

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Understanding the role of faces in person perception: Increased reliance on facial appearance when judging sociability

Bastian Jaeger^{a,b,*}, Anthony M. Evans^b, Mariëlle Stel^c, Ilja van Beest^b

^a Vrije Universiteit Amsterdam, Netherlands

^b Tilburg University, Netherlands

^c University of Twente, Netherlands

ARTICLE INFO

Keywords:

Person perception
Trait impressions
Physiognomy
Lay beliefs
Sociability

ABSTRACT

Faces play a central role in person perception. People spontaneously judge others' personality based on their facial appearance and these impressions guide many consequential decisions. When do people rely on facial appearance? In five studies ($N = 1936$, four preregistered), we test whether reliance on facial appearance depends on the goal of impression formation (i.e., on which trait dimension targets are evaluated). Trait impressions are, to a large extent, based on the resemblance of facial cues to emotional expressions. As emotional expressiveness is a central component of sociability, we hypothesized that people would more readily perceive sociability in faces. We find that facial appearance is seen as more indicative of a person's sociability than their morality or competence (Study 1, $n = 338$), and this was particularly true for sociability traits that are characterized by emotional expressions (e.g., enthusiasm, playfulness; Study 2, $n = 162$). We find the same pattern when examining the influence of facial cues on judgment and decision-making. People are more confident in the accuracy of their trait impressions when judging sociability (Study 3, $n = 527$), they value information on the facial appearance of job candidates more when looking for a sociable employee (Study 4, $n = 390$), and they view reliance on facial appearance when making hiring decisions as more appropriate and more effective when looking for a sociable employee (Study 5, $n = 519$). Together, our results provide converging evidence that people view facial appearance as especially relevant for judging a person's sociability.

People spontaneously infer personality traits from facial appearance (Klapper, Dotsch, van Rooij, & Wigboldus, 2016; Stewart et al., 2012). Even though these inferences are often inaccurate (Bonnefon, Hopfensitz, & De Neys, 2017; Jaeger et al., 2021; Rule, Krendl, Ivcevic, & Ambady, 2013; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015), people nonetheless rely on them when making a wide range of consequential decisions (Duarte, Siegel, & Young, 2012; Gomulya, Wong, Ormiston, & Boeker, 2017; Wilson & Rule, 2015; for a review, see Olivola, Funk, & Todorov, 2014). The effects of trait impressions on decision-making are not only prevalent, but also surprisingly persistent: People still rely on facial appearance when more diagnostic cues are available (Chang, Doll, Vant Wout, Frank, & Sanfey, 2010; Hooper et al., 2019; Jaeger, Evans, Stel, & van Beest, 2019; Olivola, Tingley, & Todorov, 2018; Rule, Tskhay, Freeman, & Ambady, 2014) and when they are explicitly told to discount a person's appearance (Blair, Judd, & Fallman, 2004; Jaeger, Todorov, Evans, & van Beest, 2020).

This raises the question of which factors drive reliance on facial cues.

Previous studies have addressed this question by examining characteristics of the perceiver (Ewing, Caulfield, Read, & Rhodes, 2015; Jaeger et al., 2020; Suzuki, 2016; Suzuki, Tsukamoto, & Takahashi, 2017). Studies show that many people believe in the core tenet of physiognomy (Aristotle, 1936; Lavater, 1775)—that personality traits are reflected in facial cues (Suzuki et al., 2017). These beliefs are related to how much people rely on trait impressions from faces (Jaeger et al., 2020). People who more strongly endorsed physiognomic beliefs relied more on the perceived trustworthiness of plaintiffs and defendants when making legal sentencing decisions. Other studies investigated how reliance on trait impressions changes across the life span. For example, Suzuki (2016) showed that when participants could learn about the trustworthiness of interaction partners by repeatedly observing their behavior, older adults relied more on their partners' facial appearance than younger adults.

Here, we extend previous research and go beyond examining the role of perceiver characteristics. We propose that the extent to which people

* Corresponding author at: Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.
E-mail address: bxjaeger@gmail.com (B. Jaeger).

<https://doi.org/10.1016/j.jesp.2022.104288>

Received 30 December 2020; Received in revised form 14 January 2022; Accepted 15 January 2022

Available online 29 January 2022

0022-1031/© 2022 The Author(s).

Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

rely on facial appearance also depends on the goal of impression formation. People may believe that some traits are more reflected in facial appearance than others. Crucially, if facial appearance is seen as more diagnostic for judging some personality traits, then this could influence when people rely on facial appearance. People might rely more on trait impressions (as opposed to other information) when judging a trait that is believed to be more visible in a person's facial appearance. In the present studies, we examine reliance on facial appearance for inferring different traits. We focus on three basic dimensions of person perception: morality (encompassing traits such as trustworthiness, honesty, and sincerity), sociability (encompassing traits such as warmth, friendliness, likeability), and competence (encompassing traits such as competence, intelligence, skillfulness; Goodwin, Piazza, & Rozin, 2014; Landy, Piazza, & Goodwin, 2016; Leach, Ellemers, & Barreto, 2007).

1. The perceived value of facial appearance

Are some traits believed to be more reflected in facial appearance than others? Addressing this question requires an understanding of how trait impressions from faces are formed. Theories of social perception underline the central role of emotion perception for the formation of trait impressions from faces (Jaeger & Jones, 2021; Todorov et al., 2015; Zebrowitz, 2017). Perceivers associate facial expressions, such as smiles, not only with emotional states, but also with personality traits (referred to as the temporal extension hypothesis; Caulfield, Ewing, Bank, & Rhodes, 2016; Knutson, 1996; Marsh, 2005; Secord, 1958; Sutherland, Young, & Rhodes, 2017b; Zebrowitz, 2012). A smiling person may not only be seen as happy, but also as warm, empathetic, and trustworthy. Crucially, these trait inferences are even triggered by emotionally neutral faces that merely *resemble* an emotional expression (Adams, Nelson, Soto, Hess, & Kleck, 2012; Said, Sebe, & Todorov, 2009). For example, slightly upturned corners of the mouth or raised eyebrows, which can occur due to natural variation in facial appearance, trigger ascriptions of personality traits. In short, trait impressions from faces are primarily based on perceived resemblances to emotional expressions.

If trait impressions from faces are largely based on resemblances to emotional expressions, then people might more readily perceive personality traits that are more strongly associated with facial expressions of emotions. Functional accounts of emotional expressions highlight that their primary function is to coordinate social interactions (Crivelli & Fridlund, 2018; Shariff & Tracy, 2011; Van Kleef, 2010). From this perspective, emotional expressions are a tool for navigating social relationships—a skill which forms the basis for evaluations of an individual's sociability (Landy et al., 2016). Thus, emotional expressiveness may be seen as particularly indicative of a person's sociability (vs. morality or competence). Evidence from several studies support this idea. Goodwin and colleagues (2014) surveyed a wide range of trait adjectives to test which traits best distinguish between judgments on the three dimensions. Dispositional happiness emerged as a defining feature of sociability judgments. In a similar vein, smiling (as opposed to displaying a neutral facial expression) has a positive effect on a wide range of trait judgments, but effects tend to be strongest for sociability judgments (Krumhuber et al., 2007; Mehu, Little, & Dunbar, 2007).

If perceptions of (resemblances to) emotional expressions are (a) a key determinant of trait impressions from faces and (b) especially relevant for judging an individual's sociability, then people might more readily perceive sociability (than morality or competence) in faces. Results of previous studies provide preliminary evidence in favor of this hypothesis. After attractiveness, sociability and happiness were the most frequently mentioned traits when participants could freely describe faces (Oosterhof & Todorov, 2008; see also Sutherland et al., 2017a).

2. The current studies

Here, we examine whether the extent to which people rely on facial appearance depends on which trait dimensions they are evaluating

targets on. Specifically, we hypothesize that facial appearance is seen as particularly relevant for judging a person's sociability (vs. morality or competence). We test this hypothesis in five studies ($N = 1936$, four preregistered).

First, we examine lay beliefs in the diagnostic value of facial appearance. In Study 1 ($n = 338$, preregistered), we measure physiognomic beliefs for sociability, morality, and competence. That is, we explicitly ask participants to what extent they believe that each of the three traits is reflected in a person's facial appearance. We predict that people hold stronger physiognomic beliefs for sociability (vs. morality or competence).

Next, we investigate the mechanism underlying this effect (Study 2, $n = 162$, preregistered). Emotional expressions differ in how diagnostic they are for different traits. This not only applies to broad trait dimensions (e.g., sociability vs. morality), but also to specific traits within each dimension. We test whether people hold stronger physiognomic beliefs for sociability traits that are more readily expressed through facial expressions of emotions (e.g., enthusiastic, funny) than for sociability traits that less readily expressed through facial expressions of emotions (e.g., forgiving, humble; Goodwin, 2015; Goodwin et al., 2014).

Finally, we examine the consequences of the belief that facial appearance is more indicative of a person's sociability. We hypothesize that facial cues exert a stronger influence on judgments and decisions in situations in which evaluating a person's sociability (vs. morality or competence) is central to the perceiver's goal. In short, we predict that facial cues are more influential when a target's sociability is relevant. In Study 3, we examine people's confidence in the accuracy of their impressions ($n = 527$, preregistered). Participants view photographs of emotionally neutral faces, provide impressions of sociability, morality, and competence, and indicate how accurate they think their judgments are. We predict that people are more confident in the accuracy of their impressions when judging sociability (vs. morality or competence).

