Changing the Personality of a Face: Perceived Big Two and Big Five Personality Factors Modeled in Real Photographs

Mirella Walker and Thomas Vetter
University of Basel

General, spontaneous evaluations of strangers based on their faces have been shown to reflect judgments of these persons' intention and ability to harm. These evaluations can be mapped onto a 2D space defined by the dimensions trustworthiness (intention) and dominance (ability). Here we go beyond general evaluations and focus on more specific personality judgments derived from the Big Two and Big Five personality concepts. In particular, we investigate whether Big Two/Big Five personality judgments can be mapped onto the 2D space defined by the dimensions trustworthiness and dominance. Results indicate that judgments of the Big Two personality dimensions almost perfectly map onto the 2D space. In contrast, at least 3 of the Big Five dimensions (i.e., neuroticism, extraversion, and conscientiousness) go beyond the 2D space, indicating that additional dimensions are necessary to describe more specific face-based personality judgments accurately. Building on this evidence, we model the Big Two/Big Five personality dimensions in real facial photographs. Results from 2 validation studies show that the Big Two/Big Five are perceived reliably across different samples of faces and participants. Moreover, results reveal that participants differentiate reliably between the different Big Two/Big Five dimensions. Importantly, this high level of agreement and differentiation in personality judgments from faces likely creates a subjective reality which may have serious consequences for those being perceived—notably, these consequences ensue because the subjective reality is socially shared, irrespective of the judgments' validity. The methodological approach introduced here might prove useful in various psychological disciplines.

Keywords: person perception, faces, statistical modeling, Big Five, Big Two

Individuals are often confronted with situations in which they only have very little information about the persons they have to interact with. To handle such situations, individuals have been shown to spontaneously form first impressions in an extremely fast manner (Ballew & Todorov, 2007; Bar, Neta, & Linz, 2006; Rule & Ambady, 2008). Typically, facial appearance is the most prominent source of information in such moments and thus contributes substantially to spontaneous personality judgments (e.g., Willis & Todorov, 2006). Previous work has shown that general, spontaneous face-based person evaluations can be mapped onto a two dimensional (2D) space defined by the two basic dimensions trustworthiness and dominance, signaling a person's intention and ability to cause harm (Oosterhof & Todorov, 2008).

Interestingly, once it has been decided whether a person might be willing or able to do harm, more specific personality traits might become relevant, depending on the situational context. For instance, when boarding a train for a long journey, we might have a preference to sit next to a fellow passenger who is introverted and agreeable if we want to read a book, or we might have a preference for somebody who is open to experience and extraverted if we want to have an entertaining conversation. Do individuals form judgments of extraversion, agreeableness, and openness from facial information? And do such judgments fit the 2D space of face evaluation or do individuals have more differentiated beliefs of what extraverted, agreeable, or open persons look like? Against this background, we investigate whether individuals reliably judge strangers based on their faces and whether they discriminate between different Big Two and Big Five personality dimensions. Moreover, we investigate whether these judgments go beyond the 2D space of face evaluation.

Let's assume further that the person in the above mentioned example does not stop at forming spontaneous judgments of extraversion, agreeableness, and openness based on the faces of his or her fellow passengers, but that he or she is likely to act upon these judgments. If he or she longs for privacy to read a book without being disturbed, he or she might sit next to the person signaling the most introverted and agreeable personality and avoid people signaling an open and extraverted personality. If other people on this train share their perceptions of the personalities of their fellow passengers, the introverted looking person might be more likely to be joined by readers than by talkers. Thus, one could expect that a social reality would be formed, in which passengers are treated consistently based on socially shared perceptions. Such consensual personality judgments impact the target persons' self-perception or behavior via, for example, self-fulfilling prophecies (Chen & Bargh, 1997; Darley &

Fazio, 1980). Notably, this holds true irrespective of the accuracy of such personality judgments—what is important in this example is a shared social perception. Against this background, we investigate whether individuals agree in forming Big Two and Big Five personality judgments based on faces.

To accomplish these goals, we model personality in faces. Prior research in the realm of face perception has often relied on computer-generated three dimensional (3D) heads (henceforth referred to as *3D heads*) and allowed for critical insights into underlying mechanisms (e.g., Oosterhof & Todorov, 2008; Todorov & Oosterhof, 2011). Here we take the next step and use 3D scans of real faces (henceforth referred to as *3D scans*) to statistically model perceived personality dimensions. These statistical models are then applied to manipulate perceived personality dimensions in real photographs of faces (henceforth referred to as *photographs*). This approach offers a series of advantages as is detailed later, and will allow for testing research questions in other domains that wish to employ highly controlled, natural-looking, ecologically valid stimuli.

In what follows, we first review literature on how *general, spontaneous* personality judgments based on faces are made and how they can be structured. Second, we discuss research showing that if a specific situational context is given, *context-specific* personality judgments are made based on faces. Third, we discuss preconditions and consequences of a high agreement in face-based personality judgments among different judges. Fourth, we review zero acquaintance literature providing mixed results regarding the accuracy of Big Five judgments based on faces. Fifth, we introduce the Basel Face Model (Paysan, Knothe, Amberg, Romdhani, & Vetter, 2009; http://faces.cs.unibas.ch/bfm/), which provides the technical framework in which the Big Two and the Big Five personality dimensions will be modeled.

General, Spontaneous Personality Judgments Based on Faces

A wealth of evidence indicates that individuals make spontaneous personality judgments based on faces (e.g., Bruce & Young, 1986; Willis & Todorov, 2006). Oosterhof and Todorov (2008) were the first to investigate the structure underlying such spontaneous impressions. In a first step, their participants were presented with emotionally neutral faces and asked to describe the respective persons spontaneously. Next, these descriptions were classified into 15 broad categories (e.g., attractive, unhappy, or sociable). In a subsequent study, new participants were asked to rate novel neutral faces regarding these 15 categories. Because two of the trait dimensions (i.e., egoistic and boring) did not reach acceptable interrater agreements, they were excluded from the analyses. A Principal Component Analysis (PCA) revealed that the 13 remaining personality dimensions could be represented by two basic factors, which are best described as valence/trustworthiness and power/dominance. This 2D model of face evaluation accounts for more than 80% of variance in the 13 different judgments. In a next step, trustworthiness and dominance were successfully modeled in 3D heads. This procedure revealed the facial information individuals use to make trustworthiness and dominance judgments (i.e., the physical correlates of these judgments). Faces with an extremely high salience of trustworthiness look happy, whereas faces with an extremely high salience of dominance look physically strong.

Interestingly, on the *semantic-conceptual* level, the 2D model of Oosterhof and Todorov (2008), which was built on face-based person evaluation, fits other 2D models of social perception, such as the stereotype content model with the two dimensions of warmth and competence to describe social groups (Fiske, Cuddy, & Glick, 2007; Fiske, Cuddy, Glick, & Xu, 2002), the model of interpersonal perception with the two dimensions love and dominance (Wiggins, 1979; Wiggins, Phillips, & Trapnell, 1989) or morality and competence (Wojciszke, 1994; Wojciszke, Bazinska, & Jaworski, 1998), and the Big Two personality concept with the dimensions communion and agency (Abele, Uchronski, Suitner, & Wojciszke, 2008; Abele & Wojciszke, 2007; Wiggins, 1991). Though these models differ in several critical aspects, one clear commonality emerges from their joint examination: There are two dimensions, namely valence (or morality/warmth/love/communion) and dominance (or competence/agency), which individuals rely on when referring to individuals, social groups, or to themselves. Focusing on this commonality, we selected the Big Two personality concept as one example for the empirical part of this contribution.

Context-Specific Judgments Based on Faces

Research has shown that if a specific situational context affords it, very context-specific judgments are made. Although competence does not seem relevant in evaluating whether a person might be willing and able to do harm, people use competence-cues from the faces of two opponents to judge who will win an election (Willis & Todorov, 2006). Similarly, a study investigating real sentencing decisions as a function of the degree of Afrocentric facial qualities found that more pronounced Afrocentric facial cues result in harsher sentencing (Blair, Judd, & Chapleau, 2004). These findings suggest that raters make very context-specific judgments based on faces and they agree in doing so, even if these judgments do not fit the 2D space of face evaluation. To date, such evidence is limited to isolated social judgments, such as competence or guilt.

Here we suggest that individuals also spontaneously form personality judgments such as those established in the comprehensive Big Two and Big Five personality concepts. Investigating personality judgments with respect to the Big Two personality dimensions is important because the Big Two dimensions communion and agency are semantically similar to the two basic dimensions of face evaluation, trustworthiness and dominance (Abele, Uchronski, Suitner, & Wojciszke, 2008; Abele & Wojciszke, 2007; Wiggins, 1991).

Investigating personality judgments with respect to the Big Five personality dimensions is important because the Big Five personality dimensions neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness (e.g., John & Srivastava, 1999; McCrae & Costa, 1997) affect outcomes in various contexts, such as academic performance (e.g., Poropat, 2009), parenting (Prinzie, Stams, Deković, Reijntjes, & Belsky, 2009), consumer behavior (e.g., Mulyanegara, Tsarenko, & Anderson, 2007), criminal investigations (e.g., Ono, Sachau, Deal, Englert, & Taylor, 2011), or mating (e.g., Gebauer, Leary, & Neberich, 2012). The Big Five also play an important role in different applied settings. Human resource managers, for example, use these concepts to determine whether a certain applicant is suited to the affordances of a specific job (Costa, 1996). Likewise, online dating platforms use the Big Five to determine whether two members might be meant for each other (Solomon, Russell-Bennett, &

Previte, 2012). In sum, the Big Five personality concept is very important both in theory and in practice.

Interestingly, there is some debate as to whether the Big Five personality dimensions describe personality factors at the highest level of abstraction or whether they are reducible to higher order personality factors. Some argue that the Big Five cannot be reduced to more abstract concepts that still have a descriptive value (other than mere valence) and therefore constitute the highest level of the hierarchy (Goldberg, 1993; McCrae & Costa, 1996). However, it has repeatedly been shown that the Big Five personality factors are not independent from each other (Block, 1995; Clark & Harrison, 2001; Egan, Deary, & Austin, 2000). Digman (1997) found clear and robust evidence for two higher-order factors among the Big Five dimensions, which he neutrally labeled alpha and beta. He thus noted that these two factors could be interpreted in terms of Wiggins' concepts of agency and communion (Wiggins, 1991). Against this background, the first aim of our study is to investigate whether more than two dimensions are needed to explain Big Five personality judgments based on faces.

