristics, suggestive of positive emotional expressions (i.e., greater facial trustworthiness), would serve as a valid cue to extraversion, as revealed by their positive associations not only with informant (spousal/stranger), but also self-ratings of extraversion. We also investigated whether under such circumstances, greater facial trustworthiness would be used by informants (spouse/stranger) as a cue to infer a more positive global personality profile, an effect that may weaken with longer marriage duration (hence, more knowledge of a target's personality) among spouses. Most of our hypotheses received some empirical support, which we detail next.

As predicted and in line with previous findings from the social perception literature that trustworthiness attributions determine the overall valence of interpersonal judgments (Fiske et al. 2007; Oosterhof and Todorov 2008), we found that individuals with a more trustworthy looking appearance received more positive, global personality evaluations not only from strangers, but also from their spouses. Importantly, this effect only emerged in informant, but not self-ratings of personality and, among spouses, it was impervious to relationship length.

Moreover, our data offered suggestive evidence consistent with a Dorian Gray effect of appearance on personality, because structural facial characteristics suggestive of positive emotional expression (i.e., greater trustworthiness, cf. Oosterhof and Todorov 2008) predicted not only spousal and stranger, but also, to some extent, self-ratings on this trait. Our finding thus dovetails previous studies documenting that positive emotionally expressive behaviors, such as smiling, are valid indicators of extraversion (Borkenau et al. 2009;

Naumann et al. 2009) and a lifelong predisposition to experience certain emotional states reportedly remains imprinted on the (emotionally neutral) face (Malatesta et al. 1987).

Although intriguing, we would like to advise some caution in interpreting the effect of facial trustworthiness on self-rated extraversion, because it was not robust enough to survive after we controlled for facial dominance. Indeed, rather unexpectedly, it was (low) facial dominance, rather than trustworthiness, that emerged as the strongest predictor of self-rated extraversion. Because we did not predict this effect, we are reluctant to interpret it extensively. Nevertheless, we would like to refer readers to a recent study by Lukaszewski and Roney (2011) on the origins of extraversion, which may provide a viable explanatory framework. Drawing on Tooby and Cosmides' (1990) model of personality determination, these authors propose that the extent to which individuals pursue an extraverted behavioral strategy depends on the extent to which they possess the resources necessary to manage successfully the incumbent risks. In young adulthood, physical strength is a critical asset, because at this stage the pursuit of extraverted behavioral goals entails engagement in competitive activities of some sort (e.g., competing for social attention, assertively pursuing status and influence, cf. Lukaszewski and Roney 2011). Nevertheless, extant theory and research suggest that in older adulthood, there is a shift in individuals' social goals, such that individuals tend to pursue emotionally rewarding and harmonious, rather than competitive, social interactions (e.g., Carstensen 1991). An implication of this line of argument is that more dominant looking individuals would be more likely to pursue an extraverted behavioral strategy in younger adulthood, because they look as if they possess the physical resources to implement their goals (cf. Lukaszewski and Roney 2011; Oosterhof and Todorov 2008). In contrast, in older adulthood, more submissive looking individuals may be more likely to pursue an extraverted behavioral strategy. Specifically, they may be more sought after as interaction partners, because their appearance offers the promise of smoother social exchanges. Indeed, there is some indirect support for this prediction as extraversion reportedly shows low rank-order consistency in older years (i.e., there is reduced stability in people's relative placement to each other on extraversion scores), particularly after the age of 60 (Lucas and Donnellan 2011; Specht et al. 2011). Future studies are needed to test this hypothesis.

The present research is only a first step toward understanding the mechanisms responsible for the lingering effects of facial appearance on personality evaluations, well beyond the first impressions stage. Indeed, our finding that facial appearance exerts a long-lasting effect on global personality evaluations of one's spouse deserves additional investigation, as it could have significant implications not only for personality assessment but also for close relationship research. For example, with respect to the former, John and Robins (1993) documented significantly lower self-peer, relative to peer-peer agreement, on personality traits, which are evaluative, rather than neutral in nature. Whereas these authors' explanations focused on the effect of self-evaluation bias in increasing divergence between self and peer (relative to peer-peer) personality ratings, future studies may need to take also into account the role of facial appearance in increasing convergence of informant ratings and increasing divergence of self and informant ratings, specifically, for evaluative, rather than neutral traits.

Furthermore, with respect to relational outcomes, our present findings raise the unsettling possibility that within a dyadic context, a facial appearance suggestive of greater untrustworthiness may render one more vulnerable to negative interpersonal responses, not only in the early stages of relationship formation, but also in more established partnerships, such as the ones investigated in our studies. For example, Oosterhof and Todorov (2009) provided evidence that structural facial characteristics, suggestive of emotional

expressions, interact with voluntarily expressed emotions to influence perceivers' evaluations of the targets' affective experiences (e.g., at the same affective intensity, an untrustworthy looking other is perceived as being angrier than a trustworthy looking other). Consequently, to the extent that emotionally suggestive facial features are still salient among long-term spouses, it seems plausible that a spouse with an untrustworthy looking appearance may be at an unfair disadvantage, because during relationship conflicts, s/he would be perceived as angrier and, thus, more hostile, than a trustworthy looking other.

