Who Judges a Book By Its Cover?
The Prevalence, Structure, and Correlates of Physiognomic Beliefs
Bastian Jaeger and Anthony M. Evans
Tilburg University
Mariëlle Stel
University of Twente
Ilja van Beest
Tilburg University

Draft version: 4 July 2019
Word count: 10,631
This paper is currently undergoing peer review. Comments are welcome.

Author Note
Bastian Jaeger, Anthony M. Evans, and Ilja van Beest, Department of Social Psychology, Tilburg University, The Netherlands; Mariëlle Stel, Department of Psychology of Conflict, Risk and Safety, The Netherlands.

All data, analysis scripts, materials, and preregistration documents are available at the Open Science Framework (https://osf.io/s9nj8/).

Correspondence concerning this article should be addressed to Bastian Jaeger, Department of Social Psychology, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: bxjaeger@gmail.com.
Abstract

The question of whether personality can be inferred from faces is contentiously debated. We propose that, irrespective of the actual accuracy of trait inferences from faces, lay beliefs about the manifestation of personality traits in facial features (i.e., physiognomic beliefs) have important consequences for social cognition and behavior. In five studies \((N = 3,861)\), we examine the prevalence, structure, and correlates of physiognomic beliefs. We find that belief in physiognomy is common among students (Study 1) and in a large, representative sample of the Dutch population (Study 2). Physiognomic beliefs are relatively stable over time and associated with an intuitive thinking style (Study 3). However, the strength of physiognomic beliefs varies across different personality dimensions: sociability is believed to be more reflected in facial appearance than morality or competence (Studies 1-5). Crucially, individual differences in belief strength predict how people form and use first impressions. People with stronger physiognomic beliefs are more confident in their trait inferences (Study 4) and rely more on them when making decisions (Study 5). Yet, this increased confidence is not explained by superior accuracy of personality inferences, and the endorsement of physiognomic beliefs is associated with overconfidence (Study 4). Overall, there is widespread belief in physiognomy among laypeople, and individual differences in belief strength relate to various social-cognitive processes and behaviors.

Keywords: physiognomy, lay beliefs, social perception, trait inferences, social decision-making
Who Judges a Book By Its Cover?

The Prevalence, Structure, and Correlates of Physiognomic Beliefs

The practice of physiognomy involves inferring psychological characteristics from facial (or other bodily) features. The idea dates back to Ancient Greece and enjoyed particular popularity in the 18\textsuperscript{th} and 19\textsuperscript{th} centuries (Aristotle, trans. 1936; Lavater, 1775; Woods, 2017). More rigorous scientific approaches at the beginning of the 20\textsuperscript{th} century provided evidence against many of physiognomy’s claims (e.g., Cleeton & Knight, 1924) and it is now widely regarded as pseudo-science (Todorov, 2017). However, research in the field of social perception has shown that faces play an important role in everyday impression formation: People spontaneously infer personality traits from facial appearance and these inferences guide many consequential decisions (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). This raises questions about belief in physiognomy among laypeople.

People develop and rely on lay theories—systems of implicit or explicit beliefs—to navigate the social world (Baumeister & Monroe, 2014; Boyer & Petersen, 2018). For example, research on lay personality theory has shown that people hold beliefs about the basis (Haslam, Bastian, & Bissett, 2004), malleability (Chiu, Hong, & Dweck, 1997), structure (Stolier, Hehman, Keller, Walker, & Freeman, 2018), and expression (Mehl, Gosling, & Pennebaker, 2006) of personality traits. Critically, individual differences in 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). In other words, lay beliefs about different aspects of personality are widespread and shape various social-cognitive processes and behaviors.

Here, we examine a facet of lay personality theory that has received little attention thus far: the belief that personality is reflected in facial appearance (i.e., \textit{physiognomic beliefs}). We propose that individual differences in physiognomic beliefs influence various aspects of impression formation, such as confidence in trait inferences and reliance on trait inferences in social decision-making. We examine (a) the prevalence of physiognomic beliefs, (b) their structure (i.e., which characteristics people believe are most reflected in facial features), (c) the psychological correlates of physiognomic beliefs (e.g., the relationship between physiognomic beliefs and epistemic motivation), and (d) whether physiognomic beliefs are related to accurate social perception and greater reliance on facial appearance in social decision-making.
Physiognomy and Social Perception

The core tenet of physiognomy holds that facial morphology (i.e., features of resting, non-expressive faces) is indicative of psychological characteristics and behavioral tendencies (Todorov, 2017). Early writings proposed that the size and orientation of facial features reflect their frequent use (e.g., a disagreeable person who frowns a lot will have lowered eyebrows); moreover, resemblances between humans and other animals were thought to point to shared psychological attributes (e.g., a person who looks like a lion is brave like a lion; Aristotle, trans. 1936). These speculations were not based on rigorous scientific study and many claimed links between specific facial features and personality traits were disconfirmed by empirical work at the beginning of the 20th century (Alley, 1988). For example, Cleeton and Knight (1924, p. 216) reported that the correlation between "variations in physical traits purported to reveal variations in character traits and [character] criteria was 0.000."

Even though early 20th century research found little support for physiognomy, interest in the topic has grown again in recent years. Research in the field of social perception has yielded new insights into the determinants of impression formation, showing that people spontaneously infer a variety of personality traits from resting, non-expressive faces (Todorov, Olivola, et al., 2015). In fact, people are relatively confident in their own physiognomic judgments (Ames, Kammrath, Suppes, & Bolger, 2010; Hassin & Trope, 2000) and rely on them when making a wide range of consequential decisions (Olivola, Funk, & Todorov, 2014). For instance, voting decisions are influenced by the perceived competence of political candidates (Olivola & Todorov, 2010a) and criminal sentencing decisions are influenced by the perceived trustworthiness of defendants (J. P. Wilson & Rule, 2015). People even rely on facial appearance when they have access to superior, objective information (Jaeger, Evans, Stel, & van Beest, 2019; Olivola, Tingley, & Todorov, 2018). The pervasive influence of trait impressions has again raised questions about the diagnosticity of facial features for inferring personality (Bonnefon, Hopfensitz, & De Neys, 2017; Todorov, Olivola, et al., 2015; Zebrowitz, Fellous, Mignault, & Andreoletti, 2003)

1 The term physiognomy is sometimes used more broadly to refer to judgments of various characteristics (e.g., personality, criminality, social class) based on any aspect of outward appearance. Here, we restrict our focus to the link between facial appearance and personality, which has historically received the most attention.
Can people form accurate personality impressions based on facial appearance? To address this question, studies have tested whether personality impressions from faces (i.e., trait judgments based on facial photographs) reflect a target’s actual personality (Berry, 1990; Penton-Voak, Pound, Little, & Perrett, 2006). Overall, evidence in favor of accuracy in face-based personality judgments is weak and inconsistent. Some studies find a small “kernel of truth” in trait impressions (Bonnefon, Hopfensitz, & De Neys, 2013; De Neys, Hopfensitz, & Bonnefon, 2017; Lin, Adolphs, & Alvarez, 2018; Penton-Voak et al., 2006; Satchell, Davis, Julle-Danière, Tupper, & Marshman, 2018; Slepian & Ames, 2015; Tognetti, Berticat, Raymond, & Faurie, 2013). However, others find no accuracy (Efferson & Vogt, 2013; Graham, Harvey, & Puri, 2017; Ling, Luo, & She, 2019; Rule, Krendl, Ivcevic, & Ambady, 2013) or provide theoretical arguments against accurate trait impressions (McCullough & Reed, 2016; Todorov & Porter, 2014). For example, trait impressions vary substantially across different perceivers (Hehman, Sutherland, Flake, & Slepian, 2017), contexts (Brambilla, Biella, & Freeman, 2018), and even across different images of the same target (Todorov & Porter, 2014), suggesting that they are not reliable indicators of personality. In short, the available evidence suggests that facial features are, at best, a weak indicator of personality.