We also examine reliance on facial appearance in an applied settings. We focus on the personnel selection context, as previous studies have shown that hiring decisions are influenced by the facial appearance of job candidates (Bóo, Rossi, & Urzúa, 2013; Gomulya et al., 2017; Ling, Luo, & She, 2019). In Study 4 ($n = 390$, preregistered), we test whether the perceived diagnostic value of facial appearance for making hiring decisions varies as a function of the desired personality traits of job candidates. We predict that facial photographs are valued more when looking for a sociable (vs. moral or competent) employee. In Study 5 ($n = 519$, preregistered), we test whether people evaluate human resource managers who rely on facial appearance to make hiring decisions differently, depending on which personality trait the manager is looking for in candidates. Again, we predict that reliance on facial appearance is seen as more appropriate and more effective when looking for a sociable (vs. moral or competent) employee.

All data, analysis scripts, materials, and preregistration documents are available at the Open Science Framework (<https://osf.io/cbsmw/>). We report how our sample sizes were determined, and mention all data exclusions and measures for each study.

3. Study 1: physiognomic beliefs

In Study 1, we measured physiognomic beliefs for different personality traits. We predicted that participants would believe that sociability is more reflected in facial appearance than morality or competence.

3.1. Methods

3.1.1. Participants

A total of 346 first-year psychology students from a Dutch university completed the study in return for partial course credit. The final sample size was determined by how many students participated in the study in two weeks. Eight participants with missing data were excluded from

analysis, leaving a final sample of 338 participants ($M_{age} = 19.78$, $SD_{age} = 3.37$; 84.02% female).

3.1.2. Materials and procedure

We measured physiognomic beliefs for sociability, morality, and competence with the physiognomic belief scale (Jaeger et al., 2020). Participants were prompted to “imagine seeing the passport photo of a stranger”. Beliefs for the three trait dimensions were measured with three traits for each dimension (sociability: warmth, likeability, and friendliness, McDonald's $\omega = 0.90$; competence: competence, intelligence, and skillfulness $\omega = 0.90$; morality: trustworthiness, honesty, and sincerity $\omega = 0.89$; Brambilla, Rusconi, Sacchi, & Cherubini, 2011). Participants dragged a slider to indicate how accurately they think they can judge each trait just from looking at a person's face. Slider values range from 0 (*not accurate at all*) to 100 (*extremely accurate*).¹ Dutch participants completed the study in Dutch while non-Dutch participants completed the study in English.

3.1.3. Sensitivity analysis

We conducted sensitivity analyses for our main effects of interest. We determined the smallest difference in physiognomic beliefs we were able to detect with 80% power (and $\alpha = 5\%$). A sensitivity analysis in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007; Giner-Sorolla et al., 2020) showed that we had 80% power to detect an effect of $d = 0.18$. Thus, we had sufficient power to detect even small differences in physiognomic beliefs.

3.2. Results and discussion

We compared physiognomic beliefs across the three dimensions to test whether people believe that sociability is more reflected in facial features than morality or competence. Paired t -tests showed that physiognomic beliefs were significantly stronger for sociability ($M = 41.43$, $SD = 23.96$) compared to morality ($M = 22.64$, $SD = 21.50$), $t(337) = 21.10$, $p < .001$, $d = 0.82$, and competence ($M = 21.73$, $SD = 20.19$), $t(337) = 20.41$, $p < .001$, $d = 0.87$ (see Figure 1). There was no significant difference between morality-specific beliefs and competence-specific beliefs, $t(337) = 1.14$, $p = .25$, $d = 0.04$. Thus, in line with our hypothesis, we found that sociability was believed to be more reflected in facial appearance than morality or competence.

Note. Error bars represent bootstrapped 95% confidence intervals of the mean.

4. Study 2: the role of emotional expressions

Study 1 showed that physiognomic beliefs were stronger for sociability-related traits than for morality-related or competence-related traits. We hypothesized that this is because people primarily rely on resemblances to facial expressions of emotions to form personality impressions and emotional expressions are seen as more diagnostic of a person's sociability. In other words, we would predict that the extent to which people believe that a certain trait can be accurately inferred from facial appearance depends on how much the trait is expressed through facial expressions of emotions. In Study 2, we tested this hypothesis more directly. Rather than comparing physiognomic beliefs across the three person perception dimensions (sociability, morality, competence), we examined differences in physiognomic beliefs for different sociability-related traits. Specifically, we examined sociability traits that

are more readily expressed through facial expressions of emotions (e.g., enthusiastic, funny) and sociability traits that less readily expressed through facial expressions of emotions (e.g., forgiving, humble; Goodwin, 2015; Goodwin et al., 2014). We refer to these as emotionally expressive sociability and non-emotionally expressive sociability, respectively. We predicted that people hold stronger physiognomic beliefs for traits that are more readily expressed through facial expressions of emotions.

4.1. Methods

This study was preregistered.

4.1.1. Participants

An a priori power analysis showed that a sample size of 162 participants is required to detect a small difference between conditions ($d = 0.2$) with 90% power (and $\alpha = 5\%$). We therefore recruited 162 US American crowd workers via Prolific ($M_{age} = 32.70$, $SD_{age} = 12.41$; 56.17% female, 40.74% male, 3.09% other).

4.1.2. Materials and procedure

We measured physiognomic beliefs for 16 traits that are indicative of a person's sociability (Goodwin, 2015). Participants were prompted to “imagine seeing the passport photo of a stranger”. They then dragged a slider to indicate how accurately they think they can judge each trait just from looking at a person's face. Slider values range from 0 (*not accurate at all*) to 100 (*extremely accurate*). We selected eight traits that are more readily expressed through facial expressions of emotions (warm, easy-going, happy, sociable, playful, agreeable, enthusiastic, funny; $\omega = 0.95$) and eight traits that are less readily expressed through facial expressions of emotions (kind, grateful, empathetic, helpful, humble, giving, forgiving, cooperative; $\omega = 0.95$).²

4.1.3. Sensitivity analysis

We conducted sensitivity analyses for our main effects of interest. We determined the smallest difference in physiognomic belief (comparing emotionally expressive and non-emotionally expressive traits) we were able to detect with 80% power (and $\alpha = 5\%$). A sensitivity analysis in G*Power (Faul et al., 2007; Giner-Sorolla et al., 2020) showed that we had 80% power to detect an effect of $d = 0.18$. Thus, we had sufficient power to detect even small differences in physiognomic beliefs.

4.2. Results

We predicted that people would show stronger physiognomic beliefs for traits that are more readily expressed through facial expressions. In line with this hypothesis, a paired t -test showed that physiognomic beliefs were significantly stronger for emotionally expressive traits ($M = 23.08$, $SD = 20.17$) compared to non-emotionally expressive traits ($M = 14.50$, $SD = 15.52$), $t(161) = 9.71$, $p < .001$, $d = 0.44$ (see Figure 2A).

4.2.1. Additional analyses

One of the eight emotionally expressive sociability traits that we examined also describes an emotional state with a clear facial expression (“happy”). However, even when omitting this trait from our analysis, we found that physiognomic beliefs were stronger for emotionally expressive traits ($M = 22.12$, $SD = 19.98$) compared to non-emotionally expressive traits ($M = 14.50$, $SD = 15.52$), $t(161) = 8.97$, $p < .001$, $d = 0.39$.

¹ In line with previous work (Jaeger et al., 2020), we also measured general physiognomic belief without referring to any specific personality traits with three questions (e.g., “I can learn something about a person's personality just from looking at his or her face”; $\omega = 0.76$). When assessing beliefs for specific characteristics, we also included three additional characteristics not related to a person's personality (gender, age, and attractiveness).

² A pretest with 100 participants recruited from Prolific ($M_{age} = 33.87$, $SD_{age} = 10.62$; 54% female, 45% male, 1% other) showed that facial expressions of emotions were indeed seen as more relevant for inferring the eight emotional sociability traits ($M = 5.05$, $SD = 0.77$) than the eight non-emotional sociability traits ($M = 4.10$, $SD = 0.95$), $t(99) = 12.44$, $p < .001$, $d = 1.07$.