Indeed, given its significant implications, the mechanisms underlying the resilient effect of facial trustworthiness on global spousal personality judgments are worthy of additional attention. For example, one may inquire whether facial trustworthiness serves as a powerful encoding and/or retrieval cue for behavioral evidence consistent with positive, rather than negative traits and/or whether it inhibits the successful encoding of behavioral evidence, suggestive of negative, rather positive dispositional tendencies.

Inevitably, our current research has some limitations. One limitation is the exclusive use of an elderly sample. We decided to do so because we sought to compare the effect of facial trustworthiness on informant personality ratings, provided by individuals who knew the target extensively versus those who did not know the target at all. Nevertheless, future studies investigating the effect of facial trustworthiness on younger adults' personality ratings of strangers versus close others are certainly needed. Indeed, there is some evidence that, relative to younger adults, older adults are more sensitive to subtle positive emotional cues, such as the ones driving trustworthiness attributions (cf. Petrican et al. 2013). Consequently, it is yet to be elucidated whether younger adults' global ratings of others' personality are susceptible not only to the influence of explicit positive emotional cues, as shown before (cf. Naumann et al. 2009), but also to the effect of subtler cues which determine perceptions of trustworthiness.

A second limitation of the present research is the lack of more objective behavioral criteria for assessing what may indeed be a more positive global personality profile. Recent studies documented that in response to facial appearance alone (i.e., without any additional knowledge of the target individual), people give similar trustworthiness evaluations to war criminals and war heroes, to CEOs who committed financial fraud and those who did not, as well as to students who were found cheating on an exam versus those who were not (cf. Rule et al. 2013). Consequently, we reasoned that a link between facial trustworthiness and more positive global personality ratings from strangers and spouses would reflect the informants' evaluation biases, rather than substantive variance in the targets' personality. Nevertheless, more stringent tests of the relationship between a trustworthy looking appearance and a more positive global personality profile are definitely warranted.

In sum, the present studies provided suggestive evidence that facial appearance exerts a pervasive impact on informant personality judgments, both at a global and a more trait-specific level. Whether such findings are mere illustrations of the side effects associated with mostly evolutionarily adaptive processes (Todorov et al. 2008) or whether there is indeed some (yet to be acknowledged) merit in taking some things (including oneself) at face value is yet to be determined.

Acknowledgments This research was supported by a Canadian Institutes of Health Research postdoctoral fellowship awarded to Raluca Petrican and a CIHR grant to Cheryl Grady (MOP14036).

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, *55*, 387–395.
- Ambady, N., Hallahan, M., & Rosenthal, R. (1995). On judging and being judged accurately in zero acquaintance situations. *Journal of Personality and Social Psychology*, *69*, 518–529.
- Anusic, I., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2009). The nature and structure of correlations among Big Five ratings: The Halo-Alpha-Beta model. *Journal of Personality and Social Psychology*, *97*, 1142–1156.
- Berry, D. S. (1991). Accuracy in social perception: Contributions of facial and vocal information. *Journal of Personality and Social Psychology*, *61*, 298–307.
- Berry, D. S., & Brownlow, S. (1989). Were the physiognomists right? Personality correlates of facial babyishness. *Personality and Social Psychology Bulletin*, *15*, 266–279.
- Berscheid, E. (1981). An overview of the psychological effects of physical attractiveness. In G. W. Lucker, K. A. Rubbens, & J. A. McNamara Jr (Eds.), *Psychological aspects of facial form* (pp. 1–47). Ann Arbor, MI: Center for Human Growth and Development.
- Bond, C. F., Berry, D. S., & Omar, A. (1994). The kernel of truth in judgments of deceptiveness. *Basic and Applied Social Psychology*, *15*, 523–534.
- Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality*, *43*, 702–705.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, *62*, 645–657.
- Borkenau, P., & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology*, *65*, 546–553.
- Butler, E. A., Egloff, B., Wilhelm, F. H., Smith, N. C., Erickson, E. A., & Gross, J. J. (2003). The social consequences of expressive suppression. *Emotion*, *3*, 48–67.
- Campbell, L., & Kashy, D. A. (2002). Estimating actor, partner, and interaction effects for dyadic data using PROC MIXED and HLM: A user-friendly guide. *Personal Relationships*, *9*, 327–342.
- Carstensen, L. L. (1991). Socioemotional selectivity theory: Social activity in life-span context. *Annual Review of Gerontology and*, *11*, 195–217.
- Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, *15*, 110–117.
- Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment*, *4*(1), 5–13.
- Fink, B., Neave, N., Manning, J. T., & Grammer, K. (2005). Facial symmetry and the “big-five” personality factors. *Personality and Individual Differences*, *39*, 523–529.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social perception: Warmth and competence. *Trends in Cognitive Science*, *11*, 77–83.
- Franklin, R. G., & Zebrowitz, L. A. (2013). Older adults’ trait impressions of faces are sensitive to subtle resemblance to emotions. *Journal of Nonverbal Behavior*. doi:10.1007/s10919-013-0150-4.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26–34.
- Grady, C. L., Grigg, O., & Ng, C. (2012). Age differences in default and reward networks during processing of personally relevant information. *Neuropsychologia*, *50*, 1682–1697.
- Hair, J. F., Jr, Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis with readings* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Hox, J. (2002). *Multilevel analysis techniques and applications. Quantitative methodology series*. Mahwah, NJ: Lawrence Erlbaum Associates.
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, *61*, 521–551.
- Johnson, D. F., & Pittenger, J. B. (1984). Attribution, the attractiveness stereotype and the elderly. *Developmental Psychology*, *20*, 1168–1172.
- Keightley, M. L., Winocur, G., Burianova, H., Hongwanishkul, D., & Grady, C. (2006). Age effects on social cognition: Faces tell a different story. *Psychology and Aging*, *21*, 558–572.
- Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero-acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology*, *62*, 88–97.