**Physiognomic Beliefs**

Here, we propose that, irrespective of the actual accuracy of trait inferences from faces, people may hold lay beliefs about the manifestation of personality traits in facial features (i.e., physiognomic beliefs). Crucially, widespread beliefs in physiognomy in the general population may explain the pervasive effects of facial appearance on social cognition and behavior. In a similar vein, individual differences in physiognomic beliefs may predict why some people are overconfident in the accuracy of their trait impressions and persistently rely on them when making decisions.

Little is known about lay beliefs in physiognomy. Anecdotal evidence suggests that some people hold physiognomic beliefs (Hassin & Trope, 2000; Liggett, 1974). Moreover, a recent study by Suzuki and colleagues (2017) showed that physiognomic beliefs are related to (a) other lay beliefs, such as belief in a just world and belief in the biological determinism of personality and (b) more extreme trait judgments based on facial appearance.

The present research examines the role of physiognomic beliefs in social perception more broadly. First, we examine the prevalence of physiognomic beliefs in the general population. We
assess belief in physiognomy in five samples (including a large, representative sample of the Dutch population) and explore who is more likely to endorse physiognomic beliefs. For example, we test whether physiognomic beliefs are correlated with various characteristics such as age, education level, or thinking style. Second, we investigate the heterogeneity of physiognomic beliefs across different personality dimensions. Specifically, we ask whether some personality traits are believed to be more reflected in faces than others. Third, and most importantly, we test whether individual differences in physiognomic beliefs predict how confident people are in the accuracy of their trait impressions and how much they rely on trait impressions in social decision-making.

The Current Studies

To investigate the prevalence, structure, and correlates of physiognomic beliefs, we introduce a novel scale. The scale consists of two parts with a total of 15 items. To ensure that participants envision a resting, non-expressive face, we prompt them to “imagine seeing the passport photo of a stranger”. The first part (3 items) assesses general physiognomic beliefs (e.g., “I can learn something about the person’s personality just from looking at his or her face”). The second part (12 items) assesses specific physiognomic beliefs by asking respondents to indicate how accurately they think different characteristics can be inferred from a person’s face. Our scale measures physiognomic beliefs for three fundamental dimensions underlying person perception (sociability, morality, and competence; Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Leach, Ellemers, & Barreto, 2007), as well as three additional characteristics (age, gender, and attractiveness). This allows us to test which characteristics are believed to be more reflected in faces.

2 Critically, our measure differs from the measure introduced by Suzuki and colleagues (2017), which assessed physiognomic beliefs by asking participants which characteristics (e.g., aggressiveness, cooperativeness) they think they can determine about a person from their face. It is unclear whether high scores on their scale reflect beliefs that traits can be inferred from stable, morphological features of faces—which is the claim of physiognomy (Aristotle, 1936; Lavater, 1775) and the subject of debate (e.g., Bonnefon et al., 2015; Todorov, Funk, et al., 2015)—or from other facial characteristics such as emotion expressions.

3 Research on the accuracy of trait inferences from faces has predominantly focused on basic personality dimensions such as the Big Five traits (Penton-Voak et al., 2006). Since we were interested in people’s conceptual beliefs about personality, we focused on sociability, morality, and competence. These dimensions may not capture the structure of a person’s actual personality as well as other models, but they represent the dimensions that people spontaneously use to judge others’ personality (Brambilla et al., 2011).
We report the results of five studies ($N = 3,861$; four preregistered). In Study 1 and 2, we investigate the prevalence and structure of physiognomic beliefs in a student sample ($n = 378$) and a representative sample of the Dutch population ($n = 2,624$). We estimate how many people believe that personality is reflected in facial features and test which personality dimensions are believed to be more visible in faces. Study 3 ($n = 229$) examines the relationship between physiognomic beliefs and individual differences in other lay beliefs (e.g., belief in the biological determinism of personality traits; Haslam et al., 2004) and epistemic motivation (e.g., faith in intuition; Epstein, Pacini, Denes-Raj, & Heier, 1996). We also examine the temporal stability of physiognomic beliefs.

Studies 4 and 5 investigate the relationships between physiognomic beliefs and social perception and decision-making. Study 4 ($n = 406$) tests whether people who score higher on physiognomic beliefs are more confident in the accuracy of their trait impressions and whether increased confidence can be explained by the fact that their judgments are indeed more accurate. Finally, Study 5 ($n = 224$) investigates whether people who endorse physiognomic beliefs rely more on trait inferences from faces when making social decisions.

All data, analysis scripts, and preregistration documents are available at the Open Science Framework (https://osf.io/s9nj8/). We report how our sample sizes were determined and all data exclusions and measures for each study.

**Studies 1 and 2: Prevalence and Structure**

In Studies 1 and 2, we estimated the prevalence of physiognomic beliefs and tested how belief strength varies for different personality traits. We predicted that facial appearance would be seen as more indicative of physically salient characteristics, such as gender, age, and attractiveness. People can detect gender and age from faces with high accuracy (Bruce & Young, 2012), and overall attractiveness is strongly influenced by facial appearance (Peters, Rhodes, & Simmons, 2007). We also examined differences in physiognomic beliefs across three fundamental dimensions in person perception: sociability, morality, and competence. We predicted that people would hold stronger physiognomic beliefs for sociability (compared to morality and competence). Emotion perception in resting (i.e., emotionally neutral) faces play a central role in impression formation (Said, Sebe, & Todorov, 2009). Critically, emotional expressiveness is a defining feature of sociability, making facial appearance particularly relevant for sociability judgments (Kring, Smith, & Neale, 1994; Riggio & Riggio, 2002).
We tested our predictions by administering the physiognomic belief scale in a sample of first-year psychology students (Study 1; \( n = 378 \)) and a representative sample of the Dutch population (Study 2; \( n = 2,624 \)).

**Methods**

**Participants.** In Study 1, we recruited 378 first-year psychology students from a Dutch university (\( M_{age} = 20.61, SD_{age} = 2.19; 76.46\% \) female, 23.28\% male, 0.26\% other). The majority of participants were Dutch (68.25\%) or German (19.05\%). Sample size was determined by how many students participated in the study within two weeks. A sensitivity analysis in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) showed that this sample size sample afforded us 80\% power to detect a small difference (\( d = 0.14 \)) when comparing physiognomic beliefs for sociability, morality, and competence (with \( \alpha = 5\% \)).

In Study 2, a representative sample of 2,807 Dutch participants was recruited via the LISS (Longitudinal Internet Studies of the Social Sciences) panel (Scherpenzeel & Das, 2010). The panel is based on a probability sample of Dutch households drawn from the population register. Panel members are representative of the Dutch population on indicators like gender, age, education, and income.\(^4\) Data from 183 participants (6.52\%) who had missing data for at least one question was excluded from analysis, leaving a final sample of 2,624 participants (\( M_{age} = 52.60, SD_{age} = 16.50; 52.52\% \) female, 47.48\% male). A sensitivity analysis in G*Power (Faul et al., 2007) showed that this sample size afforded us 80\% power to detect a small difference (\( d = 0.05 \)) when comparing physiognomic beliefs for sociability, morality, and competence (with \( \alpha = 5\% \)).