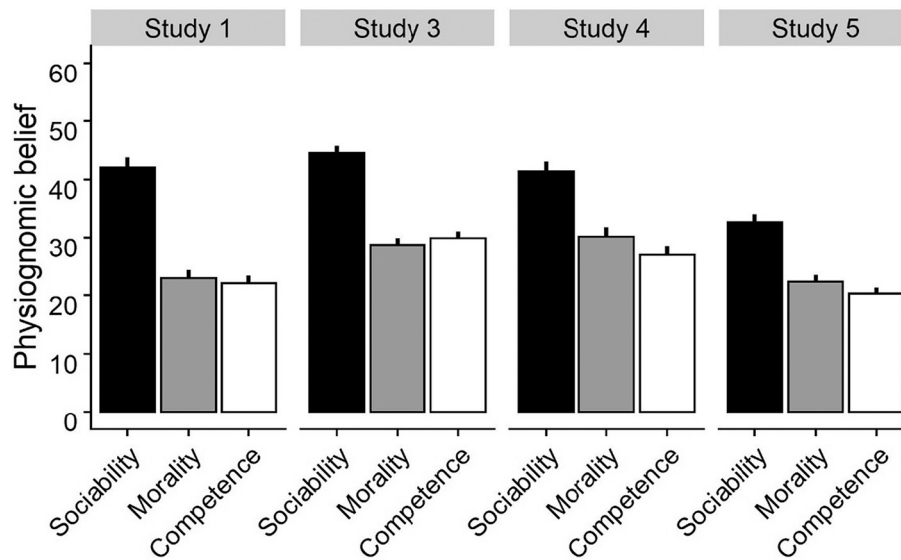


Fig. 1. Physiognomic beliefs for sociability, morality, and competence across the four studies.

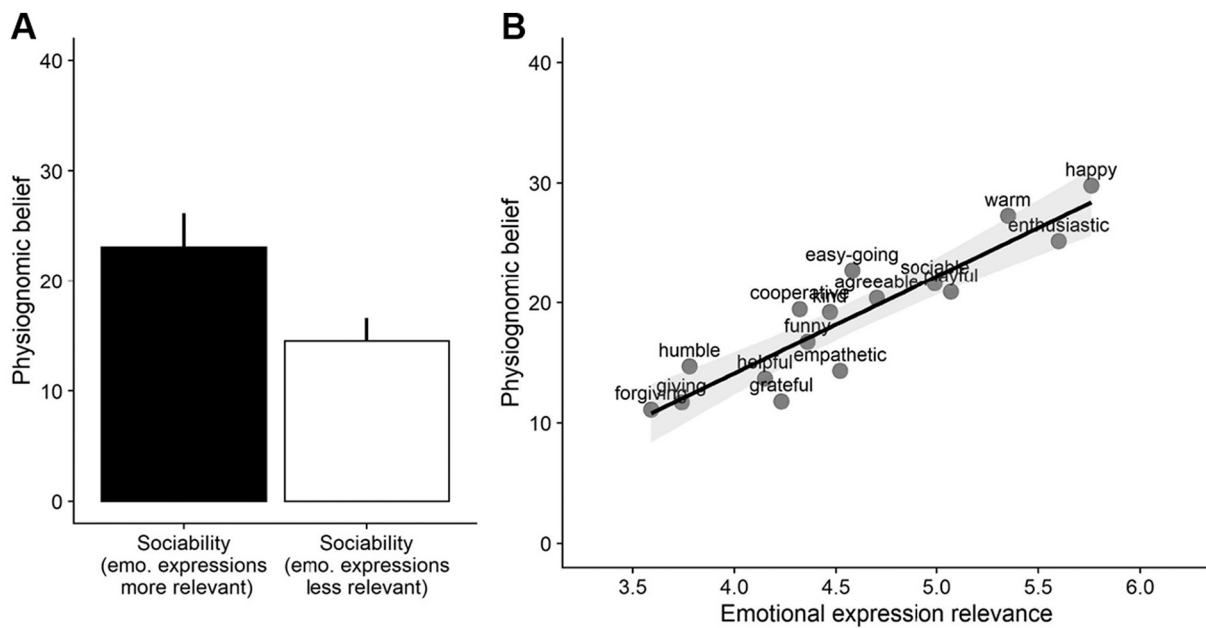


Fig. 2. Physiognomic beliefs for sociability-related traits that are more or less readily expressed through facial expressions of emotions (A) and the correlation between the extent to which a trait is expressed through emotional expressions and physiognomic beliefs for the trait (B). Note. Error bars represent bootstrapped 95% confidence intervals of the mean.

We also examined the relationship between the extent to which a trait is expressed through facial expressions of emotions (assessed in our pretest) and physiognomic beliefs (assessed in our main study) at the trait level. For each trait, we computed the average rating across all participants. There was a strong correlation between both variables, $r(14) = .91, p < .001$, showing that participants indicated stronger physiognomic beliefs for traits that are more readily expressed through facial expressions of emotions (see Figure 2B).

4.3. Discussion

Study 2 provided additional insights into why people think that facial features are particularly relevant for judging sociability (vs. morality or competence). We hypothesized that people show stronger physiognomic beliefs for sociability because trait impressions from faces are primarily

based on resemblances to facial expressions of emotions (Jaeger & Jones, 2021) and sociability-related traits are more likely to be expressed through facial expressions of emotions than morality- or competence-related traits. This view predicts that people are more confident in inferring traits from faces that are more readily expressed through emotional expressions. Results of Study 2 provided support for this hypothesis. We found that physiognomic beliefs were stronger for sociability-related traits that are more readily expressed through facial expressions of emotions (e.g., enthusiastic, funny) than for sociability-related traits that are less readily expressed through facial expressions of emotions (e.g., humble). In short, our findings suggest that facial appearance is seen as more relevant for judging traits that are more strongly associated with emotional expressions.

5. Study 3: confidence in face judgments

In Study 3, we measured physiognomic beliefs for different personality traits. We predicted that participants would believe that sociability is more reflected in facial appearance than morality or competence. We also asked participants to judge targets based on facial photographs and measured their confidence in the accuracy of their impressions. We measured perceptions of sociability, morality, and competence and predicted that participants would be more confident in their sociability judgments than their competence or morality judgments. Previous work has shown that confidence in personality inferences from faces is positively related to the extremity and speed of judgments (Ames, Kammrath, Suppes, & Bolger, 2010; Willis & Todorov, 2006). We therefore controlled for extremity and speed when comparing judgment confidence across personality dimensions.

5.1. Methods

This study was preregistered.

5.1.1. Participants

A total of 533 first-year psychology students from a Dutch university completed the study in return for partial course credit. The final sample size was determined by how many students participated in the study in two weeks. In line with our preregistered exclusion criteria, data from five participants who always indicated the same rating across all trials were excluded from analysis, leaving a final sample of 527 participants ($M_{age} = 19.60$, $SD_{age} = 2.09$; 81.21% female, 18.60% male, 0.19% other).³ The majority of participants were Dutch (73.62%) or German (13.28%).

5.1.2. Materials and procedure

We randomly selected ten Caucasian emotionally neutral images from the Chicago Face Database (five male, five female). Targets were photographed from a fixed distance in front of a uniform white background with a camera that was adjusted to their eye level and images were cropped to a size of 400×400 pixels. All participants viewed the images against a white background on a 22-in. Dell P2210 TFT monitor (1680×1050 resolution, 80 Hz refresh rate). The face stimuli were 112.8 mm high and 90.2 mm wide at an approximate viewing distance of 600 mm, subtending a visual angle of $10.7^\circ \times 8.6^\circ$. We asked participants to rate targets on three dimensions: *sociability* (*warmth, friendliness, likeability*), *morality* (*trustworthiness, sincerity, honesty*), and *competence* (*competence, intelligence, skillfulness*). Participants separately rated each image on each dimension in a random order on a scale from 1 (*not at all* [dimension]) to 9 (*extremely* [dimension]). Thus, participants completed a total of 30 trials. Next to recording rating scores, we also recorded rating speed (the time it took participants to make a rating) and rating extremity (the absolute distance of the assigned score to the midpoint of the scale). After each judgment, participants indicated their confidence in the accuracy of their judgment on a scale from 1 (*not at all confident*) to 9 (*extremely confident*). Finally, participants completed the physiognomic belief scale as in Study 1 (sociability: $\omega = 0.93$, morality: $\omega = 0.95$, competence: $\omega = 0.93$).

In line with our preregistered exclusion criteria, we excluded response times from 305 trials (1.91%) that were three standard deviations below or above the mean. Response times were \log_{10} -transformed due to their right-skewed distribution. Dutch participants completed the study in Dutch while non-Dutch participants completed the study in English.

³ We did not specify missing data as an exclusion criteria in our preregistration. Retaining participants with missing data did not influence the pattern of results.

5.1.3. Sensitivity analysis

We conducted sensitivity analyses for our main effects of interest. We determined the smallest difference in judgment confidence between the sociability, morality, and competence dimensions we were able to detect with 80% power (and $\alpha = 5\%$). We used the *simr* package (Green & Macleod, 2016) in R (R Core Team, 2021) to assess our design's sensitivity for detecting differences in confidence. The package provides power estimates for fixed effects in multilevel regression models. We varied the effect of interest in our model and calculated power at each level. This showed that our sample size allowed 80% power to detect a 0.067-point difference in confidence between sociability judgments and morality judgments and a 0.069-point difference in confidence between sociability judgments and competence judgments (on our nine-point scale). Thus, we had sufficient power to detect even small differences in physiognomic beliefs and confidence.

5.2. Results

5.2.1. Descriptive statistics

On average, participants took 6.31 s ($SD = 2.50$ s) to make a judgment and the mean confidence rating was 5.27 ($SD = 1.65$, on a scale that ranged from 1 to 9). We computed intraclass correlation coefficients (ICCs) to estimate consensus in judgments across participants (Shrout & Fleiss, 1979). Participants showed significant consensus in their judgments of sociability, $ICC(2,1) = .394$, $p < .001$, 95% CI [.234, .684], morality, $ICC(2,1) = .231$, $p < .001$, 95% CI [.123, .501], and competence, $ICC(2,1) = .380$, $p < .001$, 95% CI [.224, .672].