- Lucas, R. E., & Donnellan, B. M. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology, 4*, 847–861.
- Lukaszewski, A. W., & Roney, J. R. (2011). The origins of extraversion: Joint effects of facultative calibration and genetic polymorphism. *Personality and Social Psychology Bulletin, 37*, 409–421.
- Malatesta, C. Z., Fiore, M. J., & Messina, J. J. (1987). Affect, personality, and facial expression characteristics of older people. *Psychology and Aging, 2*, 64–69.
- Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences, 9*, 496–502.
- Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The benefits of positive illusions: Idealization and the construction of satisfaction in close relationships. *Journal of Personality and Social Psychology, 70*, 79–98.
- Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin, 35*, 1661–1671.
- Olivola, C. Y., & Todorov, A. (2010a). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Non-verbal Behavior, 34*, 83–110.
- Olivola, C. Y., & Todorov, A. (2010b). Fooled by first impressions? Re-examining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology, 46*, 315–324.
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 11087–11092.
- Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion, 9*, 128–133.
- Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgments from natural and composite facial images: More evidence for a 'kernel of truth' in social perception. *Social Cognition, 24*, 490–524.
- Petrican, R., English, T., Gross, J. J., Grady, C. L., Hai, T., & Moscovitch, M. (2013). Friend or foe? Age moderates time course-specific responsiveness to trustworthiness cues. *Journal of Gerontology: Psychological Sciences, 68*, 215–233.
- Pound, N., Penton-Voak, I. S., & Brown, W. M. (2007). Facial symmetry is positively associated with self-reported extraversion. *Personality and Individual Differences, 43*, 1572–1582.
- Rammstedt, B., & John, O. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203–212.
- Raudenbush, S., Bryk, A., & Congdon, R. (2013). *HLM 7.01 for Windows [Hierarchical linear and non-linear modeling software]*. Multivariate Software, Inc.
- Ruby, P., Collette, F., D'Argembeau, A., Peters, F., Degueldre, C., Baetee, E., et al. (2009). Perspective taking to assess self-personality: What's modified in Alzheimer's disease? *Neurobiology of Aging, 30*, 1637–1651.
- Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology, 104*, 409–426.
- Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion, 9*, 260–264.
- Schimmack, U., Oishi, S., Furr, R. M., & Funder, D. C. (2004). Personality and life satisfaction: A facet-level analysis. *Personality and Social Psychology Bulletin, 30*, 1062–1075.
- Specht, J., Egloff, B., & Schmuckler, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101*, 862–882.
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*, 25–29.
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*, 1623–1626.
- Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12*, 455–460.
- Tooby, J., & Cosmides, L. (1990). On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of Personality, 58*, 17–67.
- Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology, 78*, 546–558.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science, 17*, 592–598.

- Zebrowitz, L. A., Androletti, C., Collins, M. A., Lee, S. Y., & Blumenthal, J. (1998a). Bright, bad, babyfaced boys: Appearance stereotypes do not always yield self-fulfilling prophecy effects. *Journal of Personality and Social Psychology*, *75*, 1300–1320.
- Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998b). Appearance and personality across the lifespan. *Personality and Social Psychology Bulletin*, *24*, 736–749.
- Zebrowitz, L. A., Franklin, R. G., Hillman, S., & Boc, H. (2013). Older and younger adults' first impressions from faces: Similar in agreement, but different in positivity. *Psychology and Aging*, *28*, 202–212.
- Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). “Wide-eyed” and “crooked-faced”: Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, *22*, 1258–1269.