Agreement in Personality Judgments From Faces

Oosterhof and Todorov (2008) found that individuals exhibit strong agreement in forming personality judgments that match the 2D space of face evaluation from photographs of faces. Todorov and Oosterhof (2011) later showed that individuals exhibit strong agreement in judgments of such personality traits even when shown 3D heads that lack extrafacial features, such as hairstyle. This difference in the stimulus material seems important, because individuals can choose their own hairstyle, jewelry, or the glasses they wear and thereby express aspects of their personality. However, they are very limited when it comes to changing their facial features.

What are the preconditions that allow different individuals to perceive the same personality from a face? First, they have to use the same facial cues to infer personality judgments. Second, they need to have shared beliefs regarding the meaning of these facial cues. It is likely that such shared beliefs are learned during ontogenesis. Oosterhof and Todorov (2008) argue that spontaneous face based personality judgments are the result of an overgeneralization "of adaptive mechanisms for inferring harmful intentions and the ability to cause harm" (p. 11087).

Whereas there is evidence that individuals exhibit strong agreement in personality judgments that match the 2D space of face evaluation based on faces lacking extrafacial features, less is known regarding judgments of the Big Five personality dimensions from faces. We are only aware of one study that found interrater agreement in Big Five judgments from photographs of faces (Penton-Voak, Pound, Little, & Perrett, 2006). Because photographs show extrafacial information, we do not know whether agreement is based upon facial or extrafacial styling information. This difference is not only theoretically, but also practically relevant. There are situations that limit an individual's styling choices, because there is a dress code, for example. And these can be very consequential situations, such as job interviews.

Highly consistent judgments based on facial features alone might have various consequences for the person being judged: On the highest level of abstraction, the Big Five personality dimensions differ in valence. Some individuals might be perceived as having a more favorable personality and thus they might be approached more positively than others, which would likely impact their self-perception. In the long run, this might also result in those individuals indeed being more friendly and sociable than others, because they experience more positive social interactions (i.e., self-fulfilling prophecies; Chen & Bargh, 1997; Darley & Fazio, 1980). On a more concrete level, there are contexts in which certain personality dimensions are regarded as more favorable than others. For example, in academic contexts, agreeableness, conscientiousness, and openness are important dimensions (Poropat, 2009). A young academic with a look that doesn't signal agreeableness, conscientiousness, and openness might be discriminated against due to his or her facial appearance. Against this background we investigate whether Big Two and Big Five personality judgments derived from faces lacking any extrafacial information show high reliability across samples of different faces and participants.

Accuracy of Judgments at Zero Acquaintance

The previous section addressed whether different individuals infer personality judgments from faces reliably (henceforth referred to as *agreement*). Another question to ask is whether such judgments have predictive validity (henceforth referred to as *accuracy*). Zero acquaintance research investigates to what degree such judgments based on thin slices of the behavioral stream (Ambady, Bernieri, & Richeson, 2000) predict different outcome variables (e.g., school performance) and what cues individuals use to form such judgments (Zebrowitz & Collins, 1997). Often, zero acquaintance research relies on short video sequences. Compared with these, faces are poor stimuli because they do not convey behavioral information that could be used to infer personality. Hence, findings regarding the accuracy of personality judgments based on faces are mixed: Some findings suggest that participants are fairly accurate in some personality judgments from faces (Borkenau Brecke, Möttig, & Paelecke, 2009; Kramer & Ward, 2010; Little & Perrett, 2007; Penton-Voak et al., 2006), whereas others suggest that they are not (Pound, Penton-Voak, & Brown, 2007; Shevlin, Walker, Davies, Banyard, & Lewis, 2003). Interestingly, extraversion seems to be the personality dimension that is most accurately evaluated based on facial appearance. One reason for this might be that extraversion is a genuinely interpersonal dimension, as is agreeableness, whereas neuroticism, conscientiousness, and openness to experience are not (McCrae & Costa, 1989). In first encounters it might be more relevant whether a person is extraverted and agreeable than whether he or she is open to experience, neurotic, and/or conscientious. Imagine waiting for the bus late at night when a stranger is approaching you. In this situation it is important to know whether the person approaching you is aggressive (an instance of low agreeableness) or friendly (an instance of high extraversion). It is less important to know whether this person likes to go to art museums or tends to worry about the little things in life. Therefore, individuals might be better prepared to detect cues for extraversion and agreeableness than for openness, neuroticism, and conscientiousness. Borkenau, Brecke, Möttig, and Paelecke (2009) argue that individuals use smiling as a cue to judge extraversion, neuroticism, and agreeableness. However, because smiling is only a valid cue for extraversion, individuals might be more accurate in forming extraversion judgments than judgments of agreeableness and neuroticism.

Taken together, zero acquaintance literature shows that individuals are at most *moderately accurate* in judging others' personalities from their faces. However, this does not mean that such judgments are inconsequential. The fact that there is a high agreement between individuals and that individuals act upon their perceptions likely creates a highly consequential socially shared reality. Understanding the formation of personality judgments based on faces is therefore critical, irrespective of whether these judgments are accurate.

Visualizing Personality in Faces

Different attempts at investigating accuracy of personality judgments from facial photographs (Kramer & Ward, 2010; Little & Perrett, 2007; Penton-Voak et al., 2006) employed morphing techniques to create composite faces from the faces of participants that scored lowest and highest on the Big Five personality dimensions, thereby visualizing what a highly extraverted person typically looks like, for example. While this is an interesting approach that allows to test to what extent people's idea of what an extravert looks like overlaps with the reality of what an extravert actually looks like, this approach does not allow to model people's idea or stereotype of an extraverted person. This is a highly important distinction because when individuals judge someone's personality from appearance they rely on their ideas and stereotypes of what an extraverted person, for example, should look like and not on whether this person actually perceives themselves to be extraverted or not. The morphing approach does not only differ from the modeling approach applied here in that it visualizes actual personality dimensions instead of the ideas or stereotypes of personality dimensions, but also in it's applicability to generate face stimuli. On the one hand it does not allow for systematic and gradual manipulations of the salience of a specific personality dimension, on the other hand, it cannot be applied to novel faces.

More recent approaches for visualizing personality in faces are based on the idea that if raters agree in their personality judgments based on faces, then the facial information raters use to make these personality judgments can be statistically modeled in novel faces (e.g., Oosterhof & Todorov, 2008; Walker & Vetter, 2009). These approaches to modeling and visualizing personality in faces (Dotsch & Todorov, 2012; Mangini & Biederman, 2004; Oosterhof & Todorov, 2008; Todorov & Oosterhof, 2011; Todorov, Said, & Verosky, 2011; Walker, Jiang, Vetter, & Sczesny, 2011; Walker & Vetter, 2009) were conducted in a bottom-up way by starting with personality traits individuals spontaneously ascribe based on faces. Here we take a top-down approach by starting with the theoretically relevant and well-established Big Two and Big Five personality concepts and systematically modeling and visualizing the facial characteristics individuals use to make such judgments.

Visualizing the physical correlates of a specific psychological dimension in a novel face photograph with natural-looking results requires two different tasks (Walker et al., 2011; Walker & Vetter, 2009): The first task is to model the dimension in the statistical face space and the second is to apply this model to a real photograph. Both steps are further detailed in what follows.

Modeling Personality Dimensions

The Face Space Framework (Valentine, 1991) proposes that an individual face is mentally represented as a point in a highly dimensional space, the dimensions of which correspond to the properties that are used to encode and discriminate between faces. The distance between any two points in this space represents the perceived similarity between the corresponding faces. An early technical realization of this abstract psychological face space framework is the morphable face model (Blanz & Vetter, 1999). Within this statistical face model, the distance between any two faces is determined by their objective physical similarity. Here we use a novel morphable face model, namely the Basel Face Model, a face space derived from 200 3D scans of real faces (Paysan, Knothe, Amberg, Romdhani, & Vetter, 2009; Vetter, T. (n.d.), http://faces.cs.unibas.ch/bfm/). Each of these face scans is represented by 53490 3D vertices coding for face shape and the respective 53490 vertex colors. By performing two principal component analyses (PCAs) based on the mean adjusted shape and color representations of these scans separately, we generated a 199 dimensional space coding for face shape and a 199 dimensional space coding for color information. The 199 dimensions within each of these two face spaces are independent from each other and code for the information on which the 200 face scans maximally vary (i.e., eigenvectors). Figure 1 visualizes the *first two* shape and color components of the two face spaces.

In a reverse engineering approach we determine the direction in our face space that best represents a specific personality dimension. To this end, every face coded within these two face spaces (S_i, C_i) needs to be judged with respect to that specific personality dimension (μ_i). Then, weighted sums are computed separately for shape and color by adding up the products of judgments with the mean-free face representations:

$$\Delta S = \sum_{i=1}^{m} \mu_i(S_i - \overline{S}), \qquad \Delta C = \sum_{i=1}^{m} \mu_i(C_i - \overline{C}) \qquad (1)$$

Any of the 200 face scans can now be manipulated with respect to this personality dimension by adding or subtracting multiples of ΔS, ΔC (Blanz & Vetter, 1999; Walker & Vetter, 2009).

Applying the Models to Novel Photographs of Faces

In the previous paragraph, we introduced the procedure to model personality dimensions in the 200 3D scans the model is built upon. Importantly, our approach further allows us to model these personality dimensions in real facial photographs. Compared with 3D scans,
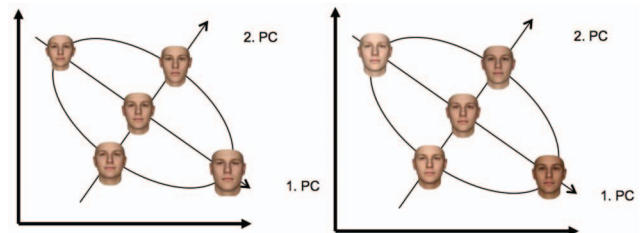

Figure 1. Visualization of the first two shape (left) and color Principal Components (PCs; right). The first shape PC codes for face size and seems to be related to gender information, the second shape PC codes for fullness of the face. The first color PC codes for brightness versus darkness of the face, whereas the second color PC codes for contrast versus evenness of the face. See the online article for the color version of this figure.

photographs contain extrafacial information, such as hairstyle or clothing. Thus, they look more natural and can serve as more ecologically valid stimulus material than 3D scans. This works as follows: First, the face on the 2D photograph has to be represented as a 3D scan by using the analysis-by-synthesis-approach (Blanz & Vetter, 1999). In particular, we reconstruct the 3D shape and surface color of the individual's face by fitting the Basel Face Model to the face on the photograph. This is done by optimizing the model coefficients along with a set of rendering parameters (e.g., coding for pose or lighting) in order to produce a 3D reconstruction of the face on the input photograph. This procedure results in a 3D representation of the shape structure and a representation of the vertex colors. This reconstructed 3D head can be manipulated by adding or subtracting multiples of ΔS, ΔC as described in the previous section. Finally, the resulting head is rendered back into the original photograph using the rendering parameters estimated in the fitting process (Blanz, Scherbaum, Vetter, & Seidel, 2004). See Figure 2 for a visualization of this procedure and Walker, M. (2015) http://mirellawalker.com/face-modeling/ for videos showing the changes in faces corresponding to the perceptions of the different Big Five and Big Two dimensions.