5.2.2. Confidence

Our main goal was to test whether participants' confidence in their trait judgments depended on which trait they were judging. We predicted that participants would be more confident in their sociability judgments (vs. competence or morality judgments). To test this prediction, we estimated a multilevel regression model with random intercepts per participant and face, in which we regressed confidence on a variable indicating which dimension was judged (sociability vs. morality vs. competence; see Figure 3). In line with our hypothesis, participants were more confident when judging sociability than when judging competence, $\beta = 0.093$, $SE = 0.023$, $t(15272) = 4.13$, $p < .001$, or morality, $\beta = 0.158$, $SE = 0.023$, $t(15272) = 7.02$, $p < .001$. Participants were also more confident when making competence (vs. morality)

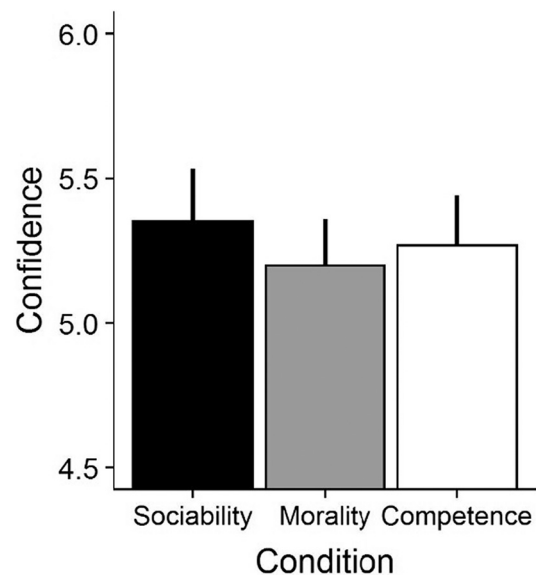


Fig. 3. Confidence when judging sociability, competence, and morality. Note. Error bars represent bootstrapped 95% confidence intervals.

judgments, $\beta = 0.065$, $SE = 0.023$, $t(15272) = 2.89$, $p = .004$.

We explored whether differences in confidence across the three dimensions would still emerge when controlling for the extremity and speed of judgments. Regressing confidence on a dummy variable indicating which dimension was judged, as well as judgment extremity and speed showed a positive effect of extremity, $\beta = 0.383$, $SE = 0.010$, $t(15069) = 39.16$, $p < .001$, and a negative effect of speed, $\beta = -0.161$, $SE = 0.010$, $t(15123) = 15.34$, $p < .001$. More importantly, there was still a significant effect of personality dimension: Participants were more confident when judging sociability than when judging competence, $\beta = 0.079$, $SE = 0.021$, $t(14967) = 3.70$, $p < .001$, or morality, $\beta = 0.094$, $SE = 0.021$, $t(14968) = 4.40$, $p < .001$. There was no difference in confidence between competence and morality judgments, $\beta = -0.005$, $SE = 0.021$, $t(15272) = 0.22$, $p = .83$.

5.2.3. Additional analyses

We also tested whether we could replicate the findings of Study 1. We compared physiognomic beliefs across the three dimensions to test whether people think that sociability is more reflected in facial features than morality or competence. Paired *t*-tests showed that physiognomic beliefs were significantly stronger for sociability ($M = 44.64$, $SD = 22.99$) compared to morality ($M = 28.82$, $SD = 21.35$), $t(526) = 23.02$, $p < .001$, $d = 1.00$, and competence ($M = 30.05$, $SD = 20.17$), $t(526) = 19.29$, $p < .001$, $d = 0.84$ (see Figure 1). Morality-specific beliefs were slightly stronger than competence-specific beliefs, but this difference was less pronounced and only marginally significant, $t(526) = 1.78$, $p = .075$, $d = 0.20$.

Although our main goal was to test for differences in judgment confidence across trait dimensions, we also investigated the link between individual differences in physiognomic belief and confidence. We calculated participants' average confidence in their judgments across the three trait dimensions. Participants who more strongly believed in physiognomy were also more confident in the accuracy of their judgments, $r(525) = .292$, $p < .001$ (see the Supplemental Materials for more detailed results).

5.3. Discussion

We found that participants were more confident in the accuracy of their sociability impressions than their morality or competence impressions. Differences in confidence across the three dimensions still emerged when controlling for the extremity and speed of judgments, which were both related to confidence. Replicating the results of Study 1, we again found differences in physiognomic beliefs across the three dimensions: Sociability was believed to be more reflected in facial appearance than morality or competence.

6. Study 4: diagnostic value of facial appearance

Results of Study 3 showed that sociability is not only believed to be more reflected in facial appearance than morality or competence, participants were also more confident in the accuracy of their sociability judgments. Going beyond impressions formation based on facial photographs, Study 4 examined the perceived diagnostic value of facial appearance in a more applied setting. Previous studies have shown that hiring decisions are influenced by the facial appearance of candidates (Bóo et al., 2013; Gomulya et al., 2017; Ling et al., 2019). In the current study, we tested if the diagnostic value of facial appearance in a personnel selection context depends on which personality trait people are looking for in job candidates. As sociability is believed to be more reflected in faces than morality or competence, we predicted that people would view information on a candidate's facial appearance as more useful if they are looking for sociable (vs. moral or competent) employee. We also measured physiognomic beliefs to test if the finding that sociability is believed to be more reflected in facial appearance than morality or competence replicates in a sample of participants from the United States.

6.1. Methods

This study was preregistered.

6.1.1. Participants

An a priori power analysis showed that a sample size of 130 participants per condition is required to detect a small-to-medium-sized difference between conditions ($d = 0.35$) with 80% power (and $\alpha = 5\%$). We therefore aimed to recruit a total of 390 participants. We recruited 430 US American MTurk workers who completed the study in return for \$0.50 each. In line with our preregistered exclusion criteria, data from 2 participants (0.47%) who indicated poor or basic English proficiency and from 38 participants (8.84%) who failed an attention check at the end of the study were excluded from analysis, leaving a final sample of 390 participants ($M_{age} = 34.64$, $SD_{age} = 9.88$; 42.56% female).

6.1.2. Materials and procedure

We asked participants to imagine that they are working in the HR department of a company that hosts various events and that they were tasked with hiring a new event planner. Participants were randomly assigned to one of three conditions, which determined which personality trait they were looking for in candidates (sociability, morality, or competence). For instance, in the sociability condition, participants read: "In the past, your company has received complaints about planners who were unfriendly and dismissive to guests. Therefore, you are looking for someone who is warm, friendly, and likeable".

To measure the perceived value of facial photographs, participants saw a list of five typical components of a job application (photograph, experience, recommendation letter, education level, grades) in a random order and were asked to rate how useful each component is for making the hiring decision on a scale from 0 (*not at all useful*) to 100 (*extremely useful*). The usefulness ratings of the applicant's photograph represented our key dependent variable. Finally, participants completed the physiognomic belief scale (sociability: $\omega = 0.93$, morality: $\omega = 0.95$, competence: $\omega = 0.95$) and three attention check items. Participants were first asked to rate several preferences on a nine point scale (e.g., "I prefer to get bonuses over steady income"), but were later instructed to indicate a specific response pattern. Participants who did not specify the correct response for any of the three questions were excluded from analysis.

6.1.3. Sensitivity analysis

We also conducted a sensitivity analysis for our main effect of interest (differences in photo usefulness ratings between the sociability, morality, and competence conditions), using G*Power (Faul et al., 2007; Giner-Sorolla et al., 2020). This confirmed that we had 80% power to detect an effect size of $d = 0.35$ when comparing ratings between conditions. For examining differences in physiognomic belief, a sensitivity analysis showed that we had 80% power to detect an effect of $d = 0.22$. Thus, our design had sufficient power to detect even small differences between personality dimensions.

6.2. Results

6.2.1. Photo usefulness

Photos were on average seen as less useful ($M = 33.15$, $SD = 28.33$) than recommendation letters ($M = 80.06$, $SD = 17.77$), past experience ($M = 66.63$, $SD = 26.49$), level of education ($M = 50.91$, $SD = 28.27$), and grades ($M = 43.15$, $SD = 27.73$), all $p < .001$, $0.35 < d < 1.98$.

To test our main hypothesis, we examined whether usefulness ratings of the photo depended on which personality trait participants were looking for in candidates. We predicted that they would rate photos as more important for more making hiring decisions when they were looking for a sociable (vs. moral or competent) candidate. A one-way ANOVA showed that usefulness ratings of photos significantly varied across the three conditions $F(2, 378) = 37.42$, $p < .001$ (see Figure 4). In

line with our hypothesis, participants perceived photos as a more useful cue when looking for a sociable candidate ($M = 48.27$, $SD = 27.42$) than when looking for a moral candidate ($M = 31.53$, $SD = 26.55$), $t(245.2) = 8.69$, $p < .001$, $d = 1.09$, or a competent candidate ($M = 20.19$, $SD = 23.93$), $t(256) = 5.00$, $p < .001$, $d = 0.62$. Participants perceived photos as a more useful cue when looking for a moral (vs. competent) candidate, $t(245) = 8.69$, $p < .001$, $d = 0.45$.