**Materials and procedure.** We developed a questionnaire that measures the belief that personality is reflected in facial features. The questionnaire consists of two parts measuring general and specific physiognomic beliefs. First, participants were prompted to imagine seeing the passport photo of a stranger. They were asked to indicate how much they agreed with three statements (e.g., *I can learn something about a person’s personality just from looking at his or her face*) on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Average scores across the three items constituted our measure of general physiognomic beliefs (Study 1: Cronbach’s \( \alpha = .74 \), Study 2: Cronbach’s \( \alpha = .63 \)). We randomized the order in which the three items were presented.

\(^4\) For more information on the LISS panel, see lissdata.nl.
Next, we asked participants how accurately they could judge a person on various characteristics from looking at that person’s passport photo. For each item, responses were indicated using a slider from 0 (not at all accurately) to 100 (extremely accurately). Participants responded to twelve items: Three sociability-related traits (warmth, friendliness, likeability), three morality-related traits (trustworthiness, sincerity, honesty), and three competence-related traits (competence, intelligence, skillfulness; Brambilla, Rusconi, Sacchi, & Cherubini, 2011). We also included three additional physically salient characteristics—gender, age, and attractiveness. We randomized the order in which the twelve items were presented. Participants’ ratings constituted our measure of specific physiognomic beliefs.

In Study 1, Dutch participants completed a Dutch version of the scale while non-Dutch participants completed an English version. In Study 2, all participants completed the Dutch version of the scale.

Results

Study 1: Student sample. In our sample of Dutch students, the average score on general physiognomic belief was just above the scale midpoint ($M = 4.13, SD = 1.23$), $t(377) = 2.03, p = .043, d = 0.10$. Around half of all participants (56.88%) believed at least somewhat in physiognomy (i.e., they scored above the midpoint of the scale), $\chi^2(1) = 6.88, p = .009$.

We calculated average physiognomic belief scores across the nine personality traits to test whether participants think that facial features are more indicative of physically salient characteristics (e.g., a person’s gender, age, and attractiveness) than a person’s personality. Personality-specific physiognomic beliefs ($M = 26.68, SD = 18.46$) were significantly lower than physiognomic beliefs for gender ($M = 88.30, SD = 15.39$), $t(377) = 55.35, p < .001, d = 2.84$, attractiveness ($M = 76.63, SD = 18.51$), $t(377) = 44.68, p < .001, d = 2.30$, and age ($M = 66.75, SD = 18.11$), $t(377) = 34.06, p < .001, d = 1.75$.

We also compared physiognomic beliefs across the three personality dimensions to test whether people think sociability is more reflected in facial features than morality or competence (see Figure 1). Physiognomic beliefs were significantly higher for sociability ($M = 37.12, SD = 21.28$) compared to morality ($M = 20.11, SD = 18.61$), $t(377) = 20.12, p < .001, d = 1.03$, and competence ($M = 22.81, SD = 19.74$), $t(377) = 15.79, p < .001, d = 0.81$ (see Figure 3). Competence-specific physiognomic beliefs were higher than morality-specific physiognomic beliefs, but this difference was less pronounced, $t(377) = 3.84, p < .001, d = 0.20$. 
**Study 2: Representative sample.** In our representative sample of the Dutch population, the average score on general physiognomic belief was again above the scale midpoint \((M = 4.17, SD = 1.08)\), \(t(2,623) = 7.96, p < .001, d = 0.16\). Around half of all participants (52.10%) believed at least somewhat in physiognomy (i.e., they scored above the midpoint of the scale), \(\chi^2(1) = 4.53, p = .033\).

As in the student sample, personality-specific physiognomic beliefs \((M = 32.12, SD = 18.70)\) were significantly lower than physiognomic beliefs for gender \((M = 75.66, SD = 19.48)\), \(t(2,623) = 83.45, p < .001, d = 1.63\), age \((M = 60.52, SD = 17.41)\), \(t(2,623) = 62.52, p < .001, d = 1.22\), and attractiveness \((M = 64.59, SD = 20.68)\), \(t(2,623) = 65.59, p < .001, d = 1.28\).

Again, physiognomic beliefs were significantly higher for sociability \((M = 38.42, SD = 21.36)\) than for morality \((M = 29.65, SD = 20.49)\), \(t(2,623) = 34.56, p < .001, d = 0.67\), or competence \((M = 28.29, SD = 18.76)\), \(t(2,623) = 34.80, p < .001, d = 0.68\) (see Figure 1). Contrary to findings from Study 1, morality-specific physiognomic beliefs were higher than competence-specific physiognomic beliefs, but the difference was again small, \(t(2,623) = 5.77, p < .001, d = 0.11\).

We also explored the relationship between physiognomic beliefs and basic demographic indicators. We regressed general physiognomic belief on gender (coded 0 for male and 1 for female), age, income (z-standardized net monthly income), and level of education (six levels, ranging from primary school to university degree). Data from 278 participants (10.59%) who indicated an income of zero and 2 respondents (0.08%) whose reported income was 23.7 and 41.4 standard deviation above the mean were excluded from analysis, leaving a sample of 2,344 participants. Women scored higher on physiognomic belief than men, \(\beta = 0.104, SE = 0.048, t(2,335) = 2.17, p = .030\), and there was a negative effect of age, \(\beta = -0.004, SE = 0.001, t(2,335) = 3.07, p < .001\). However, these effects were small: women’s physiognomic belief score was 0.097 standard deviations above men’s score and a ten-year age difference was associated with a 0.039 standard deviation decrease in belief. There was no effect of income or education (see Supplemental Materials for the full results).
Table 1

<table>
<thead>
<tr>
<th>Sample</th>
<th>n</th>
<th>α</th>
<th>M</th>
<th>SD</th>
<th>% Believers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td>378</td>
<td>.74</td>
<td>4.13</td>
<td>1.23</td>
<td>56.88</td>
</tr>
<tr>
<td>Study 2</td>
<td>2,624</td>
<td>.63</td>
<td>4.17</td>
<td>1.08</td>
<td>52.10</td>
</tr>
<tr>
<td>Study 3</td>
<td>229</td>
<td>.85</td>
<td>3.92</td>
<td>1.23</td>
<td>47.60</td>
</tr>
<tr>
<td>Study 4</td>
<td>406</td>
<td>.76</td>
<td>4.14</td>
<td>1.33</td>
<td>54.43</td>
</tr>
<tr>
<td>Study 5</td>
<td>224</td>
<td>.76</td>
<td>3.97</td>
<td>1.18</td>
<td>48.66</td>
</tr>
</tbody>
</table>

Note. General physiognomic beliefs were measured with three items that were rated on a scale from 1 (strongly disagree) to 7 (strongly agree). “% Believers” indicates the percentage of participants that scored above the midpoint of the scale.

Discussion

Belief in physiognomy was prevalent in both student and representative samples—over half of all participants at least somewhat endorsed the belief that personality is reflected in facial features. Physiognomic beliefs were higher among women and older participants, but these...
differences were small. Moreover, we found no evidence that belief endorsement varied across different levels of education or income. These results suggest that physiognomic beliefs are common across different demographic groups. We also found that people have heterogeneous beliefs about the diagnosticity of facial features for specific characteristics: Participants believed that sociability can be more accurately inferred from facial features than morality or competence. Differences in physiognomic beliefs for morality and competence, however, were small and inconsistent.