The Present Research

Aims

The present article pursues theoretical, methodological, and practical aims. The theoretical aims are threefold. First, we aim at investigating how the Big Five and the Big Two personality judgments from faces can be mapped onto the 2D space defined by the two basic dimensions of face evaluation, namely trustworthiness and dominance. We expect the Big Two dimensions to better fit the trustworthiness/dominance space than the Big Five dimensions, because of the semantic similarity between trustworthiness

and communion as well as between dominance and agency. Second, we aim at investigating whether individuals agree in their Big Two and Big Five personality judgments from faces. Third, we aim at investigating whether participants can discriminate between the different Big Two and Big Five personality dimensions.

The methodological aims of the article are twofold. First, we aim to develop statistical models of the Big Two and the Big Five personality dimensions in our face space and visualize the physical correlates of the Big Two and the Big Five dimensions. Second, we aim to validate the Big Two and the Big Five models in novel photographs of faces with different samples of participants.

Finally, the practical aim of this article is to develop valuable tools for researchers in various fields of psychology, for example, in consumer or law psychology. A consumer psychologist might be interested in the degree to which a brand is perceived as more or less exciting when the face of the model endorsing the brand is enhanced versus reduced on extraversion. Likewise, a law psychologist might be interested in the degree to which the sentence rises or drops when the culprit's face is enhanced versus reduced on agreeableness.

Overview

In Study 1, we collect Big Two and Big Five personality judgments as well as the two basic dimensions of face evaluation of the Oosterhof and Todorov 2D model (i.e., trustworthiness and dominance) based on 3D face scans. This allows us to investigate interrater agreement in these judgments and measure how well the Big Two and the Big Five personality dimensions fit in the trustworthiness/dominance space. Moreover, it allows us to investigate the physical correlates of the Big Two and the Big Five personality dimensions.

In a reverse engineering approach we then capture and visualize the physical correlates of the Big Two and the Big Five personality dimensions. In other words, we can, for example, generate novel
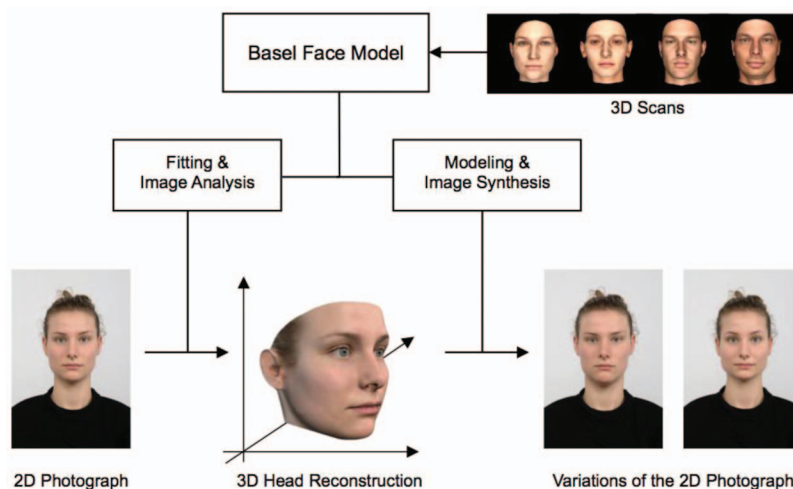

Figure 2. Visualization of the image manipulation process using the Basel Face Model: In a first step, the face on the 2D photograph (Langner et al., 2010) is reconstructed in an analysis-by-synthesis-approach (Blanz & Vetter, 1999) by fitting the Basel Face Model to the face on the photograph. In a second step, the resulting 3D head is modeled on a specific personality dimension, namely "communion" in this example. In a final step, novel images are synthesized by rendering the modeled 3D head back into the 2D photograph, resulting in natural-looking images that only vary regarding the respective personality dimension. See the online article for the color version of this figure.

faces with enhanced salience of neuroticism or extraversion and thereby show what information individuals rely on to form the respective judgments.

In Studies 2 and 3, we validate the Big Two and the Big Five models in novel photographs of faces with two different samples of participants, a European student and a U.S. M-Turk sample. These studies aim at answering two questions, namely whether manipulating the salience of the physical correlates of the Big Two and the Big Five dimensions in faces predicts personality judgments and whether the Big Two and the Big Five models generalize across samples of faces and participants.

In Study 4, we investigate whether the Big Two and the Big Five models are specific, meaning that, for example, the manipulation of extraversion more strongly influences judgments of extraversion than judgments of agreeableness. Specific models would suggest that individuals make very fine-grained personality judgments based on faces.

Study 1: Developing Statistical Models of the Big Two and the Big Five Dimensions

The aims of Study 1 are to investigate to what degree individuals agree in making Big Two and Big Five personality judgments from faces, to test how well the Big Two and the Big Five dimensions fit in the trustworthiness/dominance space, and to find the physical correlates of the Big Two and the Big Five dimensions. Therefore, we collect Big Two and Big Five personality judgments as well as trustworthiness and dominance judgments based on a set of 3D face scans. In a visual evidence-based approach toward the physical correlates of the Big Two and the Big Five dimensions, we generate vectors with maximum variability in the respective judgments and visualize the results by applying these vectors to novel faces.

Method

Participants. In total, 1,671 participants took part in this online study. They were recruited via the SoSci Panel (Leiner, 2014) and offered the chance to take part in a lottery. Of these, 598 were male, 1,066 were female, and seven did not indicate their gender. Participants had a mean age of 24.43 years (SD = 5.34).

Material and procedure. A series of 153 colored 3D scans of real faces with neutral facial expression were rendered in a frontal view (Paysan et al., 2009).

The Big Two personality dimensions were assessed with a German version of the Personal Attributes Questionnaire (PAQ, Spence, Helmreich, & Stapp, 1974; GEPAQ, Runge, Frey, Gollwitzer, Helmreich, & Spence, 1981). The Big Five personality dimensions were assessed with a German short-version of the Big Five Inventory consisting of 21 items (BFI, John, Donahue, & Kentle, 1991; BFI-K, Rammstedt & John, 2005). The items of these questionnaires were adapted to capture the personality of others instead of oneself as follows: "The person depicted is . . ." (e.g., *sociable* for the dimension extraversion). Judgments were assessed on a 5-point Likert scale ranging from 1 (*does not apply at all*) to 5 (*fully applies*). The items belonging to the same scale were presented in blocks.

Data were collected online. Participants were welcomed and instructed that the study was concerned with impression formation

and that they would spontaneously judge three persons based on their faces. In the first phase of the study personality dimensions were assessed and in the second phase of the study facial dimensions were assessed. Participants were presented with one face randomly selected from the 153 3D scans and asked to judge the person on all items from the GEPAQ and the BFI-K. Then, participants were asked to judge the person regarding trustworthiness, dominance and a series of personality traits that are not of interest in the present paper.[1] This procedure was repeated two times, so that every participant judged three persons in total. Then, the same faces were presented again in the same order and participants had to judge them on attractiveness[2] and a series of other face-related traits that are not relevant for the purpose of the present paper.[3] The order of judgments was the same for all participants. Participants were subsequently asked to provide some demographic data. Finally, they were thanked for their participation and were given the option to leave their e-mail address with the chance to win a gift voucher (experimental data and e-mail addresses were saved in separate, unconnected databases).

Results

First, focusing on the interrelation between the 2D space of face evaluation and the Big Two and the Big Five personality concepts, the different items from the BFI-K and the GEPAQ dimensions were averaged resulting in reliable scales (Cronbach's alpha ranging from .85–.98). To investigate whether the Big Two and the Big Five personality dimensions fit into the trustworthiness/dominance space, we ran regression analyses with the two basic dimensions of face evaluation (i.e., trustworthiness and dominance) as predictors and the Big Two and the Big Five personality dimensions as criteria. Results show that trustworthiness and dominance account for 69% and 90% of variance in the Big Two judgments and for 46%–81% of variance in the Big Five personality judgments (see Table 1 for all zero-order correlations, β, F, and R^2 values).

Second, focusing on the statistical modeling, we made sure that two important preconditions to successfully model these dimensions in faces were met, namely that people agree in their judgments and that the face models account for a meaningful propor-

---

[1] Additional personality traits were assessed for the purpose of generating the respective vectors to create stimuli for future research. These personality traits are not part of comprehensive personality concepts, such as the Big Two or the Big Five (e.g., violent, religious, intelligent, and honest).

[2] More recent approaches to identify basic dimensions of face evaluation differ from the approach applied by Oosterhof and Todorov (2008) in their use of photographs of real faces instead of computer-generated faces and in their use of more heterogeneous sets of faces with regard to facial expression and pose. Factor analyses based on personality judgments regarding these more diverse sets of faces revealed an additional dimension of social perception, namely youthful-attractiveness (Sutherland et al., 2013; Vernon, Sutherland, Young, & Hartley, 2014) or attractiveness-health-extraversion (Wolffhechel et al., 2014). For reasons of completeness we included attractiveness in our analyses. However, because our stimuli were highly homogeneous regarding facial expression and pose (as the stimuli used in Oosterhof & Todorov, 2008), we neither expected nor found attractiveness to significantly explain additional variance in our data.

[3] These face-based judgments were assessed to investigate the interrelation between certain face-based personality judgments as well as facial characteristics (e.g., babyfaceness/maturity or masculinity/femininity), which will be reported in a subsequent article.