6.2.2. Additional analyses

We again compared physiognomic beliefs across the three personality dimensions to test whether people think that sociability is more reflected in facial appearance than morality or competence (see Figure 1). Paired t -tests showed that physiognomic beliefs were significantly stronger for sociability ($M = 41.42$, $SD = 27.67$) compared to morality ($M = 30.21$, $SD = 26.48$), $t(389) = 15.42$, $p < .001$, $d = 0.78$, and competence ($M = 26.99$, $SD = 24.59$), $t(389) = 16.47$, $p < .001$, $d = 0.83$. Morality-specific beliefs were slightly stronger than competence-specific beliefs, but this difference was less pronounced, $t(389) = 5.72$, $p < .001$, $d = 0.29$.

We also examined whether individual differences in physiognomic belief were related to the perceived usefulness of personal photos for making hiring decisions, irrespective of which personality trait participants were looking for in applicants. Across the three conditions, participants who more strongly believed in physiognomy viewed personal photos as more useful for making hiring decisions, $r(388) = .631$, $p < .001$ (see the Supplemental Materials for more detailed results).

6.3. Discussion

Personal photos were seen as more useful when participants were looking for a sociable (vs. moral or competent) employee. In other words, facial cues were seen as more important when judging sociability was relevant to a perceiver's goal. We again replicated the finding that sociability is believed to be more reflected in facial appearance than morality or competence.

7. Study 5: appropriateness and effectiveness of relying on facial appearance

Studies 1 and 2 provided insights into how much people value facial

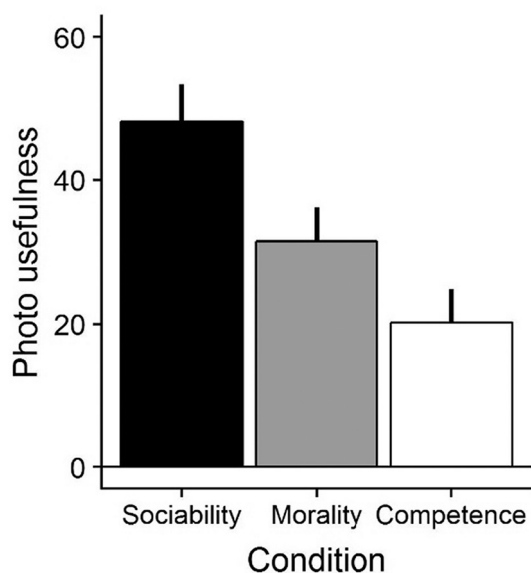


Fig. 4. Differences in rated usefulness of personal photos for making hiring decisions as a function of desired trait in the candidate. Note. Error bars represent bootstrapped 95% confidence intervals.

cues when judging different personality traits. An important question is whether these differences are driven more by beliefs about the efficacy or the moral appropriateness of relying on facial cues. In other words, do people rely on facial cues when evaluating traits related to sociability because they believe that they are more likely to form accurate impressions? Or do they believe it is problematic to try to infer other traits (i.e., traits related to competence and morality) based on superficial facial cues? To answer this question, Study 5 extended our analyses to a third-party perspective. Specifically, we tested whether attitudes towards appearance-based discrimination vary depending on which personality trait is inferred from facial appearance. If sociability is believed to be more reflected in faces than morality or competence, then people should find reliance on personal photos more acceptable and more effective when the goal is to hire a sociable (vs. moral or competent) employee.

7.1. Methods

This study was preregistered.

7.1.1. Participants

An a priori power analysis showed that a sample size of 173 participants per condition is required to detect a small-to-medium-sized difference between condition ($d = 0.35$) with 90% power (and $\alpha = 5\%$). We therefore aimed to recruit a total of 519 participants. We recruited 621 US American MTurk workers who completed the study in return for \$0.50 each. In line with our preregistered exclusion criteria, data from 7 participants who indicated poor or basic English proficiency and from 93 participants who failed at least one of two attention checks were excluded from analysis, leaving a final sample of 521 participants ($M_{age} = 35.78$, $SD_{age} = 10.93$; 45.11% female, 54.51% male, 0.38% other).

7.1.2. Materials and procedure

We asked participants to imagine a manager who is working in the HR department of a company that hosts various events. The job of the manager is to hire event planners. Participants were randomly assigned to one of three conditions, which determined which personality trait the manager is looking for in candidates (sociability, morality, or competence). For instance, in the sociability condition, participants read:

In the past, the company has received complaints about planners who were unfriendly and dismissive to guests. Therefore, the manager is looking for someone who is warm, friendly, and likeable. Since the manager is receiving many applications, he always selects a few promising candidates that are then invited for an interview. When deciding whom to invite, he looks at the candidates' photos and picks the ones that look very warm, friendly, and likeable to him.

To measure attitudes towards discrimination based on facial photographs, we asked participants to evaluate the appropriateness and effectiveness of the manager's practice. We measured perceived appropriateness with three items (appropriate, moral, ethical; $\omega = 0.95$) which were evaluated on a scale that ranged from -50 (e.g., *extremely inappropriate*) to 50 (e.g., *extremely appropriate*). We measured perceived effectiveness with three items (effective, successful, likely to achieve goal; $\omega = 0.97$) which were evaluated on a scale that ranged from -50 (e.g., *extremely ineffective*) to 50 (e.g., *extremely effective*). The order of the two measures and the order of items within each measure was randomized. Finally, participants completed the physiognomic belief scale. As in Study 4, we measured general physiognomic belief ($\omega = 0.90$) and specific physiognomic beliefs for sociability ($\omega = 0.94$), morality ($\omega = 0.94$), and competence ($\omega = 0.94$). Participants also completed the same attention check.

7.1.3. Sensitivity analysis

We conducted a sensitivity analysis for our main effects of interest

(differences in effectiveness and appropriateness ratings between sociability, morality, and competence), using G*Power (Faul et al., 2007; Giner-Sorolla et al., 2020). This showed that we had 80% power to detect an effect size of $d = 0.30$ when comparing effectiveness and appropriateness ratings between conditions. For examining differences in physiognomic belief, a sensitivity analysis showed that we had 80% power to detect an effect of $d = 0.19$. Thus, our design had sufficient power to detect even small differences in effectiveness ratings, appropriateness ratings, and physiognomic beliefs.

7.2. Results

7.2.1. Appropriateness and effectiveness ratings

On average, the manager's strategy to make hiring decisions based on photos was seen as relatively inappropriate ($M = -15.54$, $SD = 26.24$), $t(520) = 13.52$, $p < .001$, $d = 0.59$, and ineffective ($M = -14.36$, $SD = 25.18$), $t(520) = 13.02$, $p < .001$, $d = 0.57$, when compared against 0 (i.e., the midpoint of the scale which denoted neither appropriate nor inappropriate and neither effective nor ineffective). Appropriateness ratings and effectiveness ratings were strongly correlated, $r(519) = .738$, $p < .001$.

We tested whether appropriateness and effectiveness ratings of the manager's strategy depended on which personality trait the manager was looking for in candidates. We predicted that participants would view reliance on photos as more appropriate and effective when the manager is looking for a sociable (vs. moral or competent) candidate.

A one-way ANOVA showed that appropriateness ratings significantly varied across the three conditions $F(2, 518) = 11.00$, $p < .001$ (see Figure 5, left panel). In line with our hypothesis, participants perceived relying on photos as a more appropriate strategy when looking for a sociable candidate ($M = -9.77$, $SD = 26.24$) than when looking for a competent candidate ($M = -22.48$, $SD = 24.51$), $t(340) = 4.66$, $p < .001$, $d = 0.50$. However, the difference in appropriateness ratings between the sociability and morality condition ($M = -14.04$, $SD = 26.49$) was not significant, $t(341) = 1.50$, $p = .14$, $d = 0.16$. Participants perceived relying on photos as a more appropriate strategy when looking for a moral (vs. competent) candidate, $t(347) = 3.10$, $p = .002$, $d = 0.33$.

A one-way ANOVA showed that effectiveness ratings also varied significantly across the three conditions $F(2, 518) = 16.35$, $p < .001$ (see Figure 5, right panel). In line with our hypothesis, participants perceived relying on photos as a more effective strategy when looking for a sociable candidate ($M = -5.77$, $SD = 25.06$) than when looking for a moral candidate ($M = -16.67$, $SD = 23.56$), $t(338) = 4.15$, $p < .001$, $d = 0.45$, or a competent candidate ($M = -20.26$, $SD = 24.76$), $t(344) = 5.41$, $p < .001$, $d = 0.58$. There was no significant difference in effectiveness ratings between the morality and competence condition, $t(350) = 1.39$, $p = .16$, $d = 0.15$.⁴

⁴ Differences in perceived appropriateness still emerged when controlling for perceived effectiveness and vice versa. Ratings of effectiveness and appropriateness were strongly correlated. We conducted an exploratory factor analyses for the six items used to measure the two constructs and both the visual scree test and the parallel analysis suggested that a one-factor solution fit the data best. As we found slightly different results for effectiveness and appropriateness ratings, we also report separate results for each measure. When analyzing the composite score for all six items measuring effectiveness and appropriateness, a one-way ANOVA showed that composite scores significantly varied across the three conditions $F(2, 518) = 14.72$, $p < .001$ (see Figure 5, left panel). Ratings were higher when the recruiter was looking for a sociable candidate ($M = -7.77$, $SD = 23.76$) than when they were looking for a moral candidate ($M = -15.35$, $SD = 23.18$), $t(340) = 2.99$, $p = .003$, $d = 0.32$, or a competent candidate ($M = -21.37$, $SD = 23.14$), $t(342.9) = 5.40$, $p < .001$, $d = 0.58$. Ratings in the morality condition were also significantly higher than in the competence condition, $t(349.8) = 2.44$, $p = .015$, $d = 0.26$.