**Study 3: Psychological Correlates and Temporal Stability**

Our first studies provided evidence that physiognomic beliefs are relatively common, and that facial appearance is believed to be more indicative of sociability than morality or competence. In Study 3, we extended our analysis in two ways. First, we examined the psychological correlates of physiognomic beliefs. That is, we investigated the association of physiognomic beliefs with individual differences in other lay beliefs. Suzuki and colleagues (2017) found that physiognomic beliefs were positively correlated with both entity beliefs (i.e., beliefs in the fixedness and immutability of personality traits; Chiu et al., 1997) and beliefs in biological determinism (i.e., beliefs in the biological determinism of personality traits; Haslam et al., 2004). These findings suggest that physiognomic beliefs are based on the idea that a common factor (e.g., genetic makeup) influences both personality traits and facial appearance (Stirrat & Perrett, 2010). Physiognomic beliefs were also correlated with belief in a just world (Lipkus, 1991), suggesting they may also be rooted in the idea that people “get what they deserve”, with, for example, immoral people having a facial appearance that betrays their immorality to others. We investigated whether the findings by Suzuki and colleagues (2017) would replicate in a sample of British participants.

Going beyond prior studies, we also investigated whether physiognomic beliefs are related to epistemic motivation. Trait impressions from faces are formed quickly and effortlessly (Stewart et al., 2012; Willis & Todorov, 2006) and this accessibility may make trait impressions intuitively appealing. We therefore expected that endorsement of physiognomic beliefs is more prevalent among people who tend to trust their intuition (i.e., who score high on faith in intuition; Epstein et al., 1996). Relatedly, some people are more prone to override intuitive responses with more analytic and reflective responses (Frederick, 2005; Pennycook, Cheyne, Koehler, & Fugelsang, 2015). We therefore tested whether physiognomic beliefs are negatively
related to cognitive reflection. Finally, people vary in their need to form evaluative judgments (Jarvis & Petty, 1996). If people form personality impressions based on facial appearance (which is usually an easily available cue) to satisfy this tendency, endorsement of physiognomic beliefs may justify this behavior. We therefore tested whether physiognomic beliefs are correlated with individual differences in the need to evaluate (Jarvis & Petty, 1996).

Second, we tested the temporal stability of physiognomic beliefs. Participants completed the physiognomic belief scale at two time points with a four-week delay

**Methods**

**Participants.** In the study of Suzuki and colleagues (2017), correlations ranging from $r = .185$ to $r = .445$ ($n = 1,396$) were reported. We therefore aimed to recruit 227 participants, which affords 80% power to detect a correlation of $r = .185$ (with $\alpha = 5\%$). We recruited 310 British Prolific workers to complete the study in exchange for £1.25 each. In line with our preregistration, data from 79 participants (25.48%) who failed an attention check question at the end of the study and from 2 participants (0.87%) who indicated poor or basic English proficiency were excluded from analysis, leaving a final sample of 229 participants ($M_{age} = 35.62$, $SD_{age} = 11.86$; 60.26% female, 39.30% male, 0.44% other).

To measure the temporal stability of physiognomic beliefs, we re-contacted participants after four weeks. An a priori power analysis showed that a sample size of 84 participants is required to detect a medium-sized correlation ($r = .300$) between physiognomic belief scores at both time points with 80% power (and $\alpha = 5\%$). We re-contacted a total of 200 British Prolific workers to complete the second part of the study in exchange for £0.50 each. In line with our preregistration, data from 55 participants (27.50%) who failed an attention check question at the end of the study and from 16 participants (11.03%) whose responses could not be matched with data from part one were excluded from analysis, leaving a final sample of 129 participants ($M_{age} = 38.26$, $SD_{age} = 12.16$; 60.47% female, 38.76% male, 0.78% other).

**Materials and procedure.** Physiognomic beliefs were measured as described in Study 1.

Belief in the entity theory of personality was measured with eight items (e.g., "The kind of person someone is, is something basic about them, and it can't be changed very much") adapted from Levy and colleagues (1998; Study 5). Participants indicated how much they agreed with each statement on a scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*).
Following previous investigations (Haslam et al., 2004; Suzuki et al., 2017), we measured belief in biological determinism of personality traits by showing participants a list of nine personality traits (the same nine personality traits representing evaluations of sociability, morality, and competence that are used for the physiognomic belief scale). We asked them to rate how much each trait is based on biological nature (genes, brain structure, etc.) on a scale ranging from 0 (not based on biological nature) to 100 (based on biological nature).

Belief in a just world was measured with seven items (e.g., “I feel that people get what they deserve”) adapted from Lipkus (1991). Participants indicated how much they agreed with each statement on a scale ranging from 1 (strongly disagree) to 6 (strongly agree).

Faith in intuition was measured with twelve items (e.g., “I trust my initial feelings about people”) adapted from Epstein and colleagues (1996). Participants indicated how much they agreed with each statement on a scale ranging from 1 (completely false) to 5 (completely true).

Cognitive reflection was measured with the seven-item cognitive reflection test (CRT; e.g., “If you are in a race and you pass the person in second place, what place are you in?”) adapted from Thomson and Oppenheimer (2016). The CRT measures the tendency to override an intuitive, but incorrect answer with a more reflective and correct one. Participants indicated their responses in a free form text box. The test was scored by adding up the number of items that were answered correctly.

Need to evaluate was measured with sixteen items (e.g., “I form opinions about everything”) adapted from Jarvis and Petty (1996). Participants indicated to what extent each item was characteristic of them on a scale ranging from 1 (extremely uncharacteristic) to 5 (extremely characteristic).

Participants completed the six measures, and the items within each measure, in a random order. We randomized whether participants completed the physiognomic belief scale before or after the other measures. For the second part of the study which was conducted four weeks later, participants completed the physiognomic belief scale a second time.

**Results**

The average score on general physiognomic belief was just below the midpoint of our scale ($M = 3.92, SD = 1.23$), $t(228) = 0.98, p = .33, d = 0.06$ (see Table 1). Around half of all participants (47.60%) believed at least somewhat in physiognomy (i.e., they scored above the
midpoint of the scale), \( \chi^2(1) = 0.44, p = .51 \). All scales showed acceptable to excellent internal consistency (\( .72 < \alpha < .92 \); see Table 2).

**Psychological correlates.** First, we examined the relationship between physiognomic beliefs and other lay beliefs. Participants who scored higher on physiognomic beliefs also scored higher on belief in the biological determinism of personality traits, \( r(227) = .172, p = .009 \), and belief in a just world, \( r(227) = .154, p = .020 \). We did not find that physiognomic beliefs were related to belief in the entity theory of personality, \( r(227) = .093, p = .16 \).

Next, we examined how physiognomic beliefs relate to epistemic motivation. Participants who scored higher on physiognomic beliefs also scored higher on faith in intuition, \( r(227) = .409, p < .001 \). We did not find any evidence that physiognomic beliefs were related to scores on the cognitive reflection test, \( r(227) = -.048, p = .46 \). The correlation with need to evaluate was positive, but only marginally significant, \( r(227) = .125, p = .058 \).

We also explored which of the measures were uniquely related to physiognomic beliefs, by regressing physiognomic beliefs on all six measures, which were \( z \)-standardized prior to analysis. This yielded a positive effect of faith in intuition, \( \beta = 0.484, SE = 0.087, t(222) = 5.59, p < .001 \), but no significant effects of entity beliefs, \( \beta = 0.078, SE = 0.77, t(222) = 1.01, p = .31 \), beliefs in biological determinism, \( \beta = 0.045, SE = 0.080, t(222) = 0.57, p = .57 \), belief in a just world, \( \beta = 0.091, SE = 0.78, t(222) = 1.17, p = .24 \), cognitive reflection, \( \beta = 0.024, SE = 0.076, t(222) = 0.31, p = .75 \), or need to evaluate, \( \beta = -0.024, SE = 0.083, t(222) = 0.29, p = .77 \).