Table 1

Zero-Order Correlations Between Trustworthiness (T) and Dominance (D) and the Big Two and the Big Five Personality Dimensions, β, F, and R² Values for the Regression Models With Trustworthiness and Dominance as Predictors and the Big Two and the Big Five Dimensions as Criteria

Personality concept	Personality dimension	r_T	r_D	β_T	β_D	$F(2, 152)$	R^2
Big Two	Agency	.051	.656**	.609**	.997**	163.01**	.69
Big Two	Communion	.940**	−.643**	.845**	−.169**	701.48**	.90
Big Five	Neuroticism	−.091	−.503**	−.543**	−.807**	62.67**	.46
Big Five	Extraversion	.308**	.335**	.722**	.739**	66.43**	.47
Big Five	Openness	.870**	−.523**	.842**	−.051	236.76**	.76
Big Five	Agreeableness	.841**	−.731**	.629**	−.379**	309.94**	.81
Big Five	Conscientiousness	.652**	−.053	.907**	.455**	98.13**	.57

$^{**} p < .01.$

tion of variance in these judgments. First, we calculated interrater reliabilities. Due to the fully random assignment of faces to participants, cell sizes varied. Because interrater reliabilities vary as a function of participants included in the analyses and because the smallest number of participants per face was 30, we calculated interrater reliabilities with the first 30 participants per dimension to get comparable results. Results show high interrater agreement for all Big Two and Big Five dimensions (Cronbach's alpha ranging from .76–.87; see Table 2).

Second, we tested whether the Basel Face Model accounts for a meaningful proportion of variance in the personality judgments, meaning that the judgments are systematically associated with certain facial characteristics defined in our face model. We computed seven regression analyses with the positions of the 153 3D scans on the first 50 shape components of our statistical face space as predictors and the mean values of agency (communion, neuroticism, extraversion, openness, agreeableness, conscientiousness) per face as criterion. We only used the first 50 components in order to avoid overfitting. Analogously we computed seven regression analyses with the position of the faces on the first 50 color components as predictors and the mean values of the seven different personality dimensions per face as criterion. Results show that the first 50 principal components of our shape model account for 82%–90% of variance in the personality judgments, and the first 50 principal components of our color model account for 73%–84% of variance in the judgments (see Table 2).

Because interrater reliabilities are high and both our shape and color models account for a meaningful proportion of variance in the personality judgments, we modeled the Big Five and Big Two per-

sonality dimensions. First we mapped the 153 3D scans onto the dimensions of the Basel Face Model. We only used the 50 most significant shape and color PCs to describe the position of these faces in the two face spaces. We then computed the weighted sums (ΔS, ΔC) best describing the Big Five and the Big Two personality dimensions (see Modeling Personality Dimensions section). These dimensions were then used for the synthesis of new faces (Blanz et al., 2004). Figure 3a and Figure 4a illustrate the results of this procedure in one novel exemplar face and visualize the facial information people associate with the Big Two and the Big Five dimensions.
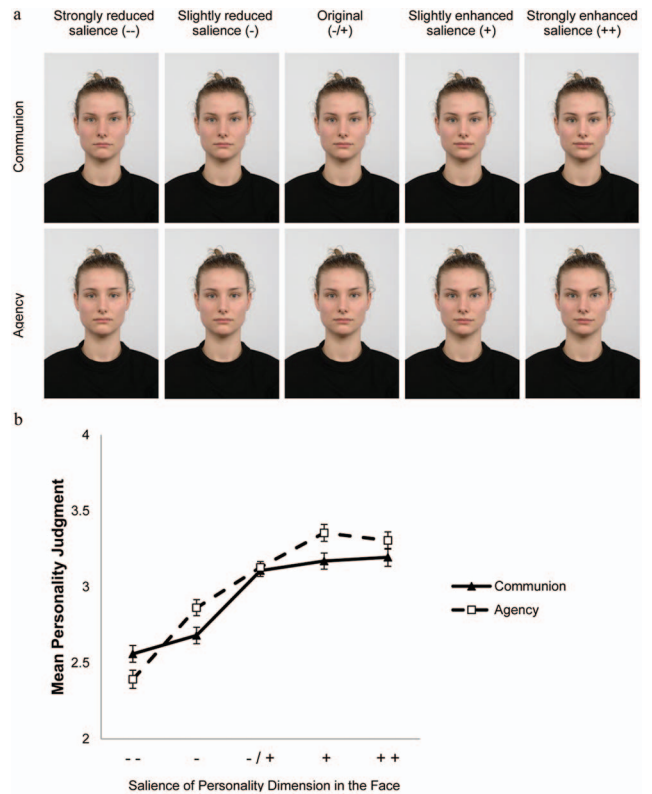

Table 2

Interrater Reliabilities (N = 30) and Explained Variance by the Shape and Color Components of the Basel Face Model for the Big Five and the Big Two Dimensions

Personality concept	Personality dimension	Reliability (α)	Shape (R^2)	Color (R^2)
Big Two	Agency	.84	.83	.74
Big Two	Communion	.87	.87	.84
Big Five	Neuroticism	.76	.82	.75
Big Five	Extraversion	.82	.84	.73
Big Five	Openness	.87	.90	.82
Big Five	Agreeableness	.87	.89	.80
Big Five	Conscientiousness	.87	.89	.76

Figure 3. (a) Visualization of the Big Two dimensions in a face from the Radboud Face Database (RaFD; Langner et al., 2010) and (b) Linear Trend Analyses for the Big Two dimensions. Error bars represent standard errors of the means. See the online article for the color version of this figure.
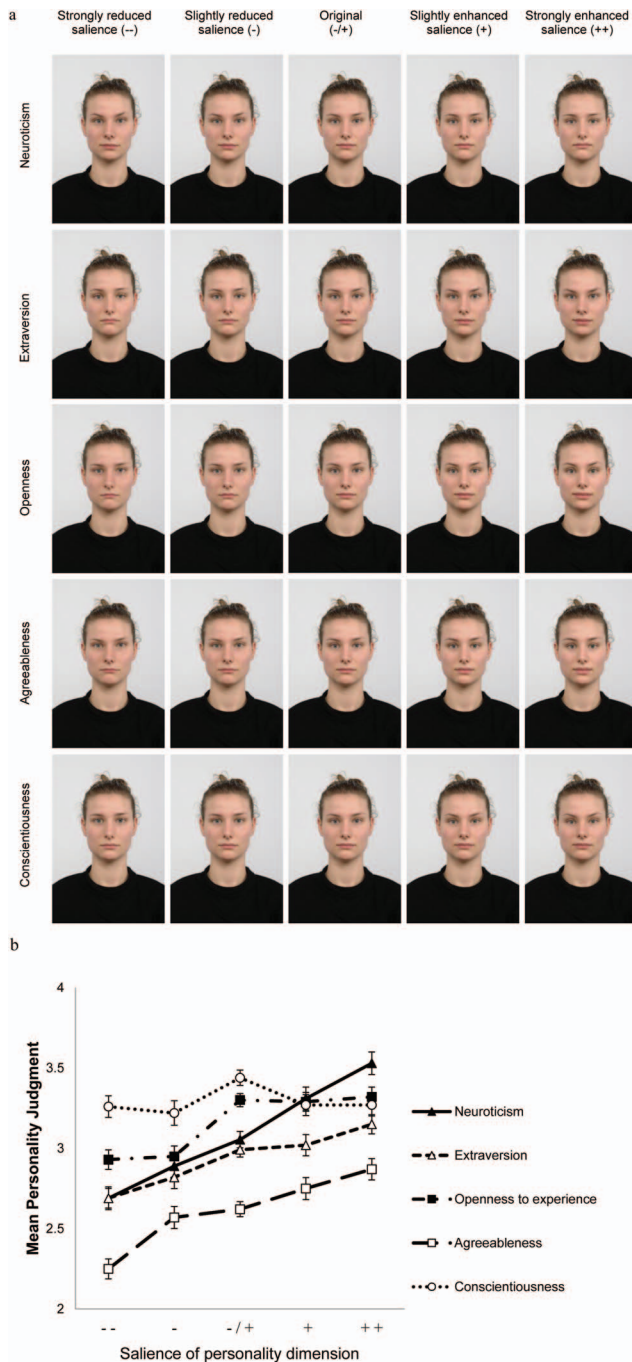
Figure 4. (a) Visualization of the Big Five dimensions in a face from the Radboud Face Database (RaFD; Langner et al., 2010) and (b) Linear Trend Analyses for the Big Five dimensions. Error bars represent standard errors of the means. See the online article for the color version of this figure.

with agency and trustworthiness almost perfectly overlaps with communion. Moreover, communion has a strong negative correlation with dominance. As shown in Table 1, trustworthiness and dominance account for more than 69% of the variance in agency judgments and 90% of the variance in communion judgments. These results provide evidence that the two basic dimensions of face evaluation and the Big Two personality dimensions are not only semantically, but also empirically akin to each other.

Results for the Big Five personality dimensions are more heterogeneous. Agreeableness and openness to experience—similar to communion—strongly overlap with trustworthiness and have a strong negative correlation with dominance. Together, the two basic dimensions of face evaluation account for 76% of variance in openness judgments and for 81% of variance in agreeableness judgments. Correlations between trustworthiness and dominance and conscientiousness, neuroticism, and extraversion are weaker. Hence, trustworthiness and dominance together account for only 46%–57% of the variance in these judgments, signaling that these three Big Five dimensions cannot as easily be explained by the dimensions trustworthiness and dominance.

Why are three out of the Big Five personality dimensions difficult to map onto the 2D space defined by trustworthiness and dominance? One reason might be that judgments of conscientiousness, neuroticism, and extraversion from faces are for some reason less reliable than judgments of agency and communion or agreeableness and openness to experience. However, this is not the case. Raters agree to the same extent in all Big Two and Big Five judgments. Moreover, the Basel Face Model accounts for equally high proportions of variance in all Big Two and Big Five judgments. In combination, these findings signal that the Big Five dimensions are significantly associated with certain facial characteristics that can be objectively described, meaning that individuals have a clear picture of what persons high or low on all five dimensions look like.

Another reason might be that more than two dimensions are needed to accurately describe the Big Five personality dimensions. Indeed, there is evidence in our data that the facial characteristics individuals use to form some of the Big Five judgments are different from the facial characteristics they use to make trustworthiness and dominance judgments. Interestingly, smiling has previously been found to impact judgments of extraversion, neuroticism, and agreeableness (Borkenau et al., 2009). Furthermore, smiling is the main cue for trustworthiness (Oosterhof & Todorov, 2008). Nevertheless, trustworthiness and dominance combined only account for approximately half of the variance in extraversion, neuroticism, and agreeableness judgments. These results suggest that three out of the Big Five personality dimensions go beyond the 2D space of face evaluation.