7.2.2. Additional analyses

We again compared physiognomic beliefs across the three personality dimensions to test whether people think that sociability is more reflected in facial appearance than morality or competence (see Figure 1). Paired t -tests showed that physiognomic beliefs were significantly stronger for sociability ($M = 32.55$, $SD = 26.16$) compared to morality ($M = 22.23$, $SD = 22.77$), $t(520) = 16.67$, $p < .001$, $d = 0.73$, and competence ($M = 20.21$, $SD = 21.74$), $t(520) = 17.45$, $p < .001$, $d = 0.76$. Morality-specific beliefs were slightly stronger than competence-specific beliefs, but this difference was less pronounced, $t(520) = 4.56$, $p < .001$, $d = 0.20$.

We also examined whether individual differences in physiognomic belief were related to the perceived appropriateness and effectiveness of relying on personal photos to make hiring decisions, irrespective of which personality trait participants were looking for in applicants. Across the three conditions, participants who more strongly believed in physiognomy perceived relying on personal photos as more appropriate, $r(519) = .530$, $p < .001$, and more effective, $r(519) = .659$, $p < .001$ (see the Supplemental Materials for more detailed results).

7.3. Discussion

Results of the current study showed that the appropriateness and effectiveness of relying on facial appearance for making hiring decisions depends on which personality trait people are looking for in applicants. Even though participants perceived reliance on facial appearance as less appropriate and less effective than reliance on other information (e.g., recommendation letters), attitudes differed depending on which personality trait recruiters were looking for in candidates. Specifically, reliance on personal photos for making hiring decisions was seen as more appropriate when looking for a sociable (vs. moral or competent) employee. Reliance on personal photos for making hiring decisions was also seen as more effective when looking for a sociable (vs. competent) employee (differences between sociability and morality were not significant). We again replicated the finding that sociability is believed to be more reflected in facial appearance than morality or competence.

8. General discussion

Many people believe in the central idea of physiognomy that a person's facial appearance is indicative of their personality (Suzuki et al., 2017). Here, we asked whether some personality traits are believed to be more reflected in facial appearance than others. Trait impressions are, to a large extent, based on the resemblance of facial cues to emotional expressions (Adams et al., 2012; Said et al., 2009). As emotional expressiveness is a central component of sociability (Kring, Smith, & Neale, 1994; Roger & Neshsoever, 1987), we hypothesized that people would more readily perceive sociability in faces than morality or competence. In line with this hypothesis, we found that sociability is believed to be more reflected in facial appearance than morality or competence and this effect emerged consistently in four large samples of Dutch (Studies 1 and 3) and U.S. American participants (Studies 4 and 5). Study 2 confirmed that this pattern of results was in fact due to the greater relevance of emotional expressions for judging sociability (vs. morality or competence). When examining physiognomic beliefs for various traits indicative of sociability, we found that physiognomic beliefs were stronger for traits that are more readily expressed through facial expressions of emotions (e.g., enthusiastic, funny) than for traits that are less readily expressed through facial expressions of emotions (e.g., forgiving, humble). In short, our findings suggest that facial appearance is seen as more relevant for judging traits that are more strongly associated with emotional expressions.

We also examined whether this pattern across personality dimensions is reflected in the influence of facial cues on judgment and decision-making. In particular, we investigated whether facial cues are more influential when a person's sociability (vs. morality or

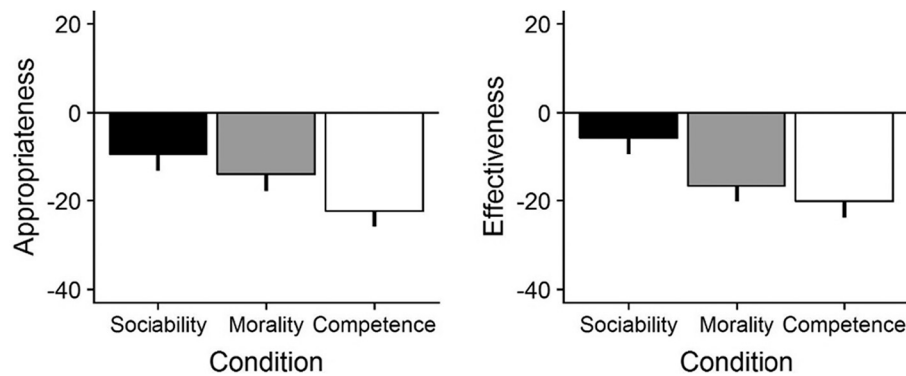


Fig. 5. Differences in rated appropriateness (left) and effectiveness (right) of relying on photos to make hiring decisions as a function of desired trait in the candidate. Note. Error bars represent bootstrapped 95% confidence intervals.

competence) is relevant. Across three studies, we manipulated which trait was relevant for participants' judgments or decisions. Study 3 showed that participants were more confident in the accuracy of their sociability judgments (vs. morality or competence judgments). In Study 4, we examined how much people value information on a person's facial appearance (i.e., a facial photograph) in a hiring context. While photographs of candidates were seen as less diagnostic than other cues (e.g., recommendation letters or past experience), photographs were valued more when looking for a sociable (vs. moral or competent) employee. A similar pattern of results was found in Study 5. Participants viewed a manager's strategy to rely on facial appearance when making hiring decisions as relatively ineffective and inappropriate. However, reliance on facial appearance was seen as more effective when looking for a sociable (vs. moral or competent) employee. Reliance on facial appearance was also viewed as more appropriate when looking for a sociable (vs. competent) employee (the comparison with morality was not significant). Together, we find converging evidence for the notion that facial appearance is seen as particularly important for judging a person's sociability.

8.1. Theoretical implications

People often rely on facial appearance even when more diagnostic information is available (Jaeger et al., 2019; Olivola et al., 2018). Overreliance on facial cues is problematic because it can lead to worse outcomes for the decision-maker (Olivola & Todorov, 2010b) and to systematic biases against people with a certain facial appearance (Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Wilson & Rule, 2015). To explain the widespread influence of facial appearance, previous studies have mostly focused on individual characteristics of perceivers, examining who relies more or less on facial appearance (Ewing et al., 2015; Suzuki, 2016; Suzuki et al., 2017). The present findings suggest that the extent to which facial appearance influences decision-making outcomes not only depends on who is making the decision, but also in what context the decision is being made. People may be more likely to rely on trait impressions from faces when a person's sociability (vs. morality or competence) is relevant for their decision.

It should be noted that our results do not imply that people *only* rely on facial appearance when others' sociability is relevant. When investigating effects of facial appearance on decision-making, prior studies have mostly focused on impressions of trustworthiness and competence, which have been shown to predict voting (Olivola & Todorov, 2010a; Todorov, Mandisodza, Goren, & Hall, 2005), hiring (Gomulya et al., 2017), and sentencing decisions (Jaeger et al., 2020). Our results do not stand in conflict with these findings, as we consistently find that people see some value in facial appearance as an indicator of targets' morality and competence. However, the current results do suggest that the perceived diagnostic value of facial appearance is even larger for

judging targets' sociability.

Going beyond differences across trait dimensions, our results also extend recent work on individual differences in reliance on facial appearance. Similar to Suzuki et al. (2017), we find that many people believe in the core tenet of physiognomy—that personality traits are reflected in facial cues. Our analyses showed that individuals who more strongly endorsed physiognomic beliefs were more confident in the accuracy of their trait impressions (Study 3), valued personal photos more for making hiring decisions (Study 4), and viewed reliance on facial appearance to make hiring decisions as more appropriate and more effective (Study 5). Thus, individual differences in physiognomic beliefs may explain why some people rely more on trait impressions from faces than others.

Studies in the field of social perception often highlight the efficient and automatic way in which people form trait impressions from faces (Bonnenon, Hopfensitz, & De Neys, 2013; Engell, Haxby, & Todorov, 2007; Willis & Todorov, 2006). However, these findings should not be interpreted as evidence for procedural uniformity in impression formation. In fact, recent studies have shown that there are various top-down influences on social perception (Freeman & Ambady, 2010; Hehman, Stolier, Freeman, Flake, & Xie, 2019). For instance, lay beliefs about the conceptual overlap between specific personality traits (e.g., whether trustworthiness correlates with sociability) predict the extent to which face-based impressions of the traits overlap (e.g., whether trustworthiness impressions correlate with sociability impressions; Stolier, Hehman, & Freeman, 2018a; Stolier, Hehman, Keller, Walker, & Freeman, 2018b). In a similar vein, our results show that the extent to which people think about and rely on facial cues in judgment and decision-making depends on both individual differences (i.e., endorsement of physiognomic beliefs) and situational differences (i.e., which trait dimension targets are evaluated on).