---

5 The correlation with belief in a just world was no longer significant when correcting for multiple comparisons (see Supplemental Materials).
6 The faith in intuition scale includes two items that directly refer to the accuracy of appearance-based impressions (“My initial impressions of people are almost always right” and “I believe I can judge character pretty well from a person's appearance”). Physiognomic beliefs were still correlated with faith in intuition when these two items were omitted, \( r = .345, p < .001 \).
Table 2
Descriptive statistics and correlations for all personality variables.

<table>
<thead>
<tr>
<th>Measure</th>
<th>$M$</th>
<th>$SD$</th>
<th>$\alpha$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PB</td>
<td>3.92</td>
<td>1.23</td>
<td>.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. BET</td>
<td>3.30</td>
<td>0.90</td>
<td>.92</td>
<td>.093</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. BBD</td>
<td>43.61</td>
<td>18.57</td>
<td>.88</td>
<td>.172***</td>
<td>.229***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. BJW</td>
<td>2.98</td>
<td>0.81</td>
<td>.87</td>
<td>.154*</td>
<td>.017</td>
<td>.104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. FI</td>
<td>3.56</td>
<td>0.60</td>
<td>.85</td>
<td>.409***</td>
<td>.054</td>
<td>.253***</td>
<td>.183**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. CRT</td>
<td>3.73</td>
<td>1.91</td>
<td>.72</td>
<td>-.048</td>
<td>.019</td>
<td>-.037</td>
<td>.016</td>
<td>-.177**</td>
<td></td>
</tr>
<tr>
<td>7. NE</td>
<td>3.48</td>
<td>0.60</td>
<td>.82</td>
<td>.125†</td>
<td>.079</td>
<td>.116†</td>
<td>-.125†</td>
<td>.370***</td>
<td>-.009</td>
</tr>
</tbody>
</table>

Note. PB = Physiognomic belief, BET = Belief in the entity theory of personality, BBD = Belief in the biological determinism of personality traits, BJW = Belief in a just world, FI = Faith in intuition, CRT = Cognitive reflection, NE = Need to evaluate.

† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Temporal stability. The average score on general physiognomic belief when measured four weeks later was just below the midpoint of our scale ($M = 3.99$, $SD = 1.18$), $t(128) = 0.12$, $p = .91$, $d = 0.01$. Around half of all participants (48.84%) believed at least somewhat in physiognomy (i.e., scored above the midpoint of the scale), $\chi^2(1) = 0.03$, $p = .86$. We found a strong correlation between general physiognomic belief scores at both time points, $r(127) = .644$, $p < .001$, and between specific physiognomic beliefs at both time points, $r(127) = .609$, $p < .001$.

Discussion

The current study provided new insights into who believes in physiognomy. In line with Suzuki and colleagues (2017), we found that people who score high on physiognomic belief were more likely to endorse the belief that personality traits are determined by biological factors. This suggests that physiognomic beliefs may be rooted in the idea that biological factors (e.g., genetic makeup) determine both personality and facial appearance. We did not find evidence that physiognomic beliefs were related to entity beliefs, and the correlation with belief in a just world was weak.

$^7$ Correlation coefficients ranged from .43 for age-specific belief to .72 for competence-specific belief.
We did find a correlation between physiognomic beliefs and a measure of intuitive thinking style. Trait inferences from faces are formed spontaneously, quickly, and effortlessly and can serve as input for intuitive judgments or decisions (Freeman & Johnson, 2016; Jaeger et al., 2019). Thus, the intuitive accessibility of trait inferences may make them particularly appealing for individuals who tend to follow their intuitions. In line with this view, people who scored high on faith in intuition endorsed physiognomic beliefs more strongly. Moreover, predicting physiognomic beliefs with all individual differences measured here showed only a unique effect of intuitive thinking style. At the same time, we did not find a correlation with scores on the cognitive reflection test. This result suggests that reliance on physiognomic beliefs stems from the preference to rely on intuition, rather than an inability to override intuitive response.

**Study 4: Confidence, Accuracy, and Meta-Accuracy**

In Study 4, we tested how individual differences in physiognomic beliefs are related to actual and predicted accuracy of trait impressions from faces. We aimed to address three questions. First, we examined whether physiognomic beliefs are related to confidence in the accuracy of trait impressions. Research shows that people are relatively confident in the accuracy of their physiognomic judgments, in spite of their generally low actual accuracy (Ames et al., 2010; Biesanz et al., 2011). We propose that this confidence may result from subjective beliefs that faces are a good indicator of personality. Second, we examined whether physiognomic beliefs are related to the actual accuracy of trait impressions. People might endorse physiognomic beliefs because their physiognomic judgments are indeed more accurate. Third, we examined whether superior judgment accuracy explains the relationship between physiognomic beliefs and confidence. People who believe in physiognomy may be justifiably more confident in their trait impressions because their impressions are more accurate. Alternatively, physiognomic beliefs may influence confidence irrespective of judgment accuracy. This would imply that increased confidence by people who believe in physiognomy is not normatively justified. To test these competing accounts, we examined whether the relationship between physiognomic beliefs and confidence is mediated by judgment accuracy.

To answer these questions, we replicated a previous study on the accuracy of trait impressions from faces (Lin et al., 2018). In this study, participants rated the corruptibility of government officials who had a clean record or who had been found guilty of political
corruption. Lin and colleagues (2018) found that accuracy in corruptibility detection based on facial photographs was significantly above chance and people varied in how accurate their judgments were. Here, we gathered corruptibility ratings of the same photo stimuli to measure accuracy in corruptibility judgments. We measured confidence by asking participants to estimate how many individuals they would classify correctly. Participants also completed the physiognomic belief scale. Given that participants specifically judged corruptibility, which is conceptually similar to trustworthiness (Lin et al., 2018), we analyzed their trustworthiness-specific physiognomic beliefs.

**Methods**

**Participants.** An a priori power analysis showed that a sample size of 193 participants is required to detect a small-to-medium-sized correlation between physiognomic beliefs and confidence in trait judgments \((r = .20)\) with 80% power (and \(\alpha = 5\%\)). We therefore aimed to recruit at least 193 participants, with the final sample size being determined by how many students participated in the study in two weeks. In total, we recruited 512 first-year psychology students from a Dutch university who completed the study in return for partial course credit. In line with our preregistration, data from 101 participants (19.73%) who indicated poor or basic English proficiency, from 3 participants (0.01%) who always indicate the same rating (corruptible or not corruptible) across all trials, and from 1 participant (0.002%) whose response time was faster than 100 milliseconds on at least 10% of all trials were excluded, leaving a final sample of 406 participants \((M_{\text{age}} = 20.01, SD_{\text{age}} = 2.17; 80.54\% \text{ female}, 19.21\% \text{ male}, 0.25\% \text{ other})\).

**Materials and procedure.** We used an image set of 72 US government officials created by Lin and colleagues (2018, Study 1). Half of the politicians were convicted of political corruption, whereas the other half had clean records. The images were obtained from personal websites, news articles, or Wikipedia. The faces were converted to gray-scale, cropped to a uniform size, and shown against a uniform background.

Participants saw the 72 images in a randomized order and were asked to indicate whether they thought the politician in each photo was corruptible (i.e., untrustworthy, dishonest, selfish) or not (response options were “yes” or “no”). Next, participants were asked whether they had recognized any of the individuals and, in case they answered affirmatively, whom they had recognized. None correctly identified any politicians. We measured confidence in the accuracy
of impressions by asking participants to rate how often they think they made the right judgment on a scale from 0% of the time to 100% of the time. Participants were reminded that, given the two-alternative forced choice design, 50% accuracy would be expected by chance. Finally, participants completed the physiognomic belief scale (Cronbach’s $\alpha = .76$). All participants completed the study in English.