Study 2: Validating the Models of the Big Two Concept

Study 1 showed that individuals agree in their Big Two personality judgments and that the Basel Face Model explains most of the variance in these judgments, meaning that two preconditions to successfully model these personality dimensions are met. Building on this evidence, Study 2 investigates whether the Big Two models generalize across samples of faces and participants. We hypothesize that (a) different samples of participants (i.e., European psychology students vs. U.S. M-Turk workers) have consensual beliefs on what the Big Two dimensions look like; and (b) that linearly enhancing the

Discussion

In Study 1 we investigated whether judgments of the Big Two and the Big Five personality dimensions can be mapped onto the 2D space defined by the dimensions trustworthiness and dominance. As expected, dominance has a strong positive correlation

salience of a personality dimension in novel faces leads to a linear increase in the ascription of the respective personality dimension.

Method

Participants and design. Two different samples of participants were recruited. Data from the European sample were collected in an online study that was announced on a Facebook page for German-speaking psychology students. A sample of 201 European participants took part in the study and had the chance to win an Amazon gift voucher worth 15 Euros. Of these, 159 were female, 40 were male, and two did not indicate their gender. Their mean age was 25.30 ($SD = 6.50$). Data from the U.S. sample were collected via Amazon Mechanical Turk. A sample of 181 U.S. participants completed the study and each participant was paid $0.25 for each trial, corresponding to an hourly salary of approximately $7. Of these, 77 were female and 104 were male. Participants had a mean age of 28.92 years ($SD = 9.24$).

The design was a 3 (Salience of Personality Dimension: reduced vs. original vs. enhanced; within-participants) × 2 (Degree of Manipulation: weak vs. strong; between-participants) × 2 (Cultural Background: European vs. U.S.; between-participants) mixed design.

Materials and procedure. First, we randomly selected six male and six female faces from the Radboud Face Database (RaFD; Langner et al., 2010). We analyzed the 3D shape and the 2D color information of these faces. Specifically, we actively reconstructed the 12 faces by linearly combining the 200 3D scans of the Basel Face Model. We then varied the salience of communion and agency in the 12 reconstructed 3D heads by adding and subtracting the vectors as determined in Study 1 in five increments (i.e., strongly reduced, slightly reduced, original, slightly enhanced, and strongly enhanced) along each dimension. The resulting heads were then rendered back into the original photograph (see Figure 3a). As a result, we had nine versions (i.e., one original, four variations in communion, and four variations in agency) for each of the 12 faces taken from the RaFD.

The dependent variables were the ascriptions of agency and communion. For practical reasons, we used the two items that scored highest on the agency scale (i.e., "The person depicted is self-confident" and "The person depicted makes decisions easily") and the two items that scored highest on the communion scale (i.e., "The person depicted is sympathetic" and "The person depicted is cordial") in Study 1 to assess the two dimensions. Answers were given on a 5-point Likert scale ranging from 1 (*does not apply at all*) to 5 (*fully applies*). Participants were randomly assigned to one of two conditions. One condition contained weak manipulations (i.e., slightly reduced, original, and slightly enhanced salience of personality dimension) and the other contained strong manipulations (i.e., strongly reduced, original, and strongly enhanced salience of personality dimension). Every participant saw 12 different faces. One male and one female face were presented with reduced salience of agency, one male and one female face were presented with original salience of agency, and one male and one female face were presented with enhanced salience of agency, and analogously for communion. Participants were presented with one face together with one item per screen and had to evaluate the depicted person. Male faces were presented together with one item from the agency and communion scale, whereas female faces were presented with the other item. The combination of faces and salience of personality trait was counterbal-

anced. After the last trial, participants were asked to indicate how carefully they had answered the questionnaire, whether there was any reason not to use their data, and to explicitly name that reason. Lastly, they were asked to provide some demographic data. Depending on the sample, participants were provided with a code in order to get paid for participation (U.S. sample) or they were redirected to an external survey where they could leave their e-mail address with the chance to win a gift voucher worth 15 Euros (European sample).

Results

To test our assumption that the agency and communion models generalize across samples of participants with different cultural and education backgrounds/occupations, data were submitted to two separate 3 (Salience: reduced vs. original vs. enhanced; within-participants) × 2 (Degree of Manipulation: weak vs. strong; between-participants) × 2 (Cultural Background: European vs. U.S.; between-participants) mixed-factorial ANOVAs with agency and communion as dependent variables. No interaction effect involving participants' cultural background reached statistical significance; for agency $F_{max}(2, 377) = 1.56$, $p_{min} = .212$, $\eta^2 = .008$ and for communion $F_{max}(2, 377) = 1.13$, $p_{min} = .325$, $\eta^2 = .006$. Therefore, we collapsed data over participants' cultural background for all subsequent analyses; see Footnote[4] for the results of the ANOVAs.

To test our hypothesis that individuals are perceived to be more agentic with increasing salience of agency in their face, we ran a mixed linear trend analysis, which strongly supports the hypothesis, $t(759.58)[5] = 13.11$, $p < .001$, $r = .43$, 95% CI [.33, .44]. Analogously, a linear trend analysis for communion revealed that individuals are perceived to be more communal with increasing salience of communion in their face, $t(760) = 9.87$, $p < .001$, $r = .34$, 95% CI [.24, .35] (see Figure 3b).

Discussion

Results from Study 2 indicate that linearly enhancing the salience of agency (communion) in a face leads to a linear increase in the ascription of agency (communion) to the respective person. Therefore, we can conclude that our method to manipulate both Big Two dimensions in novel faces is successful. Moreover, re-

---

[4] Two separate 3 (Salience: reduced vs. original vs. enhanced; within-participants) × 2 (Degree of Manipulation: weak vs. strong; between-participants) mixed-factorial ANOVAs with agency and communion as dependent factors were run. For agency, we found a main effect of salience and a main effect of degree of manipulation. Persons were judged to be more agentic with increasing salience of agency in their faces, $F(2, 379) = 99.98$, $p < .001$, $\eta^2 = .345$. Weak manipulations, on average, resulted in higher ascriptions of agency, $F(1, 380) = 10.41$, $p = .001$, $\eta^2 = .027$. These main effects were qualified by a significant interaction of salience and degree of manipulation indicating that negative manipulations more strongly affected judgments than positive manipulations, $F(2, 379) = 13.31$, $p < .001$, $\eta^2 = .066$. For communion, we found a main effect of salience, indicating that persons were judged to be more communal with increasing salience of communion in their faces, $F(2, 379) = 68.16$, $p < .001$, $\eta^2 = .265$. Neither the main effect of degree of manipulation, $F(1, 380) = .072$, $p = .001$, $\eta^2 = .789$ nor the interaction of salience and degree of manipulation reached statistical significance, $F(2, 379) = 1.69$, $p = .178$, $\eta^2 = .009$.

[5] Fractional degrees of freedom occur because these statistics do not have exact t-distributions. Degrees of freedom stem from a Satterthwaite approximation.

sults generalize across samples of faces and participants, meaning that the face models work equally well for European students and U.S. Amazon M-Turk workers.

Study 3: Validating the Models of the Big Five Concept

Study 2 focused on the Big Two. Analogously, Study 3 now focuses on the Big Five and investigates whether the Big Five models generalize across samples of faces and participants. We hypothesize that (a) different samples of participants (i.e., European psychology students vs. U.S. M-Turk workers) have consensual beliefs on what the Big Five dimensions look like; and (b) that linearly enhancing the salience of a personality dimension in novel faces leads to a linear increase in the ascription of the respective personality dimension.

Method

Participants and design. Two different samples of participants were recruited. Data from the European sample were collected in an online study announced on a Facebook page for German-speaking psychology students. A sample of 160 participants took part in the study and had the chance to win one of two Amazon gift vouchers worth 15 Euros. Of these, 115 were female and 45 were male with a mean age of 24.87 (SD = 6.27). Data from the U.S. sample were collected via Amazon Mechanical Turk. A sample of 144 participants completed the study and each participant was paid $0.50, corresponding to an hourly salary of approximately $7. Of these, 78 were female, 65 were male, and one did not indicate his or her gender. Participants had a mean age of 32.83 years (SD = 10.98).

The design was a 3 (Salience of Personality Dimension: reduced vs. original vs. enhanced; within-participants) × 2 (Degree of Manipulation: weak vs. strong; between-participants) × 2 (cultural background: European vs. U.S.; between-participants) mixed design.

Materials and procedure. Materials and procedure were similar to those in Study 2, differing only (a) in regard to the quantity of faces judged by a participant (30 vs. 12); (b) in regard to the dimensions that were manipulated in faces, namely neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness; and correspondingly (c) in the dependent variable measured with the two items loading most strongly on the five K-BFI scales as determined in Study 1.

Results

To test our assumption that the Big Five models generalize across samples of participants with different cultural and education background, data were submitted to five separate 3 (Salience: reduced vs. original vs. enhanced; within-participants) × 2 (Degree of manipulation: weak vs. strong; between-participants) × 2 (Cultural Background: European vs. U.S.; between-participants) mixed-factorial ANOVAs for the dependent variables neuroticism, extraversion, openness, agreeableness and conscientiousness. As expected, no interaction effect involving participants' cultural background reached statistical significance; for neuroticism $F_{max}(2, 298) = 1.00$, $p = .370$, $\eta^2 = .007$, for extraversion $F_{max}(2,$

298) = 1.79, $p = .168$, $\eta^2 = .012$, for openness $F_{max}(2, 298) = 0.66$, $p = .516$, $\eta^2 = .004$, for agreeableness $F_{max}(2, 298) = 0.83$, $p = .437$, $\eta^2 = .006$, and for conscientiousness $F_{max}(1, 298) = 1.27$, $p = .262$, $\eta^2 = .004$. For all subsequent analyses we collapsed data over cultural background; see Footnote[6] for the results of the ANOVAs.

To test our hypothesis that linearly enhancing the salience of a personality trait in a face leads to a linear increase in the ascription of that trait, we ran five mixed linear trend analyses. Regarding neuroticism, extraversion, openness, and agreeableness, we found support for our hypothesis. Linearly enhancing the salience of these four dimensions in faces resulted in higher ascriptions of the respective personality dimensions, for neuroticism $t(602.64) = 9.43$, $p < .001$, $r = .36$, 95% CI [.28, .42], for extraversion $t(602.01) = 5.68$, $p < .001$, $r = .23$, 95% CI [.12, .25], for openness $t(602.28) = 5.95$, $p < .001$, $r = .24$, 95% CI [.13, .25], and for agreeableness $t(602.26) = 6.92$, $p < .001$, $r = .27$, 95% CI [.17, .31]. Only the results for conscientiousness failed to reach statistical significance, $t(599.34) = .35$, $p = .727$, 95% CI [−.06, .08]. A closer look at the results revealed that in the negative domain (i.e., strongly reduced conscientiousness, slightly reduced conscientiousness, and original face), the linear trend also reached statistical significance, $t(600.88) = 2.84$, $p = .005$, $r = .12$, 95% CI [.03, .17] (see Figure 4b).