8.2. Limitations and future directions

Why is facial appearance seen as especially diagnostic for judging sociability (vs. other personality traits)? We suggested that this is due to the fact that trait impressions are primarily formed based on emotion cues (i.e. facial cues that resemble emotion expressions). As emotional expressiveness is a central component of sociability, people more readily perceive sociability in faces. An alternative (albeit non-mutually exclusive) explanation holds that people rely more on facial appearance when judging sociability because sociability can be more accurately inferred than other personality traits. Previous research using stranger rating paradigms—in which participants make personality judgments based on limited information—suggests that judgment accuracy is highest for extraversion (which is conceptually similar to sociability) compared to other Big Five personality traits (Kenny & West, 2008). However, the evidence is less clear for judgments based on static images of

emotionally neutral faces. While some studies found that the accuracy of extraversion impressions is above chance (Borkenau, Brecke, Möttig, & Paelecke, 2009; Kramer & Ward, 2010; Naumann, Vazire, Rentfrow, & Gosling, 2009; Penton-Voak, Pound, Little, & Perrett, 2006), others did not (Ames et al., 2010; Jones, Kramer, & Ward, 2012; Shevlin, Walker, Davies, Banyard, & Lewis, 2003). In general, there is no clear evidence that extraversion judgments are *more* accurate than other trait judgments. More research is needed to determine the accuracy of different personality trait judgments from facial appearance and to what extent this pattern is reflected in people's beliefs about the accuracy of their own trait judgments.

Previous findings on the meta-accuracy of trait impressions from faces also speak against the idea that people rely more on sociability judgments because they are actually more accurate than judgments of other personality traits. This would require people to know how accurate their judgments across various trait dimensions are. However, people are generally not aware of how accurate their judgments are—that is, they show poor meta-accuracy (Ames et al., 2010; Jaeger et al., 2021). People who are more confident in the accuracy of their impressions are on average not more accurate (i.e., inter-personal meta-accuracy is poor) and people's confidence does not seem to track their accuracy across judgments (i.e., intra-personal meta-accuracy is poor). Together, these findings cast doubt on the idea that people rely more on facial appearance when judging sociability because sociability impressions are more accurate.

More research is also needed to examine whether the present findings replicate in more externally valid contexts. For example, previous studies have investigated the effect of facial appearance on organizational decision-making by examining the outcomes of real-life hiring procedures (Gomulya et al., 2017; Ling et al., 2019). Future studies could take a similar approach and test if candidates' facial appearance is more consequential when their sociability is more relevant for the position in question. Reliance on facial appearance and beliefs about the appropriateness of this strategy may also vary across different countries. For example, in some countries, people are prohibited from including a photograph in their job application. Finally, Study 3 showed that people were more confident in the accuracy of their impressions when judging sociability, but these results were based on a relatively small and demographically homogenous sample of targets. Future studies should examine how the goal of impression formation (i.e., on which dimensions targets are judged) shapes the impressions formation process using larger and more diverse samples.

Another promising avenue for future research is a more thorough investigation of the structure, antecedents, and consequences of physiognomic beliefs. Lay beliefs shape many aspects of social perception and behavior. For example, research on lay personality theory has shown that people hold beliefs about the basis (Haslam, Bastian, & Bissett, 2004) and malleability (Chiu, Hong, & Dweck, 1997) of personality traits and that these beliefs predict outcomes related to impression formation (Haslam et al., 2004), information search (Plaks, Stroessner, Dweck, & Sherman, 2001), and stereotyping (Levy, Stroessner, & Dweck, 1998). Here, we showed that people also hold beliefs about the extent to which different traits are reflected in facial appearance. We conducted exploratory factor analyses on participants' physiognomic belief scores for the nine different personality traits, which showed that a one-factor solution fit the data best (see Supplemental Materials). This suggests that there are important person-level (as well as trait-level) factors shaping physiognomic belief. Additional studies are needed to better understand how this general belief in physiognomy shapes social perception and behavior (for first attempts, see Jaeger et al., 2020; Suzuki et al., 2017).

9. Conclusion

People spontaneously infer personality traits from faces and rely on their inferences to make a wide range of social decisions (Olivola et al.,

2014; Todorov et al., 2015). Here, we asked whether the perceived diagnostic value of facial appearance varies across different personality dimensions. Are some traits believed to be more reflected in facial appearance than others? We found that beliefs in the manifestation of personality traits in facial appearance are strongest for sociability (vs. morality or competence). We also found that people are more confident in the accuracy of their trait impressions when judging sociability (vs. morality or competence). In a hiring context, people value information on a person's facial appearance more when looking for a sociable (vs. moral or competent) employee, reliance on facial appearance to select job candidates is seen as more appropriate when looking for a sociable (vs. moral or competent) employee, and reliance on facial appearance is seen as more effective when looking for a sociable (vs. competent) employee (differences between sociability and morality were not significant). Overall, the current results provide converging evidence that people see facial appearance as especially relevant for judging a person's sociability. Put differently, facial cues weigh heavy in judgment and decision-making, especially when judging another person's sociability.

10. Open practices

All data, materials, analysis scripts, and preregistration documents are available at the Open Science Framework (<https://osf.io/cbsmw>).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jesp.2022.104288>.

References

- Adams, R. B., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: A mechanism for impression formation? *Cognition & Emotion*, 26(3), 431–441. <https://doi.org/10.1080/02699931.2012.666502>
- Ames, D. R., Kammrath, L. K., Suppes, A., & Bolger, N. (2010). Not so fast: The (not-quite-complete) dissociation between accuracy and confidence in thin-slice impressions. *Personality and Social Psychology Bulletin*, 26(2), 264–277. <https://doi.org/10.1177/0146167209354519>
- Aristotle. (1936). *Physiognomics*. In T. E. Page, E. Capps, W. H. D. Rouse, A. Post, & E. H. Warmington (Eds.), *Minor works*. Harvard University Press.
- Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology*, 87(6), 763–778. <https://doi.org/10.1037/0022-3514.87.6.763>
- Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General*, 142(1), 143–150. <https://doi.org/10.1037/a0028930>
- Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2017). Can we detect cooperators by looking at their face? *Current Directions in Psychological Science*, 26(3), 276–281. <https://doi.org/10.1177/0963721417693352>
- Bóo, F. L., Rossi, M. A., & Urzúa, S. S. (2013). The labor market return to an attractive face: Evidence from a field experiment. *Economics Letters*, 118(1), 170–172. <https://doi.org/10.1016/j.econlet.2012.10.016>
- 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(4), 703–706. <https://doi.org/10.1016/j.jrp.2009.03.007>
- Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, 41(2), 135–143. <https://doi.org/10.1002/ejsp.744>
- Caulfield, F., Ewing, L., Bank, S., & Rhodes, G. (2016). Judging trustworthiness from faces: Emotion cues modulate trustworthiness judgments in young children. *British Journal of Psychology*, 107(3), 503–518. <https://doi.org/10.1111/bjop.12156>
- Chang, L. J., Doll, B. B., Vant Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, 61(2), 87–105. <https://doi.org/10.1016/j.cogpsych.2010.03.001>
- Chiu, C., Hong, Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology*, 73(1), 19–30. <https://doi.org/10.1037/0022-3514.73.1.19>
- Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in Cognitive Sciences*, 22(5), 388–399. <https://doi.org/10.1016/j.tics.2018.02.006>
- Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *Review of Financial Studies*, 25(8), 2455–2483. <https://doi.org/10.1093/rfs/hhs071>
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy. *Psychological Science*, 17(5), 383–386. <https://doi.org/10.1111/j.1467-9280.2006.01716.x>

- Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, 19(9), 1508–1519. <https://doi.org/10.1162/jocn.2007.19.9.1508>
- Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Appearance-based trust behaviour is reduced in children with autism spectrum disorder. *Autism*, 19(8), 1002–1009. <https://doi.org/10.1177/1362361314559431>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, 42(1), 226–241. <https://doi.org/10.3758/BRM.42.1.226>
- Giner-Sorolla, R., Carpenter, T., Lewis, N. A., Jr., Montoya, A. K., Abersohn, C. L., Bostyn, D. H., & Soderberg, C. (2020). Power to Detect What? Considerations for Planning and Evaluating Sample Size. <https://osf.io/9bt5s/>.
- Gomulya, D., Wong, E. M., Ormiston, M. E., & Boeker, W. (2017). The role of facial appearance on CEO selection after firm misconduct. *Journal of Applied Psychology*, 102(4), 617–635. <https://doi.org/10.1037/apl0000172>
- Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, 24(1), 38–44. <https://doi.org/10.1177/0963721414550709>
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106(1), 148–168. <https://doi.org/10.1037/a0034726>
- Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7, 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin*, 30, 1661–1673. <https://doi.org/10.1177/0146167204271182>
- Hehman, E., Stolier, R. M., Freeman, J. B., Flake, J. K., & Xie, S. Y. (2019). Toward a comprehensive model of face impressions: What we know, what we do not, and paths forward. *Social and Personality Psychology Compass*, 13(2), 1–16. <https://doi.org/10.1111/spc3.12431>
- Hooper, J. J., Sutherland, C. A. M., Ewing, L., Langdon, R., Caruana, N., Connaughton, E., ... Rhodes, G. (2019). Should I trust you? Autistic traits predict reduced appearance-based trust decisions. *British Journal of Psychology*, 110(4), 617–634. <https://doi.org/10.1111/bjop.12357>
- Jaeger, B., Evans, A. M., Stel, M., & van Beest, I. (2019). Explaining the persistent influence of facial cues in social decision-making. *Journal of Experimental Psychology: General*, 148(6), 1008–1021. <https://doi.org/10.1037/xge0000591>
- Jaeger, B., & Jones, A. L. (2021). Which facial features are central in impression formation? *Social Psychological and Personality Science*. <https://doi.org/10.1177/19485506211034979>
- Jaeger, B., Oud, B., Williams, T., Krumbhuber, E. G., Fehr, E., & Engelmann, J. B. (2021). Can people detect the trustworthiness of strangers based on their facial appearance?. <https://psyarxiv.com/ayqeh/>.
- Jaeger, B., Todorov, A. T., Evans, A. M., & van Beest, I. (2020). Can we reduce facial biases? Persistent effects of facial trustworthiness on sentencing decisions. *Journal of Experimental Social Psychology*, 90, Article 104004. <https://doi.org/10.1016/j.jesp.2020.104004>
- Jones, A. L., Kramer, R. S. S., & Ward, R. (2012). Signals of personality and health: The contributions of facial shape, skin texture, and viewing angle. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1353–1361. <https://doi.org/10.1037/a0027078>
- Kenny, D. A., & West, T. V. (2008). Zero acquaintance: Definitions, statistical model, findings, and process. In N. Ambady, & J. J. Skowronski (Eds.), *First impressions* (pp. 129–146). Guilford Press.
- Klapper, A., Dotsch, R., van Rooij, I., & Wigboldus, D. H. J. (2016). Do we spontaneously form stable trustworthiness impressions from facial appearance? *Journal of Personality and Social Psychology*, 111(5), 655–664. <https://doi.org/10.1037/pspa0000062>
- Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, 20(3), 165–182. <https://doi.org/10.1007/bf02281954>
- Kramer, R. S. S., & Ward, R. (2010). Internal facial features are signals of personality and health. *Quarterly Journal of Experimental Psychology*, 63(11), 2273–2287. <https://doi.org/10.1080/17470211003770912>
- Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness: Development and validation of the emotional expressivity scale. *Journal of Personality and Social Psychology*, 66(5), 934–949. <https://doi.org/10.1037/0022-3514.66.5.934>
- Krumbhuber, E. G., Manstead, A. S. R. R., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, 7(4), 730–735. <https://doi.org/10.1037/1528-3542.7.4.730>
- Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When it's bad to be friendly and smart: The desirability of sociability and competence depends on morality. *Personality and Social Psychology Bulletin*, 42(9), 1272–1290. <https://doi.org/10.1177/0146167216655984>
- Lavater, J. C. (1775). *Essays on physiognomy: Designed to promote the knowledge and the love of mankind*. William Tegg and Co.
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93(2), 234–249. <https://doi.org/10.1037/0022-3514.93.2.234>
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74(6), 1421–1536. <https://doi.org/10.1037/0022-3514.74.6.1421>
- Ling, L., Luo, D., & She, G. (2019). Judging a book by its cover: The influence of physical attractiveness on the promotion of regional leaders. *Journal of Economic Behavior and Organization*, 158, 1–14. <https://doi.org/10.1016/j.jebo.2019.01.005>
- Marsh, A. A. (2005). Why do fear and anger look the way they do? Form and social function in facial expressions. *Personality and Social Psychology Bulletin*, 31(1), 73–86. <https://doi.org/10.1177/0146167204271306>
- Mehu, M., Little, A. C., & Dunbar, R. I. M. (2007). Duchenne smiles and the perception of generosity and sociability in faces. *Journal of Evolutionary Psychology*, 5, 183–196. <https://doi.org/10.1556/JEP.2007.1011>
- Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin*, 35(12), 1661–1671. <https://doi.org/10.1177/0146167209346309>
- Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, 18(11), 566–570. <https://doi.org/10.1016/j.tics.2014.09.007>
- Olivola, C. Y., Tingley, D., & Todorov, A. (2018). Republican voters prefer candidates who have conservative-looking faces: New evidence from exit polls. *Political Psychology*, 39(5), 1157–1171. <https://doi.org/10.1111/pops.12489>
- Olivola, C. Y., & Todorov, A. (2010a). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, 34(2), 83–110. <https://doi.org/10.1007/s10919-009-0082-1>
- Olivola, C. Y., & Todorov, A. (2010b). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, 46(2), 315–324. <https://doi.org/10.1016/j.jesp.2009.12.002>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087–11092. <https://doi.org/10.1073/pnas.0805664105>
- 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(5), 607–640. <https://doi.org/10.1521/soco.2006.24.5.607>
- Plaks, J. E., Stroessner, S. J., Dweck, C. S., & Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology*, 80(6), 876–893. <https://doi.org/10.1037/0022-3514.80.6.876>
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.r-project.org/>.
- Roger, D., & Nesselrover, W. (1987). The construction and preliminary validation of a scale for measuring emotional control. *Personality and Individual Differences*, 8(4), 527–534. [https://doi.org/10.1016/0191-8869\(87\)90215-7](https://doi.org/10.1016/0191-8869(87)90215-7)
- 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(3), 409–426. <https://doi.org/10.1037/a0031050>
- Rule, N. O., Tskhay, K. O., Freeman, J. B., & Ambady, N. (2014). On the interactive influence of facial appearance and explicit knowledge in social categorization. *European Journal of Social Psychology*, 44(6), 529–535. <https://doi.org/10.1002/ejsp.2043>
- Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, 9(2), 260–264. <https://doi.org/10.1037/a0014681>
- Secord, P. (1958). Facial features and inference processes in interpersonal perception. In R. Tagiuri, & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 300–315). Stanford University Press.
- Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, 20(6), 395–399. <https://doi.org/10.1177/0963721411424739>
- Shevlin, M., Walker, S., Davies, M. N. O., Banyard, P., & Lewis, C. A. (2003). Can you judge a book by its cover? Evidence of self-stranger agreement on personality at zero acquaintance. *Personality and Individual Differences*, 35(6), 1373–1383. [https://doi.org/10.1016/S0191-8869\(02\)00356-2](https://doi.org/10.1016/S0191-8869(02)00356-2)
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Stewart, L. H., Ajina, S., Getov, S., Bahrami, B., Todorov, A., & Rees, G. (2012). Unconscious evaluation of faces on social dimensions. *Journal of Experimental Psychology: General*, 141(4), 715–727. <https://doi.org/10.1037/a0027950>
- Stolier, R. M., Hehman, E., & Freeman, J. B. (2018a). A dynamic structure of social trait space. *Trends in Cognitive Sciences*, 22(3), 197–200. <https://doi.org/10.1016/j.tics.2017.12.003>
- Stolier, R. M., Hehman, E., Keller, M. D., Walker, M., & Freeman, J. B. (2018b). The conceptual structure of face impressions. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.1807222115>
- Sutherland, C. A. M., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. (2017a). Facial first impressions across culture: Data-driven modelling of Chinese and British perceivers' unconstrained facial impressions. *Personality and Social Psychology Bulletin*, 44(4), 521–537. <https://doi.org/10.1177/0146167217744194>
- Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017b). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, 108(2), 397–415. <https://doi.org/10.1111/bjop.12206>
- Suzuki, A. (2016). Persistent reliance on facial appearance among older adults when judging someone's trustworthiness. *The Journals of Gerontology Series B: Psychological*

- Sciences and Social Sciences*, 73(4), 573–583. <https://doi.org/10.1093/geronb/gbw034>
- Suzuki, A., Tsukamoto, S., & Takahashi, Y. (2017). Faces tell everything in a just and biologically determined world. *Social Psychological and Personality Science*, 10(1), 62–72. <https://doi.org/10.1177/1948550617734616>
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, 66(1), 519–545. <https://doi.org/10.1146/annurev-psych-113011-143831>
- Van Kleef, G. A. (2010). The emerging view of emotion as social information. *Social and Personality Psychology Compass*, 4(5), 331–343. <https://doi.org/10.1111/j.1751-9004.2010.00262.x>
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7), 592–598. <https://doi.org/10.1111/j.1467-9280.2006.01750.x>
- Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, 26(8), 1325–1331. <https://doi.org/10.1177/0956797615590992>
- Zebrowitz, L. A. (2012). Ecological and social approaches to face perception. In G. Rhodes, A. Calder, M. Johnson, & J. V. Haxby (Eds.), *Oxford handbook of face perception*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199559053.013.0003>
- Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science*, 26(3), 237–242. <https://doi.org/10.1177/0963721416683996>