**Results**

The average score on general physiognomic belief was just above the midpoint of our scale ($M = 4.14, SD = 1.13$), $t(405) = 2.53, p = .012, d = 0.13$ (see Table 1). Around half of all participants (54.43%) believed at least somewhat in physiognomy (i.e., they scored above the midpoint of the scale), $\chi^2(1) = 3.02, p = .082$.

**Accuracy and meta-accuracy.** First, we investigated whether face-based corruptibility judgments were accurate. We examined individual-level accuracy by testing whether the percentage of times participants made a correct judgment (i.e., judging a corrupt politician as corruptible or a politician with a clean record as incorruptible) was higher than 50% with a one-sample one-tailed $t$-test. The average individual-level accuracy was 57.35% ($SD = 5.93$), which was significantly higher than chance, $t(405) = 193.16, p < .001, d = 9.59$. A total of 356 participants (87.58%) identified more than 50% of politicians correctly. Thus, participants’ impressions were somewhat accurate.

Were participants aware of the accuracy of their judgments? On average, participants expected their judgments to be correct on 50.39% ($SD = 18.60$) of all trials, which was significantly below their actual accuracy, $t(405) = 7.54, p < .001, d = 0.37$. The correlation between individual accuracy levels and predicted accuracy levels was not significant and close to zero, $r(404) = .018, p = .72$ (see Figure 2). In sum, we found no evidence for meta-accuracy and on average, participants underestimated their accuracy.
Figure 2. Correlation between predicted and actual accuracy in corruptibility judgments. The diagonal line represents perfect meta-accuracy (i.e., predicted accuracy corresponds to actual accuracy). Data points to the left of the line represent participants who underestimated accuracy and data points to the right of the line represent participants who overestimated accuracy. The histograms in the margins show the distributions of predicted and actual accuracy.

Physiognomic belief. Our next set of analyses investigated the relationships between physiognomic beliefs, individual-level judgment accuracy, and confidence in judgment accuracy. We expected participants scoring higher on physiognomic beliefs to be more confident in the accuracy of their judgments. In fact, endorsement of physiognomic beliefs was positively correlated with predicted judgment accuracy, $r(404) = .285$, $p < .001$ (see Figure 3, left panel).

Next, we examined whether people who scored higher on physiognomic beliefs were actually more accurate in their judgments. There was a positive correlation between physiognomic beliefs and judgment accuracy, $r(404) = .167$, $p < .001$ (see Figure 3, right panel). Was increased confidence by people scoring high on physiognomic beliefs justified? If better judgment accuracy explains the positive relationship between physiognomic beliefs and confidence in accuracy, then we would expect that accuracy mediates the link between physiognomic beliefs and confidence. However, there was no significant indirect effect of physiognomic belief on confidence via accuracy, $\beta = -0.050$, $p = .49$, 95% CI [-0.452, 0.200].
We also regressed confidence and accuracy on physiognomic beliefs (in separate models) and compared the strength of the effects with a z-test. This showed that the effect of physiognomic beliefs on confidence ($\beta = 5.135$, $SE = 0.890$, $t(404) = 5.97$, $p < .001$) was significantly stronger than the effect of physiognomic beliefs on accuracy ($\beta = 0.992$, $SE = 0.292$, $t(404) = 3.40$, $p < .001$), $z = 4.62$, $p < .001$. Participants who scored one standard deviation higher on physiognomic beliefs expected to be 5.32 percentage points more accurate, but were only 0.99 percentage points more accurate. In sum, the increased confidence of people scoring high on physiognomic beliefs was not warranted given their actual judgment accuracy. The increase in confidence was disproportionately larger compared to the increase in actual accuracy.

Finally, we explored whether physiognomic beliefs were related to overconfidence in face-based impressions. We subtracted actual accuracy from predicted accuracy to create a variable indicating by how much participants overestimated their accuracy. Physiognomic beliefs were positively correlated with overestimation, $r(404) = .222$, $p < .001$. Thus, people scoring higher on physiognomic beliefs were more likely to be overconfident in their judgment accuracy.

**Discussion**

We found that people who believe in physiognomy were more accurate in judging the corruptibility of government officials based on facial photographs. Physiognomic beliefs were
also related to confidence in judgments: Participants who more strongly endorsed physiognomic beliefs expected their corruptibility judgments of politicians to be more accurate. Crucially, we found that the increase in confidence was disproportionately larger than the increase in judgment accuracy. In other words, physiognomic beliefs were related to overconfidence in trait impressions. Moreover, the correlation between accuracy and confidence was non-significant and close to zero. This shows that the increased confidence of people endorsing physiognomic beliefs is not normatively justifiable, as it is not based on their actual judgment accuracy.

**Study 5: Reliance on Judgments**

We proposed that physiognomic beliefs may help explain why people overrely on face-based personality judgments (Olivola et al., 2014). Results from Study 4 were in line with this view, showing that people who scored higher on physiognomic beliefs were too confident in the accuracy of their judgments. In Study 5, we extended these findings by examining whether people who score higher on physiognomic beliefs also rely more on face-based trait impressions in social decision-making. More specifically, we tested whether physiognomic beliefs relate to reliance on trait impressions when people could also rely on more valid information instead.

We tested this hypothesis in the context of the trust game (Berg, Dickhaut, & McCabe, 1995; Snijders & Keren, 1999). This interaction between two players, a *trustor* and a *trustee*, reflects the essential structure of trust-based social exchange. Participants played a series of trust games and, in each round, they saw a photo of their interaction partners. This allowed us to measure how much participants relied on the perceived facial trustworthiness of their interaction partner when making trust decisions. In real life, people can often rely on a variety of cues when making decisions and the persistent influence of facial trustworthiness is particularly surprising in cases where people could rely on more valid cues instead (Olivola et al., 2018; Rezlescu, Duchaine, Olivola, & Chater, 2012). We therefore varied a second cue that participants could rely on: the trustee’s temptation to choose betrayal, which actually predicts the likelihood of reciprocation (Evans & Krueger, 2014).

We expected that individual differences in physiognomic beliefs would be related to reliance on facial trustworthiness, but not to reliance on temptation. That is, we predicted that people who believe in physiognomy would rely more on their face-based trustworthiness judgments, but not necessarily more on other cues unrelated to face-based judgments.
Methods

Participants. We recruited 243 first-year psychology students from a Dutch university who completed the study in return for partial course credit. Due to a technical error, some face images were not displayed for 19 participants and we therefore excluded their data from analysis, leaving a final sample of 224 participants ($M_{age} = 20.45$, $SD_{age} = 2.41$; 75.45% female, 24.11% male, 0.45% other).

The majority of participants were Dutch (43.75%) or German (30.36%). Sample size was determined by how many students participated in the study within two weeks. A sensitivity analysis in G*Power (Faul et al., 2007) showed that this sample size afforded us 80% power to detect a small correlation ($r = .186$) between physiognomic beliefs and reliance on facial trustworthiness (with $\alpha = 5\%$).

Materials and procedure. The experiment was administered online. Participants first learned about and then played a series of 24 trust games in the role of the trustor. On each trial, participants saw a photo of their supposed interaction partner next to the decision tree. We selected 12 photos of Caucasian Dutch adults (six females and six males) with a forward gaze from the Radboud Faces Database (Langner et al., 2010). To introduce variance in perceived trustworthiness, half of the selected faces displayed a neutral expression and half a happy facial expression (Evans & van de Calseyde, 2017; Krumhuber et al., 2007).

We also varied the trustee’s temptation, the economic incentive to choose betrayal. We defined temptation as the difference between the trustee’s gain in case of betrayal ($T$) and reciprocation ($R_2$) divided by the value of betrayal: $(T - R_2)/T$ (Evans & Krueger, 2014). Each photo was shown twice, once when temptation was low (0.33) and once when temptation was high (0.60). These values correspond to a 50% (low temptation) and 150% (high temptation) increase in payoffs for the trustee in case betrayal is chosen over reciprocation.