Discussion

Overall, results from Study 3 indicate that reducing (enhancing) the salience of neuroticism, extraversion, openness, and agreeableness in faces is reflected in lower (higher) neuroticism, extraversion, openness, and agreeableness judgments. Only one out of the Big Five dimensions, namely conscientiousness, seems to be difficult to model. Interestingly, based on the interrater reliabilities and the proportion of explained variance examined in Study 1, this dimension looked as promising as the others. Participants agreed in their conscientiousness judgments and the Basel Face Model explained a meaningful proportion of variance in these judgments. However, when applied to novel faces, manipulations worked for low levels of conscientiousness, but not for high levels, which might reflect a negativity bias (i.e., the phenomenon that negative information is weighted more than positive information; e.g., Kanouse, 1984). Because these results generalize across both samples, they seem to be systematic. Obviously, more research is needed to understand why the conscientiousness-vector does not

---

[6] Five separate 3 (Salience: reduced vs. original vs. enhanced; within-participants) × 2 (Degree of Manipulation: weak vs. strong; between-participants) mixed-factorial ANOVAs with the Big Five dimensions as dependent factors revealed main effects of salience for all five dimensions; for neuroticism $F(2, 300) = 48.43$, $p < .001$, $\eta^2 = .244$, for extraversion $F(2, 300) = 15.66$, $p < .001$, $\eta^2 = .095$, for openness $F(2, 300) = 27.06$, $p < .001$, $\eta^2 = .153$, for agreeableness $F(2, 300) = 22.33$, $p < .001$, $\eta^2 = .130$, and for conscientiousness, $F(1, 299) = 4.54$, $p = .011$, $\eta^2 = .029$. For neuroticism and agreeableness this main effect was qualified by a significant interaction of salience and degree of manipulation indicating that personality judgments were more extreme for stronger manipulations, for neuroticism, $F(2, 300) = 6.52$, $p = .001$, $\eta^2 = .042$ and for agreeableness, $F(2, 300) = 8.69$, $p < .001$, $\eta^2 = .055$. No main effect of degree of manipulation reached statistical significance, $F_{max}(1, 300) = 1.091$, $p = .297$, $\eta^2 = .004$.

fully generalize to novel faces. On the basis of our data, we can only speculate as to why the results concerning conscientiousness are different from the results concerning the other four dimensions. One reason might be that conscientiousness is not a genuinely interpersonal dimension. Accomplishing a task thoroughly does not involve the presence of others. Moreover, it is not of primarily importance to know whether a stranger we meet generally accomplishes tasks thoroughly. At first glance, this argument holds true for openness and neuroticism as well. However, openness and neuroticism judgments have relatively well-defined physical correlates (i.e., open eyes/attentive look and a sad facial expression, respectively), and may therefore be more clearly perceived. Such well-defined correlates are missing for conscientiousness.

Interestingly our results nicely match results from a study using judgments of composite faces of individuals scoring very low and very high on the Big Five personality dimensions (Kramer & Ward, 2010): Participants in this study recognize the composites averaged across the faces of individuals scoring high on all dimensions but conscientiousness. By generating composite faces, extrafacial features are no longer interpretable because clothes and hairlines become blurry. It has been previously shown in different studies that such extrafacial features (e.g., a formal dress) are valid and necessary cues to judge an individual's conscientiousness (Albright, Kenny, & Malloy, 1988; Borkenau & Liebler, 1992; Naumann, Vazire, Rentfrow, & Gosling, 2009). Because our face model does not consider any extrafacial features, the critical cues for judging conscientiousness might be absent in our model.

Taken together, we can conclude that our method to manipulate novel faces is successful in at least four Big Five dimensions. Again, the results generalize across samples of faces and participants from different cultural, educational, and occupational backgrounds.

Study 4: Testing the Specificity of the Big Two and Big Five Models

In Study 4, we investigate whether the Big Two and the Big Five vectors specifically impact the respective personality judgments, meaning that, for example, the extraversion vector most strongly impacts judgments of extraversion and not judgments of openness, agreeableness, neuroticism, or conscientiousness. If the vectors are specific for the respective personality dimensions, this indicates that individuals do not just evaluate others on an abstract level (e.g., regarding their intention and ability to do harm) and then translate this general evaluation into more fine-grained judgments. Rather, dimension specific models support the notion that individuals make very differentiated personality judgments.

Specific models would result, for example, in higher ascriptions of extraversion to a face with enhanced extraversion than to a face with enhanced openness, which is a dimension highly correlated with extraversion (see Table 3). Such highly correlated dimensions naturally result in vectors pointing in similar directions within the statistical Face Space. Consequently, participants need to detect subtle nuances in faces in order to differentiate between the different dimensions. Because the nuances in which some of the faces presented in Study 4 differ are so subtle, Study 4 is a very strict test of our models (see Figure 5).

Method

Participants and design. Again, two different samples of participants were recruited. Data from the European sample were collected in an online study announced on a Facebook page for German-speaking psychology students. A sample of 65 participants took part in the study and had the chance to win an Amazon gift voucher worth 10 Euros. Of these, 48 were female and 17 were male, with a mean age of 26.23 ($SD = 7.87$). Data from the U.S. sample were collected via Amazon Mechanical Turk. A sample of 51 participants took part in the study and each participant was paid $0.60, corresponding to an hourly salary of approximately $7. Of these, 23 were female, 28 were male, and the mean age was 35.25 years ($SD = 11.42$).

The design consisted of one independent variable, namely cultural background of participants and the dependent variable was the percentage of correct choices.

Materials and procedure. Materials were similar to the materials used in Studies 2 and 3 except that (a) faces were manipulated more strongly on the respective dimensions, (b) only four faces were used, and (c) the items were reframed as detailed below. Two male and two female faces were randomly selected from the RaFD (Langner et al., 2010). We enhanced and reduced the salience of all the seven personality dimensions from the Big Five and the Big Two personality concepts in all four faces. We then paired those faces as follows: Every pair consisted of two faces from the same identity (i.e., same original photograph), but were manipulated on different personality dimensions. For positively correlated dimensions (e.g., openness and extraversion) we used the versions with enhanced salience on both dimensions. For negatively correlated dimensions (e.g., openness and neuroticism), we used versions manipulated in opposite directions (e.g., openness enhanced, neuroticism reduced for one pair and openness reduced, neuroticism enhanced for the second pair). Every Big Five personality dimension was combined with every other Big Five dimension resulting in 10 combinations of Big Five personality dimensions. In addition, agency was combined with communion, resulting in one combination of Big Two dimensions. Together, this procedure resulted in 11 combinations of personality dimensions. Every combination of dimensions was presented four times, because two different dimensions and two different items per dimension were involved.

This procedure resulted in a total of 44 trials per participant. The items were taken from Studies 2 and 3, and reframed into questions. Participants were presented, for example, with a face pair consisting of an openness-enhanced and an extraversion-enhanced face and were asked to indicate which of the two better fits a typical characteristic of one of the manipulated dimensions (e.g., Which person has a more vivid imagination?). Faces were presented at a random position (i.e., left vs. right) and pairs were presented in a random order.

After the last trial, participants were asked to indicate how carefully they had answered the questionnaire and whether there was any reason not to use their data and to explicitly name that reason. Depending on the sample, participants were provided with a code in order to get paid for participation (U.S. sample) or they were redirected to an external survey where they could leave their

Table 3
Correlations (r, N = 153) of Personality Dimensions Separately for the Big Two and the Big Five Personality Dimensions in Study 4

Personality concept	Personality dimension	1	2	3	4	5	6	7
Big Two	1. Agency	1	−.02					
Big Two	2. Communion		1					
Big Five	3. Neuroticism			1	−.77**	−.04	−.04	−.21**
Big Five	4. Extraversion				1	.29**	.27**	.28**
Big Five	5. Openness					1	.75**	.71**
Big Five	6. Agreeableness						1	.34**
Big Five	7. Conscientiousness							1

** $p < .01$.

e-mail address in order to take part in the lottery (European sample).

Results

First, we tested whether participants from different cultural backgrounds differed regarding their performance, which was not

Figure 5. Pairs of faces with the highest-correlating personality dimensions presented in Study 4 (left: neuroticism [reduced], right: extraversion [enhanced]). Original faces stem from the Radboud Face Database (RaFD; Langner et al., 2010). See the online article for the color version of this figure.

the case, $t(114) = 1.31$, $p = .193$, 95% CI [−.01, .06]. Therefore, data were collapsed over participants' cultural background.

Overall, participants chose the face with enhanced salience of the respective personality dimension in 64% of all cases ($SD = 10\%$), which is significantly above chance level, $t(115) = 15.50$, $p < .001$, $d = 1.44$, 95% CI [.13, .16]. Results for the 11 combinations of personality dimensions (see Table 4) reveal that in only one out of 11 cases, namely when participants had to discriminate between extraversion and neuroticism, they failed to choose the correct face above chance level ($M = 44\%$, $SD = 23\%$), $t(115) = −3.08$, $p = .003$, $d = .29$, 95% CI [−.11, −.02]. A closer look reveals that if extraversion had been enhanced in one face and neuroticism reduced in the other, participants chose the correct face in 57% of all cases ($SD = 35\%$), $t(115) = 2.25$, $p = .026$, $d = .21$, 95% CI [.01, .14]. However, if neuroticism had been enhanced in one face and extraversion reduced in the other, participants chose the correct face in only 30% of all cases ($SD = 32\%$), $t(115) = −6.76$, $p < .001$, $d = .63$, 95% CI [−.26, −.14].

The similarity between two dimensions limits the distinguishability between two correspondingly manipulated faces. Therefore, we tested whether the strength of correlations between two vectors (see Table 3) predicts distinguishability between them (see Table 4). A regression analysis revealed that the correlation coefficients for pairs of vectors predicted the accuracy (i.e., percentage of correct answers) with which participants discriminated between them, $F(1, 1275) = 73.88$, $p < .05$, $\eta^2 = .05$, $R^2 = .06$, signaling that participants differentiate between two different dimensions to the degree that these dimensions differ from each other.