After indicating their trust decisions, participants were shown each face again and asked to rate how trustworthy they think the person in the photo is on a scale from 1 (not at all trustworthy) to 9 (extremely trustworthy). We used the average trustworthiness rating of each face across all participants as our measure of facial trustworthiness. Prior to analysis, we rescaled the cue variables (i.e., facial trustworthiness and temptation) to range from -0.5 to 0.5. Thus, for

---

8 Participants were drawn from the same subject pool as participants in Study 1. We could not check for potential overlap between samples as no identifying information was collected in either study.
the two cues, a one-unit increase denotes a change from the lowest average trustworthiness rating to the highest and a change from low to high temptation. Finally, participants completed the physiognomic beliefs scale ($\alpha = .76$). All participants completed the study in English.

**Results**

The average score on general physiognomic belief was below the midpoint of our scale ($M = 3.97$, $SD = 1.18$), $t(223) = 0.32$, $p = .75$, $d = 0.02$ (see Table 1). Around half of all participants (48.66%) believed at least somewhat in physiognomy (i.e., they scored above the midpoint of the scale), $\chi^2(1) = 0.11$, $p = .74$. Average trustworthiness ratings of the faces ranged, from 4.04 to 6.61 on our 9-point scale ($M = 5.21$, $SD = 0.88$). We computed intraclass correlation coefficients (ICCs) to estimate consensus in ratings across participants (Shrout & Fleiss, 1979). Participants showed significant consensus in their ratings, $ICC(2, 1) = .261$, $p < .001$, 95% CI [.132, .597]. The average trust rate across all trials was 46.09% and participants took on average 5.43 seconds ($SD = 3.50$) to make a decision. Eight participants (3.57%) never trusted whereas six participants (2.68%) always trusted.

**Reliance on facial trustworthiness and temptation.** First, we checked whether participants relied on facial trustworthiness and temptation when making trust decisions. We estimated a multilevel regression model with random intercepts and slopes per participant in which we regressed trust decisions (coded 0 for no trust and 1 for trust) on facial trustworthiness and temptation. This revealed a positive effect of facial trustworthiness, $\beta = 1.608$, $SE = 0.158$, $z = 10.17$, $p < .001$, $OR = 4.99$. The individual with the lowest trustworthiness rating was trusted 29.96% of the time while the individual with the highest trustworthiness rating was trusted 62.11% of the time. There was also a negative effect of temptation, $\beta = -0.463$, $SE = 0.090$, $z = 5.12$, $p < .001$, $OR = 0.63$. Participants trusted 50.55% of the time when temptation was low and 42.07% of the time when temptation was high. Thus, participants relied on both cues when making trust decisions.

**Physiognomic beliefs and cue reliance.** Next, we tested our main hypothesis that physiognomic beliefs are correlated with reliance on facial trustworthiness. We extracted the participant-specific effects of temptation and facial trustworthiness (i.e., the random slopes) from our multilevel regression models as an indicator of how much each participant relied on the two cues. Both cue reliance variables were $z$-standardized. Results showed a positive correlation between physiognomic beliefs and reliance on facial trustworthiness, $r(222) = .142$, $p = .034$, but
no significant correlation with reliance on temptation, $r(222) = .014$, $p = .83$. Exploratory analyses (see Supplemental Materials) showed that there were also no significant correlations with reliance on other facial features (i.e., facial expression, gender, attractiveness). Moreover, regressing physiognomic beliefs on all cue reliance variables showed only a significant effect of reliance on facial trustworthiness. Thus, participants who scored higher on physiognomic beliefs relied more on facial trustworthiness, but not more on other cues, when making trust decisions.

Discussion

In line with previous studies, we found that participants relied on the facial appearance of their interaction partners when deciding whom to trust (Jaeger et al., 2019; van’t Wout & Sanfey, 2008). More importantly though, we found that how much participants relied on trustworthiness impressions from face was related to their endorsement of physiognomic beliefs. Specifically, belief that trustworthiness is reflected in facial features predicted to what extent trust decisions were driven by the facial appearance of interaction partners. Physiognomic beliefs were only correlated with reliance on the perceived trustworthiness of interaction partners and not with reliance on other cues such as the interaction partner’s facial expression, attractiveness, gender, or their temptation (i.e., their incentive to betray trust).

General Discussion

The goal of the current investigation was to provide insights into the belief that personality is reflected in facial appearance (i.e., physiognomic beliefs). The accuracy of face-based personality inferences has been studied since the time of Ancient Greece (Aristotle, trans. 1936) and remains a subject of contentious debate (Bonnefon, Hopfensitz, & De Neys, 2015; Todorov, Funk, & Olivola, 2015). Here, we argue that irrespective of their actual accuracy, people hold beliefs about the diagnosticity of facial features for inferring personality. Crucially, individual differences in the endorsement of physiognomic beliefs may help explain why and under what conditions people are confident in their personality impressions, or rely on them to make decisions. To test these hypotheses, we introduced a scale to measure physiognomic beliefs and examined their prevalence, structure, and correlates.

First, our results suggest that belief in physiognomy is relatively widespread. Across all studies, around half of all participants at least somewhat endorsed the belief that personality is reflected in facial features. Physiognomic beliefs were prevalent among psychology students at a Dutch university (Studies 1, 4, and 5), in the general Dutch population (Study 2), and among
participants from the United Kingdom recruited through a crowdsourcing platform (Study 3). A
closer analysis of our representative sample of the Dutch population showed that physiognomic
beliefs were slightly more prevalent among women and younger participants, but these
differences were small. We found no significant differences across different levels of education
or income. Together, these results suggest that belief in physiognomy is common across different
demographic groups.

Second, people hold heterogeneous beliefs about the manifestation of different traits in
faces. To map physiognomic beliefs across a variety of characteristics, we measured beliefs in
three fundamental dimensions underlying person perception: sociability, morality, and
competence (Brambilla et al., 2011). Participants believed that sociability is more reflected in
faces than morality or competence and this pattern replicated in all five studies (see Figure 1).
Differences in physiognomic beliefs for morality and competence were small and inconsistent
across studies.

Third, physiognomic beliefs are related to an intuitive thinking style. Personality
impressions from faces are formed spontaneously, quickly, and effortlessly (Klapper, Dotsch, &
Wigboldus, 2017; Stewart et al., 2012; Willis & Todorov, 2006). We therefore hypothesized that
people who tend to trust their intuitions would be more likely to endorse physiognomic beliefs.
Results from Study 3 supported this prediction. Physiognomic beliefs were also correlated with
beliefs in the biological determinism of personality traits (replicating results of Suzuki et al.,
2017), but this relationship was less pronounced and non-significant when controlling for faith in
intuition. These results suggest that physiognomic beliefs may be rooted in the quick and
efficient processing of faces which makes trait inferences from faces intuitively accessible.

Fourth, physiognomic beliefs shape how impressions from faces are formed and
subsequently used in decision-making. People who more strongly endorsed physiognomic beliefs
were more confident in the accuracy of their face-based personality judgments (Study 4) and
relied more on face-based personality judgments when deciding whom to trust (Study 5). In fact,
trustworthiness-specific physiognomic beliefs were correlated with reliance on an interaction
partner’s perceived trustworthiness, but not with reliance on other facial or non-facial cues (e.g.,
the interaction partner’s attractiveness or their economic incentive to betray trust).