Discussion

In Study 4 we used a very strict test to investigate whether the Big Two and the Big Five vectors are specific for the respective personality dimension. Overall, the results indicate that every vector triggers the corresponding personality trait best. In particular, in 10 out of the 11 combinations, when asked to select the face that looked more extreme regarding a specific personality dimension, participants selected the face that was enhanced on that respective dimension. This finding seems to signal that only the dimensions extraversion and neuroticism are hard to distinguish if they are manipulated in opposite directions. As Table 3 reveals, these two dimensions are the two most highly correlated dimensions, which naturally limits their discriminability. However, more detailed analyses reveal that participants are able to detect the face with enhanced salience of extraversion if extraversion is enhanced in one face and neu-
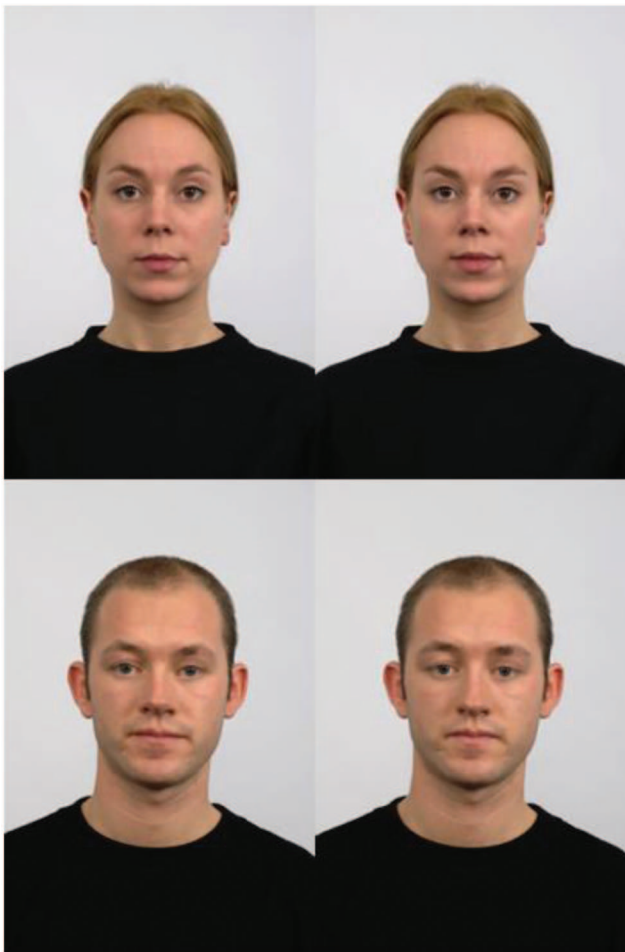
Table 4
Means (Standard Deviations) Indicating the Percentage of Correct Answers in Study 4

Personality concept	Personality dimension	1	2	3	4	5	6	7
Big Two	1. Agency		74 (23)**					
Big Two	2. Communion							
Big Five	3. Neuroticism				44 (23)	68 (24)**	81 (25)**	57 (22)**
Big Five	4. Extraversion					57 (24)**	73 (26)**	55 (26)*
Big Five	5. Openness						63 (22)**	63 (22)**
Big Five	6. Agreeableness							76 (21)**
Big Five	7. Conscientiousness							

* $p < .05$. ** $p < .01$.

roticism is reduced in the other, whereas they are not able to detect the face with enhanced neuroticism if neuroticism is enhanced in one face and extraversion is reduced in the other. This might indicate that the vectors are in fact specific for the two dimensions, but that the items intended to capture neuroticism do not constitute the best measures.

General Discussion

Previous research on general, spontaneous person evaluation based on faces has found that judgments can be mapped in a 2D space defined by the two basic dimensions trustworthiness and dominance, signaling the person's intention and ability to cause harm (Oosterhof & Todorov, 2008). In this paper, we went beyond such general, spontaneous judgments and investigated whether individuals reliably judge strangers regarding the Big Two and the Big Five personality dimensions, whether they differentiate between these dimensions, and whether the Big Two and the Big Five judgments can be mapped onto the 2D space of face evaluation or whether they go beyond that 2D space. Moreover, we aimed at modeling the Big Two and the Big Five personality dimensions in our face space, apply them to novel photographs of faces and investigate whether they generalize across different samples of faces and participants. In what follows, we briefly review the main results and discuss implications for theory and practice.

Theoretical Aspects

On a theoretical level, we first investigated whether the 2D space defined by the two basic dimensions of face evaluation, namely trustworthiness and dominance, is adequate to describe judgments of the Big Two and the Big Five personality dimensions from real faces. Results from Study 1 suggest that the two basic dimensions of face evaluation highly overlap with the Big Two personality dimensions agency and communion. So, Study 1 provides evidence that the concept of the Big Two and the two basic dimensions of face evaluation are not only conceptually, but also empirically akin to each other. With regard to the Big Five personality dimensions, the two basic dimensions of face evaluation explain substantially more variance in openness to experience and agreeableness judgments than in neuroticism, extraversion, and conscientiousness judgments. Although one of the facial cues for neuroticism, extraversion, and conscientiousness judgments is smiling (Borkenau et al., 2009), a facial characteristic also relevant for trustworthiness judgments (Oosterhof & Todorov, 2008), ap-

proximately half of the variance in these judgments remains unexplained by trustworthiness and dominance. Two implications ensue: First, it is evident that the Big Five capture more than the two basic dimensions of face evaluation. Second, it will be interesting for future research to explore which facial characteristics are used to make these personality judgments.

The contribution's second theoretical goal was to investigate whether different judges agree in their Big Two and Big Five judgments from faces. Results from Study 1 provide evidence that individuals exhibit strong agreement in their Big Five and Big Two judgments from faces. Although individuals are at most moderately accurate in forming personality judgments from faces, different individuals use the same facial cues to judge a specific personality dimension and they have shared beliefs regarding the meaning of these facial cues. In other words, people have consensual beliefs on what, for example, an extraverted or an agreeable person looks like. This even holds true for those dimensions that are not genuinely interpersonal (i.e., neuroticism, openness to experience, and conscientiousness).

Results from Studies 2 to 4 revealed that even participants from different cultural and educational/occupational backgrounds and with different ages have consensual beliefs on what the Big Two and the Big Five dimensions look like. This finding suggests that individuals from different cultures learn the same associations between certain facial cues and personality dimensions. One might argue that both the European and the U.S. sample are Western samples, and therefore are more similar than, for example, an American and an Asian sample. However, results from cross-cultural research have shown that Europeans are more collectivistic and less individualistic than U.S. Americans (Oyserman, Coon, & Kemmelmeier, 2002).

The third theoretical goal was to investigate whether individuals can discriminate between the different Big Two and Big Five dimensions. Results from Study 4 provided evidence that individuals make very differentiated Big Two and Big Five personality judgments. Even though the different dimensions of one personality concept are not orthogonal, every vector triggers the corresponding dimension strongest. Not surprisingly, the higher the correlations between two vectors are (i.e., applying these vectors to novel faces produces similar results), the more difficult it is to discriminate between them. This finding further supports the notion that individuals make very fine-grained personality judgments from faces.

Taken together, individuals have both highly consensual and differentiated images of what persons scoring high or low on the

Big Two and the Big Five personality dimensions look like. Results regarding the Big Five personality dimensions might be especially intriguing, because these dimensions—in contrast to the Big Two dimensions—have been shown here to go beyond the basic dimensions of face evaluation, namely trustworthiness and dominance (Oosterhof & Todorov, 2008).

Such consensual beliefs regarding the Big Five dimensions can have serious consequences. First of all, the Big Five dimensions are not neutral with regard to valence. Neuroticism, for example, is perceived less favorably than extraversion. If an individual's face signals that he or she is neurotic, this might shape his or her social interactions. Others might approach them less positively than they would approach a person signaling extraversion. Consequently, this might affect the individual's self-perception and in turn also his or her behavior (e.g., via a self-fulfilling prophecy; Chen & Bargh, 1997; Darley & Fazio, 1980). Second, the Big Five are ubiquitous in a variety of applied settings. Therefore, face-based Big Five personality judgments might have very direct consequences as well. Imagine, for example, a human resource manager looking for an employee that is suited to the affordances of a specific job (e.g., agreeableness and openness). After inviting the most promising candidates to job interviews (a procedure that might already be biased due to the photograph on the CV, which is common in many European countries), the human resource manager meets the candidates in person and quickly forms first impressions of them. These impressions might be strongly influenced by the facial appearance of the candidates, because this information is salient in the first moments of an encounter. Moreover, these first impressions are likely to influence processing of any information that is subsequently received regarding the candidates. A lot of inconsistent information is necessary to change the initial impression. Therefore, the final decision of who gets the job is likely to be influenced by the facial appearance of the candidates. And again, this outcome might affect the target person's self-perception and behavior. Especially if first encounters do not reveal a lot of other information about an individual, and if perceivers know what personality they are looking for, the interpretation of facial qualities is likely to impact decisions. Similarly, such context-specific personality judgments based on faces might also influence the friends we make or the partners we choose. And they might impact highly consequential decisions in other domains, such as in criminal sentencing.

Methodological Aspects

On a methodological level, the high interrater agreement in the Big Two and Big Five judgments together with the finding that our face model accounts for meaningful proportions of variance in all Big Two and Big Five judgments signals that chances to successfully model these dimensions are high. In a backward engineering approach we modeled the Big Two and Big Five personality judgments in our face space and applied the models to novel faces in order to visualize their physical correlates. Results of Study 2 and 3 revealed that systematically enhancing or reducing the salience of these dimensions in novel faces resulted in more or less extreme judgments on the respective dimension. Only conscientiousness was difficult to model—at least in the positive domain. More research is needed to understand why the conscientiousness-vector did not fully generalize to novel faces or novel samples of participants.

Applied Aspects

On a practical level, the here presented method to model the Big Five and the Big Two personality dimensions in novel faces might prove highly useful in future research, because it allows the systematic modeling of how a person is perceived with regard to a certain fundamental personality dimension, while all facial qualities that are independent from this dimension remain unchanged. Moreover, this method allows for gradual modeling on these personality dimensions. In other words, both the direction and the degree to which a face is modeled can be exactly defined. Furthermore, stimuli generated with this approach look natural and thus are highly ecologically valid. Therefore, the Big Five and the Big Two models presented so far might be useful in research in different areas. One could investigate, for example, whether the salience of agreeableness in a culprit's face impacts sentencing decisions. And if so, to what extent the salience of agreeableness needs to be enhanced in order, for example, to result in a prison sentence reduced by one year. Or, one could investigate, whether a human resource manager tends to select the most agreeable-looking applicant among equally qualified applicants if the job particularly requires agreeableness. And if so, what degree of facial agreeableness possibly makes up for not being as qualified as the other candidates.

Coda

Taken together, the present findings bridge a gap between personality psychology and face perception research by isolating and modeling perceptions of the Big Two and the Big Five personality dimensions independent of any other face information. This can be done in any novel face with natural-looking results. Moreover, these models generalize across two samples that differ regarding several variables. Against the background that personality judgments based on faces have far-reaching consequences in many domains, the potential to model the most established personality dimensions in faces in a controlled way and with natural-looking results might render these models useful tools for researchers in various psychological disciplines investigating the impact of personality on diverse outcome variables.

References

Abele, A. E., Uchronski, M., Suitner, C., & Wojciszke, B. (2008). Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. *European Journal of Social Psychology, 38,* 1202–1217. http://dx.doi.org/10.1002/ejsp.575

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93,* 751–763. http://dx.doi.org/10.1037/0022-3514.93.5.751

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55,* 387–395. http://dx.doi.org/10.1037/0022-3514.55.3.387

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. *Advances in Experimental Social Psychology, 32,* 201–271. http://dx.doi.org/10.1016/S0065-2601(00)80006-4

Ballew, C. C., II, & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 17948–17953. http://dx.doi.org/10.1073/pnas.0705435104

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6,* 269–278. http://dx.doi.org/10.1037/1528-3542.6.2.269

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15,* 674–679. http://dx.doi.org/10.1111/j.0956-7976.2004.00739.x

Blanz, V., Scherbaum, K., Vetter, T., & Seidel, H. P. (2004). Exchanging faces in images. *Computer Graphics Forum, 23,* 669–676. http://dx.doi.org/10.1111/j.1467-8659.2004.00799.x

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, 19,* 187–194. http://dx.doi.org/10.1145/311535.311556

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117,* 187–215. http://dx.doi.org/10.1037/0033-2909.117.2.187

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,* 703–706. http://dx.doi.org/10.1016/j.jrp.2009.03.007

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62,* 645–657. http://dx.doi.org/10.1037/0022-3514.62.4.645

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327. http://dx.doi.org/10.1111/j.2044-8295.1986.tb02199.x

Chen, M., & Bargh, J. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 560,* 541–560. http://dx.doi.org/10.1006/jesp.1997.1329

Clark, L. A., & Harrison, J. A. (2001). Assessment instruments. In W. J. Livesley (Ed.), *Handbook of personality disorders: Theory, research, and treatment* (pp. 277–306). New York, NY: Guilford Press.

Costa, P. T. (1996). Work and personality: Use of the NEO-PI-R in industrial/organisational psychology. *Applied Psychology, 45*(July 1994), 225–241. http://dx.doi.org/10.1111/j.1464-0597.1996.tb00766.x

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35,* 867–881. http://dx.doi.org/10.1037/0003-066X.35.10.867

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73,* 1246–1256. http://dx.doi.org/10.1037/0022-3514.73.6.1246

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological & Personality Science, 3,* 562–571. http://dx.doi.org/10.1177/1948550611430272

Egan, V., Deary, I., & Austin, E. (2000). The NEO-FFI: Emerging British norms and an item-level analysis suggest N, A and C are more reliable than O and E. *Personality and Individual Differences, 29,* 907–920. http://dx.doi.org/10.1016/S0191-8869(99)00242-1

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11,* 77–83. http://dx.doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82,* 878–902. http://dx.doi.org/10.1037/0022-3514.82.6.878

Gebauer, J. E., Leary, M. R., & Neberich, W. (2012). Big two personality and big three mate preferences: Similarity attracts, but country-level mate preferences crucially matter. *Personality and Social Psychology Bulletin, 38,* 1579–1593. http://dx.doi.org/10.1177/0146167212456300

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26–34. http://dx.doi.org/10.1037/0003-066X.48.1.26

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The "Big Five" Inventory – Versions 4a and 54.* Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research, 2,* 102–138.

Kanouse, D. E. (1984). Explaining negativity biases in evaluation and choice behavior: Theory and research. *Advances in Consumer Research Association for Consumer Research, 11,* 703–708.

Kramer, R. S. S., & Ward, R. (2010). Internal facial features are signals of personality and health. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 63,* 2273–2287. http://dx.doi.org/10.1080/17470211003770912

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud faces database. *Cognition and Emotion, 24,* 1377–1388. http://dx.doi.org/10.1080/02699930903485076

Leiner, D. J. (2014). *SoSci survey.* Retrieved from https://www.soscisurvey.de

Little, A. C., & Perrett, D. I. (2007). Using composite images to assess accuracy in personality attribution to faces. *British Journal of Psychology, 98,* 111–126. http://dx.doi.org/10.1348/000712606X109648

Mangini, M., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science, 28,* 209–226. http://dx.doi.org/10.1207/s15516709cog2802_4

McCrae, R. R., & Costa, P. T. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51–87). New York: Guilford.

McCrae, R. R., & Costa, P. T., Jr. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology, 56,* 586–595. http://dx.doi.org/10.1037/0022-3514.56.4.586

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist, 52,* 509–516. http://dx.doi.org/10.1037/0003-066X.52.5.509

Mulyanegara, R. C., Tsarenko, Y., & Anderson, A. (2007). The Big Five and brand personality: Investigating the impact of consumer personality on preferences towards particular brand personality. *Journal of Brand Management, 16,* 234–247. http://dx.doi.org/10.1057/palgrave.bm.2550093

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin, 35,* 1661–1671. http://dx.doi.org/10.1177/0146167209346309

Ono, M., Sachau, D. A., Deal, W. P., Englert, D. R., & Taylor, M. D. (2011). Cognitive ability, emotional intelligence, and the big five personality dimensions as predictors of criminal investigator performance. *Criminal Justice and Behavior, 38,* 471–491. http://dx.doi.org/10.1177/0093854811399406

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 11087–11092. http://dx.doi.org/10.1073/pnas.0805664105

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin, 128,* 3–72.

Paysan, P., Knothe, R., Amberg, B., Romdhani, S., & Vetter, T. (2009). A 3D face model for pose and illumination invariant face recognition. In *Advanced video and signal based surveillance* (pp. 296–301). IEEE. http://dx.doi.org/10.1109/AVSS.2009.58

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,* 607–640. http://dx.doi.org/10.1521/soco.2006.24.5.607

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135,* 322–338. http://dx.doi.org/10.1037/a0014996

Pound, N., Penton-Voak, I. S., & Brown, W. M. (2007). Facial symmetry is positively associated with self-reported extraversion. *Personality and Individual Differences, 43,* 1572–1582. http://dx.doi.org/10.1016/j.paid.2007.04.014

Prinzie, P., Stams, G. J., Deković, M., Reijntjes, A. H., & Belsky, J. (2009). The relations between parents' Big Five personality factors and parenting: A meta-analytic review. *Journal of Personality and Social Psychology, 97,* 351–362. http://dx.doi.org/10.1037/a0015823

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica, 51,* 195–206. http://dx.doi.org/10.1026/0012-1924.51.4.195

Rule, N. O., & Ambady, N. (2008). First impressions. Peeking at the neural underpinnings. In N. Ambady & J. Skowronski (Eds.), *First impressions* (pp. 35–56). New York, NY: Guilford Press.

Runge, T. E., Frey, D., Gollwitzer, P. M., Helmreich, R. L., & Spence, J. T. (1981). Masculine (instrumental) and feminine (expressive) traits: A comparison between students in the United States and West Germany. *Journal of Cross-Cultural Psychology, 12,* 142–162. http://dx.doi.org/10.1177/0022022181122002

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,* 1373–1383. http://dx.doi.org/10.1016/S0191-8869(02)00356-2

Solomon, M. R., Russell-Bennett, R., & Previte, J. (2012). *Consumer behavior. Buying, having, being.* Frenchs Forest, Australia: Pearson Higher Education AU.

Spence, J. T., Helmreich, R. L., & Stapp, J. (1974). The Personal Attributes questionnaire: A measure of sex role stereotypes and masculinity-femininity. *Journal Supplement Abstract Service, American Psychological Association, 4,* 1–42.

Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition, 127,* 105–118. http://dx.doi.org/10.1016/j.cognition.2012.12.001

Todorov, A., & Oosterhof, N. N. (2011). Modeling social perception of faces. *IEEE Signal Processing Magazine, 28,* 117–122. http://dx.doi.org/10.1109/MSP.2010.940006

Todorov, A., Said, C. P., & Verosky, S. C. (2011). Personality impressions from facial appearance. In A. Calder, J. V. Haxby, M. Johnson, & G. Rhodes (Eds.), *Handbook of face perception* (pp. 631–652). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199559053.013.0032

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43,* 161–204.

Vernon, R. J. W., Sutherland, C. A., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences of the United States of America, 111,* E3353–E3361. http://dx.doi.org/10.1073/pnas.1409860111

Vetter, T. (n.d.). *Morphface.* Retrieved from http://faces.cs.unibas.ch/bfm/

Walker, M. (2015). *Face modeling.* Retrieved from http://mirellawalker.com/face-modeling

Walker, M., Jiang, F., Vetter, T., & Sczesny, S. (2011). Universals and cultural differences in forming personality trait judgments from faces. *Social Psychological & Personality Science, 2,* 609–617. http://dx.doi.org/10.1177/1948550611402519

Walker, M., & Vetter, T. (2009). Portraits made to measure: Manipulating social judgments about individuals with a statistical face model. *Journal of Vision, 9,* 1–13. http://dx.doi.org/10.1167/9.11.12

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 395–412. http://dx.doi.org/10.1037/0022-3514.37.3.395

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In W. M. Grove & D. Cicchetti (Eds.), *Thinking Clearly about Psychology* (pp. 89–113). Minneapolis: Minnesota Press.

Wiggins, J. S., Phillips, N., & Trapnell, P. (1989). Circular reasoning about interpersonal behavior: Evidence concerning some untested assumptions underlying diagnostic classification. *Journal of Personality and Social Psychology, 56,* 296–305. http://dx.doi.org/10.1037/0022-3514.56.2.296

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17,* 592–598. http://dx.doi.org/10.1111/j.1467-9280.2006.01750.x

Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology, 67,* 222–232. http://dx.doi.org/10.1037/0022-3514.67.2.222

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin, 24,* 1251–1263. http://dx.doi.org/10.1177/01461672982412001

Wolffhechel, K., Fagertun, J., Jacobsen, U. P., Majewski, W., Hemmingsen, A. S., Larsen, C. L., . . . Jarmer, H. (2014). Interpretation of appearance: The effect of facial features on first impressions and personality. *PLoS ONE, 9,* e107721. http://dx.doi.org/10.1371/journal.pone.0107721

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review, 1,* 204–223. http://dx.doi.org/10.1207/s15327957pspr0103_2