Finally, the increased weighing of personality inferences by people scoring higher on
physiognomic beliefs is not due to superior judgment accuracy. We asked whether people who
endorse physiognomic beliefs were justifiably more confident because of their superior judgment accuracy (Study 4). To this end, we replicated a recent study by Lin and colleagues (2017) that demonstrated accuracy in corruptibility judgments of politicians based on face images. We found that corruptibility judgments by people scoring higher on physiognomic beliefs were indeed slightly more accurate, suggesting that individual differences in physiognomic beliefs may reflect superior accuracy in inferring personality from faces. However, mediation analyses showed that judgment accuracy did not account for the link between physiognomic beliefs and confidence. The absolute increase in confidence was also disproportionately larger than the increase in accuracy: A one standard deviation increase in physiognomic beliefs was related to a one percentage point increase in accuracy, but a five percentage point increase in predicted accuracy. In other words, the observed advantage in accuracy did not justify the increase in confidence and, on average, people who endorsed physiognomic beliefs were more likely to overestimate the accuracy of their judgments.

**Lay Personality Theory**

People hold beliefs about the basis (Haslam et al., 2004), malleability (Chiu et al., 1997), structure (Stolier, Hehman, Keller, et al., 2018), and expression (Mehl et al., 2006) of personality traits. We add to this work by showing that people also hold beliefs about the manifestation of personality traits in facial appearance. While endorsement of this belief varied across individuals, there was considerable consistency in belief structure across different personality dimensions. Across all five studies, physiognomic beliefs were strongest for sociability compared to morality and competence. People vary in their absolute belief in physiognomy, but beliefs about the relative expression of different personality dimensions in faces is, to a large extent, shared.

Sociability and morality are often subsumed under the label of warmth (Fiske, Cuddy, Glick, & Xu, 2002) or communion (Abele & Wojciszke, 2007). However, judgments of sociability and morality show several important differences and researchers have argued that they should be treated as separate dimensions of person evaluation (Brambilla et al., 2011; Goodwin, 2015; Landy, Piazza, & Goodwin, 2016). For instance, morality information more strongly determines the formation (Goodwin, Piazza, & Rozin, 2014; Leach et al., 2007) and updating (Brambilla, Carraro, Castelli, & Sacchi, 2019) of impressions. We showed that sociability and morality also displayed divergent patterns in the domain of lay personality theory.
Social Perception

Dominant theories on impression formation from faces have mostly focused on how different facial features elicit trait inferences (Oosterhof & Todorov, 2008; Zebrowitz, 2017), which has produced a long list of cues that people use to infer personality from faces (e.g., Jaeger, Wagemans, Evans, & van Beest, 2018; Said, Sebe, & Todorov, 2009; Sutherland, Young, & Rhodes, 2016). This approach reflects the view that social perception is mostly a reflexive, stimulus-driven processes in which the presence of certain facial cues automatically trigger personality inferences (Engell, Haxby, & Todorov, 2007; van’t Wout & Sanfey, 2008; Winston, Strange, O’Doherty, & Dolan, 2002). However, recent studies have highlighted that there are many top-down processes that influence social perception (Brambilla et al., 2018; Freeman & Johnson, 2016). For example, beliefs about the extent to which personality traits correlate intrapersonally (e.g., whether trustworthy people tend to be sociable) influence the correlation between personality trait impressions from faces (e.g., the overlap in impression of trustworthiness and sociability; Stolier, Hehman, & Freeman, 2018; Stolier, Hehman, Keller, et al., 2018). In a similar vein, our results showed that beliefs about the manifestation of personality traits in faces influences confidence in and reliance on trait impressions. Thus, the processing of personality trait information from faces is moderated by individual differences in lay personality theory.

Despite their poor predictive validity, trait inferences from faces influence a wide range of consequential decisions, such as criminal sentencing, voting, and personnel selection (Olivola et al., 2014) and people still rely on facial appearance when better information is available (Jaeger et al., 2019; Olivola & Todorov, 2010b; Rezlescu et al., 2012). The influence of facial stereotypes can lead to worse outcomes for decision-makers, and to systematic discrimination against people of a certain appearance. Although the effects of this bias are well-documented, little is known about the mechanisms underlying it. Yet, addressing the bias—for example, by implementing interventions that reduce reliance on facial appearance—requires knowledge about its cognitive underpinnings (Forscher et al., 2019; T. D. Wilson & Brekke, 1994). The current results suggest that widespread influence of facial stereotypes may be explained by lay beliefs in the diagnosticity of facial appearance as an indicator of personality. As a consequence, people are overconfident in the accuracy of their face-based trait inferences and rely on them too much when making decisions. Future studies could test whether changing physiognomic beliefs by
educating people about the low predictive validity of their impressions reduces reliance on facial appearance.

**Limitations and Future Directions**

Our findings point to a widespread belief in the core tenet of physiognomy (Aristotle, 1936; Lavater, 1775) that personality traits are reflected in facial morphology. But did participants actually envision a resting, non-expressive face or is their belief rooted in the perceived informational value of dynamic features of faces, such as emotion expressions? This distinction may ultimately be inconsequential because people readily perceive emotion expressions in resting faces (Adams, Nelson, Soto, Hess, & Kleck, 2012; Said et al., 2009). We tried to ensure that participants imagine a resting face by prompting them to imagine seeing the passport photo of a stranger. In both the Netherlands and the United Kingdom, the countries of origin of most participants, people are required to maintain a neutral expression in passport photos. Future studies could investigate the role of emotional expressions, for example, by comparing confidence in personality judgments of neutral and expressive faces.

Future research could also explore how physiognomic beliefs vary across different social groups. The majority of our studies relied on student samples from a Dutch university, which constraints the generalizability of our findings. We did find that belief in physiognomy was common in a sample of British participants and in a representative sample of the Dutch population. Moreover, Study 2 showed that belief in physiognomy varied little across different demographic indicators (i.e., gender, age, education, and income). Nonetheless, research with participants from more diverse cultural backgrounds is needed to investigate who believes in physiognomy. In addition, future studies could leverage the rich and openly accessible data of the LISS panel—from which we recruited a representative sample of Dutch participants for Study 2—to map how different psychological or socio-economic variables relate to physiognomic beliefs.

Another question that remains unanswered is whether individual differences in physiognomic beliefs predict which cues people rely on to form impressions. For example, trustworthiness impressions are based on a wide variety of interrelated facial cues, such as width-to-height ratio (Stirrat & Perrett, 2010), resemblance to emotion expressions (Saïd et al., 2009), and sexual dimorphism (Gladstone & O’Connor, 2014). Laypeople seem to agree that trustworthiness is reflected in facial appearance, but do they use the same cues to infer
trustworthiness? If physiognomic beliefs are rooted in a greater ability to infer personality from faces—and Study 4 provided some support for this—then people who believe in physiognomy should not only rely on similar cues, but they should rely on the cues that are actually valid indicators of the trait in question. These questions could be addressed in a lens model framework (Brunswik, 1956), by testing whether physiognomic beliefs are related to cue utilization (i.e., how much people rely on different cues) and cue validities (i.e., how well these cues predict the criterion being judged).

**Conclusion**

We showed that people hold lay beliefs about the manifestation of personality traits in facial appearance (i.e., physiognomic beliefs). While people differ in their absolute endorsement of physiognomy, beliefs about how much different personality dimensions are reflected in facial features are largely shared. We also find that individual differences in physiognomic beliefs are related to various aspects of social perception. People who score high on physiognomic beliefs are more confident in the accuracy of their personality judgments (but this cannot be explained by their superior judgment accuracy) and they rely more on their personality judgments when making decisions. In sum, our results show that physiognomic beliefs are widespread and associated with a range of social-cognitive processes and behaviors.
References


